Friendship Formation and Smoking Initiation Among Teens

by
Yun(Sean) Zhang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2007

Approved by:

Donna B. Gilleskie, Advisor

David K. Guilkey, Reader

Han Hong, Reader

Sandra Campo, Reader

John S. Akin, Reader

ABSTRACT

YUN(SEAN) ZHANG: Friendship Formation and Smoking Initiation Among Teens
(Under the direction of Donna B. Gilleskie)

In this research, I use a unique data set to examine the effect of peer influence on teen smoking initiation. First, I develop a game theoretic model where friendship network and smoking decisions are modeled as the equilibrium outcome of a Bayesian Nash game. A unique feature of my model is that individuals choose both teens' friends and smoking decisions simultaneously to maximize utility. Second, I develop an empirical strategy that allows me to estimate the structural equations that arises out of the theoretical model. Identification depends on instrumental variables that exogenously shift peer smoking norms through either friendship probabilities or individual smoking probabilities. I apply my estimator to The National Longitudinal Study of Adolescent Health. Estimation results suggest that peer influence is an important determinant of teen smoking. I 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 although peer influence, as a social multiplier, amplifies the cigarette tax deterrent effect on smoking, it primarily promotes smoking.

To my parents, Luanfeng Wang and Bin Zhang.

To my wife, Xiaofang Zhang.

To my son, Ian Chan Zhang, and my daughter, Zoe Chan Zhang.

ACKNOWLEDGMENTS

I would like to thank my advisor, Donna Gilleskie, as well as David Guilkey, Han Hong, Sandra Campo, John Akin and Sergio Parreiras. All errors are my own. I am especially indebted to my wife for her everlasting love, support, and faith in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 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 (e.g., Kandel, 1978; van Roosmalen and McDaniel, 1989, 1992; Crane, 1991; Graham et al., 1991; Brooks-Gunn et al., 1993; Bauman and Ennet, 1994, 1996). 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.[1] Peer effects are a double-edged sword. On one hand, they motivate an agent to smoke when the agent has smoker friends; on the other hand, they discourage an agent from smoking when the agent has nonsmoker friends. This implies that both the direction and the magnitude of smoking peer effects on an agent depend on the agent's peer smoking norm.[2] We further note that an agent's peer smoking norm is endogenous because an agent chooses her friends. Thus, evaluating the overall impact of peer effects on teen smoking entails understanding teen friendship. So far, little is known about friendship formation among teens and, consequently, whether peer effects promote or contain teen smoking is still unclear.

The presence of a cigarette tax further complicates peer effects on smoking. Peer effects lead to a so called "social multiplier effect" (Sheinkman, 2006; Hoxby, 2000; Epple and

[1] 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.

[2] 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 governing friendship outcomes.

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 on teen smoking. Given this research context, this paper investigates two research questions. First, do peer effects on teen smoking initiation exist? 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, I model high school students' smoking decisions under peer effects in 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. Different from single-agent utility maximization problems, in a game an agent's actions on all outcomes of interest (say, smoking and peer selection) 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 economist) heterogeneity leading to an identification concern.

To address the concern over identification, my 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 endogenous peer smoking norms and exogenous characteristics affect an agents' latent index of smoking linearly and additively in a behavioral specification, the corresponding reduced-form representation of endogenous actions is complicated.[3] As such, an agent's peer smoking norm is a complicated function of exogenous characteristics. This suggests that the economist wants to instrument endogenous peer norms through flexibly-specified 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, I instrument endogenous peer smoking norms in three steps sequentially. First, I instrument individual equilibrium smoking probabilities using a nonparametric bagged tree classifier. Second, I instrument friendship through a flexibly specified logit model. Last, I recover instrumented peer smoking norms using the instrumented individual smoking probabilities and the instrumented friendship probabilities. Both parental characteristics and school sizes are used as instrumental variables and both of them could be correlated with school-level unobserved heterogeneity. Therefore, in estimating the behavioral model in the second stage, I further control for school-level unobserved heterogeneity through school fixed effects.[4]

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 me 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).

Estimation results indicate that peer effects on smoking initiation are significant and homogenous among teens in different grades. In all grades (7 to 12), a one percent increase in the peer smoking norm causes about the same amount of increase in the probability of smoking initiation. Interestingly, although the observed smoking rate among white teens is significantly higher than that among black teens, after controlling for peer influence, smoking initiation

[3]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).

[4]Parents choose schools for their kids, 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.

rates among black teens are found to be 7 percentage points higher than that among white teens. 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 gap. This finding provides another potential explanation for the racial smoking rate gap puzzle (US Dept of Health and Human Services, 1998).

Without controlling for peer influence, the marginal tax deterrent effect on the smoking initiation rate is -3.86 percentage points per 10 cents with a standard error of 1.81. After controlling for peer effects, the state cigarette tax deterrent effect falls slightly to -3.68 percentage points per 10 cents with a larger standard error of 2.29 percentage points. Such a drop in tax effect indicates that social interaction amplifies the tax deterrent effect on smoking initiation in the field. This study compares different specifications that attempt to control for peer influence and 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 completely 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, I 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 reflecting 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, it promotes teen smoking initiation more severely. Combined, peer influence is a significant promoting factor for teen smoking initiation.

This study makes several contributions to the literature. First, it provides a theoretical framework to endogenize both the friendship decision and other decisions (e.g., smoking). Previous theoretical models (Manski, 1995; Brock and Durlauf, 2001; Bajari et al., 2006) begin by considering a reference group of people (say N people) in which an agent's utility on an action is affected by *all* the other N-1 agents' actions.[5] 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, my model explains not only strategic decision on actions other than peer selection among members in a reference group but also how agents choose friends strategically to form the reference group.

In this paper, the reader can gain insights 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 the reader exactly which elements are absorbed into the error term. Furthermore, the reader will see why in the first stage of the estimation of the game, the economist is willing to trade model interpretability for flexibility.

Policy simulations in the peer effect studies (e.g. Norton, 1998; Krauth, 2006; Lundborg, 2006; Cooley, 2007) only allow agents to choose actions given their friendships. This practice is flawed because it ignores that fact that a rational agent should update her friendships while choosing other actions in order to maximize her expected utility. This study contributes to the literature by conducting policy simulations that allow agents to choose friends while also choosing other actions (e.g., smoking).

[5] 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.

Chapter 2

Background

Traditionally, economists have focused on the effects of price, income, addiction, and various regulations on cigarette consumption. In studying those effects, economists 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 utility (i.e., peer influence). In a widely cited work on social interaction, Manski (1995) argues that an agent's decision may affect other agents' decisions through three different interactions: preferences, constraints, and expectations. In teen smoking, 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. Some economists have noticed the potential existence of peer influence. Lewitt et al. (1981) argued that peer effects may increase the magnitude of price elasticity in cigarette consumption. Becker (1992) showed that addiction and peer effects both would increase the magnitude of the price elasticity.

In the peer effects literature, empirical researchers have proposed various strategies to correct for endogeneity bias in 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.[1] Such a data requirement is too demanding to

[1] To illustrate this point, let us consider a smoking game with only $N = 10$ agents. To simplify the problem,

be practical often times. So far, few studies have modeled peer effects in a game. Bajari et al. (2006) model peer effects on stock recommendation. The nature of the financial market determines that major stock appraising firms do not choose their peers. Therefore, peer selection is not a concern in that study. Krauth (2006) models teens' smoking decisions in a pure strategy game of complete information. Krauth (2006) 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.[2]

Some studies use exogenous peer arrangements (e.g., random assignment of roommates) in 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. Expectedly, when agents assigned to the same experimental group are quite homogenous and exchanging friendship signals across

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 decisions and 2 possible smoking decisions, so the size of the entire joint decision space is $[2 \times (2^9)]^{10}$. This expression is a number greater than the total population on the earth. Therefore, even if the economist 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 have the peer selection decision recorded; what Add Health records is realizations of equilibrium friendship probabilities.

[2] 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 economist). 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. 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, arguably using instrumental variables to purge off correlation between the peer norm and unobserved heterogeneity is still necessary.

groups is costly, then peer 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 within 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 exploits instrumental variables that directly affect peers' outcome of interest but not the individual's outcome to purge off the correlation between one's peer norm and her unobserved heterogeneity (Evans, Oates and Schwab, 1992; Gaviria and Raphael, 2001; Hoxby, 2000; Ioannides and Zabel, 2002). This strategy is not thorough because 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 is still a variable that is conditional on endogenous friendships. Realizing this shortcoming, I instrument an agent's friendships as well as the agent's peers' smoking decisions to purge off confounding due to peer section.

Estimating peer effects entails understanding how agents make friends with each other. Therefore, I first present a theoretical model that endogenous outcomes of interest and friendship formation (Chapter 3). In Section 4, I describe the empirical model that accounts for endogenous peer norms (both friendship selection and peer smoking behavior) in the decision to initiate smoking. The unique data are described in Chapter 5. Chapter 6 discusses estimation results. Section 7 concludes with a simulated policy change.

Chapter 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.[1] The decision process in the game is the following. At the beginning of a period, each agent first 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 shock 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, equilibrium friendship probabilities among agents are completely governed by exogenous public information and a common prior.[2]

3.1 Peer Selection and Friendship Production

Consider a self-containing group of players $N_t = \{1, 2, ..., n_t\}$ in period t ($t = 1, 2, ..., T$). N_t is self-containing in the sense that every player in N_t only makes friends with a subset of

[1] Although my major 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) and multiple exogenous dimensions (e.g., family income, gender).

[2] Though the common prior is known to every player in the game, I do not include it as a piece of public information consistently.

N_t. I use subscript '$-i$' to represent all players in N_t except player i. I assume that peers are equally important. Players in N_t are characterized by a $n_t \times K$ matrix z_t recording exogenous characteristics and predetermined actions. At the decision moment in period t, z_t is known to all members in N_t.

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

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

$$s_{i,t} = [s^1_{i,t} \dots s^j_{i,t} \dots s^{n_t}_{i,t}]'$$

. If player i decides to send a friendship signal to player $j \in N_t \backslash \{i\}$ then $s^j_{i,t} = 1$; otherwise $s^j_{i,t} = 0$. The cost associated with peer selection action $s_{i,t}$ is reflected in player i's budget constraint. 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 $s'_{1,t}$, $s'_{2,t}$,...,$s'_{i,t}$,...,$s'_{n_t,t}$, 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,

$$s_t = \begin{bmatrix} s'_{1,t} \\ . \\ s'_{i,t} \\ . \\ s'_{n_t,t} \end{bmatrix} = \begin{bmatrix} 0 & \dots & s^j_{1,t} & \dots & s^{n_t}_{1,t} \\ . & \dots & . & \dots & . \\ s^1_{i,t} & \dots & 0 & \dots & s^{n_t}_{i,t} \\ . & \dots & . & \dots & . \\ s^1_{n_t,t} & \dots & s^j_{n_t,t} & \dots & 0 \end{bmatrix} \tag{3.1}$$

A player does not send a friendship signal to herself; therefore, $s^{i}_{i,t} = 0 \; \forall i \in N_t$.

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 ($\zeta^{j}_{i,t}$) is either 0 (agent i regards agent j as a peer) or 1 (agent i does not regard agent j as a peer). Since one does not consider herself as a peer, $\zeta^{i}_{i,t} = 0 \; \forall i \in N_t$. Because friendship can be asymmetric, $\zeta^{j}_{i,t}$ may not equal to $\zeta^{i}_{j,t}$. More specifically,

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

A generic element of matrix ζ_t, $\zeta^{j}_{i,t}$, is the output of the friendship network production function $\zeta(.)$. That is,

$$\zeta^{j}_{i,t} = \zeta(s^{j}_{i,t}, s^{i}_{j,t}; z_{i,t}, z_{j,t}) = 1(\zeta^{*}(s^{j}_{i,t}, s^{i}_{j,t}; z_{i,t}, z_{j,t}) \geq \zeta^{*}_{\mathbf{c}}(z_{i,t}, z_{j,t})) \tag{3.2}$$

where $\zeta^{*}(s^{j}_{i,t}, s^{i}_{j,t}; z_{i,t}, z_{j,t})$ is the latent index governing whether agent i regards agent j as a peer. $\zeta^{*}_{\mathbf{c}}(z_{i,t}, z_{j,t})$ is the friendship cutoff value that varies as characteristics of agent i and agent j vary.

Let us assume that friendship between two agents is affected only by their friendship signals to each other and their exogenous characteristics. Then, $z_{l,t} \; \forall l \in N_t \backslash \{i, j\}$ and any element in s_t other than $s^{j}_{i,t}$ and $s^{i}_{j,t}$ are not inputs of $\zeta(s^{j}_{i,t}, s^{i}_{j,t}; z_{i,t}, z_{j,t})$ in equation 3.2.

Because friendship can be asymmetric, the sequencing of input arguments in function $\zeta^{*}(.)$, $\zeta^{*}_{\mathbf{c}}(.)$, and $\zeta(.)$ matters. In general,

$$\zeta^{*}_{\mathbf{c}}(z_{i,t}, z_{j,t}) \neq \zeta^{*}_{\mathbf{c}}(z_{j,t}, z_{i,t}) \tag{3.3}$$

$$\zeta^{*}(s^{j}_{i,t}, s^{i}_{j,t}; z_{i,t}, z_{j,t}) \neq \zeta^{*}(s^{i}_{j,t}, s^{j}_{i,t}; z_{j,t}, z_{i,t}) \tag{3.4}$$

$$\zeta(s^j_{i,t}, s^i_{j,t}; z_{i,t}, z_{j,t}) \neq \zeta(s^i_{j,t}, s^j_{i,t}; z_{j,t}, z_{i,t}) \tag{3.5}$$

I make the following three assumptions, regarding the friendship network production function $\zeta(.)$.

Assumption 1: $\forall s^i_{j,t}, \forall z_{i,t}$ and $\forall z_{j,t}$, if $s^j_{i,t} = 0$ then $\zeta^j_{i,t} = 0$.

This assumption states that if agent i does not send a friendship signal to agent j then for sure agent i does not regard 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 $\zeta^j_{i,t} = 1$ then $s^i_{j,t} = 1$.

