

NBER WORKING PAPER SERIES

ASYMMETRIC SOCIAL INTERACTION IN ECONOMICS:  
CIGARETTE SMOKING AMONG YOUNG  
PEOPLE IN THE UNITED STATES, 1992-1999

Jeffrey E. Harris  
Beatriz González López-Valcárcel

Working Paper 10409  
<http://www.nber.org/papers/w10409>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
April 2004

Supported in part by grant PR2002-0134 from the Ministerio de Educación Cultura y Deporte, España. This work is the authors' sole responsibility, and does not necessarily represent the opinions of the Massachusetts Institute of Technology, the University of Las Palmas de Gran Canaria, or any other organization. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by Jeffrey E. Harris and Beatriz González López-Valcárcel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Asymmetric Social Interaction in Economics:  
Cigarette Smoking Among Young People in the United States, 1992-1999  
Jeffrey E. Harris and Beatriz González López-Valcárcel  
NBER Working Paper No. 10409  
April 2004  
JEL No. I12

**ABSTRACT**

We analyzed cigarette smoking among people aged 15 - 24 in approximately 90,000 households in the 1992 - 1999 U.S. Current Population Surveys. We modeled social influence as an informational externality, in which each young person's smoking informs her peers about its "coolness." The resulting "family smoking game," with each sibling's smoking endogenous, may have multiple equilibria. We found that the pro-smoking influence of a fellow smoker markedly exceeded the deterrent effect of a non-smoking peer. The phenomenon of asymmetric social influence has implications for financial markets, educational performance, criminal behavior, and other areas of inquiry where peer influence is important.

Jeffrey E. Harris  
Department of Economics  
MIT  
Cambridge, MA 02139  
and NBER  
jeffrey@mit.edu

Beatriz González López-Valcárcel  
Department of Quantitative Methods in Economics  
University of Las Palmas de Gran Canaria  
Campus de Tafira  
35017 Las Palmas  
Spain  
bvalcarcel@dmc.ulpgc.es

## I. INTRODUCTION

Economists have devoted considerable theoretical and empirical attention to the study of social interactions, but few have inquired whether such interactions might be asymmetric.

Glaeser, Sacerdote, and Scheinkman [1996] observed, for example, that rates of larceny and auto theft in a neighborhood influence the propensity of young residents to commit such crimes. One might further ask, could the deterrent effect of a decrease in the neighborhood crime rate be absolutely larger or smaller than the crime-enhancing effect of an increase in the crime rate? Clark and Oswald [1996] found that workers' job satisfaction was significantly negatively correlated with their co-workers' income. One might similarly ask, could the disutility of a high-income coworker be greater or smaller than the utility of a low-income colleague? Kindleberger [1989], Bikhchandani, Hirshleifer, and Welch [1992, 1998] and Lux [1998] and others have studied the "contagion of opinion" in financial markets, where investors follow the behavior of other fellow traders to form expectations about future prices. One might further inquire, do investors make different inferences when they see their peers buy, rather than sell, securities? Only in the literature on school performance, it appears, have Sacerdote [2001], Hoxby [2002], and Winston and Zimmerman [2003] formally inquired whether a low achiever gains more by being in a class (or dorm room) with a high achiever than a high achiever loses by being with a low achiever.

Numerous authors have similarly observed a young person's decision to smoke is influenced by peers' smoking practices. (See, for example, Flay et al. [1994], Wang et al. [1995], Engels et al. [1997], Norton, Lindrooth, and Ennett [1998], DeCicca, Kenkel, and Mathios [2000], Wang, Eddy, and Fitzhugh [2000], Gaviria and Raphael [2001], Alexander et al. [2001], Simons-Morton et al. [2001], Norton, Lindrooth, and Ennett [2003].) Yet no one, to our knowledge, has explicitly raised the possibility such peer influence might be asymmetric, too.

Our task in this paper is to do just that. We focus sharply on social interactions within the household unit, rather than larger peer groups, such as those at school or work. We find, in fact, that the pro-smoking influence of one fellow smoker markedly exceeds the deterrent effect of a non-smoking peer.

The implications of asymmetry in social interactions are extensive. If the positive effect of a high-performing peer exceeds the negative effect of a low performer, then mixing good and bad workers or students might enhance overall group performance. Conversely, if the deterrent effect of a non-smoking peer is weaker than the pro-smoking effect of a fellow smoker, then rules that segregate smokers will reduce overall smoking rates. What is more, when peer effects are nonlinear, tracking or clustering may create positive externalities only when the number of peers reaches a critical mass. (See Hoxby [2002].)

It is well known that social interactions can create multiplier effects (Glaeser, Sacerdote, and Scheinkman [2002]), and such a phenomenon has not been lost on researchers who have studied teenage smoking (Lewit, Coate, and

Grossman [1981], Harris and Chan [1999], Gruber [2000], and Krauth [2001]). In the presence of positive social externalities, an increase in the price of cigarettes (or, for that matter, a change in any exogenous factor that influences smoking rates) will have a larger aggregate effect than the sum of the effects on each individual separately. This “social multiplier” effect has, in fact, been used to explain why teenagers seem to have higher price elasticities of demand than adult smokers.

But if social interactions are asymmetric, then social multipliers can be asymmetric, too. In that case, an increase in the price of cigarettes could have a different absolute effect on the rate of teenage smoking than a decrease in price. Such a phenomenon, in fact, may underlie the observation that teenage smoking rates rebounded sharply after the 1993 premium brand price war, but have declined only sluggishly in times of price hikes (Gruber [2001]). In the context of financial markets, asymmetric social multipliers might offer an explanation for slow, modulated booms and rapid busts.

In the next section, we define the scope of our problem. In Section III, we propose a simple model of informational externalities that guides our empirical analysis. In Section IV, we analyze the resulting “family smoking game,” in which each young person’s smoking is endogenous, and where multiple Nash equilibria can obtain. Section V then details our econometric methods. Finally, sections VI and VII present our results and conclusions.

## II. THE PROBLEM

Table 1 shows the prevalence of current cigarette smoking among persons aged 15–24 years in relation to the smoking status of other young people, aged 15–24, who reside in the same household. The cross-tabulation is based upon 122,010 individual responses to the Tobacco Use Supplements (TUS) that were appended to the U.S. Current Population Surveys during three waves: September 1992 – May 1993; September 1995 – May 1996; and September 1998 – May 1999. (See U.S. Bureau of the Census [1996, 1998].) Each cell in Table 1 shows the proportion who currently smoke, as well as the sample size, the number of households, and the log-odds of current smoking.

[INSERT TABLE 1 ABOUT HERE.]

As Table 1 shows, a young person (call her “Bea”) who resides in a household with no other young people has a 19 percent probability of currently smoking. Although the pattern is not completely consistent in the largest households, where the sample sizes are small, the inclusion of nonsmoking housemates appears to decrease Bea’s probability of smoking, while the addition of more smokers appears to increase her probability of smoking. Thus, if only one young nonsmoker also resides in the household, then Bea’s probability of smoking falls to 12 percent, and if there is a second young nonsmoker in residence, then Bea’s probability of smoking falls further to 8 percent. On the other hand, if only one young smoker resides in the household, then Bea’s probability of smoking rises from 19 to 50 percent, and if a second smoker is

resident, then Bea's probability of smoking rises further to 63 percent. What is more, the quantitative effect of including one more smoking housemate appears to exceed the effect of including one more nonsmoking housemate. For example, if Bea started out with no young housemates, then adding two nonsmokers would reduce her log-odds of smoking by 0.99, while adding two smokers would increase her log-odds by 1.98.

Although the data in Table 1 suggest the presence of strong social influences within the household, we need to be wary of the possibility of *self-selection*, that is, Bea simply chose to cohabit with housemates who had compatible smoking habits. In the TUS sample, however, only 4 percent of young persons residing in the same household were unrelated. In all likelihood, most young people had little choice but to live with their parents and other siblings. For convenience, in what follows, we shall use the terms "sibling" and "young housemate" interchangeably.

Quite apart from self-selection, we need to be concerned that the observed concordance of smoking among young people is simply the result of other common, observable family influences. It is well known, for example, that smoking prevalence is inversely related to family income (e.g., Harris and Chan [1999]). Since variations in retail cigarette prices are largely determined by state excise tax rates, residents of the same family are likely to face nearly identical prices. The presence or absence of smoking adults may similarly influence the

smoking practices of young people in a household. In some households, smoking is not permitted.

It is also possible that unobserved common influences, or *common shocks*, within the same household underlie the patterns observed in Table 1. While the TUS explicitly asked respondents whether smoking was permitted in the home, there may be an unstated family rule that nobody should smoke in the presence of an infirm, elderly relative. Or perhaps the underage siblings in the same household may obtain cigarettes from the same source in the schoolyard.

Even if we properly took account of self-selection, observed characteristics and unobserved common shocks, we would still be faced with the conundrum of two-way causation, what Manski [1993] has called the *reflection problem*. Suppose that 18-year-old Bea has a 19-year-old brother Jeff who lives in the same household and does not smoke cigarettes. We observe that Bea likewise does not smoke. Now consider an otherwise identical two-sibling household, endowed with the same income, facing the same retail cigarette prices, caring for the same sick grandmother. In this clone household, 18-year-old Bea has a 19-year-old brother Jeff who does smoke cigarettes, and we observe that Bea likewise smokes. Do we conclude that Jeff's smoking practices influenced Bea's? Or is it the other way around?

If we could perform such an experiment, then we could at least measure the two-way social interaction between Bea's and Jeff's smoking. Beyond that, we would have to impose strong restrictions to infer one-way causation. We



could posit, for example, that if Jeff began to smoke before Bea, then Jeff caused Bea to smoke, but not vice versa. Such a restriction is dubious because Bea's smoking, even if she began after Jeff started, may also be keeping Jeff from quitting.

When we observe households with three or more young people, however, we can impose restrictions that follow naturally from the organization of the data in Table 1. Suppose that we observed another household, with Bea and Jeff as above, which differed from the first clone in that there was a third sibling, 20-year-old Pepe. The combined influence of Bea's two housemates on Bea's smoking, we posit, depends only on the number of smoking siblings, and not their identity. In that case, all other things equal, Bea's probability of smoking when Jeff does smoke and Pepe doesn't is the same as her probability of smoking when Jeff doesn't and Pepe does. When we impose such a restriction, and when we use the data on both the two-sibling and three-sibling households together, then as we will see below, we can identify the one-way causal effect of one young person's smoking on all the remaining sibs.

Even if we could confront the reflection problem, there is still one more knot to untie. Return to our two-sibling household with only Bea and Jeff. Let's assume that if Jeff smokes, then with 100% probability Bea will smoke, too. Likewise, if Bea smokes, then with certainty Jeff will smoke, too. In that case, there are two equally logical outcomes: both Bea and Jeff do not smoke; both Bea and Jeff do smoke. In this example, we know everything about the household,

including the one-way effect of Jeff's smoking on Bea and vice versa. Yet we cannot make an unequivocal, unique prediction about who smokes and who doesn't. This indeterminacy is what Tamer [2002, 2003] has called the problem of *incoherence*. In effect, we have a static binary game between Bea and Jeff, in which there are two Nash equilibria. And in families with more than two siblings, it is not hard to imagine that there can be more than two such equilibria. As we will see below, one way to address the problem of incoherence is to specify additional rules that determine which equilibrium will obtain.