This assumption states that 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 even spend some time or other resources on the former.

Assumption 3: $\forall s^j_{i,t}, \forall z_{i,t}$ and $\forall z_{j,t}$, $\zeta^*(s^j_{i,t}; s^i_{j,t} = 1,\ z_{i,t}, z_{j,t}) \geq \zeta^*(s^j_{i,t}; s^i_{j,t} = 0; z_{i,t}, z_{j,t})$.

This assumption states that given agent i's peer selection decision and agent i and agent j's exogenous characteristics, if agent j sends a friendship signal to agent i, the latent index value of 'agent i nominates agent j as a friend' increases.

Assumption 1 and Assumption 2 jointly imply the following:

$$\zeta^j_{i,t} = 1 \Rightarrow s^j_{i,t} = 1 \text{ and } s^i_{j,t} = 1.$$

This says, consider any two generic agents, as long as one of them regards the other as a peer then these two agents have exchanged friendship signals.

The number of agents regarded by agent i as friends in period t is $\sum_{j \in N_t \setminus \{i\}} \zeta^j_{i,t}$.

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 below:

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

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 categories. For example, if A_i is the smoking action set and the smoking action is binary, then $A_i = \{0, 1\}$ and $J_1 = 2$.

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; i.e., $d_{i,t} \in D_i$. 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 decisions player i can choose from at decision moment t. Since any two different players, say player i and player j, 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 I define $N_D = N_{D_i}\ \forall i \in N_t$.

Let $\overline{a}_{i,t} = [\ \overline{a}_{i,t}^1 \quad \overline{a}_{i,t}^2 \quad ... \quad \overline{a}_{i,t}^m \quad ... \quad \overline{a}_{i,t}^M\ ]'$ denote player i's $M \times 1$ peer norm vector in period t. We note player i's peer norm is determined by both her peer selection decision ($s_{i,t}$) and other $n_t - 1$ players' M decisions other than peer selection. An element of $\overline{a}_{i,t}$ can be expressed as below:

$$\overline{a}_{i,t}^m = \chi(\zeta(s_t, z_t), a_{-i,t}^m) = \frac{\sum\limits_{j \in N_t \setminus \{i\}} \zeta_{i,t}^j a_{j,t}^m}{\sum\limits_{j \in N_t \setminus \{i\}} \zeta_{i,t}^j} \tag{3.6}$$

Next, I lay out the main elements of this game.

3.3 Preference

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

$$\max_{d_{i,t}} \sum_{m=1}^{M} u_m(a_{i,t}^m; z_{i,t}) + 1[\sum_{j \in N_t \setminus \{i\}} \zeta_{i,t}^j > 0] \sum_{m=1}^{M} u_m^p(|a_{i,t}^m - \overline{a}_{i,t}^m|; z_{i,t}) \tag{3.7}$$

$$s.t.\ C(d_{i,t}; d_{-i,t}, z_{i,t}) - y_{i,t} \leq 0$$

where $u_m(a_{i,t}^m; z_{i,t})$ captures the payoff associated with the m^{th} action $(m = 1, ..., M)$ other than peer selection. If player i regards at least one person as a friend, then she receives an additional component of utility. $u_m^p(|a_{i,t}^m - \overline{a}_{i,t}^m|; z_{i,t})$ is the peer influence utility derived from the m^{th} action $(m = 1, ..., M)$ other than peer selection given that player i regards one or more people as friend(s). $|a_{i,t}^m - \overline{a}_{i,t}^m|$ is player i's deviation from the m^{th} peer norm. Player i may maximize her payoff by choosing to not regard anyone as a friend. If so, then $1[\sum_{j \in N_t \setminus \{i\}} \zeta_{i,t}^j > 0] = 0$. Price variables are contained in $z_{i,t}$. $y_{i,t}$ is per-period income. ($y_{i,t}$ is also an element of $z_{i,t}$.) $C(d_{i,t}; d_{-i,t}, z_t)$ 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 (1995) point that social interaction could work through not only preferences (utility function) but also the budget constraint. For example, suppose Jack dislikes Tom. 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.

3.4 Information Structure

At the beginning of period t, n_t players collectively receive a (action-dependent) $n_t \times N_D$ private shock (private information) matrix ε_t. That is,

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{i,t}^{'} \\ . \\ \varepsilon_{i,t}^{'} \\ . \\ \varepsilon_{n_t,t}^{'} \end{bmatrix} = \begin{bmatrix} \varepsilon_{i,t}^{1} & .... & \varepsilon_{i,t}^{h} & .... & \varepsilon_{i,t}^{N_D} \\ . & ... & . & ... & . \\ \varepsilon_{i,t}^{1} & .... & \varepsilon_{i,t}^{h} & .... & \varepsilon_{i,t}^{N_D} \\ . & ... & . & ... & ... \\ \varepsilon_{n_t,t}^{1} & .... & \varepsilon_{n_t,t}^{h} & .... & \varepsilon_{n_t,t}^{N_D} \end{bmatrix}_{n_t \times N_D}$$

$$h = 1, 2, ..., N_D$$

where $\varepsilon_{i,t}^{h}$ is player i's private shock if she chooses the h^{th} decision from her decision choice set D_i. Column vector $\varepsilon_{i,t} = [\ \varepsilon_{i,t}^{1} \ \ .... \ \ \varepsilon_{i,t}^{h} \ \ .... \ \ \varepsilon_{i,t}^{N_D}\ ]'$ 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, $f(\varepsilon_{i,t})$, 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 are consistent with common prior; following Bayes' rule, we have

$$f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) = \frac{f(\varepsilon_{i,t}, \varepsilon_{-i,t})}{\displaystyle\int f(\varepsilon_{i,t}, \varepsilon_{-i,t}) d\varepsilon_{-i,t}} \ \forall i \in N_t \tag{3.8}$$

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.

3.5 Allowed Decision Space

Social interactions among agents' budget constraints implies 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, I 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. I refer to such a subspace as an Allowed Decision Space (ADS). I characterize an ADS in Appendix 1.

Hereafter, I abuse notation and use $D = \prod_{i\in N_t} D_i$ to denote an ADS to ease the notational burdern.

Consider an ADS, $D = \prod_{i\in N_t} D_i$, 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 3.7, we obtain action-specific payoff functions as below:

$$V(d_{i,t}, d_{-i,t}, \varepsilon_{i,t}, z_t, u_m(.), u_m^p(.)) = \sum_{m=1}^{M} V_m(d_{i,t}; z_t\ d_{-i,t}) + \varepsilon_{i,t}^{d_{i,t}} \quad \forall d_t \in D \tag{3.9}$$

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

$$I_t = \{z_t, V(.),\ \zeta(.)\} \in \vartheta_{pub}.$$

3.6 Bayes Game and Pure Strategy Bayes Nash Equilibrium (PSBNE)

At decision moment t, the following fundamentals govern the n_t players' decisions:

- Players: $N_t = \{1, 2, ..i.., n_t\}$

- Public information: $I_t = \{z_t, V(.), \zeta(.)\} \in \vartheta_{pub}$
- Private shock set (player type space) : $E = \prod_{i \in N_t} E_i$
- Priors: $f_{-i|i} = (f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}))_{i \in N_t}$
- an ADS: D

In essence, n_t players make decisions in the following Bayes game (G_t) where

$$G_t = \left\langle N_t,\ E,\ f_{-i|i},\ D,\ I_t \right\rangle .$$

A player's strategy profile $(d_{i,t}(I_t, \varepsilon_{i,t}))$ is a correspondence between (both public and private) information and the individual decision

$$d_{i,t}(I_t, \varepsilon_{i,t}) : \vartheta_{pub} \times E_i \longmapsto D_i$$

At a Nash equilibrium, no player has an incentive to alter her strategy as long as other players do not.

Therefore, a PSBNE is a collection of n_t mappings from the information set to the decision set, i.e.,

$$(d^*_{i,t}(I_t, \varepsilon_{i,t}))_{i \in N_t} \text{ where } d^*_{i,t}(.) : \vartheta_{pub} \times E_i \longmapsto D_i$$

such that

$$d^*_{i,t}(I_t, \varepsilon_{i,t}) \in \arg\max_{d_{i,t} \in D_i} \int_{\varepsilon_{-i,t} \in E_{-i}} V(d_{i,t}, (d^*_{j,t})_{j \in N_t \setminus \{i\}}) f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) d\varepsilon_{-i,t} \tag{3.10}$$

$\forall i \in N_t$ and $\forall \{I_t, \varepsilon_{i,t}\} \in \vartheta_{pub} \times E_i$

where

$D_i = \bigcup_{\varepsilon_{i,t} \in E_i} d^*_{i,t}(I_t, \varepsilon_{i,t})$,

$f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t})$ is individual priors consistent with a common prior restricted to E_i, and

$f_{-i|i}(\varepsilon_{-i,t}|\varepsilon_{i,t}) = \frac{f(\varepsilon_{i,t},\varepsilon_{-i,t})}{\int f(\varepsilon_{i,t},\varepsilon_{-i,t})d\varepsilon_{-i,t}}$ and $f(.)$ is the joint density function of private shocks restricted to E,

i.e., $f(\varepsilon_t) = \frac{f(\varepsilon_t)}{\int 1(\varepsilon_t \in E) f(\varepsilon_t) d\varepsilon_t}$.

Let us define $d^*(.) : \vartheta_{pub} \times E \longmapsto D$

$$d^*((I_t, \varepsilon_{i,t})_{i \in N_t}) \equiv (d^*_{i,t}(I_t, \varepsilon_{i,t}))_{i \in N_t}$$

By stacking equation 3.10 over all players in N_t, the right hand side of equation 3.10 can be written as a function $\Psi(.) : D^* \mapsto D^*$. We get

$$d^*((I_t, \varepsilon_{i,t})_{i \in N_t}) = \Psi(d^*((I_t, \varepsilon_{i,t})_{i \in N_t});\ I_t,\ N_t,\ f_{-i|i})\ \forall \{I_t, \varepsilon_{i,t}\} \in \vartheta_{pub} \times E_i \tag{3.11}$$

According to equation 3.11, a PSBNE $d^*((I_t, \varepsilon_{i,t})_{i \in N_t})$ 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, I 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 $\pi_i^*(.) : D_i^* \longmapsto \pi_i^*$ denote the correspondence between equilibrium decision elements and their probabilities. Then

$$\pi_i^*(d^*) = \int\limits_{\varepsilon_{i,t}} 1(d^*_{i,t}(I_t, \varepsilon_{i,t}) = d^*) f_{\varepsilon_{i,t}}(\varepsilon_{i,t}) d\varepsilon_{i,t} \quad \forall i \in N_t, \forall d^* \in D_i$$

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

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

and 1[.] is an indicator function where

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

We note that the following equation holds:

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

Chapter 4

Empirical Strategy

In the presence of peer effects, schoolmates are playing a game. This implies that the ideal approach is to estimate all schoolmates' actions on peer selection and smoking jointly. Doing so means that the economist should regard a cross section of peer selection actions and smoking actions made by all schoolmates as a single observation. The analysis sample (sample 1 in Table 5.1) only contains 124 schools, therefore, I could not implement this ideal approach.[1] Another data limitation is that I do not observe students' peer selection actions in Add Health; I only observe the friendship outcomes. Given these data limitations, I have chosen to estimate smoking peer effects in a single agent style in the sense that I estimate agents' smoking actions agent by agent rather than jointly across all agents in a reference group (i.e., a school in this study). In addition, my empirical strategy should avoid estimating peer selection actions and smoking actions jointly within an agent because I do not observe peer selection actions.

In what follows, I first derive an (structural form) empirical specification that is consistent with the theoretical model and is implementable under the data limitations mentioned above. In addition, I discuss identification concerns in the empirical specification. I also explain how to instrument endogenous peer smoking norms using two separate reduced-form analyses in sequence. Next, I explain how to estimate the structural form empirical specification in two stages. The first stage involves instrumenting the peer norm flexibly through three flexible (reduced-form) steps sequentially to control for endogeneity bias caused by both peer selection

[1] Implementing this ideal approach means that I only have 124 observations.

and smoking decision simultaneity. The second stage estimates the structural form empirical specification for smoking initiation using the instrumented peer norm.

4.1 Derivation of Empirical Specification and Identification

I assume players only choose a smoking action and a peer selection action in the smoking game. Hence, a generic player's decision is $d_i = [a_i \ s_i]'$ where a_i is the smoking action and s_i is the peer selection action. At a PSBNE, a generic player i expects that others players adopt the equilibrium strategy $d^*_{-i}(.)$. Therefore, I parameterize player i's decision-specific payoff (equation 3.9) as follows:

$$\begin{aligned}
&V([a_i \ s_i]', d^*_{-i}(\varepsilon_{-i}, I), \varepsilon_i, z, u_{smoke}(.)) \\
&= u_{smoke}(d_i; z, d^*_{-i}) + \varepsilon_i^{d_i} \\
&= z\beta(a_i; I) - \delta a_i - \gamma \left| a_i - \overline{a}_{-i}(s_i; z, d^*_{-i}) \right| + \varepsilon_i^{d_i}
\end{aligned} \tag{4.1}$$

where

(1) $d^*_{-i}(\varepsilon_{-i}, I)$ is other players' equilibrium decisions.

(2) z is all players' exogenous characteristics in the game; $z = \begin{bmatrix} z'_i \\ z'_{-i} \end{bmatrix}$

(3) δ is the biological smoking disutility;

(4) $\beta(a_i; I)$ is the parameter vector corresponding to z. According to equation 3.9, $\beta(a_i; I)$ is smoking–action-dependent;

(5) $\gamma \geqslant 0$ is the smoking peer effects parameter. The larger the γ, the heavier the punishment for deviation from peer smoking norm;

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

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

I do not observe students' peer selection actions (s_i) in Add Health. To proceed with estimation, I first derive player i's *ex post* expected smoking-action-specific payoff based on equation 4.1 below.[2]

Player i's *ex post* expected smoking-action-specific payoff is

$$\begin{aligned} & EV(\varepsilon_i, I, F(.); a_i) \\ & = \underset{s_i}{Max} \left[z\beta(a_i; I) - \delta a_i - \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \overline{a}_{-i}(s_i; z, d^*_{-i}) dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{d_i} \right] \\ & = z\beta(a_i; I) - \delta a_i - \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \overline{a}_{-i}(s_i; z, d^*_{-i}) dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{[a_i \ s^*_i(a_i)]'} \end{aligned} \tag{4.2}$$

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

Substituting equation 3.6 and equation 4.1 into equation 4.2 yields

$$\begin{aligned} & EV(\varepsilon_i, I, F(.); a_i) \\ & = z\beta(a_i; I) - \delta a_i - \gamma \left| a_i - \int\limits_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum\limits_{j \in N_t \setminus \{i\}} \zeta^j_i(s^{j*}_i(a_i), s^{i*}_j, z_i, z_j) a^*_j(\varepsilon_j, I)}{\sum\limits_{j \in N_t \setminus \{i\}} \zeta^j_i(s^{j*}_i(a_i), s^{i*}_j, z_i, z_j) a^*_j(\varepsilon_j, I)} dF(\varepsilon_{-i}|\varepsilon_i) \right| + \varepsilon_i^{[a_i \ s^*_i(a_i)]'}. \end{aligned} \tag{4.3}$$

Defining the *ex post* expected peer smoking norm as $E(\overline{a}^*_{-i}|\varepsilon_i, I, F(.))$, we have

$$\begin{aligned} E(\overline{a}^*_{-i}|\varepsilon_i, I, F(.)) = \int\limits_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum\limits_{j \in N_s \setminus \{i\}} \zeta(s^*_i(\varepsilon_i, I), s^*_j(\varepsilon_j, I); I) a^*_j(\varepsilon_j, I)}{\sum\limits_{j \in N_s \setminus \{i\}} \zeta(s^*_i(\varepsilon_i, I), s^*_{-i}(\varepsilon_{-i}, I); I)} dF(\varepsilon_{-i}|\varepsilon_i) \\ E(\overline{a}^*_{-i}|\varepsilon_i, I, F(.)) \in [0, 1]. \end{aligned} \tag{4.4}$$

[2] *ex post* means after the realization of private shock (ε_i); *ex ante* means prior to the realization of private shock (ε_i).

Then player i's *ex post* expected smoking-action-specific payoff is

$$EV(\varepsilon_i, I, F(.); a_i) = z\beta(a_i; I) - \delta a_i - \gamma \left| a_i - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.)) \right| + \varepsilon_i^{a_i} \tag{4.5}$$

where, $\left| a_i - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.)) \right| \geqslant 0$ is the deviation from the *ex post* expected peer smoking norm; if player i smokes then $\left| a_i - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.)) \right| = 1 - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.))$; otherwise, $\left| a_i - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.)) \right| = E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.))$.

We note even if the number of players in the game is modestly large (say 50 players), estimating parameters related to exogenous terms containing z_{-i} is hard because there is a large number of different combinations 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, I focus on the component of $z\beta(a_i; I)$, which is purely related to z_i. To do so, I assume that $z\beta(a_i; I)$ can be decomposed as

$$z\beta(a_i; I) = z_i \beta_{self}(a_i; I) + f(z_i, z_{-i}) \beta_{interact}(a_i; I) \tag{4.6}$$

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 some or all elements in z_{-i}; $\beta_{interact}(a_i; I)$ is the parameter vector corresponding to z_{-i}. Let us define

$$\upsilon_i^{a_i}(I) \equiv f(z_i, z_{-i}) \beta_{interact}(a_i; I). \tag{4.7}$$

We note that $\upsilon_i^{a_i}(I)$ is smoking-action-specific because $\beta_{interact}(a_i; I)$ is smoking-action-specific. Substituting equation 4.6 and equation 4.7 into equation 4.5, we get the smoking-action-specific *ex post* expected payoff as below

$$EV(\varepsilon_i, I, F(.); a_i) = z_i \beta_{self}(a_i; I) - \delta a_i - \gamma \left| a_i - E(\overline{a}^*_{-i} | \varepsilon_i, I, F(.)) \right| + \varepsilon_i^{a_i} + \upsilon_i^{a_i}(I) \tag{4.8}$$

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

$$\Delta EV(\varepsilon_i, I, F(.)) \equiv EV(\varepsilon_i, I, F(.); a_i = 1) - EV(\varepsilon_i, I, F(.); a_i = 0) \tag{4.9}$$

A player's smoking decision is

$$a^*_{1,i} = 1(\Delta EV(\varepsilon_i, I, F(.)) > 0).$$

Substituting equation 4.8 into equation 4.9, we have

$$\begin{aligned} &\Delta EV(\varepsilon_i, I, F(.)) = -\gamma - \delta + z_i[\beta_{self}(a_i = 1; I) - \beta_{self}(a_i = 0; I)]+ \\ &2\gamma E(\overline{a}^*_{-i}|\varepsilon_i I, F(.)) + \varepsilon^1_i - \varepsilon^0_i + \upsilon^1_i(I) - \upsilon^0_i(I). \end{aligned} \tag{4.10}$$

The constant term "$-\gamma - \delta$" in the equation above is collinear with the intercept; therefore, the biological smoking disutility parameter (δ) cannot be identified separately from the intercept. Without loss of generality, I can rewrite the equation above as

$$\begin{aligned} &\Delta EV(\varepsilon_i, I, F(.)) = z_i[\beta_{self}(a_i = 1; I) - \beta_{self}(a_i = 0; I)]+ \\ &2\gamma E(\overline{a}^*_{-i}|\varepsilon_i, I, F(.)) + \varepsilon^1_i - \varepsilon^0_i + \upsilon^1_i(I) - \upsilon^0_i(I). \end{aligned} \tag{4.11}$$

The economist does not have full observability on public information (I). Let us decompose $I = [I^o \ I^u]$; z_i is a component of I, so, correspondingly, $z_i = [z^o \ z^u]$; superscript "o" means observed by the economist and superscript "u" means unobserved by the economist. In the game, common prior $F(.)$, I^u, and private shocks are unobserved by the economist. Define

$$\begin{aligned} \Delta\beta_{self} &\equiv \beta_{self}(a_i = 1; I) - \beta_{self}(a_i = 0; I) \\ \Delta\varepsilon_i &\equiv \varepsilon^1_i - \varepsilon^0_i \\ \Delta\upsilon_i(I) &\equiv \upsilon^1_i(I) - \upsilon^0_i(I) \end{aligned} \tag{4.12}$$

Assume that $z_i\Delta\beta_{self}$ can be decomposed as

$$z_i\Delta\beta_{self} = z^o_i\Delta\beta^o_{self} + z^u_i\Delta\beta^u_{self}.$$

Then $\Delta EV(\varepsilon_i, I, F(.))$ can be rewritten into

$$\Delta EV(\varepsilon_i, I, F(.)) = z_i^o \Delta\beta_{self}^o + 2\gamma E(\overline{a}_{-i}^* | \varepsilon_i, I, F(.)) + \Delta\varepsilon_i + \Delta\upsilon_i(I) + z_i^u \Delta\beta_{self}^u. \qquad (4.13)$$

The economist does not know private shock ε_i, therefore, she cannot use the private-shock-specific *ex post* expected equilibrium peer norm (i.e., $E(\overline{a}_{-i}^* | \varepsilon_i, I, F(.))$) in equation 4.13) as a predictor of smoking initiation. In order to proceed with estimation, she can, alternatively, use the *ex ante* equilibrium peer norm to proxy $E(\overline{a}_{-i}^* | \varepsilon_i, I, F(.))$ because she can recover the *ex ante* equilibrium peer norm from the data. Player i's *ex ante* equilibrium peer norm is

$$E(\overline{a}_{-i}^* | I, F(.)) = \int E(\overline{a}_{-i}^* | \varepsilon_i, I, F(.)) dF_{\varepsilon_i}(\varepsilon_i). \qquad (4.14)$$

Define

$$\tau(\varepsilon_i, I, F(.)) = 2\gamma [E(\overline{a}_{-i}^* | \varepsilon_i, I, F(.)) - E(\overline{a}_{-i}^* | I, F(.))]. \qquad (4.15)$$

Substituting equation 4.15 into equation 4.13, we get

$$\begin{aligned} \Delta EV(\varepsilon_i, I, F(.)) &= z_i^o \Delta\beta_{self}^o + 2\gamma E(\overline{a}_{-i}^* | I, F(.)) \\ &+ \Delta\varepsilon_i + \Delta\upsilon_i(I) + z_i^u \Delta\beta_{self}^u + \tau(\varepsilon_i, I, F(.)). \end{aligned} \qquad (4.16)$$

An identification concern regarding the peer effect (γ) arises if correlation exists between $E(\overline{a}_{-i}^* | I, F(.))$ and either of the four unobserved components: $\Delta\varepsilon_i$, $\Delta\upsilon_i(I)$, $z_i^u \Delta\beta_{self}^u$, and $\tau(\varepsilon_i, I, F(.))$. 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 due to two reasons. First, $E(\Delta\upsilon_i(I), \Delta\upsilon_j(I)) \neq 0$ and $E(\tau(\varepsilon_i, I, F(.)), \tau(\varepsilon_j, I, F(.))) \neq 0$ in general; second, if private shocks are correlated then $E(\Delta\varepsilon_i, \Delta\varepsilon_j) \neq 0$.

(2) correlation between $E(\overline{a}_{-i}^* | I, F(.))$ and $\Delta\upsilon_i(I)$ is expected because both of them are functions of public information (I);

(3) correlation between $E(\overline{a}_{-i}^* | I, F(.))$ and $\tau(\varepsilon_i, I, F(.))$ is expected because both of them are functions of public information (I) and the common prior ($F(.)$).

My interest is to examine peer effects on smoking among schoolmates; therefore, it is reasonable to assume that school-level unobservables account for the correlation of unobserved heterogeneity across students attending the same school. As such, I decompose a student's unobservables into two components; one is school-level unobservables and the other is individual unobservables. The final empirical specification is:

$$\begin{aligned} \Delta EV_{i,s} &= X_{i,s}\alpha + \gamma \widehat{E}(\overline{a}^*_{-i,s}; F(.), I) + \mu_s + \varepsilon_{i,s} \\ a^*_{i,s} &= 1(\Delta EV_{i,s} > 0);\ E(\overline{a}^*_{-i,s}) \in [0,1] \end{aligned} \tag{4.17}$$

where

(1) s indexes schools;

(2) i indexes students; $i \in N_s = \{1, 2, ..., n_s\}$

(3) $X_{i,s}$ is student i's observed exogenous characteristics;

(4) α is the parameter vector corresponding to $X_{i,s}$;

(5) γ is the peer effect parameter;[3]

$E(\overline{a}^*_{-i,s}; F(.), I)$ is student i's *ex ante* equilibrium peer smoking norm; it is a function of public information and the function form of the common prior;

μ_s is the school-level unobserved heterogeneity; this term allows for students' unobservables to be correlated at school level, $\mu_s = \Delta v_i(I)$;

$\varepsilon_{i,s}$ is the individual unobserved heterogeneity;

$a^*_{i,s}$ is observed equilibrium smoking action;

$E(\mu_s, \mu_{s'}) = 0$, $E(\varepsilon_{i,s}, \varepsilon_{j,s}) = 0$, $E(\varepsilon_{i,s}, \varepsilon_{j,s'}) = 0$, and $E(\mu_s, \varepsilon_{i,s}) = 0$.

4.2 Estimation

In this subsection, I discuss estimation of the empirical specification in equation 4.17. The identification concern is that $E(\overline{a}^*_{-i,s}; F(.), I)$ may be correlated with (school-level and/or individual) unobserved heterogeneity. To address this concern, in the first stage of estimation,

[3]Compared with equation 4.16, I drop "2" in front of γ for simplicity. This simplification rescales the peer effect estimate and is innocuous.

I predict the *ex ante* equilibrium peer norm ($E(\overline{a}^*_{-i,s}; F(.), I)$) using instrumental variables that affect an agent's peer smoking norm and are uncorrelated with the agent's unobserved heterogeneity ($\varepsilon_{i,s}$). I also use school fixed effects to futher control for the potential correlation between the instrumented peer norm and school-level unobserved heterogeneity (μ_s) in the second stage of estimation.

Before further discussion, let us first examine the functional form of the *ex ante* equilibrium peer smoking norm ($E(\overline{a}^*_{-i,s}; F(.), I)$). Substituting equation 3.6 into equation 4.14 yields the reduced form representation of $E(\overline{a}^*_{-i,s}; F(.), I)$ in equation 4.18 below. Apparently, exogenous inputs (I and $F(.)$) of the game affect the *ex ante* equilibrium peer smoking norm nonlinearly and interactively.[4] 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 peers' smoking

[4]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. (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

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

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. γ is peer effect behavioral parameter.

Solving for $\pi_1$and π_2 in terms of exogenous characteristics, we can get a reduced form representation as

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

where, R is the reduced form representation. In this example, $R(.)$ 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 $R(.)$ nonlinearly and interactively even if they affect the latent outcome in a linear and additive pattern in the behavioral outcome process.