Before we squarely confront the above-noted problems of common shocks, reflection and incoherence, we first need to specify a model that elucidates why Bea's smoking decision might actually depend on Jeff's. For this purpose, we could simply make the ad hoc assumption that Bea's utility from smoking depends on the group average smoking rate (e.g., Brock and Durlauf [2001], Lindbeck, Nyberg, and Weibull [1999]). Or we could specifically assume that Bea gets greater utility from smoking when she and Jeff smoke together. Such notions of joint production especially make sense in the context of illicit drug use, where heroin users share syringes and other paraphernalia, or where cocaine users congregate in crack houses. We might point out, for example, that Jeff, as an experienced smoker, teaches the technology of consumption to Bea the novice (Jones [1994].) Alternatively, we could also model peer influence as a network externality, a phenomenon that seems apt for drug users, too. In what follows, we devise a model of peer influence that relies on informational

externalities in a manner similar to that used by analysts of financial markets (Bikhchandani, Hirshleifer, and Welch [1992, 1998]).

### III. A LEARNING MODEL

We think of each young person as uncertain whether smoking cigarettes is, on the one hand, a safe, non-addicting and socially acceptable thing to do, or on the other hand, a dangerous, addicting and socially passe practice. Such a young person draws inferences from her peers' smoking to resolve this uncertainty. Although her peer group is likely to be much larger than her household, her siblings' practices provide important data that she relies upon to revise her prior beliefs about the likelihood that smoking is, in fact, the "cool" thing to do. In what follows, we use "cool" as a shorthand term for safe, non-addicting and socially acceptable, while "uncool" is short-hand for dangerous, addictive and passe.

A stylized version of the decision faced by such a young person, whom we call "Bea," is shown in Figure 1 below. At the square choice node, Bea decides whether or not to smoke. We assume that Bea does not take account of the possibility that her own smoking may influence its social desirability in her peer group. Thus, her subjective probability  $\phi$  that smoking is "cool" does not depend on her own smoking decision. Since we observe only the smoking practices of individual household members and not those of her larger peer group, we refrain from any rational-expectations assumptions that relate Bea's

subjective probability  $\phi$  to the actual prevalence of smoking in the larger peer group. Since our focus is on social interactions, we sidestep questions as to whether Bea makes a myopic short-run decision or formulates a long-run lifetime plan. (See Becker and Murphy [1988], Becker, Grossman, and Murphy [1994].)

[INSERT FIGURE 1 ABOUT HERE.]

We model uncertainty about the “coolness” of smoking by the circular chance nodes emanating from both the “smoke” and “not smoke” branches in the figure. Associated with each (action, state of nature) pair is a terminal utility. Thus  $V_1$  is the utility of smoking when it turns to be “cool,” while  $V_2$  is the utility of smoking if it turns out to be “uncool.”  $V_3$  is the utility of refraining from smoking when it turns to be “cool,” while  $V_4$  is the utility of not smoking when it turns out to be “uncool.” We assume that  $B = V_1 - V_3 > 0$ . That is, Bea experiences a positive benefit  $B$  from smoking when it is “cool” to smoke. Conversely, we assume that  $C = V_4 - V_2 > 0$ . That is, she experiences a cost  $C$  of lighting up when it is “uncool” to do so.

The difference between the expected utility of smoking,  $\phi V_1 + (1 - \phi) V_2$ , and the expected utility of not smoking,  $\phi V_3 + (1 - \phi) V_4$ , is

$$(1) \quad \phi V_1 + (1 - \phi) V_2 - \{ \phi V_3 + (1 - \phi) V_4 \} = \phi(B + C) - C$$

Bea smokes so long as this difference in expected utility is non-negative. Let us denote Bea's smoking behavior by the binary variable  $y_1$ , where  $y_1 = 1$  if Bea smokes and  $y_1 = 0$  if not. Then Bea smokes so long as her subjective probability  $\phi$  is no less than a threshold value  $\tilde{\phi}$ , that is,

$$(2) \quad \begin{aligned} y_1 &= 1 \text{ if } \phi - \tilde{\phi} \geq 0, \text{ where } \tilde{\phi} = \frac{C}{B+C}; \text{ and} \\ y_1 &= 0 \text{ otherwise} \end{aligned}$$

The pair  $(\phi, \tilde{\phi})$  thus describes Bea's subjective beliefs about the coolness of smoking ( $\phi$ ) and her perceptions of its relative costs and benefits ( $\tilde{\phi}$ ). Denoting

$$\mu = \ln\left(\frac{\phi}{1-\phi}\right) \text{ and } \tilde{\mu} = \ln\left(\frac{\tilde{\phi}}{1-\tilde{\phi}}\right) \text{ as the log-odds transformations of } \phi \text{ and } \tilde{\phi},$$

respectively, we can rewrite the condition (2) as

$$(3) \quad \begin{aligned} y_1 &= 1 \text{ if } \mu - \tilde{\mu} \geq 0, \text{ where } \tilde{\mu} = \ln C - \ln B; \text{ and} \\ y_1 &= 0 \text{ otherwise} \end{aligned}$$

We now inquire how information about a peer's smoking practices can influence Bea's own decision to smoke. In particular, we assume that Bea has a point prior distribution on  $\phi$  and, in Bayesian fashion, updates this prior when she learns whether or not her brother Jeff smokes cigarettes. Let  $\lambda$  denote the probability that Jeff would smoke if it were "cool" to do so, and let  $\omega$  denote the probability that Jeff would refrain from smoking if it were "uncool." If we think of Jeff's smoking as a diagnostic test of the "coolness" of smoking, then  $\lambda$  is the

test's sensitivity, while  $w$  represents the test's specificity. When the sensitivity  $\lambda$  is high, Bea thinks, "My brother Jeff would definitely smoke if it were the cool thing to do." When specificity  $w$  is high, the Bea thinks, "My brother Jeff wouldn't be caught dead with a cigarette if it weren't cool." When  $\lambda + w > 1$ , Jeff's smoking is informative about its "coolness." By contrast, when  $\lambda + w = 1$ , Jeff's smoking is entirely uninformative, and knowledge of Jeff's smoking would have no effect on Bea's prior  $\phi$ .

If Jeff in fact smokes, then by Bayes' rule, the log-odds of Bea's posterior probability that smoking is "cool" becomes

$$(4) \quad \mu + \ln \lambda - \ln(1 - \omega)$$

If Jeff does not smoke, then the log-odds of Bea's posterior probability that smoking is "cool" becomes

$$(5) \quad \mu + \ln(1 - \lambda) - \ln \omega$$

Let us similarly denote Jeff's smoking behavior by the binary variable  $y_2$ , where  $y_2 = 1$  if Jeff smokes and  $y_2 = 0$  if not. Then, in light of the information about Jeff's smoking, Bea's decision to smoke can be described by the rule

$$(6) \quad \begin{aligned} y_1 &= 1 \text{ if } \mu - \bar{\mu} + \ln\left(\frac{\lambda}{1 - \omega}\right)y_2 + \ln\left(\frac{1 - \lambda}{\omega}\right)(1 - y_2) \geq 0; \text{ and} \\ y_1 &= 0 \text{ otherwise} \end{aligned}$$

When Jeff's smoking is informative, that is,  $\lambda + \omega > 1$ , the coefficient  $\ln\left(\frac{\lambda}{1-\omega}\right)$  will be strictly positive, which means that his smoking ( $y_2 = 1$ ) will reinforce Bea's decision to smoke. Similarly, when Jeff's smoking is informative, the coefficient  $\ln\left(\frac{1-\lambda}{\omega}\right)$  will be strictly negative, which means that Jeff's abstaining ( $y_2 = 0$ ) will reinforce Bea's decision to refrain.

Jeff's influence on Bea, however, has a built-in asymmetry. That is, the reinforcing effect of Jeff's smoking,  $\ln\left(\frac{\lambda}{1-\omega}\right)$ , is not necessarily equal to and of opposite sign to the dissuading effect of Jeff's abstaining,  $\ln\left(\frac{1-\lambda}{\omega}\right)$ . When the former is stronger than the latter, we have  $\ln\left(\frac{\lambda}{1-\omega}\right) + \ln\left(\frac{1-\lambda}{\omega}\right) > 0$ . So long as Jeff's smoking is informative ( $\lambda + \omega > 1$ ), this condition is equivalent to  $\omega > \lambda$ , that is, the specificity of Jeff's smoking exceeds its sensitivity. Conversely, the dissuading influence of Jeff's abstaining exceeds the reinforcing influence of his smoking when  $\omega < \lambda$ . Only when  $\omega = \lambda$  would symmetric social influence prevail.

Generalizing across an arbitrary number  $n > 1$  of young persons in a household, we now let  $(\mu_i, \mu_i)$  denote the prior beliefs and preferences of young person  $i$ , while we let  $y_i$  be the binary variable representing that person's smoking practices. For now, we make the simplifying assumption that each

young person views the smoking practices of each of her siblings as independent information about the “coolness” of smoking. Thus, if Bea had another brother, Pepe, she would not dismiss the fact of Pepe’s smoking as uninformative simply because he might copy his brother Jeff.

Once young person  $i$  observes the smoking habits of all of her siblings  $\{y_j; j \neq i\}$ , her posterior probability that smoking is “cool” is (in log-odds form)

$$(7) \quad \mu_i + \{\ln \lambda - \ln(1 - \omega)\} \sum_{j \neq i}^n y_j + \{\ln(1 - \lambda) - \ln \omega\} \sum_{j \neq i}^n (1 - y_j)$$

We write  $\kappa = \ln\left(\frac{\lambda}{1 - \omega}\right) > 0$  and  $\alpha = \ln\left(\frac{1 - \lambda}{\omega}\right) < 0$  for compactness of notation.

(The inverse mapping is  $\lambda = \frac{\exp(\kappa) - \exp(\kappa + \alpha)}{\exp(\kappa) - \exp(\alpha)}$  and  $\omega = \frac{\exp(\kappa) - 1}{\exp(\kappa) - \exp(\alpha)}$ .) We

also denote  $u_i = (\mu_i - \bar{\mu}_i)$ . Then each young person’s smoking is described by the rule

$$(8) \quad \begin{aligned} y_i &= 1 \text{ if } u_i + \kappa \sum_{j \neq i}^n y_j + \alpha \sum_{j \neq i}^n (1 - y_j) \geq 0; \text{ and} \\ y_i &= 0 \text{ otherwise} \end{aligned}$$

Letting  $\theta = \kappa - \alpha$ , we can rewrite the rule (8) as

$$(9) \quad y_i = 1 \text{ if } u_i + \alpha(n - 1) + \theta \sum_{j \neq i}^n y_j \geq 0; \text{ and } y_i = 0 \text{ otherwise}$$



for all  $i = 1, \dots, n$ . Since  $\kappa > 0$  and  $\alpha < 0$ , we know that  $\theta > 0$ . The coefficients  $\kappa$  and  $\alpha$ , respectively, gauge the effects of importing a young smoker or importing a young nonsmoker into the household. By contrast, the coefficient  $\theta$  measures the net effect of a young person's switching from a nonsmoker to a smoker in a household with a fixed number  $n$  of young people.

In equation (8), the term  $u_i = (\mu_i - \bar{\mu}_i)$  captures each young person's prior beliefs  $(\mu_i)$  about the "coolness" of smoking, as well as the perceived net benefits of smoking  $(-\bar{\mu}_i = \ln B_i - \ln C_i)$ . The former can depend on other social influences, such as the smoking practices of adults and peers outside the household, or any other information about the perceived dangers or acceptability of smoking. The latter can depend on the price of cigarettes, income, age, gender, and other personal or household characteristics. In anticipation of the empirical analysis below, we therefore specify the linear link function

$$(10) \quad u_i = X_i\beta + Z\gamma + \varepsilon_i$$

for all  $i = 1, \dots, n$ , where  $X_i$  is a vector of observable personal characteristics,  $Z$  is a vector of observable household characteristics,  $(\beta, \gamma)$  is a parameter vector, and  $\varepsilon_i$  is a disturbance term that captures unobservable factors. We have

$$(11) \quad \begin{aligned} y_i &= 1 \text{ if } X_i\beta + Z\gamma + \alpha(n-1) + \theta \sum_{j \neq i}^n y_j + \varepsilon_i \geq 0; \text{ and} \\ y_i &= 0 \text{ otherwise} \end{aligned}$$

for all  $i = 1, \dots, n$ . For now, we need assume only that the unobserved factors are random variables  $(\varepsilon_1, \dots, \varepsilon_n)$  with a common mean of zero and a joint cumulative distribution function  $F$ . In what follows, it will be helpful to abbreviate  $W_i = X_i\beta + Z\gamma + \alpha(n-1)$  for all  $i = 1, \dots, n$ , so that (11) becomes

$$(12) \quad y_i = 1 \text{ if } W_i + \theta \sum_{j \neq i}^n y_j + \varepsilon_i \geq 0; \text{ and } y_i = 0 \text{ otherwise}$$

So long as we assume that each young person views the smoking practices of each of her siblings as independent information about the “coolness” of smoking, the parameters  $\alpha$  and  $\theta$  in equation (11) will remain constant and independent of the total number of siblings  $n$  in the family. However, if Bea recognizes the interdependence of the smoking practices of her brothers Jeff and Pepe, then the additional observation that Pepe smokes may be more or less informative than the initial observation that Jeff smoked. In that case, we expect the parameter  $\theta$  to vary with  $n$ . While we continue to assume that  $\alpha$  and  $\theta$  are constant in the upcoming discussion of the equilibria of the system (12), we shall relax this assumption in our econometric analyses later on.

#### IV. THE FAMILY SMOKING GAME

We can think of the system (12) as delineating the strategies of each player in an  $n$ -person static, binary, simultaneous-move game. In this “family smoking

game,” as we will call it, each young person’s decision to smoke (that is,  $y_j = 1$ ) enhances the propensity of the others to follow suit (that is,  $y_i = 1$ ). Our family smoking game differs from the market entry games studied by Bresnahan and Reiss [1990], Berry [1992], and others, where the entry of a rival firm  $j$  into the market (that is,  $y_j = 1$ ) makes it less profitable for firm  $i$  to be in the market (that is,  $y_i = 1$ ). It also differs from the binary game studied by Tamer [2003], in which  $y_j = 1$  makes it more likely that  $y_i = 1$ , but  $y_i = 1$  makes it less likely that  $y_j = 1$ . However, like the market entry game and Tamer’s game, the family smoking game can have multiple Nash equilibria in pure strategies.

To illustrate the point, consider a family with two young people, where Bea is young person #1, while her brother Jeff is young person #2. The system (12) becomes

$$(13) \quad \begin{aligned} y_1 &= 1 \quad \text{if} \quad W_1 + \theta y_2 + \varepsilon_1 \geq 0; \quad y_1 = 0 \quad \text{otherwise} \\ y_2 &= 1 \quad \text{if} \quad W_2 + \theta y_1 + \varepsilon_2 \geq 0; \quad y_2 = 0 \quad \text{otherwise} \end{aligned}$$

When  $\varepsilon_1 \geq -W_1$ , smoking cigarettes is a dominant strategy for Bea, that is, she will smoke whether or not Jeff does. Conversely, when  $\varepsilon_1 < -W_1 - \theta$ , abstaining from smoking is a dominant strategy for her. But when  $-W_1 - \theta > \varepsilon_1 \geq -W_1$ , Bea will choose to smoke only if Jeff does, too. In the region of the  $(\varepsilon_1, \varepsilon_2)$  plane where  $-W_1 > \varepsilon_1 \geq -W_1 - \theta$  and  $-W_2 > \varepsilon_2 \geq -W_2 - \theta$ , there will be two

equilibria:  $(y_1, y_2) = (0,0)$ ; and  $(y_1, y_2) = (1,1)$ . The situation is illustrated in Figure 2 below, which plots the different pure-strategy equilibria in the  $(\epsilon_1, \epsilon_2)$  plane.

[INSERT FIGURE 2 ABOUT HERE.]

The 3-sibling family smoking game, involving Bea, Jeff, and Pepe (sibling #3), entails similar rules:

$$\begin{aligned}
 (14) \quad y_1 &= 1 \quad \text{if} \quad W_1 + \theta(y_2 + y_3) + \epsilon_1 \geq 0; y_1 = 0 \quad \text{otherwise} \\
 y_2 &= 1 \quad \text{if} \quad W_2 + \theta(y_1 + y_3) + \epsilon_2 \geq 0; y_2 = 0 \quad \text{otherwise} \\
 y_3 &= 1 \quad \text{if} \quad W_3 + \theta(y_1 + y_2) + \epsilon_3 \geq 0; y_3 = 0 \quad \text{otherwise}
 \end{aligned}$$

While we do not display all possible equilibria in the  $(\epsilon_1, \epsilon_2, \epsilon_3)$  plane for this game, we do show in Figure 3 the equilibria in the  $(\epsilon_2, \epsilon_3)$  subspace where  $-W_1 - \theta > \epsilon_1 \geq -W_1 - 2\theta$ , that is, where Bea smokes only when both of her siblings smoke, too.

[INSERT FIGURE 3 ABOUT HERE.]

In the lower right rectangle in the figure, where the disturbances  $(\epsilon_1, \epsilon_2, \epsilon_3)$  satisfy  $\{-W_1 - \theta > \epsilon_1 \geq -W_1 - 2\theta; \epsilon_2 \geq -W_2; -W_3 - 2\theta > \epsilon_3\}$ , smoking is a dominant strategy for Jeff (sibling #2), while abstinence is a dominant strategy for Pepe (sibling #3). Bea (sibling #1) will therefore abstain, too, and the sole equilibrium is  $(y_1, y_2, y_3) = (0,1,0)$ . In the rectangle

immediately above that one, where  $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$  instead satisfies

$$\{-W_1 - \theta > \varepsilon_1 \geq -W_1 - 2\theta; \varepsilon_2 \geq -W_2; -W_3 - \theta > \varepsilon_3 \geq -W_3 - 2\theta\},$$

smoking remains Jeff's dominant strategy. However, Bea will smoke only if Pepe does, too, and vice versa. Thus, we have two equilibria in pure strategies:

$(y_1, y_2, y_3) = (0, 1, 0)$  and  $(1, 1, 1)$ . In the central rectangle where no sibling has a dominant strategy, the two Nash equilibria are:  $(y_1, y_2, y_3) = (0, 0, 0)$  and  $(1, 1, 1)$ .

In market entry games, the number of entrants, that is, the quantity  $\sum_{i=1}^n y_i$ , is uniquely determined even when there are multiple equilibria. Thus, in the 2-firm market entry game, there is a region of the  $(\varepsilon_1, \varepsilon_2)$  plane where both  $(y_1, y_2) = (1, 0)$  and  $(y_1, y_2) = (0, 1)$  are equilibria. In that case, we can still say that there will be exactly one entrant into the market, even if we cannot specify which firm. In the family smoking game, however, we cannot uniquely determine the number of smokers. In the center rectangle in Figure 2, there can either be 0 or 2 smokers, and in the center rectangle of Figure 3, there can be 0 or 3 smokers.

While the 2- and 3-sibling games have at most two pure-strategy equilibria, this is not generally the case. Consider a 4-sibling smoking game in which both Bea (#1) and Jeff (#2) will smoke if just one sibling does, too (that is,  $-W_i > \varepsilon_i \geq -W_i - \theta; i = 1, 2$ ), while Pepe (#3) and Silvia (#4) will each smoke only if at least 3 siblings do, too (that is,  $-W_i - 2\theta > \varepsilon_i \geq -W_i - 3\theta; i = 3, 4$ ). Then

there are three pure-strategy equilibria:  $(y_1, y_2, y_3, y_4) = (0,0,0,0)$ ;  $(1,1,0,0)$ ; and  $(1,1,1,1)$ .

In those regions of the  $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  space where multiple pure-strategy equilibria prevail, the family smoking game does admit a unique mixed strategy equilibrium. For example, when  $(\varepsilon_1, \varepsilon_2)$  belongs to the central rectangle in

Figure 2, the mixed strategy equilibrium entails Bea's smoking with probability

$-\frac{W_2 + \varepsilon_2}{\theta}$  and Jeff's smoking with probability  $-\frac{W_1 + \varepsilon_1}{\theta}$ . When  $n > 2$  and no

player has a dominant strategy, the mixed strategy equilibrium entails sibling  $i$

smoking with probability  $\frac{1}{n\theta} \left( (W_i + \varepsilon_i) - \sum_{j \neq i}^n (W_j + \varepsilon_j) \right)$ . In the context of youth

smoking initiation, we find it difficult to devise a realistic interpretation for such

mixed strategies. Nor do we see any appeal in a Stackelberg solution, as

explored in other contexts by Kooreman [1994], Bjorn and Vuong [1997],

Hiedemann [1998], and Chao [2002]. If Jeff is older than Bea, then his smoking

can exert a greater influence on her than vice versa, but this does not imply that

Jeff "pre-commits" to smoking. As discussed below, sensitivity analyses of

alternative model specifications indicated that the net influence parameter  $\theta$  did

not depend significantly on a sibling's age.