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

$$
\begin{aligned}
E(\overline{a}^*_{-i,s}; F(.), I) &= \int\limits_{\varepsilon_i}\int\limits_{\varepsilon_{-i}|\varepsilon_i} \frac{\sum\limits_{j\in N_s\setminus\{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_j, I); I) a_j^*(\varepsilon_j, I)}{\sum\limits_{j\in N_s\setminus\{i\}} \zeta(s_i^*(\varepsilon_i, I), s_{-i}^*(\varepsilon_{-i}, I); I)} dF(\varepsilon_{-i}|\varepsilon_i) dF(\varepsilon_i) \\
&= \int\limits_{\varepsilon} \frac{\sum\limits_{j\in N_s\setminus\{i\}} \zeta(s_i^*(\varepsilon_i, I), s_j^*(\varepsilon_j, I); I) a_j^*(\varepsilon_j, I)}{\sum\limits_{j\in N_s\setminus\{i\}} \zeta(s_i^*(\varepsilon_i, I), s_{-i}^*(\varepsilon_{-i}, I); I)} dF(\varepsilon) \\
&= \int\limits_{\varepsilon} \frac{\sum\limits_{j\in N_s\setminus\{i\}} \zeta^*(\varepsilon_i, \varepsilon_j, I) a_j^*(\varepsilon_j, I)}{\sum\limits_{j\in N_s\setminus\{i\}} \zeta^*(\varepsilon_i, \varepsilon_j, I)} dF(\varepsilon) \qquad (4.18)
\end{aligned}
$$

The economist is unable to recover the *ex ante* equilibrium peer smoking norm based on the two *ex post* quantities ($\zeta^*(\varepsilon_i, \varepsilon_j, I)$ and $a_j^*(\varepsilon_j, I)$) because she does not observe private shocks. To proceed with estimation, the economist 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 ($p(\zeta_i^{j*}(I; F(.)) = 1)\ \forall j \in N_s\setminus\{i\}$) and her schoolmates' *ex ante* smoking probabilities ($p(a_j^*(I; F(.)) = 1)\ \forall j \in N_s\setminus\{i\}$) because both of them can be recovered from the data. As such, I predict a student's *ex ante* equilibrium peer smoking norm in three sequential steps.[5] In a school (say school s), I first instrument all students' *ex ante* equilibrium smoking probabilities ($p(a_i^*(I; F(.)) = 1)\ \forall i \in N_s$); next, I instrument all students' *ex ante* equilibrium friendship probabilities ($p(\zeta_i^{j*}(I; F(.)) = 1)_{i=1}^{n_s}$ $\forall j \in N_s\setminus\{i\}$ and $\forall i \in N_s$); in the third step, I construct a student's instrumented *ex ante* equilibrium peer smoking norm ($\widehat{E}(\overline{a}^*_{-i,s}; F(.), I)$) from her schoolmates' instrumented *ex ante* equilibrium smoking probabilities ($\widehat{p}(a_j^*(I; F(.)) = 1)\ \forall j \in N_s\setminus\{i\}$) and her instrumented *ex*

[5]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 specify appropriate interaction terms. It turns out that even when I instrument the *ex ante* equilibrium peer norms in three steps, the interaction terms involved are fairly complicated.

ante friendship probabilities ($\widehat{p}(\zeta_i^{j*}(I; F(.)) = 1)\ \forall j \in N_s \backslash \{i\}$) according to equation 4.19:

$$\widehat{E}(\overline{a}^*_{-i,s}; F(.), I) = \frac{\sum\limits_{j \in N_s \backslash \{i\}} \left[\widehat{p}(\zeta_i^{j*}(I; F(.)) = 1) \times \widehat{p}(a_j^*(I; F(.)) = 1)\right]}{\sum\limits_{j \in N_s \backslash \{i\}} \widehat{p}(\zeta_i^{j*}(I; F(.)) = 1)} \tag{4.19}$$

$$\forall j \in N_s \backslash \{i\} \text{ and } \forall i \in N_s$$

where, $\widehat{E}(\overline{a}^*_{-i}; F(.), I)$ is student i's instrumented *ex ante* equilibrium peer smoking norm;

$p(\zeta_i^{j*}(I; F(.)) = 1)$ is the *ex ante* equilibrium probability that student i regards student j as a friend;

$\widehat{p}(\zeta_i^{j*}(I; F(.)) = 1)$ is the instrumented *ex ante* equilibrium probability that student i regards student j as a friend;

$p(a_j^*(I; F(.)) = 1)$ is schoolmate j's *ex ante* equilibrium smoking probability;

$\widehat{p}(a_j^*(I; F(.)) = 1)$ is schoolmate j's instrumented *ex ante* equilibrium smoking probability. To obtain instrumented *ex ante* equilibrium individual smoking probabilities in the first step, I use a bagged tree classifier which is, in essence, an ensemble of 100 fully-grown tree classifiers created from 100 bootstrap samples (Breiman 1996).[6] The bagged tree classifier includes the following exogenous characteristics as explanatory variables: grade level, race, gender, age, parental smoking, the highest education of parents, family income, and state cigarette tax. We note that these variables directly affect a student's schoolmates' *ex ante* smoking probabilities ($p(a_j^*(I; F(.)) = 1)$), and in turn, the student's *ex ante* equilibrium peer smoking norm but should be uncorrelated with the student's individual unobserved heterogeneity (i.e., μ_s in equation 4.17).[7]

To implement the second step, I use a flexibly specified logit model to predict *ex ante* equi-

[6] In this study, a logit model does not perform well when even only a few interaction terms are added. A quasi-perfect separation problem emerges and logit estimation cannot proceed. Interaction terms causing the quasi-perfect separation problem are covariates that explain outcomes so well that in one realization (one wave of Add health) variations in binary smoking actions within categories defined by those interaction terms disappear. Thus, the quasi-perfect separation problem is a drawback with the logit model specification. This concern further motivates me to use the bagged tree classifier.

[7] The predicted smoking probability for a student may still be correlated with school-level unobserved heterogeneity (μ_s). I will come back to this concern soon.

librium friendship probabilities ($p(\zeta_i^{j*}(I; F(.)) = 1)$). Three reasons motivate me to adopt a logit model rather than a bagged tree classifier in this second step. First, I 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 (Bernheim, 1994) provides guidance for specifying a logit model, so this step is less "data mining" oriented than the first step. Third, the large number of pairwise observations makes the computational cost in a bagged tree classifier overly expensive. To be specific, I use student i's directional deviations from a generic schoolmate j in gender, race, age, grade level, family income, parental smoking, and parents' highest education as explanatory variables for student i's friendship with the generic schoolmate j.[8] I also use the instrumented *ex ante* equilibrium individual smoking probabilities from the first step to create directional deviations in smoking. It is worth mentioning that because the instrumented *ex ante* equilibrium individual smoking probabilities from the first step are functions of exogenous characteristics such as whether the students' parents smoke or not, a student's created 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.

After modeling *ex ante* equilibrium individual smoking probabilities and *ex ante* equilibrium friendship probabilities, I then construct the instrumented *ex ante* equilibrium peer smoking norm according to equation 4.19 as the third step.

[8] To illustrate direction deviations, let us consider directional deviation in a categorical variable, say, gender, between two agents, say, 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"; "male-to-female" is not equal to "female-to-male". Friendship between two agents may not be symmetric; this is the reason why I use directional deviations as instrumental variables.

As mentioned above, one's schoolmates' parental characteristics 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, I control for school-level unobserved heterogeneity by including school fixed effects in the second stage of estimation.

Chapter 5

Data

The theoretical model indicates that an agent's friendships directly affect her peer norm. Therefore, to estimate peer effects, the economist should observe an agent's friends in the data.[1] 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 three waves of in-home surveys fielded in 1995, 1996, and 2001. I only use the wave I in-home survey data collected between April 1995 and December 1995 because the other two 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 gender.[2] 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.[3] After deleting respondents with missing values, I obtain

[1]Observing agents' peer selection actions is unnecessary if the economist does not intend to estimate the peer selection action and smoking action jointly. In fact, as explained before, even if the economist 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 economist cannot observe the friendship network in the data, then the economist has to subjectively assign friends to an agent. For example, the economist 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 her classmates.

[2]Due to administrative error, a small fraction (<5%) of students in those schools were asked to nominate one friend of each gender.

[3]When respondents are randomly selected from a school, even if a student is asked to nominate all her

17,844 students (Sample 4 in Table 5.1). Among the 17,844 students, there are 5,774 students (Sample 4.1) who were asked to nominate five schoolmate friends of each gender.[4] To model friendship formation, I 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 i regards j as a friend or not and student i's directional deviations from schoolmate j in terms of their exogenous characteristics that affect their peer selection actions. In the section of empirical strategy, I explained the construction of these pairwise directional deviations and estimation of friendship formation in more detail.

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, I dichotomize the smoking decision. I 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

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 N by 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

$$\begin{aligned} R_{missing} &= \frac{N^2(1 - r^2) - N(1 - r)}{(N^2 - N)} \\ &= \frac{N(1 - r^2) - (1 - r)}{(N - 1)} \end{aligned}$$

As N approaches infinity, we have

$$\begin{aligned} \lim_{N \to +\infty} R_{missing} &= \lim_{N \to +\infty} \frac{N(1 - r^2) - (1 - r)}{(N - 1)} \\ &= 1 - r^2 \end{aligned}$$

[4]Ideally, the economist 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 my knowledge.

affects an agent's current smoking decision through nicotine tolerance and dependence 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 residence, I am 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, I restrict the analysis sample 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. Hence, this selection applies only to the analysis of the smoking initiation decision. The full sample of respondents from each school with complete data (sample 4 of Table 5.1) is used in the analysis of friendship formation and construction of the peer smoking norm. Table 5.1 details derivation of the sample used in estimation.

Add Health provides a rich set of measurements on respondents's characteristics. For my purposes, the following measurements are of particular interest due to their potential impact on smoking. They are gender, age, race, grade, religious orientation, family income, parental smoking behavior and highest education of parents. All the $13,924$ respondents in sample 4.2 of Table 5.1 have no smoking history. About 12% of the $13,924$ initiated smoking in wave I. Table 5.2 lists the summary statistics for this sample. Figure 5.1 (created from Sample 4 in Table 5.1) 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 5.2 (created from sample 1 in Table 5.1) presents the distribution of the number of schoolmate friends. 72.81% of respondents (15,103 out of 20,745) in Add Health wave I in-home survey had at least one schoolmate friend. On average, a respondent had 1.76 friends.

Table 5.1: Sample Derivation

Sample number and selection criterion	Sample size
1. wave I respondent	20,745
2. with at least one parent	19,903
3. has measurements on family income	18,245
4. with complete state cigarette tax information	17,844
4.1. asked to nominate 5 schoolmate friends of each gender	5,774[1]
4.2. never smoking regularly before wave I	13,924[2]

[1] Using the 5,774 observations, I generate 2,987,761 directional pairwise observations to model friendship formation.

[2] Both sample 4.1 and sample 4.2 are subsets of sample 4.

Table 5.2: Summary Statistics ($N = 1,3924$)

Variables	Mean (Std. Dev.)	Min	Max
Smoking initiation	0.117(0.003)	0	1
State Cig. Tax (in 10 cents)	3.240(1.595)	0.25	7.5
No. of Smoking Parents			
0	0.384(0.004)	0	1
1	0.250(0.004)	0	1
2	0.367(0.004)	0	1
Highest Parents' Education			
College	0.334(0.004)	0	1
High school	0.499(0.004)	0	1
Less than high School	0.167(0.003)	0	1
Family Income by grade (in 10K dollars)	3.455(4.857)	0	99.9
Grade Level			
7	0.149(0.356)	0	1
8	0.139(0.346)	0	1
9	0.179(0.383)	0	1
10	0.194(0.396)	0	1
11	0.180(0.384)	0	1
12	0.158(0.364)	0	1
Age	16.028(1.729)	12	21
Female	0.504(0.500)	0	1
Race			
White	0.582(0.493)	0	1
Black	0.266(0.442)	0	1
Asian	0.076(0.265)	0	1
Other	0.076(0.265)	0	1
Religion			
Not Religious	0.148(0.335)	0	1
Unimportant	0.056(0.229)	0	1
Important	0.340(0.474)	0	1
Very important	0.457(0.498)	0	1

Figure 5.1: Individual Smoking vs. Peer Smoking

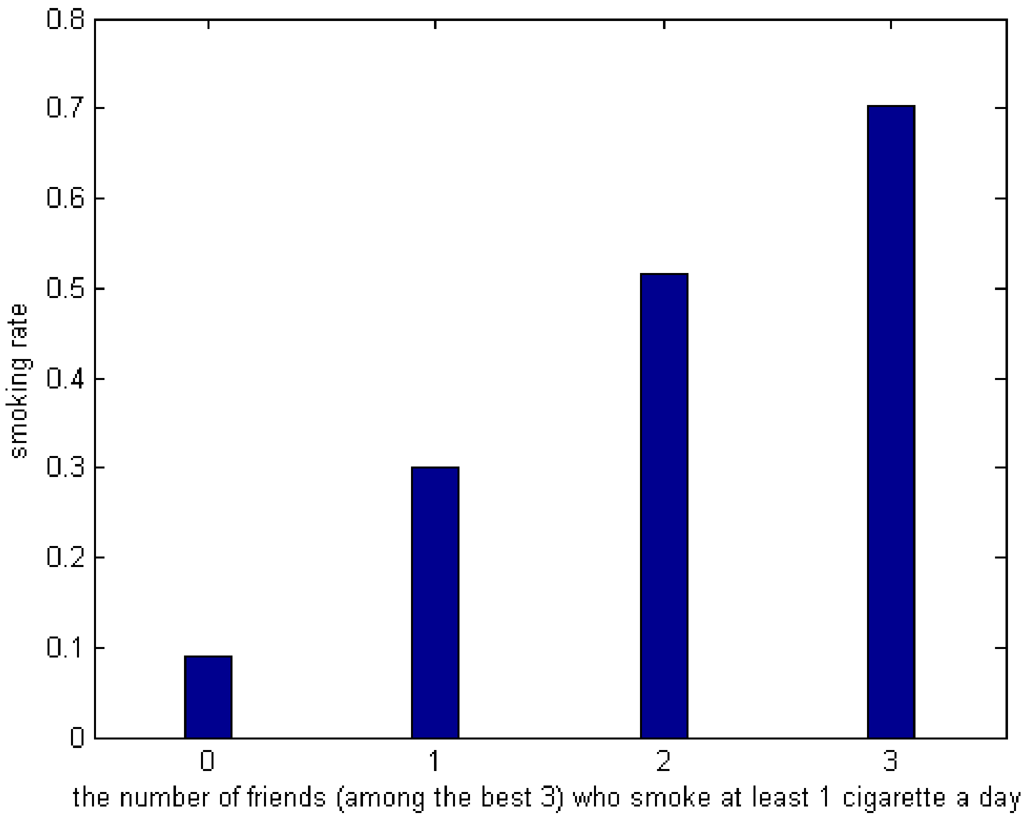
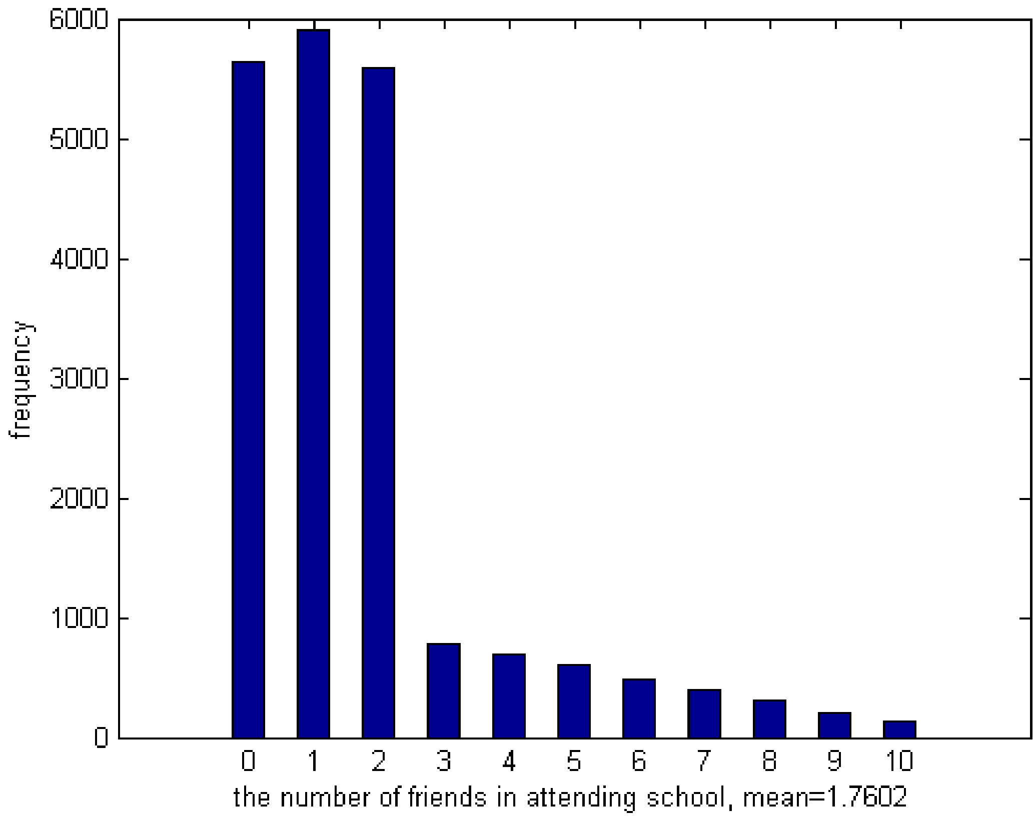


Figure 5.2: Number of Schoolmate Friends



Chapter 6

Results

6.1 Estimation of Instrumented *ex ante* Equilibrium Smoking Probabilities

The distribution of instrumented *ex ante* equilibrium smoking probabilities from the bagged tree classifier is shown in Figure 6.1. We can see that the median of the instrumented smoking probabilities among smokers (0.7024) is much larger than that among nonsmokers (0.0268). This is an indication that the bagged tree classifier does a good job in differentiating between smokers and nonsmokers. I further examine the performance of the bagged 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).[1] Figure 6.2 and Figure 6.3 compares the performance of the flexibly specified logit model and the bagged (n=100) tree classifier in terms of ROC and reliability, respectively. We can see clearly that in Figure 6.2, the area under the ROC curve of smoking probabilities obtained from the bagged tree classifier are much larger than that of smoking probabilities obtained from the flexibly specified logit model. This indicates that the fit in the bagged tree classifier is better than that in the logit model.

Figure 6.3 compares the bagged tree classifier and the flexibly specified logit model in terms of reliability.[2] We can see that the dots in the bagged tree classifier panel are clustering around

[1]The logit model includes the same exogenous explanatory variables as in the bagged tree. I manually added interactions terms into the logit model. The logit model already has quasi-perfect separation problem, so it already reaches the limit of flexibility in logit framework.

[2]For a dot (x, y) in Figure 6.3, y is the average smoking rate of students falling into an estimated smoking

the 45 degree line in Figure 6.3 over the whole range of the estimated smoking probabilities. For the logit model, an apparent problem is that the maximum estimated smoking probability is less than 0.75. This suggests that the flexibly specified logit model underestimates the individual smoking probability because sample 4 (see Table 5.1) includes many students with previous smoking histories whose smoking probabilities should be close to 1.

6.2 Estimation of Friendship Probabilities

Table 6.1 presents coefficient estimates of the directional friendship logit model. We can see as school size increases friendship probability drops.[3] This finding is consistent with the intuition that as school size increases a student's probability of making friends with an arbitrary schoolmate drops holding others constant. Conformities in grade, gender, age, and race contribute to friendship significantly. Income conformity does not contribute to friendship except for friendships among teens from low income families. As we can see in Table 6.1, directional smoking deviations are significant predictors for friendships. This is consistent with the theoretical implication that smoking actions and peer selection actions are interdependent. Figure 6.4 presents the reliability diagram of the estimated directional friendship probabilities. We can see that dots are closely clustering along the 45 degree line over the whole range indicating that the logit model performs well.

6.3 Estimation of the Structural Smoking Initiation Probabilities

With the instrumented *ex ante* equilibrium smoking probabilities and the instrumented *ex ante* equilibrium directional probabilities in hand, I then construct the instrumented *ex*

probability bin. y is calculated from observed data. x is the center of the estimated smoking probability bin. If estimation is good, we should expect x to be close to y. Therefore, dots should cluster closely to a 45 degree line.

[3]To achieve a higher degree of goodness of fit, I include the square of school size as an explanatory variable. If I do not include these square terms, the coefficient estimates of all the first order terms are negative values and are statistically significant.

ante equilibrium peer smoking norms based on equation 4.19. Figure 6.5 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. Apparently, the instrumented *ex ante* equilibrium individual smoking probabilities are positively correlated with the instrumented *ex ante* equilibrium peer smoking norms.

For comparison purposes, I present in Table 6.2 the coefficient estimates from six different specifications of the smoking initiation probability. Each Specification differently controls for the peer smoking norm. The corresponding marginal effect estimates are presented in Table 6.3. The first specification (no peer effects) does not include peer effects. Specification 1 assumes that peer influence does not exist. It controls for school-level unobserved heterogeneity by using school fixed effects. Specification 2 and Specification 3 model peer influence by using school norms as explanatory variables (Norton, 1998 and Lundborg, 2006). If the economist does not have detailed friendship network data on hand but knows the respondents' school memberships, then the economist may run Specification 2 and Specification 3 to capture peer influence. It is worth mentioning that both Specification 2 and Specification 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 uses observed smoking actions to construct school-level peer norms.[4] peer Specification 3 uses instrumented ex ante equilibrium smoking probabilities. Specification 4, Specification 5 and Specification 6 exploit the detailed peer composition information in Add Health in modeling peer influence. They differ in their efforts to correct for bias caused by the endogenous peer norms. In Specification 4, the peer norms are constructed based on observed friendships and observed smoking actions. In Specification 5, the peer norms are constructed based on observed friendships and instrumented ex ante equilibrium smoking probabilities. Therefore, both Specification 4 and Specification 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

[4] I am able to construct such school-level peer norms using the Add Health data that provides smoking behavior of all students in a school. However, those data are typically unavailable. Rather, researchers use reported school-level smoking average to represent peer smoking norms.

not. Some people did not nominate any schoolmate as a friend, therefore, in Specification 4 and Specification 5, only 4,268 out of the 13,924 observations can be used in analysis because for those students who did not nominate any schoolmate as a friend, the peer norms are undefined. Specification 6 is the preferred model because it controls for both the endogenous friendships and the endogenous smoking actions.

An obvious pattern in Table 6.3 is that the grade effect differs dramatically across different model specifications. Before further discussion on the interpretation of coefficients on grade levels, let us first examine what grade level really measures. Recall that all observations in the analysis samples (either N=13,924 or N=4,268) are free of smoking history. Therefore, a respondent's grade level is, in fact, perfectly collinear with the respondent's left censored survival time when we interpret 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 grade 7 depending on research convenience. 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 economist) 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. Expectedly, the more the time-invariant deterrents on smoking initiation are controlled, the less severe is the bias. As explained, grade level is perfectly collinear with the left censored survival time. Hence, we should expect that a marginal change in grade level from 7 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. From Table 6.3, we see that estimation results in Specification 4, 5 and 6 are consistent with such an expectation. In all these three specifications, as grade level increases (from grade

7), the predicted smoking initiation rate drops significantly. However, estimation results in Specification 1, 2 and 3 are inconsistent with such an expected pattern. For example, in 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, the smoking initiation rate increases by 3.30 (1.22) percentage points.

Comparing Specification 2, 3 and 6 in Table 6.3, we can see that peer influence estimated in Specification 2 (0.18 (0.03)) and Specification 3 (0.16 (0.03)) are much smaller than that that in Specification 6 (1.07 (0.09)). This is expected because it is reasonable to believe that a person's chosen friends influence the person more heavily than the person's schoolmates.

Comparing Specification 4, 5 and 6 in Table 6.3, we see a pattern that peer influence estimates in the former two specifications (0.11 (0.01) and 0.19 (0.04), respectively) are 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 4.17). 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 of the person's friends average smoking actions should be. From an econometric perspective, using a person's friends' average smoking actions as an explanatory variable in Specification 4 and 5, in essence, adds a measurement error onto her expected peer norm.[5] Since an average respondent in Add Health has only 1.76 friends, the variances of the measurement errors among the 4,268 respondents in Specification 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.[6]

[5]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 1 friend.

[6]The author notices that Specification 4 and 5 use an analysis sample different than that used in Specification 6. This difference may explain the smaller peer influence estimates in Specification 4 and 5. To investigate this possibility, the author checked the smoking rate (12.4%) in the analysis sample used in Specification 4 and 5 and that (11.7%) used in Specification 6 and found that they are very close. Therefore, the author believes

The preferred model (Specification 6) shows that peer influence is large and significant. A one percentage point increase in the peer norm causes the smoking initiation rate to increase by 1.07 (0.10) percentage points. All other models 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. It should be cautioned that Specification 6 itself cannot provide useful policy implications related to the peer norm because it is typically impossible for policy makers to exogenously assign peer norms to teens because teens choose their friends after all. I will discuss how to do policy simulation based on both the smoking equation and the friendship formation equation below.

Estimation results suggest that peer influence amplifies the direct tax deterrent effect on smoking initiation. In Specification 1 (no peer effect), 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. This differential in tax effect implies that the social multiplier effect amplifies the tax deterrent effect in the field because the tax effect estimated in Specification 1 is the combination of the direct tax effect estimated in Specification 6 and a social multiplier tax effect.

In specification 1 (no peer effects), compared to white teens with a smoking initiation rate of 13.76 (0.43) percentage, black teens are less likely to initiate smoking by 4.86 (0.81) percentage points. However, in Specification 6 (preferred model), after taking peer influence into account, being black increases an individual's smoking probability by 7.01 (1.55) percentage points. This finding indicates that friendship sorting based on racial conformity explains why black teens have a lower smoking rate than white teens. In specification 1, compared to having parents who have college education, having parents who have only a high school

that the differences in the analysis sample do not explain why the peer influence estimate in Specification 6 is larger than those in Specification 4 and 5.

education, the teen's smoking probability is increased by 1.25 (0.60) percentage points, and having parents who have an education level less than high school, the teen smoking probability is increased by 1.69 (0.82) percentage points In both specification 1 and specification 6, we can see that family does not matter too much. For 8th to 12th graders, family income effect was insignificant. For 7th graders in Specification 1, family income has a statistically significant effect, but is still trivial in a practical sense (-0.59 (0.23) percentage points).

Figure 6.1: Distribution of Estimated Smoking Probabilities(from bagged tree)