## V. ECONOMETRIC METHODS

As Tamer [2002, 2003] has explained, the structural parameters of models such as that specified in equations (11) and (12) can sometimes be identified even when the model itself is rendered “incomplete” by the existence of multiple equilibria. In the case of the family smoking game, with sufficiently rich data on the observables  $\{X, Z, y\}$ , we should in principle be able to distinguish between exogenous effects  $(\beta, \gamma)$ , and the endogenous effect  $\theta$ . Moreover, if we study households of different sizes  $n$ , then we can identify the parameter  $\alpha$  as well.

Although Tamer [2002, 2003] has suggested several novel strategies for estimation, we resort to a more practical approach. In particular, we assume that in those regions of the  $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$  space where multiple pure-strategy equilibria prevail, the observed equilibrium is the outcome of a random trial governed by mixing probabilities that we estimate as incidental parameters.

The case of the two-sibling family illustrates our approach. Let us assume that the disturbances  $(\varepsilon_1, \varepsilon_2)$  are joint normally distributed with zero means, unit variances, and correlation coefficient  $\rho$ . The probability that  $(\varepsilon_1, \varepsilon_2)$  lies in the center region of Figure 2, where there are multiple equilibria, is

(15)

$$R = \Phi(W_1 + \theta, W_2 + \theta, \rho) - \Phi(W_1 + \theta, W_2, \rho) - \Phi(W_1, W_2 + \theta, \rho) + \Phi(W_1, W_2, \rho)$$

where  $\Phi(u, v, \rho)$  is the corresponding bivariate cumulative normal cumulative distribution function. We assume that when the pair  $(\varepsilon_1, \varepsilon_2)$  lies in the center region in Figure 2, the observed equilibrium is the outcome of a Bernoulli trial, with probability  $\pi$  of the all-smoker equilibrium  $(1,1)$ , and probability  $1 - \pi$  of the all-nonsmoker equilibrium  $(0,0)$ . If both Bea and Jeff are observed to smoke, that is,  $(y_1, y_2) = (1,1)$ , then the contribution of such an observation to the likelihood function is

$$(16) \quad \ell(1,1) = \Phi(W_1 + \theta, W_2 + \theta, \rho) - (1 - \pi)R$$

If both Bea and Jeff do not smoke, that is,  $(y_1, y_2) = (0,0)$ , then the contribution of such an observation to the likelihood function is

$$(17) \quad \ell(0,0) = \Phi(-W_1, -W_2, \rho) - \pi R$$

If Bea smokes but Jeff does not, that is,  $(y_1, y_2) = (1,0)$ , then the contribution of such an observation to the likelihood function is

$$(18) \quad \ell(1,0) = \Phi(W_1, -W_2 - \theta, -\rho)$$

Similarly, if Bea does not smoke but Jeff does, that is,  $(y_1, y_2) = (0,1)$ , then the contribution of such observation to the likelihood function is

$$(19) \quad \ell(0,1) = \Phi(-W_1 - \theta, W_2, -\rho)$$



With sufficiently rich data, we can estimate not only the structural parameters but also the incidental parameters ( $\rho, \pi$ ) by maximizing the likelihood function defined in (15) through (19).

The main drawback of the mixing-probability approach is that the number of incidental parameters increases nonlinearly with family size. For example, in the 3-sibling family described in part in Figure 3, we would in principle need one distinct mixing parameter for each of the four regions in which there are two equilibria. In this 3-sibling case, we can economize on incidental parameters by assuming that there is a single mixing parameter  $\pi$ , which represents the probability that the equilibrium with the larger number of smokers, that is, the larger value of  $y_1 + y_2 + y_3$ , will prevail. With four or more siblings, however, the number of potential incidental parameters can become unwieldy, and the choice of economizing restriction is less obvious.

Fortunately, as shown in Table 1, the vast majority of households in our data set contained no more than 3 young people. In both 2- and 3-sibling households, as we shall see below, our maximum likelihood estimates of the structural parameters were robust with respect to alternative specifications of the mixing probabilities, even though such mixing probabilities could not be estimated with precision. Put differently, at the maximum likelihood estimates, the density of the region of multiple equilibria (which, in the 2-family case, corresponds to the quantity  $R$  in equation (15)) turned out to be quite small.

Expanding on the specification in equation (11), we denote the smoking behavior of young person  $i$  in household  $h$ , which contains  $n_h$  young people, by the binary variable  $y_{ih}$ , where

$$(20) \quad y_{ih} = 1 \text{ if } X_{ih}\beta + Z_h\gamma + \alpha(n_h - 1) + \theta Y_{ih} + v_h + \delta_{ih} \geq 0, \text{ and} \\ y_{ih} = 0 \text{ otherwise,}$$

where  $Y_{ih} = \sum_{j \neq i}^{n_h} y_{jh}$  denotes the number of siblings of young person  $i$  who currently smoke; each  $v_h$  is independently identically normally distributed with mean 0 and variance  $\rho \in [0,1]$ ; each  $\delta_{ih}$  is independently identically normally distributed with mean 0 and variance  $1 - \rho$ ; and  $\{\theta, \alpha, \beta, \gamma, \rho\}$  are unknown parameters. Under this random-effects error structure, each combined disturbance term  $\varepsilon_{ih} = v_h + \delta_{ih}$  is normally distributed with zero mean and unit variance. Moreover,  $E[\varepsilon_{ih} \varepsilon_{jh}] = \rho$  for  $i \neq j$ , that is, the disturbances  $\{\varepsilon_{ih}\}$  may be correlated within the household, even if they are independent across households.

Our estimation and testing strategy proceeded from the particular to the general. We first estimated the structural model (20) for the subset of 2-sibling families alone under the assumption, detailed above, that in cases of two possible equilibria, there was a probability  $\pi$  that the equilibrium with the larger value of  $Y_{ih}$  would prevail. With fixed  $n_h = 2$ , this case permitted us to estimate the

structural parameter  $\theta$ , but not the parameter  $\alpha$ , which is indistinguishable from the constant term implicit in the vector  $(X_{ih}, Z_h)$ .

Second, we estimated the structural model (20) for the subset of 3-sibling families alone under the equivalent assumption that in cases of two possible equilibria, there was a probability  $\pi$  that the equilibrium with the larger value of  $Y_{ih}$  would prevail. In the 3-sibling analysis, the likelihood function is analogous to that shown for the 2-sibling family in equations (15) through (19), and the relevant multivariate probit integrals can be computed by standard quadrature methods. Again, with fixed  $n_h = 3$ , this case provided us with a separate estimate of the parameter  $\theta$ , but not the parameter  $\alpha$ .

Third, our maximum likelihood analyses of 2-, and 3-sibling households, which constituted the vast majority of multi-sibling households, gave imprecise point estimates of the parameters  $\pi$  and  $\rho$ . We therefore investigated the robustness of our estimates of the structural parameters  $(\beta, \gamma, \theta)$  with respect to the values of  $\pi$  and  $\rho$ . For subsequent models that included households with 4 or more siblings, whose estimation was computationally burdensome, we then specified the fixed values of  $\pi$  and  $\rho$  that yielded the most conservative estimates of peer influence.

Fourth, in pooled analyses of all households with 1 to 4 siblings, we were able to identify not only the parameter  $\theta$ , but also the parameter  $\alpha$ . From this information, we could also recover the parameter  $\kappa = \theta + \alpha$ . Since  $\kappa > 0$  and

$\alpha < 0$ , the null hypothesis of symmetric social influence would mean that  $\kappa + \alpha = 0$  or, equivalently,  $\theta + 2\alpha = 0$ . To test this null hypothesis against the two-sided alternative, we computed the ratio of the likelihood for the constrained version of model (20), where  $\theta + 2\alpha = 0$ , divided by that for the unconstrained version of (20).

Fifth, in our pooled samples of households with 4 or fewer siblings, we formally tested the hypotheses that the peer-influence parameters  $\theta$  and  $\alpha$  depended on family size. In particular, we assumed the linear link functions:  $\theta = \theta_0 + \theta_1 (n_h - 1)$  and  $\alpha = \alpha_0 + \alpha_1 (n_h - 1)$ , so that the specification in equation (20) becomes

$$(21) \quad y_{ih} = 1 \text{ if } X_{ih}\beta + Z_h\gamma + \alpha_0 (n_h - 1) + \alpha_1 (n_h - 1)^2 + \theta_0 Y_{ih} + \theta_1 (n_h - 1) Y_{ih} + v_h + \delta_{ih} \geq 0, \\ \text{and } y_{ih} = 0 \text{ otherwise,}$$

From the parameter estimates in (21), we could then recover the dependence of the parameter  $\kappa$  on family size as well, that is,  $\kappa = (\theta_0 + \alpha_0) + (\theta_1 + \alpha_1)(n_h - 1)$ . In this context, to test the null hypothesis of symmetric social influence against the two-sided alternative, we computed the ratio of the likelihood for the constrained version of model (21), where both  $\theta_0 + 2\alpha_0 = 0$  and  $\theta_1 + 2\alpha_1 = 0$ , divided by that for the unconstrained version.

Sixth, for our pooled samples, we estimated the reduced-form version of the model (20), in which we assumed that both  $\theta = 0$  and  $\alpha = 0$ . That is,

$$(22) \quad y_{ih} = 1 \text{ if } X_{ih}\beta + Z_h\gamma + v_h + \delta_{ih} \geq 0, \text{ and} \\ y_{ih} = 0 \text{ otherwise,}$$

This special case corresponds simply to a random-effects probit model in which the parameter  $\rho$  is estimated. Comparison of the estimated parameters  $(\beta, \gamma)$  from the reduced form model in equation (22) with those from the structural model (20) permitted us to assess the “social multiplier” associated with within-family peer influence. In this context, we focused specifically on the contrast between the structural effect and the reduced-form effect of an increase in the real price of cigarettes.

Finally, we performed a number of sensitivity analyses. We analyzed a pooled sample of families with 2 to 4 siblings, excluding households with only one young person. We studied an alternative model to (21) where the peer-influence parameter  $\theta$  depended on a young person’s age. We also estimated the model of equation (20) under the dubious assumption that the variable  $Y_{ih}$ , the number of siblings of young person  $i$  who currently smoke, was exogenous.

## VI. EMPIRICAL RESULTS

In separate analyses of 2- and 3-sibling households, our estimates of the peer-influence parameter  $\theta$  were always positive and statistically significant. In

12,632 two-sibling households alone, we obtained  $\hat{\theta} = 0.521$  with a 95% confidence interval of  $[0.488, 0.554]$ , while in 1,786 three-sibling households alone, we observed  $\hat{\theta} = 0.445$  with a 95% confidence interval of  $[0.277, 0.613]$ . In both cases, the maximum likelihood estimates of the probability of multiple equilibria were quite small. In the 2-sibling case, the estimated probability that the disturbances  $(\varepsilon_1, \varepsilon_2)$  were contained in central rectangle in Figure 2 was only 0.013. In the 3-sibling case, the combined probability that the disturbances  $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$  were contained in one of the analogous regions in Figure 3 was 0.040. Empirically, the existence of multiple equilibria did not turn out to be a statistically significant problem but, as a result, our data were uninformative about the values of the parameters  $\pi$  and  $\rho$ . While the point estimates were  $\hat{\pi} = 1$  and  $\hat{\rho} = 0$ , the corresponding 95% confidence intervals included virtually all values in the feasible intervals (that is,  $0 \leq \pi \leq 1$  and  $0 \leq \rho \leq 1$ ). Put differently, the likelihood function was nearly flat when projected onto the  $(\pi, \rho)$  subspace.