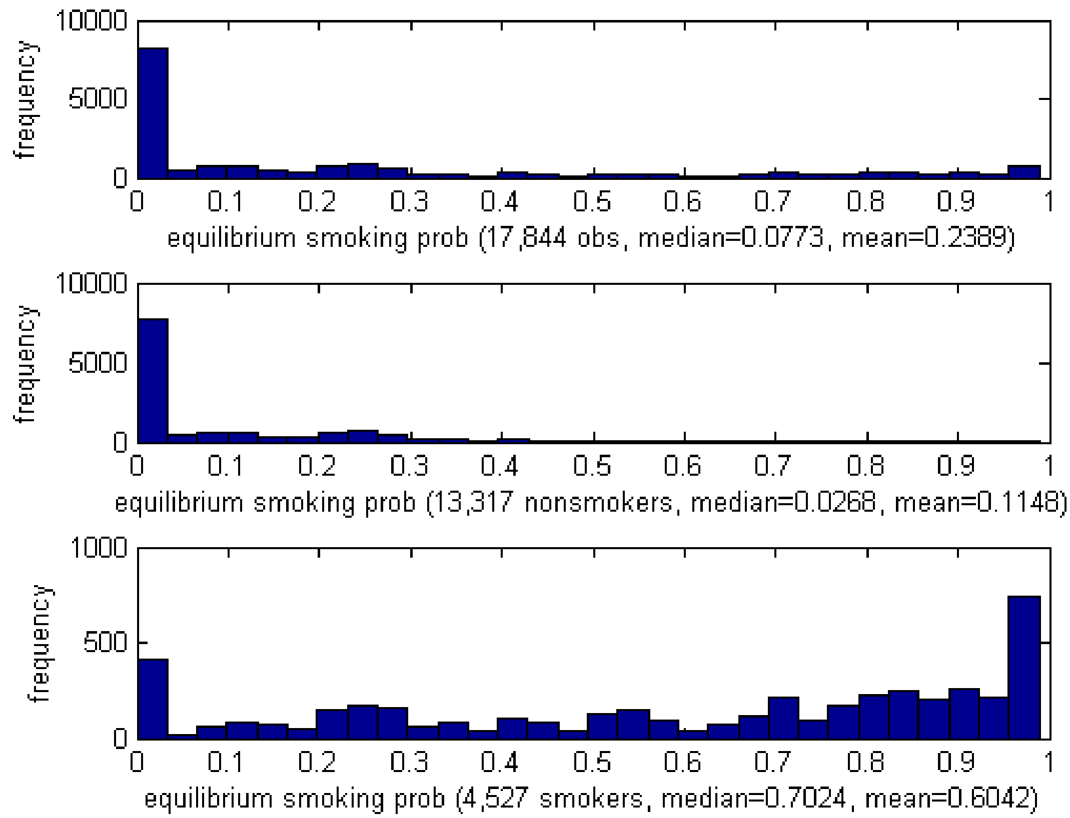


Figure 6.2: ROC Comparison Bagged Tree vs. Logit

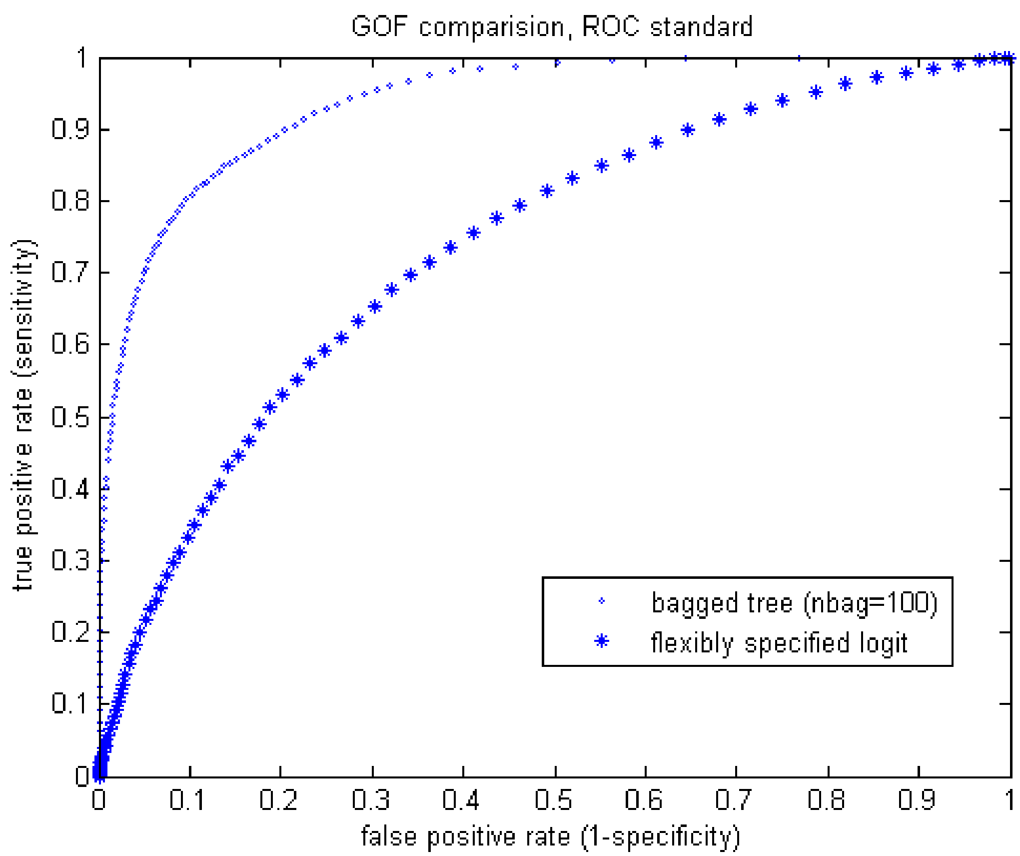


Figure 6.3: Reliability Comparison: Bagged Tree vs. Logit

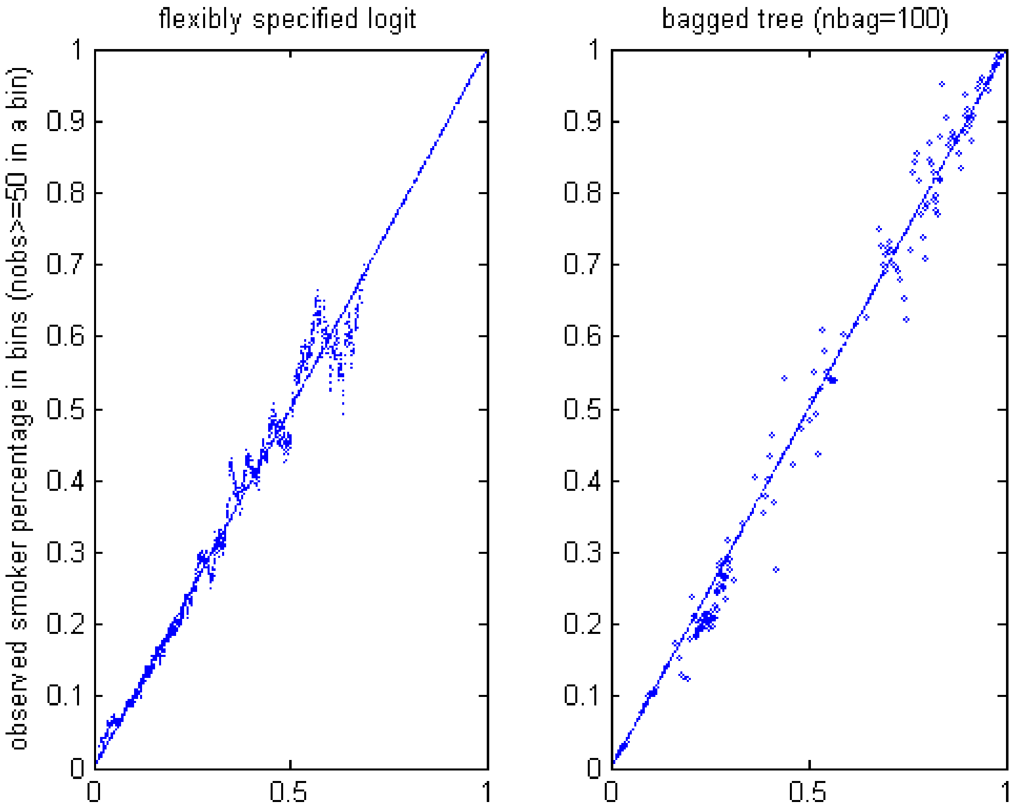


Figure 6.4: Reliability: Logit Directional Friendship

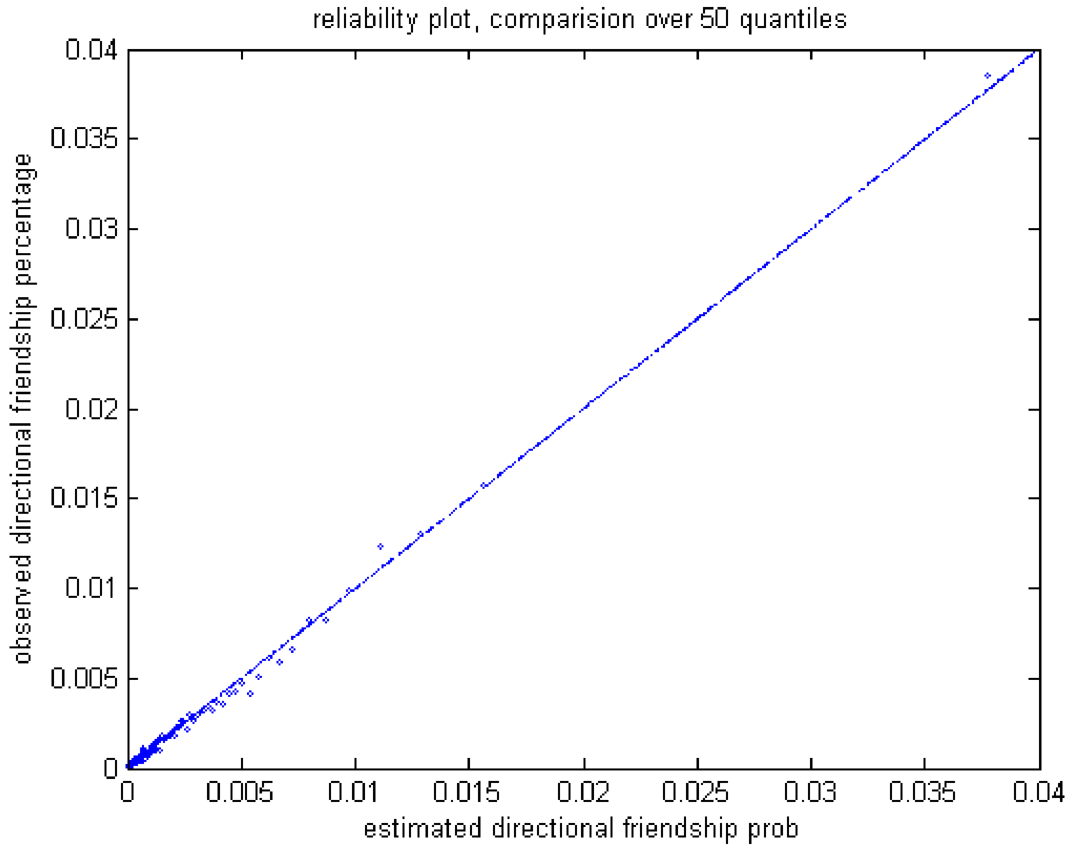


Table 6.1: Friendship logit model (N=2,987,761 created from 5,774 students)

Explanatory Variable	Coefficient (std err)
School size	
School size (<= 100)	0.065(8.44e-3)***
School size square	-5.76e-4(7.17e-5)***
School size (100 <size<= 200)	0.030(3.73e-3)***
School size square (100 <size<= 200)	-1.41e-4(1.45e-5)***
School size square(> 200)	2.83e-3(5.52e-4)***
School size square (> 200)	-1.91e-6(2.72e-7)
Grade deviation	.
7-upper grader	-1.872(0.149)***
8-8	2.50e-2(0.106)***
8-upper graders	-1.460(0.178)***
8-lower graders	-1.791(0.169)***

Continued on next page...

... table 6.1 continued

Explanatory Variable	Coefficient (std err)
9-9	-0.414(0.151)***
9-upper graders	-1.951(0.162)***
9-lower graders	-1.252(0.267)***
10-10	-0.525(0.149)***
10-upper graders	-1.814(0.155)***
10-lower graders	-1.985(0.174)***
11-11	-0.477(0.149)***
11-upper graders	-1.708(0.160)***
11-lower graders	-1.872(0.159)***
12-12	-0.472(0.150)***
12-lower graders	-1.980(0.163)***
Gender deviation	
Male-female	-0.320(0.027)***
Female-male	-0.381(0.037)***
Female-female	0.107(0.031)***
Age deviation	
Nonnegative age deviation (age<= 14)	-0.484(0.141)***
Negative age deviation (age<= 14)	-0.457(0.057)***
Nonnegative age deviation (14 <age<= 17)	-0.407(0.038)***
Negative age deviation (14 <age<=17)	-0.350(0.024)***
Nonnegative age deviation (age> 17)	-0.417(0.031)***
Negative age deviation (age> 17)	-0.283(0.064)***
Racial deviation	
White-Black	-1.842(0.130)***
White-Asian	-1.399(0.111)***
White-Other races	-0.357(0.076)***
Black-White	-1.957(0.165)***
Black-Black	0.107(0.064)*
Black-Asian	-2.850(0.251)***
Black-Other races	-2.000(0.209)***
Asian-White	-1.420(0.153)***
Asian-Black	-3.118(0.239)***
Asian-Asian	0.589(0.066)***
Asian-Other races	-1.507(0.129)***
Other races-White	-0.389(0.081)***
Other races-Black	-1.916(0.206)***
Other races-Asian	-1.238(0.161)***
Other races-Other races	2.31e-2(8.82e-2)

Continued on next page...

... table 6.1 continued

Explanatory Variable	Coefficient (std err)
Family income deviation	
Nonnegative income deviation (self income<= 25K)	-7.29e-2(3.75e-2)***
Negative income deviation (self income<= 25K)	-2.62e-2(7.80e-3)
Nonnegative income deviation (25K<self income<= 50K)	-3.67e-3(1.17e-3)
Negative income deviation (25K<self income<= 50K)	3.22e-3(5.86e-3)
Nonnegative income deviation (50K<self income75K)	-1.02e-2(9.50e-3)
Negative income deviation (50K<self income<= 75K)	-2.62e-3(7.92e-3)
Nonnegative income deviation (self income> 75K)	-4.52e-3(4.16e-3)
Negative income deviation (self income> 75K)	1.32e-3(8.77e-3)
IV smoking probability deviation	
Magnitude of IV smoking probability deviation (deviation > 0)	-0.641(0.115)***
Magnitude of IV smoking probability deviation (deviation < 0)	-0.297(0.096)***

Note: ***indicates significance at the 1% level; **5% level; *10% level

Figure 6.5: Trend Plot: IV individual Smoking Probability vs. IV Peer Smoking Norm