Despite the uncertainty in  $\pi$  and  $\rho$ , we found that the estimates of the structural parameters (including, for example, the effect of cigarette prices) were nonetheless quite robust. The robustness for the peer-influence parameter  $\theta$  is specifically depicted in Figure 4 below. In the Figure, we re-estimated the model (including the parameters  $(\beta, \gamma, \theta, \rho)$ ) for various fixed values of the mixing parameter  $\pi$ . In both 2- and 3-sibling households, the estimates  $\hat{\theta}$  displayed only a relatively small degree of dependence on  $\pi$ . Moreover, the estimated

values of  $\theta$  were lowest when the parameter  $\pi$  was fixed at the maximum likelihood estimate of  $\hat{\pi} = 1$ .

[INSERT FIGURE 4 ABOUT HERE.]

Since the assumption that  $\pi$  and  $\rho$  equaled their respective point estimates  $\hat{\pi} = 1$  and  $\hat{\rho} = 0$  gave conservative estimates of the key parameter  $\theta$ , we retained this assumption in our subsequent analyses. In particular, for 308 four-sibling families alone, we obtained  $\hat{\theta} = 0.396$  with a 95% confidence interval of  $[0.318, 0.473]$ , while in 50 five-sibling households alone, we observed  $\hat{\theta} = 0.241$  with a 95% confidence interval of  $[0.080, 0.402]$ .

The column identified as equation (20) in Table 2 shows our estimates of the parameters  $\theta$  and  $\alpha$  in equation (20), as well as the estimated coefficient  $\gamma$  for cigarette price, in a pooled sample of 49,898 households with 1 to 4 siblings. Given maximum likelihood estimates of  $\hat{\theta} = 0.479$  and  $\hat{\alpha} = -0.150$ , we computed  $\hat{\kappa} = \hat{\theta} + \hat{\alpha} = 0.329$ . As Table 2 further shows, the null hypothesis of symmetric peer influence, namely,  $\kappa + \alpha = 0$ , was strongly rejected ( $P < 10^{-6}$ ). At the sample means of the right-hand side variables, each additional smoking sibling raised the probability of smoking by an estimated 7.6 percent, while each non-smoking sibling lowered the probability by an estimated 3.5 percent. Thus, the pro-smoking influence of a sibling who smokes is more than twice the deterrent effect of a non-smoking sibling.

The results for equation (20) in Table 2 permitted us to recover the remaining parameters in our learning model of Section III. In particular, the sensitivity  $\lambda$  of a peer's smoking as a test of its "coolness" was an estimated 0.366 (approximate 95% confidence interval, 0.335–0.397), while the specificity  $\mathbf{w}$  was an estimated 0.737 (approximate 95% confidence interval, 0.701–0.772). Put differently, if smoking is "uncool," then Bea believes that there is about a 74% probability that her brother Jeff will not smoke. But if smoking is "cool," then there is only about a 37% probability that Jeff will smoke.

[INSERT TABLE 2 ABOUT HERE.]

In our analysis of subsets of households of different size, we found that the estimated values of  $\theta$  were inversely related to the number of young people. We therefore investigated a model, specified in equation (21), wherein both the peer-influence parameters  $\theta$  and  $\alpha$  depended on family size through the linear link functions  $\theta = \theta_0 + \theta_1 (n_h - 1)$  and  $\alpha = \alpha_0 + \alpha_1 (n_h - 1)$ . The parameter estimates for the resulting model, based upon the pooled sample of 49,898 households with 1 to 4 siblings are also shown in Table 2. We found that while the parameter  $\theta$  declined with the number of siblings, the parameter  $\alpha$  did not. Thus, as the number of siblings in the household increased, the pro-smoking influence of any one smoking sibling was diluted, while the anti-smoking influence of a non-smoking sibling was not clearly dependent on family size. In particular, in a two-sibling family (that is,  $n_h = 2$ ), the estimated values of  $\theta$  and



$\alpha$  were 0.513 and  $-0.127$ , respectively, and the corresponding values of the sensitivity  $\lambda$  and specificity  $\omega$  were 0.297 and 0.798, respectively. In a four-sibling family, the estimated values of  $\theta$  and  $\alpha$  were 0.387 and  $-0.153$ , respectively, and the corresponding values of the sensitivity  $\lambda$  and specificity  $\omega$  were 0.442 and 0.650, respectively. As in equation (20), our likelihood ratio test strongly rejected the null hypothesis of symmetric peer influence in the pooled sample of families all sizes. Moreover, the values of  $\theta$  in models where the parameters were estimated separately for subsets of households of different sizes were indistinguishable from the values estimated from equation (21) in the pooled sample of households of all sizes. (Results not shown.)

By way of sensitivity analysis, we found that the influence parameter  $\theta$  did not vary significantly with each sibling's age. (Results not shown.) When we used standard random-effects probit methods to estimate equations (20) and (21) under the dubious assumption that the number of smoking siblings  $Y_{ih}$  was exogenous, we obtained estimates of  $\theta \approx 0.90$  and  $\alpha \approx -0.22$ , values that (in absolute terms) significantly overstated those shown in Table 2. (Detailed results not shown.)

The right-most column of Table 2 shows the estimate of the coefficient for cigarette price in the reduced-form model of equation (22). The estimated coefficient of price in the reduced form model was 1.6 times the value derived from the structural models of (20). At the sample means of the right-hand side variables, each \$1-increase in the real price of a carton of cigarettes (or 10-cent

increase in the price per pack) reduced the probability of smoking in the reduced-form model (20) by 0.0060, while it reduced the probability of smoking in the structural model (21) by only 0.0038. Thus, we found that the marginal deterrent effect of a price increase on the probability of smoking, as derived from the reduced form model, was likewise about 60 percent greater than the marginal effect of price derived from the structural model. This comparison gives us an estimate of the social multiplier effect of peer influence within the household.

The full set of parameter estimates for equations (20) through (22) in Table 2 are shown in the Results Appendix. In particular, we found that the effect of a household adult's smoking is only about two-thirds that of a younger peer's smoking. In separate sensitivity analyses, we rejected the possibility that older siblings had more influence than younger ones. (Results not shown.) We also re-estimated our pooled analyses with a subset of 14,781 families with 2, 3, or 4 siblings, but the results were not markedly different from those reported here. (Results also not shown.)

## VII. DISCUSSION AND CONCLUSIONS

Our analyses of smoking practices among young people aged 15–24 in a large nationwide sample of U.S. households polled during the 1990s strongly confirmed the hypothesis of asymmetric social influence. We found that the pro-smoking influence of a sibling who smokes is more than twice the deterrent

effect of a non-smoking sibling. Moreover, the pro-smoking influence of any particular smoking peer was partially diluted in households with larger numbers of young people, but even in households with up to 5 young residents, the pro-smoking social influence of a smoker continued to dominate the deterrent effect of a nonsmoker.

Our results can be interpreted in terms of a model of information externalities, in which each young person, who is uncertain about the risks and benefits of smoking, learns about such risks and benefits from peers' smoking practices. In particular, a young person employs the information contained in each of her peer's smoking practices to update her prior probability that smoking is safe, non-addictive, and the socially acceptable, in short, the "cool" thing to do. In this process of Bayesian updating, each peer's smoking becomes, in effect, a diagnostic test with its own sensitivity and specificity. In this context, we estimated that, on average, the sensitivity of a peer's smoking was 37 percent, while the specificity was 74 percent. Put differently, young Bea believes that if smoking were in fact "uncool," then there would be only about a one-in-four chance that her brother Jeff would smoke anyway. But if smoking were in fact "cool," then she believes that there would be a nearly two-thirds chance that Jeff could still be a nonsmoker.

The dilution of a smoker's influence in large families can likewise be expressed in terms of sensitivity and specificity. If Jeff were Bea's only sibling, then we estimated that Bea would attach about a 30-percent sensitivity and an

80-percent specificity to his smoking. By contrast, if Jeff were just one of Bea's three siblings, then the sensitivity of Jeff's smoking would rise to about a 44 percent, while the specificity would fall to about 65 percent. In the medical literature, repeated testing alters an individual test's accuracy when the tests are not independent, often because the populations of persons with and without the disease are heterogeneous (Dendukuri and Joseph [2001], Qu and Hadgu [1998], Torrance-Rynard and Walter [1997]). In the present context, our findings concerning diluted influence suggest that young people in fact understand the interdependence of their peers' smoking practices. Thus, if Pepe already smokes, Jeff's smoking is less influential for Bea because she recognizes that Jeff is smoking, at least in part, because Pepe is.

Our model of asymmetric information externalities could be applied to many other fields, especially financial markets, where traders infer good and bad news from the upward and downward price movements of securities (Bikhchandani, Hirshleifer, and Welch [1992, 1998] and Lux [1998]). It has obvious applications in other contexts, including local crime rates (Glaeser, Sacerdote, and Scheinkman [1996]) and school performance (Sacerdote [2001], Hoxby [2002], and Winston and Zimmerman [2003]). In the classroom, for example, the behavior of good students who "set an example" may be highly informative about the value of regular studying and attendance, while the behavior of poor students may be uninformative. In the present context, our analysis cautions those who might distinguish too sharply between "peer

influence” and “knowledge” theories of youth smoking initiation (U.S. Department of Health and Human Services [1994]).

Comparing the results from a reduced form model to our structural estimates, we estimated that the aggregate deterrent effect of an increase in cigarette price on youth current smoking prevalence was approximately 60 percent greater than the sum of the deterrent effects on each individual separately. Our analysis thus provides an important explanation for the apparently large price responsiveness of youth smoking found by numerous researchers (e.g., Chaloupka and Pacula [1999], Chaloupka and Wechsler [1995], Emery, White, and Pierce [2001], Harris and Chan [1999], Lewit and Coate [1982]). Although the static equilibrium model employed in this paper (the “family smoking game”) allowed for asymmetric peer influence, it did not by itself generate asymmetric social multipliers. While we might consider more complex static models, or even dynamic models of youth smoking, we leave the modeling and estimation of such asymmetric social multipliers to later work. Nonetheless, asymmetric social multipliers may be an important source of asymmetric responses of demand to increases and reductions in cigarette taxes and other measures intended to reduce youth smoking.

Our structural model of peer influence entailed a system of simultaneous equations in discrete endogenous variables. While this approach permitted us to address a number of problems in the analysis of peer effects, particularly the problem of reflection (Manski [1993]), such models are known to be susceptible

to the problem of multiple equilibria. Fortunately, in our applied work with households of 2 and 3 young people, which constituted the vast majority of those in our sample, those regions within the disturbance space where multiple equilibria could obtain had too small a probability to influence our estimates of the key structural parameters. Our results thus offer an illustration of a theorem by Tamer [2002] that when the data are sufficiently rich, parameters can be identified even when the model is, strictly speaking, incomplete.

Other researchers have attempted to apply non-Nash equilibrium concepts to the analysis of small-group data (e.g., Chao [2002], Bjorn and Vuong [1997], and Kooreman [1994]). In the context of a couple's decision about contraception, studied by Chao [2002], for example, there is a natural asymmetry between man and woman (and the contraceptive pill and male condom) that lends itself to a Stackelberg specification, in which one "player" can commit to moving first. In our context, one might conjecture that an older sibling could pre-commit to cigarette smoking by virtue of its addictive propensity. However, there is evidence that addiction to cigarettes develops gradually along a continuum during youth (Harris and Chan [1999]) and, in any case, we found no evidence that older siblings had a greater peer influence than younger ones.

Our empirical analysis is subject to a number of limitations. We were able to attach specific cigarette prices from the ACCRA database to only about 64 percent of the households polled about smoking in the Current Population Survey. However, we have no evidence that the excluded households differed

materially with respect to any factor that might determine youth smoking rates. We are conscious that state-level cigarette prices may be endogenous because the overall prevalence of smoking among adults (as well as young people) in a state may influence the political acceptability of a hike in the state cigarette excise taxes. However, we see the ACCRA prices, which are specific to metropolitan areas and non-metropolitan counties, as less susceptible to such criticism.

We are likewise aware that the Tobacco Use Supplements to the CPS during the 1990s contained more than 30 percent of proxy responses. Since the CPS is designed so that one proxy typically answers for multiple respondents in the same household, we see potential biases in proxy responses primarily as a “common shock” within the household, which we specifically addressed in our mixed-effects error specification. What is more, we assumed that while youth smoking was endogenous within the household, the smoking practices of adults were exogenous. While there is a literature on the influence of parental quitting on youth smoking (e.g., Farkas et al. [1999]), less is known about the reverse. Thus, the peer influence of adults in the household could have mediated some of the social influence among young people that we measured here. Moreover, our data did not permit us to address the global influence of peer groups outside the home. We have confirmed only that, at least within the household, peer effects can be highly local.

In our econometric analyses of 2- and 3-sibling households, we were able to compute maximum likelihood estimates even when the disturbance space

contained regions with multiple equilibria, and even when the disturbances were not independent. In these small households, the number of such regions and the variety of multiple equilibria were limited, and the multivariate probit integrals could be computed reliably by quadrature methods. However, as the number of players in the “family smoking game” increased beyond three, the rapidly growing computational burden forced us to impose restrictions on the possible equilibria. For larger groups with very large numbers of possible equilibria, simulation methods, such as those suggested by Krauth [2001] and Tamer [2002, 2003] may prove necessary.

Nonetheless, our results validate the use of statistical methods that properly account for the endogeneity of peer behaviors. Thus, when we estimated our structural model under the dubious assumption that siblings’ smoking practices were exogenous, we obtained estimates of peer influence that substantially exceeded those obtained when we treated such behaviors as endogenous. Our results thus illustrate why economists cannot ignore the problem of reflection, originally articulated by Manski [1993], in studies of peer effects.

The implications of asymmetry in social interactions for the study of youth smoking and other fields of economic inquiry are substantial. If the pro-smoking social influence of a smoker exceeds the deterrent influence of her nonsmoking counterpart, then policies to segregate smokers will reduce overall smoking rates. If the effects of price changes are asymmetric, then the price war



in the early 1990s in the highly concentrated U.S. cigarette market may have enhanced youth smoking, dollar for dollar, more than the price increases effectuated by excise tax hikes or settlement payments later in the same decade. What is more, those who would seek to deter youth smoking through higher cigarette taxes, antismoking messages, or other restrictions need to ensure that such policies are sustained.

## TABLES

Table 1. Prevalence of Current Cigarette Smoking among Young People,  
Aged 15-24, in Relation to Current Smoking Status of Other Young  
Household Members, Aged 15-24<sup>a,b,c</sup>

|  |   | Number of Siblings, Aged 15-24, Who Do Not Smoke |        |       |       |
|--|---|--|--------|-------|-------|
|  |   | 0  | 1      | 2     | 3     |
| Number of Siblings,<br>Aged 15-24, Who Smoke | 0 | 19%  | 12%    | 8%    | 5%    |
|  |   | 62,817   | 37,561 | 7,646 | 1,724 |
|  |   | 62,817   | 18,781 | 2,549 | 431   |
|  |   | -1.45  | -1.99  | -2.44 | -2.94 |
|  | 1 | 50%  | 34%    | 28%   | 17%   |
|  |   | 8,859  | 1,746  | 366   | 72    |
|  |   | 4,430  | 582    | 92    | 14    |
|  |   | 0.00   | -0.66  | -0.94 | -1.59 |
|  | 2 | 63%  | 49%    | 54%   | 0%    |
|  |   | 808  | 201    | 39    | 16    |
|  |   | 269  | 50     | 8     | 3     |
|  |   | 0.53   | -0.04  | 0.16  | -     |
|  | 3 | 73%  | 53%    | 100%  | -     |
|  |   | 121  | 30     | 4     |       |
|  |   | 30   | 6      | 1     |       |
|  |   | 0.99   | 0.12   | -     |       |

## Notes to Table 1

- a. Source: Combined Tobacco Use Supplements of the CPS during September 1992 – May 1993, September 1995 – May 1996, and September 1998 – May 1999. (U.S. Bureau of the Census [1996, 1998]).
- b. The four entries in each cell are, respectively: the percentage of young people, aged 15–24, who are current smokers; the total number of young people in the cell; the total number of households in the cell; and the log odds of the prevalence of current smoking, that is,  $\ln\left(\frac{p}{1-p}\right)$ , where  $p$  is smoking prevalence.
- c. A current smoker is a respondent who answered “some days” or “every day” to the question, “Do you smoke cigarettes some days, every day, or not at all?” Among 126,352 respondents with known smoking status, we excluded 3,809 who resided in families where the smoking status of at least one other sibling was unknown. Not shown are the data for another 533 respondents who resided in families with more than 3 nonsmoking siblings or more than 3 smoking siblings.

Table 2. Estimates of the Peer-Influence Parameters and the Coefficient of Cigarette Price in a Pooled Sample of 49,898 Households with 1 to 4 Siblings.

| Parameter                         | Description   | Social Interaction  |   | No Social Interaction      |
|-----------------------------------|---|---|---|----------------------------|
|                                   |   | Eq. (20)  | Eq. (21)  | Eq. (22)                   |
| $\gamma$                          | Coefficient of Price                                  | -0.016<br>[-0.023, -0.010]                                | -0.016<br>[-0.023, -0.010]                                | -0.026<br>[-0.036, -0.016] |
| $\theta$                          | Coefficient of $Y_h$<br>(# Smoking Sibs in Household) | 0.479<br>[0.455, 0.503]                                   |   |                            |
| $\alpha$                          | Coefficient of $n_h - 1$ (# Sibs in Household)        | -0.150<br>[-.170, -0.130]                                 |   |                            |
| $\theta_0$                        | Intercept of $\theta = \theta_0 + \theta_1(n_h - 1)$  |   | 0.576<br>[0.517, 0.635]                                   |                            |
| $\theta_1$                        | Slope of $\theta = \theta_0 + \theta_1(n_h - 1)$      |   | -0.063<br>[-0.098, -0.027]                                |                            |
| $\alpha_0$                        | Intercept of $\alpha = \alpha_0 + \alpha_1(n_h - 1)$  |   | -0.114<br>[-.161, -0.067]                                 |                            |
| $\alpha_1$                        | Slope of $\alpha = \alpha_0 + \alpha_1(n_h - 1)$      |   | -0.013<br>[-.037, 0.011]                                  |                            |
| $\kappa$                          | $\theta + \alpha$                                     | 0.329<br>[0.308, 0.350]                                   |   |                            |
| $\kappa_0$                        | $\theta_0 + \alpha_0$                                 |   | 0.462<br>[0.410, 0.514]                                   |                            |
| $\kappa_1$                        | $\theta_1 + \alpha_1$                                 |   | -0.076<br>[-0.104, -0.048]                                |                            |
| LLR* Test for Symmetric Influence |   | LLR= -52.75<br>$\chi^2 = 105.5$<br>(1df)<br>$P < 10^{-6}$ | LLR=-61.92<br>$\chi^2 = 123.84$<br>(2df)<br>$P < 10^{-6}$ |                            |

\*LLR = log Likelihood Ratio

## FIGURES

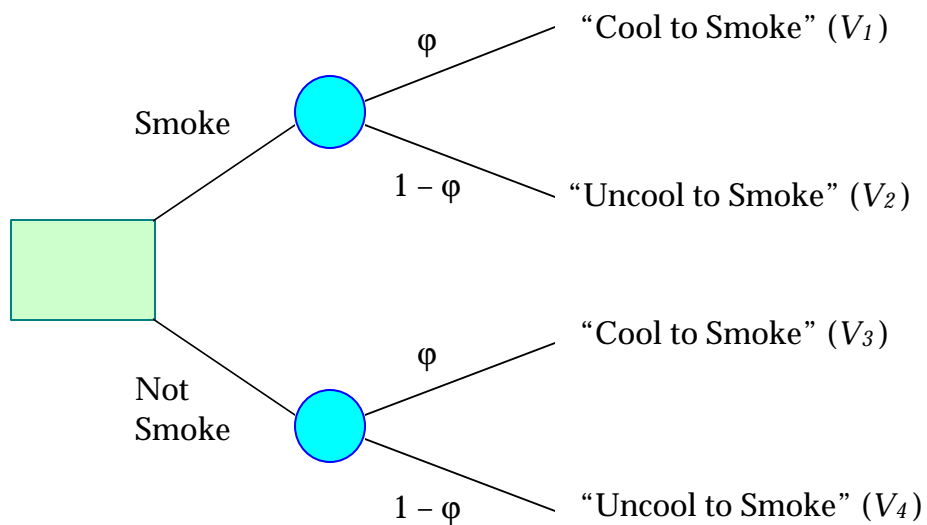
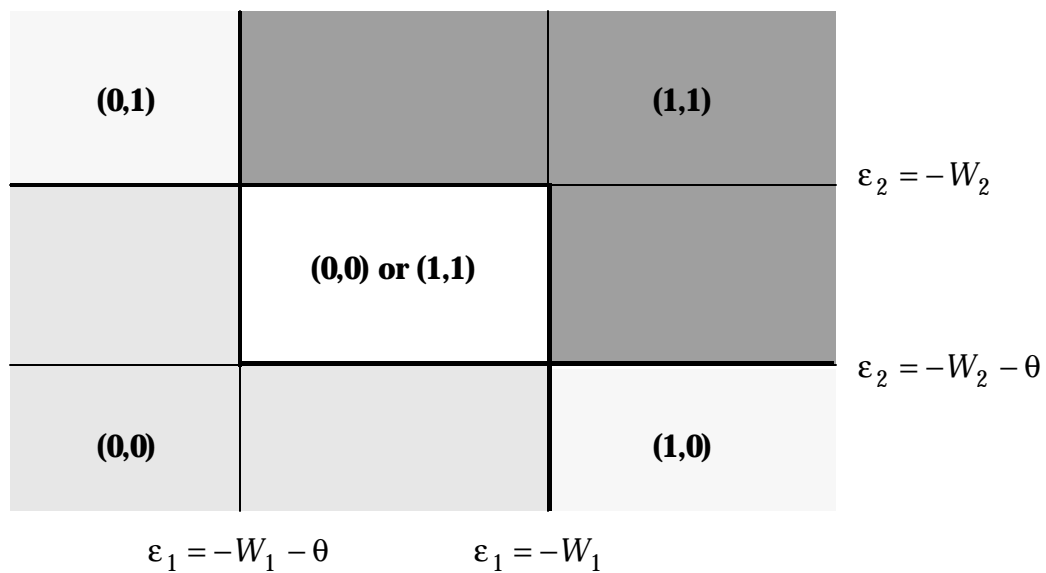