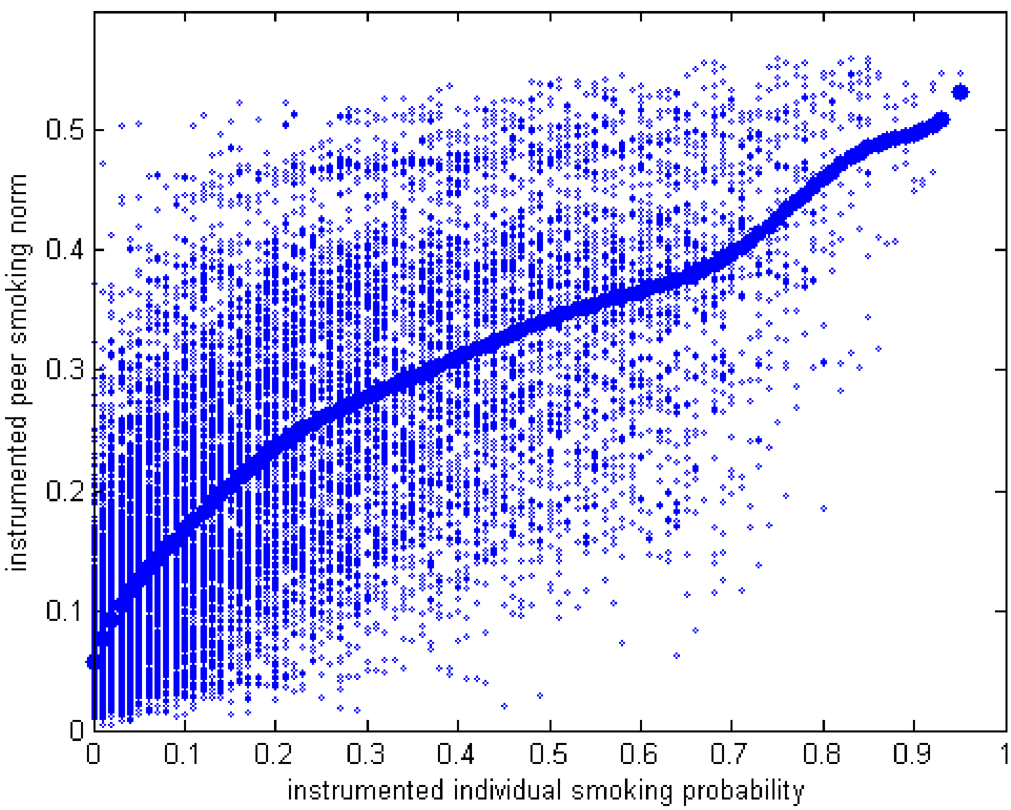


Table 6.2: Smoking Initiation Logit Model

	Six Different Specifications					
	Spec.1	Spec.2	Spec.3	Spec.4	Spec.5	Spec.6
	(N=13,924)	(N=13,924)	(N=13,924)	(N=4,268)	(N=4,268)	(N=13,924)
	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
Smoking Norm(unit is 0.01)						
7	NA	1.750(0.889)*	1.541(1.004)	1.163 (0.373)***	3.150(1.254)***	12.932(1.699)***
8	NA	-0.206(0.822)	-0.761 (0.948)	0.915(0.315)***	1.144(0.821)	9.684(1.424)***
9	NA	1.726(0.591)***	2.002(0.609)***	1.483(0.315)***	1577(0.695)**	10.145(1.096)***
10	NA	1.849(0.618)***	1.74(0.572)***	1.289(0.268)***	2.589(0.634)***	10.485(1.111)***
11	NA	1.780(0.624)***	1.519(0.599)***	0.797(0.263)***	1.194(0.556)**	9.179(0.924)***
12	NA	3.450(0.691)***	2.861(0.577)***	1.091(0.291)***	1.817(0.654)***	10.676(1.001)***
State Cig. Tax by Grade(unit in 10 cents)						
7	-0.620(0.248)***	-0.115(4.61e-2)**	-0.111(4.61e-2)**	-0.939(0.546)*	-0.874(0.528)	-0.505(0.328)*
8	-0.532(0.246)**	-1.77e-2(4.31e-2)	-1.57e-2(4.32e-2)	-0.891(0.544)*	-0.841(0.524)	-0.478(0.321)
9	-.509(0.244)**	-0.113(3.77e-2)***	-0.112(3.77e-2)***	-0.674(0.537)	-0.598(0.517)	-0.456(0.313)
10	-0.392(0.243)*	-3.74e-3(3.71e-2)	-3.43e-3(3.71e-2)	-0.489(0.535)	-0.416(0.516)	-0.414(0.314)
11	-0.407(0.243)*	-2.68e-2(3.86e-2)	-2.75e-2(3.86e-2)	-0.552(0.535)	-0.493(0.516)	-0.448(0.314)
12	-0.361(0.245)	4.13e-2(4.30e-2)	3.86e-2(4.30e-2)	-0.526(0.534)	-0.454(0.515)	-0.509(0.314)*
No. of Smoking Parents(comparison:0)						
1	0.260(0.071)***	0.259(0.071)***	0.258(0.072)***	0.347(0.130)**	0.370(0.128)***	0.255(0.074)***
2	0.342(0.065)***	0.343(0.064)***	0.347(0.068)***	0.373(0.125)***	0.397(0.123)***	0.298(0.068)***
Family Income by Grade(unit in 10K dollars)						
7	-7.21e-2(2.80e-2)***	-7.68e-2(2.74e-2)***	-7.82e-2(2.75e-2)***	-4.63e-2(4.80e-2)	-0.874(0.528)*	-5.80e-2(2.82e-2)**
8	1.11e-3(1.39e-2)	4.74e-4(1.33e-2)	1.12e-3(1.33e-2)	1.29e-2(1.80e-2)	-0.841(0.525)*	6.57e-3(1.45e-2)
9	6.67e-3(1.52e-2)	9.75e-3(1.46e-2)	8.90e-3(1.47e-2)	-1.84e-2(3.53e-2)	-0.598(0.517)	1.28e-2(1.89e-2)

Continued on next page...

... table 6.2 continued

	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
10	-2.41e-3(9.99e-3)	-1.55e-3(9.55e-3)	-1.41e-3(9.49e-3)	-1.84e-2(3.53e-2)	-0.416(0.516)	-1.35 e-3(1.27e-2)
11	1.26e-3(1.11e-2)	1.69e-2(1.09e-2)	1.73e-2(1.09e-2)	-2.45e-2(2.12e-2)	-0.493(0.516)	6.39e-3(1.30e-2)
12	1.81e-2(1.32e-2)	1.79e-2(1.30e-2)	1.81e-2(1.28e-2)	1.67e-2(1.02e-2)	-0.454(0.515)	9.05e-3(1.58e-2)
Highest Parents' Education(comparison: college)						
High Schl	0.127(0.061)*	0.115(0.061)*	0.115(0.061)*	0.268(0.113)**	0.278(0.112)**	0.129(0.063)**
< High Schl	0.169(0.080)*	0.158(.079)**	0.155(0.079)**	0.254(0.148)*	0.262(0.147)*	0.185(0.083)**
Race						
Black	-0.510(0.091)***	-0.399(0.075)***	-0.391(0.076)***	-0.725(0.193)***	-0.631(0.195)**	0.621(0.124)***
Asian	-0.529(0.138)***	-0.407(0.125)***	-0.405(0.127)***	-0.522(0.230)**	-0.526(0.228)**	-0.168(0.138)
Other Races	-0.092(0.115)	-0.014(0.106)***	0.014(0.107)	-0.432(0.236)	-0.458(0.223)	0.074(0.119)
Female	0.052(0.055)	0.045(0.054)	0.046(0.054)	-2.61e-2(0.101)	-4.42e-2(0.100)	0.028(0.056)
Age	0.074(0.041)*	0.066(0.040)*	0.063(0.038)*	0.171(8..21e-2)*	0.139(0.079)*	9.93e-3(0.043)
Grade(comparison:7)						
8	-0.214(0.235)	0.156(0.316)	-0.057(0.276)	-8.80e-2(0.480)	2.82e-2(0.491)	-0.190(0.283)
9	-0.581(0.319)*	-0.234(0.330)	-0.330(0.298)	-1.350(0.654)**	-1.285(0.667)	-0.922(0.393)**
10	-1.030(0.341)***	-0.650(0.348)*	-6.393(0.301)**	-2.073(0.679)***	-2.328(0.685)***	-1.201(0.447)***
11	-0.979(0.358)***	-0.564(0.361)	-0.521(0.317)*	-1.953(0.713)***	-1.953(0.714)*	-0.971(0.458)**
12	-1.377(0.382)***	-1.517(.408)***	-1.252(0.363)***	-2.263(0.776)***	-2.372(0.777)**	-1.261(0.471)***
Religion(comparison:textrmnot religious at all)						
Unimp.	0.230(0.117)*	0.215(0.115)*	0.210(0.118)*	0.435(0.219)**	0.361(0.217)*	0.210(0.123)*
Imp.	-8.45e-3(0.080)	-0.015(0.077)	-0.020(0.079)	0.144(0.155)	0.112(0.154)	-0.012(0.083)
Very imp.	-0.389(0.084)***	-0.393(0.080)***	-0.402(0.081)***	-0.191(0.165)	-0.208(0.163)	-0.382(0.086)***

***indicates significance at the 1% level; **5% level; *10% level

[a] In this specification, I do not use school fixed effects because the set of school dummy variables is perfectly collinear with (IV) school norms.

[b] In this specification, the standard errors are obtained from 5,000 bootstraps.

Continued on next page...

... table 6.2 continued

	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]

[c] In this specification, I could not bootstrap the standard errors. Due to small sample size (N=4,268) and the large number of school dummy variables, quasi-perfect separation problem appears in almost every bootstrap samples.

Table 6.3: Marginal Effects on Smoking Initiation Rate

	Six Different Specifications[a,b,c,d,e,f]					
	Spec.1	Spec.2	Spec.3	Spec.4	Spec.5	Spec.6
	(N=13,924)	(N=13,924)	(N=13,924)	(N=4,268)	(N=4,268)	(N=13,924)
	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
Smoking Norm(unit is 0.01)						
Peer Effect	NA	0.18(0.03)***	0.16(0.03)***	0.11(0.01)***	0.19(0.04)***	1.07(0.09)***
	NA	11.78(0.27)	11.79(0.27)	13.47(0.51)	13.68(1.41)	12.10(0.28)
7	NA	0.15(0.07)**	0.13(0.09)	0.11(0.03)***	0.30(0.12)**	1.12(0.16)***
	NA	9.35(0.64)	9.36(0.64)	12.61(1.45)	12.61(1.44)	9.78(0.64)
8	NA	-0.02(0.09)	-0.08(0.11)	0.11(0.04)***	0.14(0.10)	1.07(0.16)***
	NA	12.70(0.76)	12.72(0.75)	16.78(1.57)	16.87(1.60)	13.19(0.74)
9	NA	0.18(0.06)***	0.21(0.06)***	0.16(0.03)***	0.18(0.08)**	1.08(0.12)***
	NA	12.22(0.65)	12.23(0.66)	14.64(1.34)	14.68(1.34)	12.51(0.65)
10	NA	0.19(0.06)***	0.18(0.06)***	0.12(0.03)***	0.25(0.06)***	1.11(0.12)***
	NA	11.87(0.61)	11.86(0.61)	12.26(1.02)	12.24(1.02)	12.11(0.62)
11	NA	0.20(0.07)***	0.17(0.07)**	0.08(0.03)***	0.12(0.06)**	1.02(0.11)***
	NA	12.84(0.66)	12.86(0.66)	12.85(1.06)	12.83(1.07)	13.07(0.67)
12	NA	0.35(0.07)***	0.28(0.07)***	0.11(0.03)***	0.19(0.07)***	1.06(0.11)***
	NA	11.49(0.68)	11.49(0.68)	13.10(1.23)	13.15(1.24)	11.69(0.66)
State Cig. Tax	-3.86(1.81)**	-0.35(0.17)**	-0.35(0.17)**	-4.70(3.82)	-4.27(3.93)	-3.68(2.29)*
(unit: 10 cents)	12.10(0.28)	11.78(0.27)	11.79(0.27)	13.47(0.51)	13.68(1.41)	12.22(0.32)
By grade						
7	-4.01(1.34)***	-0.91(0.35)***	-0.86(0.36)**	-6.16(3.07)**	-5.86(3.13)*	-3.24(1.93)*
	9.75(0.65)	9.35(0.64)	9.36(0.64)	12.61(1.45)	12.61(1.44)	9.83(0.69)
8	-4.64(1.90)**	-0.19(0.47)	-0.16(0.46)	-7.68(4.11)*	-7.40(4.15)*	-4.08(2.53)*

Continued on next page...

... table 6.3 continued

	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
	13.16(0.76)	12.70(0.76)	12.72(0.75)	16.78(1.57)	16.87(1.60)	13.28(0.81)
9	-4.27 (1.83)**	-1.14(0.37)***	-1.11(0.36)***	-5.32(4.07)	-4.85(4.28)	-3.74(2.39)
	12.49(0.66)	12.22(0.65)	12.23(0.66)	14.64(1.34)	14.68(1.34)	12.57(0.68)
10	-3.30(1.95)*	-0.03(0.38)	-0.02(0.38)	-3.36(3.94)	-2.74(4.02)	-3.35(2.43)
	12.11(0.62)	11.87(0.61)	11.86(0.61)	12.26(1.02)	12.24(1.02)	12.16(0.64)
11	-3.66(2.05)*	-0.27(0.40)	0.30(0.42)	-3.98(4.02)	-3.52(4.09)	-3.35(2.43)
	13.09(0.65)	12.84(0.66)	12.86(0.66)	12.85(1.06)	12.83(1.07)	13.14(0.69)
12	-2.96(1.93)*	0.41(0.44)	0.38(0.43)	-3.87(4.14)	-3.30(4.21)	-3.86(2.17)*
	11.72(0.68)	11.49(0.68)	11.49(0.68)	13.10(1.23)	13.15(1.24)	11.77(0.67)
No. Smk Parents						
0	10.15(0.42)	9.84(0.41)	9.83(0.42)	11.16(0.82)	11.00(0.82)	10.40(0.44)
1	2.52(0.71)***	2.50(0.68)***	2.49(0.70)***	3.41(1.29)***	3.73(1.31)***	2.47(0.71)***
2	3.41(0.65)***	3.41(0.64)***	3.47(0.65)***	3.71(1.25)***	4.02(1.25)***	2.93(0.67)***
Family Income	-0.03(0.06)	-0.01(0.06)	-0.02(0.06)	-0.09(0.12)	-0.09(0.12)	-0.01(0.07)
(unit in 10K dollars)	12.10(0.28)	11.78(0.27)	11.79(0.27)	13.47(0.51)	13.68(1.41)	12.22(0.32)
7	-0.59(0.23)***	-4.46(1.24)	-0.63 (0.22)***	-3.33(3.69)	3.10(3.72)	-0.46(0.23)
	9.75(0.65)	9.35(0.64)	9.36(0.64)	12.61(1.45)	12.61(1.44)	9.83(0.69)
8	0.01(0.15)	0.14(1.47)	0.02(0.15)	1.64(2.29)	1.80(2.34)	0.07(0.16)
	13.16(0.76)	12.70(0.76)	12.72(0.75)	16.78(1.57)	16.87(1.60)	13.28(0.81)
9	0.07(0.16)	1.19(1.64)	0.09(0.15)	-1.54(3.49)	1.44(3.65)	0.13(0.20)
	12.49(0.66)	12.22(0.65)	12.23(0.66)	14.64(1.34)	14.68(1.34)	12.57(0.68)
10	-0.02(0.10)	-0.10(0.95)	-0.01(0.10)	-2.01(1.75)	-2.58(1.80)	-0.01(0.13)
	12.11(0.62)	11.87(0.61)	11.86(0.61)	12.26(1.02)	12.24(1.02)	12.16(0.64)
11	0.13(0.12)	1.98(1.33)	0.19(0.12)	1.85(1.92)	1.57(1.91)	0.06(0.14)
	13.09(0.65)	12.84(0.66)	12.86(0.66)	12.85(1.06)	12.83(1.07)	13.14(0.69)
12	0.18(0.13)	1.92(1.44)	0.18(0.13)	-0.38(2.41)	-0.17(2.43)	0.09(0.15)

Continued on next page...