Figure 1. Young Person's Smoking Decision Tree

Figure 2. Equilibria in the  $(\epsilon_1, \epsilon_2)$  Plane in the 2-Sibling Family Smoking Game

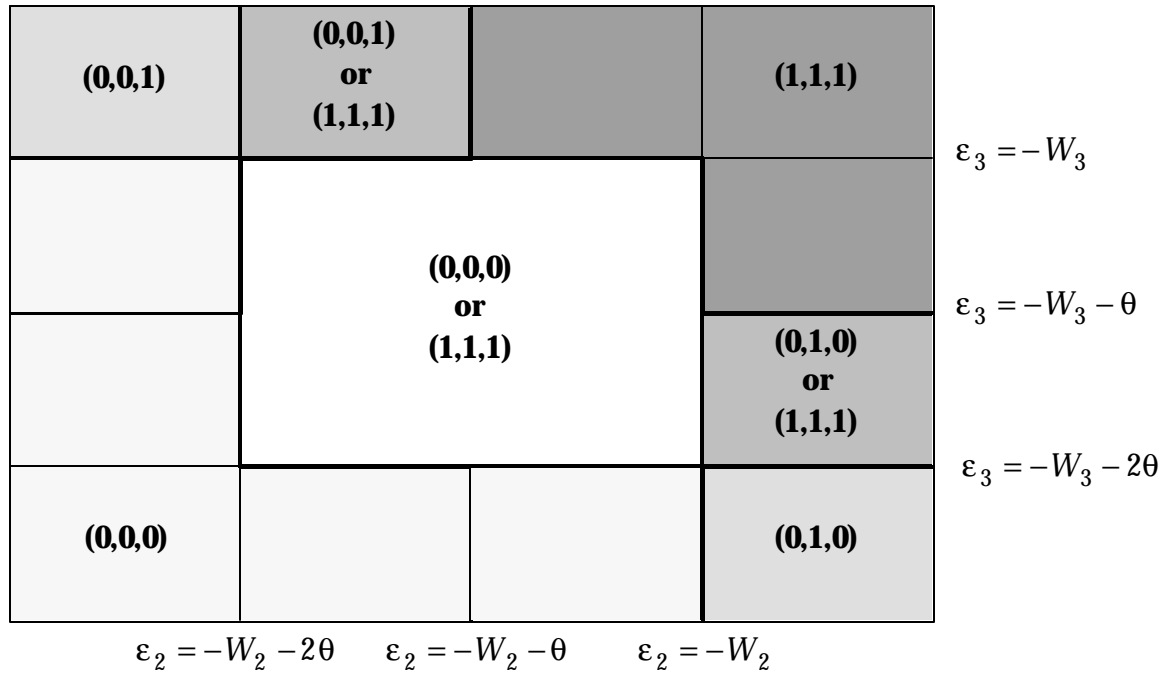


Figure 3. Equilibria in the  $(\varepsilon_2, \varepsilon_3)$  Plane in the 3-Sibling Family Smoking Game When  $-W_1 - \theta > \varepsilon_1 \geq -W_1 - 2\theta$

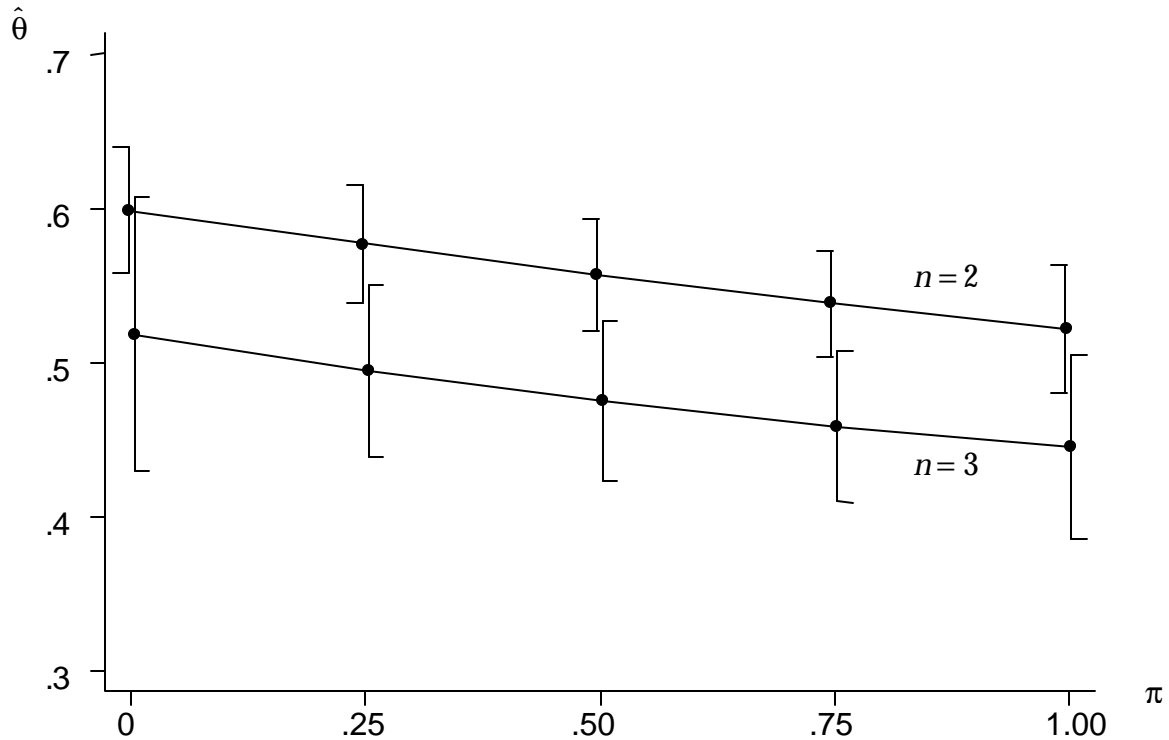


Figure 4. Relation between the Mixing Parameter  $\pi$  and the Estimate of  $\theta$  for Households with 2 and 3 Siblings. (Error bars represent 95% confidence intervals.)

## DATA APPENDIX

We analyzed data from the Tobacco Use Supplements (TUS) to the Current Population Survey (CPS), a nationally representative household-based survey of the U.S. population aged 15 years or more (U.S. Bureau of the Census [1996, 1998] ). The TUS has an extended history of use in the economic analysis of cigarette smoking. (See, e.g., Marcus et al. [1989], Shopland et al. [1996], Gerlach et al. [1997], Arday et al. [1997], Ohsfeldt, Boyle, and Capilouto [1998], Hersch [2000], Sweeney et al. [2000], and Shopland et al. [2001]. For TUS-based studies of adolescent and young-adult smoking, see also Cummings and Shah [1995], Gilpin et al. [1999], Farkas et al. [1999], Anderson and Burns [2000], and Gilpin et al. [2000].) During the 1990s, the TUS was administered during nine monthly installments of the CPS, which we grouped into three successive “waves” as follows: (1) September 1992, January 1993 and May 1993; (2) September 1995, January 1996 and May 1996; and (3) September 1998, January 1999 and May 1999.

To determine an individual’s current smoking status, the TUS asked each self-respondent or, if unavailable, a proxy respondent, “Has ... smoked at least 100 cigarettes in his/her entire life?” If the answer was affirmative, then the TUS further queried, “Does ... now smoke every day, some days, or not at all?” A current smoker answered “yes” to the first question and either “every day” or “some days” to the second, while a non-smoker answered either “no” to the first question or “not at all” to the second.

Out of 1,037,644 individual TUS records for all three waves combined, a total of 145,783 (or 14 percent) represented persons aged 15–24 years. Of these, 126,352 records (or 87 percent) gave informative responses to the two aforementioned questions about smoking status. (Virtually all of the remaining 19,431 non-responses came from proxy respondents.) This data set of 126,352 responses among young person served as the basis for the calculations shown in Table 1.

Among the 126,352 young people with known current smoking status, the overall prevalence of current smoking was 18.6 percent, with 14.2 percent smoking “every day” and 4.4 percent smoking “some days.” The prevalence of current smoking was 21.2% among self-respondents and 14.3% among proxy respondents, who constituted 36.9 percent of the responses with known current smoking status. Of the 126,352 young people with known current smoking status, we then excluded 3,809 individuals (3 percent) who resided in households where the current smoking status of one or more siblings was unknown. That left a universe of 122,543 respondents in 81,662 households where each young person’s smoking status was known.

Because the CPS was a household-based survey, we were thus able observe the smoking practices of all young people residing in each living unit. In contrast to other studies where only one subject self-reports the smoking habits of his peers, we could thus make inferences about the reciprocal influences of the peer group on each of its members. The Teenage Attitudes and Practices Survey,



analyzed by Krauth [2001], asked the respondent, “How many of your four best friends smoke?” The respondent’s “best friends,” however, were unlikely to be in the sample. Both the U.K. General Household Survey and the British Health and Lifestyle Survey, analyzed by Jones [1989, 1994], asked the respondent to report whether there are “other smokers” in the household, but it is unclear whether these “other smokers” participated as well. By contrast, the National Education Longitudinal Study, a school-based analyzed by Gaviria and Raphael [2001], asked each tenth-grader about his own smoking, drinking and illegal drug use. Gaviria and Raphael then inquired whether a student’s risky behaviors were related to the average for those schoolmates, from 5 to 43 in number, who also participated in the survey. Farkas et al. [1999] used the 1992–1993 wave of the TUS to test the hypothesis that 15-17-year-olds were less likely to smoke when their parents had quit. In the Health and Retirement Survey, Lahiri and Song [2000] related the respondent’s smoking status to his spouse’s smoking.