... table 6.3 continued

	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
	11.72(0.68)	11.49(0.68)	11.49(0.68)	13.10(1.23)	13.15(1.24)	11.77(0.67)
Highest Parents' Education						
College	11.19(0.46)	10.95(0.46)	10.95(0.45)	11.72(0.85)	11.65(0.84)	11.22(0.46)
High Schl	1.25(0.60)**	1.13(0.60)*	1.14(0.59)*	2.65(1.11)**	2.77(1.14)**	1.26(0.62)**
Less High Schl	1.69(0.82)**	1.59(0.83)*	1.58(0.81)**	2.51(1.55)*	2.69(1.53)*	1.86(0.85)**
Race						
White	13.76(0.43)	12.96(0.38)	12.95(0.38)	15.67(0.82)	15.55(0.82)	11.25(0.37)
Black	-4.86(0.81)***	-3.80(0.68)***	-3.74 (0.69)***	-6.78(1.62)***	-6.11(1.75)***	7.03(1.55)***
Asian	-4.97(1.14)***	-3.83(1.05)***	-3.79(1.06)***	-5.21(2.07)**	-5.20(2.10)***	-1.42(1.17)
Other Races	-0.97(1.25)	-0.13(1.15)	0.11(1.15)	-4.38(2.26)**	-4.62(2.19)**	0.71(1.17)
Gender						
Male	13.77(0.43)	12.02(0.39)	12.02(0.40)	13.33(0.74)	13.26(0.74)	12.31(0.41)
Female	-0.53(0.55)	-0.46(0.54)	-0.46(0.55)	0.26(1.03)	0.44(1.02)	-0.29(0.55)
	0.79(0.44)	0.69(0.42)	0.65(0.42)	1.87(0.93)**	1.68(0.93)*	0.11(0.43)
Age	12.10(0.28)	11.78(0.27)	11.79(0.27)	13.47(0.51)	13.68(1.41)	12.22(0.32)
Grade						
7	12.40(1.62)	11.68(1.42)	11.56(1.47)	18.91(3.96)	23.83(4.78)	26.99(3.65)
8	3.30(1.22)***	2.07(1.35)	1.76(1.47)	2.28(2.75)	-1.89(3.85)	-8.27(3.22)***
9	0.13(1.67)	0.34(1.39)	0.28(1.43)	-4.10(3.85)	-9.23(4.56)**	-14.71(3.58)***
10	-0.92(1.91)	-0.43(1.60)	-0.27(1.66)	-6.41(4.21)	-11.81(4.93)**	-15.85(3.71)***
11	-0.37(2.20)	0.19(1.91)	0.39(1.96)	-7.31(4.69)	-12.34(5.40)**	-17.09(3.94)***
12	2.40(2.36)	-2.05(2.08)	-1.78(2.13)	-8.98(5.14)*	-13.99(5.81)**	-18.17(4.00)***
Religion						
Not Relig. at all	13.62 (0.75)	13.36(0.72)	13.42(0.73)	13.47(1.39)	13.75(1.44)	13.67(0.78)
Not imp.	2.83(1.46)**	2.64(1.44)*	2.61(1.46)*	5.12(2.66)**	4.30(2.68)*	2.54(1.47)*
Imp.	-0.10(0.89)	-0.17(0.86)	-0.22(0.88)	1.51(1.66)	1.22(1.64)	-0.14(0.94)

Continued on next page...

... table 6.3 continued

	No peer effect	Schl norm[a]	IV Schl norm[a,b]	Peer norm	IV peer norm[c]	Preferred[b]
Very import.	-3.84(0.86)***	-3.88(0.83)***	-3.96(0.84)***	-1.87(1.65)	-2.08(1.67)	-3.75(0.89)***

[a] A marginal effect is calculated as the average of observation-wise marginal effects on the smoking initiation probability in the analysis sample.

[b] Marginal effects are measured in percentages. For example, overall marginal peer effects is 1.07(0.10)*** in Specification 6; this means that an one percentage point increase in peer norm causes a 1.07percentage point increase in smoking imitation rate with a standard error of 0.10.

[c] Standard errors of marginal effects are obtained through bootstrapping method. The presence of interaction terms in the six specifications in the table makes it inconvenient to write down the analytical expression of a marginal effect, so I do not use the delta method to calculate standard errors. Alternatively, I randomly sample a large number of model coefficients (I use 5,000) from their estimated joint distribution then obtain means and standard errors of marginal effects from the 5,000 different models.

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

[e] All numbers in italics are the predicted smoking initiation rates at which marginal effects are evaluated. For example, in Specification 1, the overall marginal tax effect is evaluated at 12.10(1.28); this means the smoking initiation rate is expected to drop to 12.10-3.86=8.24 percent after a 10-cent increase in cigarette tax.

[f] For continuous variables such as tax, income and etc., 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 which all observations are assumed to be a member of the comparison 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 becomes a 7th grader without a smoking history.

Chapter 7

Policy Simulations

Consider a policy intervention that causes a variation in observed public information (ΔI^o). Since a Nash equilibrium strategy is a function of public information, such a policy intervention affects students' actions on both smoking and peer selection and, in turn, their peer smoking norms.

Consider a generic school s with n_s students under a policy intervention, ΔI^o. Let the n_s by 1 vector $E(\overline{a}^*_{-i,s}; F(.), I + \Delta I^o)$ denote the n_s students' *ex ante* equilibrium peer smoking norms. We note the following three theoretical inferences. First, according to equation 4.19, the *ex ante* peer norm vector is completely determined by a $n_s \times (n_s - 1)$ by 1 *ex ante* equilibrium directional friendship probability vector, $p(\zeta_i^{j*}(I + \Delta I^o; F(.)) = 1)$, and a n_s by 1 *ex ante* individual equilibrium smoking probability vector, $p(a^*_{s,j}(I + \Delta I^o; F(.)) = 1)$. Second, the *ex ante* individual smoking probability vector is partially determined by an *ex ante* equilibrium smoking norm vector due to peer effects on smoking. Third, the *ex ante* equilibrium friendship vector is partially determined by the *ex ante* individual smoking probability vector that affects directional deviations in smoking probabilities among schoolmates. Hence, the *ex ante* equilibrium peer smoking norm vector, the *ex ante* equilibrium 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, I 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, I do not have instrumental variables (e.g., lagged cigarette price and lagged state cigarette tax) to explain lagged smoking behavior. This data limitation motivates me 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 I 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.

Table 7.1 shows the basic information in the two schools used in policy simulation. The first school has 55 students and the second school, 63. State cigarette tax is 28 cents in the first school and 75 cents in the second school. The original equilibrium smoking rates in the first school and the second school are 5.45% and 33.33%, respectively.[1] In the policy simulation, I perturb cigarette taxes by adding an additional amount of tax to the original state cigarette tax. For comparison purposes, I solve for equilibrium smoking rates with and without peer effects at each tax level. In simulating equilibrium smoking rates without peer effects, I set the coefficient estimates corresponding to grade-specific peer effects in the estimated smoking equation (Table 6.2) and the coefficient estimates corresponding to directional deviations in smoking dimension in the friendship equation (Table 6.1) to zeros. Parameters used in simulation are randomly drawn from the estimated distribution of smoking equation parameters and the estimated distribution of friendship equation parameters.

Figure 7.1 presents the simulation results in the two schools. The upper panel is the first school; the lower panel is the second school. Overall, with or without peer effects, as tax in-

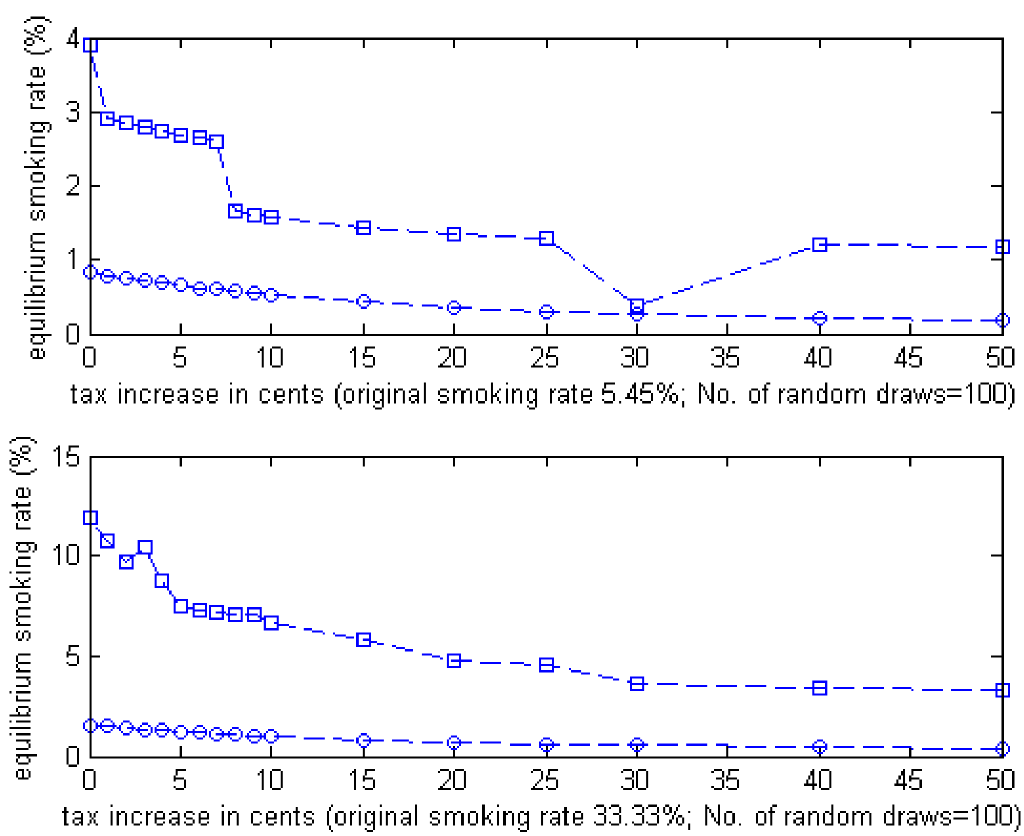
[1] The original equilibrium smoking rate is just the smoking rate observed in the Add Health data.

creases the smoking rate decreases. Interestingly, in the presence of peer effects, tax increases in certain ranges may cause "abnormal" increases in the smoking rate (e.g., tax increase=40 cents, in the first school). We note if the friendship network does not change as taxes increase, then a tax increase should monotonically pull down the smoking rate. These "abnormal" increases in the smoking rate reflect that a variation in the cigarette tax motivates agents to update their smoking decision and their friendships as well. 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 smoking rate in the presence of peer effects. This indicates peer effects significantly amplify tax deterrent effects. Meanwhile, in both schools the presence of peer effects significantly increases the smoking 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. Also we note under peer effects, at certain tax thresholds, the smoking rate drops abruptly (e.g., tax increase=7 cents in the first school). This indicates that at those tax thresholds the social multiplier effect is particularly large (herding behavior appears).

Table 7.1: Two Schools Used in Policy Simulation

School	No. of students	1995 state cig. Tax	Smoking rate
1	55	28	5.45
2	63	75	33.33

Figure 7.1: Equilibrium Smoking Rate: Peer Effects vs. No Peer Effects



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 ($d_{-i,t}$).

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

Proposition 1. *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*

$$\widetilde{D}_i = \bigcup_{\widetilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\widetilde{d}_{-i,t}) \ \forall i \in N_t \tag{7.1}$$

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

where $\bigcup_{\widetilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\widetilde{d}_{-i,t})$ 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 $\widetilde{D}_{-i} = \prod_{j \in N_t \setminus \{i\}} \widetilde{D}_j$

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

We note, $\widetilde{d}^k_t \in \prod_i \widetilde{D}_i \Rightarrow \widetilde{d}^k_{-i,t} \in \widetilde{D}_{-i}$. In turn, $\widetilde{d}^k_{-i,t} \in \widetilde{D}_{-i}$ and $\widetilde{D}_i = \bigcup_{\widetilde{d}_{-i,t} \in \widetilde{D}_{-i}} B_i(\widetilde{d}_{-i,t}) \Rightarrow B_i(\widetilde{d}^k_{-i,t}) \subseteq \widetilde{D}_i$. 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 $\forall k = 1, 2, ..., K(\prod_i \widetilde{D}_i)$, $\widetilde{d}_{i,t} \in \prod_i \widetilde{D}_i$ satisfies n_t budget constraints simultaneously. **Q.E.D.**

Bibliography

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

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

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

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

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

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

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

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

[9] 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

[10] 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.

[11] Chaloupka. F. AND K. Warner, (2000). The Economics of Smoking.

[12] 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.

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

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

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

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

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

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

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

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

[21] 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.

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

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

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

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

[26] 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.

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

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

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

[30] 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.

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

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

[33] Mackay, J. AND Eriksen M. (2002). The Tobacco Atlas. World Health Organziation.

[34] Manski, C. (1995). Economic Analysis of Social Interactions. *Journal of Economic Perspectives* 14, 115-136.

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

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

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

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

[39] Scheinkman, J. (2006). Social Interactions. working paper

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

[41] 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.