Although the CPS contained information on each subject’s gender, age, educational attainment, work status, ethnicity, income, and other individual and family characteristics, it did not include data on cigarette prices. For this purpose, we used data from contemporary surveys of retail cigarette prices collated by ACCRA [2000], formerly the American Chamber of Commerce Researchers’ Association. Other studies relying on ACCRA cigarette price data have included Chaloupka and Wechsler [1995] and Arcidiacono, Sieg, and Sloan

[2001]. The ACCRA data represented the retail prices, including all applicable excise and sales taxes and promotional discounts, of a specific premium-priced brand of cigarettes. Although discount cigarettes claimed a market share exceeding 25 percent during 1992–1999, such non-premium brands were rarely the choice of young smokers (Centers for Disease Control [1994, 2000]).

In contrast to the annual state-level prices compiled originally by the Tobacco Institute (TI) and updated by Orzechowski and Walker [2002], the ACCRA prices were sampled quarterly for metropolitan statistical areas (MSAs) and non-metropolitan areas of population 50,000 or more within each state. We were thus able to match each respondent with the retail price prevailing in the metropolitan or non-metropolitan area of the household's location and in the calendar quarter in which the survey was performed. (For households residing outside of MSAs, the CPS recorded only the state of residence. Those households were assigned the average price for all non-metropolitan areas in the state.) Our tests of the concordance of the TI state-level prices and ACCRA retail prices gave correlation coefficients between 0.89 for 1992 and 0.95 for 1998. For a given year, the between-state variability of the ACCRA prices was slightly lower than that of the TI prices. This finding may reflect the more complete inclusion of promotional discounts in the ACCRA retail price data.

We did not make any adjustments for cross-border shopping. In attempts to take account of this phenomenon, some researchers have included neighboring states' prices (Baltagi and Levin [1986]). Others have computed a

weighted-average of the local price and the lowest nearby price (Chaloupka [1991]). Still others have limited the sample to respondents living at least 20 miles from states with lower taxes (Lewit and Coate [1982], Wasserman et al. [1991], Chaloupka and Wechsler [1995], Pacula and Chaloupka [1999]), or have made other adjustments (Chaloupka and Pacula [1999]). Studies of teenagers and college students have found little effect of such price adjustments on the estimated demand for cigarettes, perhaps because such young people have lower mobility.

Because the geographic areas covered by the ACCRA price survey did not entirely overlap those covered by the CPS, we could not assign a price to each young person in the TUS sample. In particular, ACCRA provided price data on only 194 out of the 280 MSAs that the CPS surveyed. Out of 122,543 subjects who lived in households where each young person's smoking status was known, we could therefore assign prices to 78,038 (or 63.7 percent). The coverage was 62.8 percent of the CPS-surveyed residents of MSAs and 69.1 percent of the residents of non-metropolitan areas. The prevalence of current smoking among the 78,038 remaining subjects in our pooled cross-sectional sample was 18.5 percent, while the smoking prevalence among the excluded individuals was 18.6 percent.

Of the 78,038 young people with known smoking status, known sibling smoking status and known ACCRA-based price, we excluded an additional 10,570 subjects (or 13.5%) who had missing data on any of the remaining

explanatory variables used in the analysis. This left a total of 67,468 complete observations in our analytical sample. In that sample, 34,557 individuals (or 51.2%) lived in households with no other young people (that is,  $n_h = 1$  in equation (20) above). A total of 25,743 individuals (or 38.2%) lived in 2-sibling households; 5,519 (or 8.2%) lived in 3-sibling households; 1,275 (or 1.9%) lived in 4-sibling households; 264 (or 0.4%) lived in 5-sibling households; and 110 (or 0.2%) lived in households with 6 or 7 young people.

The variables employed in our analysis are shown in Appendix Table A1. In addition to current smoking status (corresponding to  $y_{ih}$  in equation (20) above), we have classified the explanatory variables as individual-level (corresponding to  $X_{ih}$ ), and household-level (corresponding to  $Z_h$ ). The latter subset included variables included area-level prices and wave-specific temporal variables. Attempting to distinguish between age-related, cohort and contemporaneous effects, we tried various combinations of respondent's age and dummy variables for the respondent's year of birth and the date of interview. Other variables, such as sex, academic achievement, employment status, ethnicity, are well documented in the literature. Other researchers have measured the influence of adult smoking (Farkas et al. [1999], Lahiri and Song [2000]), as well as restrictions on smoking at home (Proescholdbell, Chassin, and MacKinnon [2000], Farkas et al. [2000]).

In Table A1, we specifically included a measure of the 3-year lagged price of cigarettes for young people aged 18–24 years in the 1995-96 wave, who would have been aged 15–21 in the 1992-93 wave. (See Douglas and Hariharan [1994] for a similar use of past cigarette prices that prevailed during the prime years of smoking initiation.) We did not include a measure of future prices, as might be suggested by theories of “rational” or non-myopic addiction (Becker, Grossman, and Murphy [1994]). As noted by Pacula and Chaloupka [1999] and Gruber [2001], young people appear to make myopic smoking decisions because they have high discount rates or do not fully appreciate the future health consequences of smoking. In any case, it is difficult to see how a non-myopic teenager interviewed in the first two months of wave 1 (September 1992 – Jan 1993) could have forecast the industry-wide price war that brought retail prices down later in 1993. Nor is it obvious how a “rational” teenager interviewed in waves 1 or 2 could have forecast the subsequent avalanche of lawsuits, culminating in the Master Settlement Agreement of 1998, that pushed prices up in wave 3.

Table A1. Sample Means and Standard Deviations of Variables for the 67,468 Individuals, Aged 15–24, in Whom All Variables were Observed

| Variable  | Sample Mean | Sample Standard Deviation |
|---|-------------|---------------------------|
| <i>Individual-level variables</i>   |             |                           |
| Current smoker (binary)   | 0.183       |                           |
| Age (years)   | 19.332      | 2.942                     |
| Sex female (binary)   | 0.514       |                           |
| Married (binary)  | 0.125       |                           |
| Did not finish high school (binary)   | 0.086       |                           |
| Full-time student (binary)  | 0.429       |                           |
| Currently working (binary)  | 0.519       |                           |
| Unemployed (binary)   | 0.076       |                           |
| No parent at home (binary)  | 0.352       |                           |
| Restrictions on smoking at work (binary)  | 0.181       |                           |
| Proxy respondent (binary)   | 0.359       |                           |
| <i>Household-level variables</i>  |             |                           |
| Number of siblings (integer)  | 0.626       | 0.770                     |
| Black (binary)  | 0.122       |                           |
| Asian (binary)  | 0.030       |                           |
| Hispanic (binary)   | 0.122       |                           |
| Number of persons aged >24 in household who smoke (integer)   | 0.428       | 0.663                     |
| ln (Real family annual income)¶   | 9.706       | 0.984                     |
| Real price per 10-pack carton ¶   | 13.105      | 2.549                     |
| 1995-96 Wave (binary)   | 0.265       |                           |
| 1998-99 Wave (binary)   | 0.312       |                           |
| Lagged real price §   | -0.030      | 0.778                     |
| ¶ Deflated to constant 1992 dollars based on the Consumer Price Index.  |             |                           |
| § For respondents who were ≥ 18 years old in the 1995-96 Wave, this variable is computed as $\left( \frac{p_{t-3} - p_t}{p_{t-3}} \right)$ , where $p_t$ is the current real price and $p_{t-3}$ is the 3-year lagged real price. For all other respondents, the variable equals 0. |             |                           |

## RESULTS APPENDIX

Table B1 shows the complete results for the three models (that is, equations 20, 21 and 22) described in Table 2 in the main text. Each cell contains the corresponding maximum likelihood parameter estimate, below which is shown the asymptotic standard error in parentheses.

Table B1. Detailed Estimates of Equations (20), (21) and (22)

| Variable or Parameter                                       | Social Influence  |                        | No Social Influence |
|---|-------------------|------------------------|---------------------|
|   | Eq. (20)          | Eq. (21)               | Eq. (22)            |
| Age (years)   | .0454<br>(.0028)  | .0459<br>(.0028)       | .0659<br>(.0041)    |
| Sex female (binary)   | -.1424<br>(.0125) | -.1411<br>(.0125)      | -.2022<br>(.0175)   |
| Married (binary)  | -.2350<br>(.0195) | -.2441<br>(.0196)      | -.3437<br>(.0301)   |
| Did not finish high school (binary)                         | .5114<br>(.0204)  | .5119554<br>(.0203588) | .7128<br>(.0301)    |
| Full-time student (binary)                                  | -.2973<br>(.0143) | -.2974<br>(.0143)      | -.4057<br>(.0207)   |
| Currently working (binary)                                  | .2688<br>(.0160)  | .2667<br>(.0160)       | .3808<br>(.0231)    |
| Unemployed (binary)   | .4338<br>(.0230)  | .4319<br>(.0230)       | .5979<br>(.0331)    |
| No parent at home (binary)                                  | .3251<br>(.0179)  | .3239<br>(.0179)       | .5196<br>(.0272)    |
| Restrictions on smoking at work (binary)                    | -.2197<br>(.0181) | -.2199<br>(.0181)      | -.2993<br>(.0256)   |
| Proxy respondent (binary)                                   | -.1846<br>(.0147) | -.1855<br>(.0147)      | -.2524<br>(.0210)   |
| Black (binary)  | -.6331<br>(.0226) | -.6304<br>(.0226)      | -.9636<br>(.0360)   |
| Asian (binary)  | -.3627<br>(.0409) | -.3603<br>(.0409)      | -.5463<br>(.0619)   |
| Hispanic (binary)   | -.5827<br>(.0219) | -.5789<br>(.0219)      | -.8524<br>(.0343)   |
| Number of persons aged >24 in household who smoke (integer) | .3139<br>(.0093)  | .3154<br>(.0093)       | .4852<br>(.0155)    |
| ln (Real family annual income) <sup>¶</sup>                 | -.1305<br>(.0070) | -.1304<br>(.0070)      | -.2112<br>(.0112)   |
| Real price per 10-pack carton <sup>¶</sup>                  | -.0164<br>(.0032) | -.0163<br>(.0032)      | -.0261<br>(.0050)   |
| 1995-96 Wave (binary )                                      | .0127<br>(.0191)  | .0141<br>(.0191)       | .0280<br>(.0285)    |
| 1998-99 Wave (binary)                                       | .1000<br>(.0158)  | .1006<br>(.0158)       | .1487<br>(.0244)    |



(Table B1 continued)

| Variable or Parameter                                     | Social Influence  |                   | No Social Influence |
|---|-------------------|-------------------|---------------------|
|   | Eq. (20)          | Eq. (21)          | Eq. (22)            |
| Lagged real price §                                       | -.0832<br>(.0602) | -.0762<br>(.0603) | -.0355<br>(.0826)   |
| Constant term   | -.4525<br>(.0924) | -.4745<br>(.0926) | -.4454<br>(.1398)   |
| $\theta$ (eq. 20) or $\theta_0$ (eq. 21)                  | .4791<br>(.0124)  | .5758<br>(.0181)  |                     |
| $\theta_1$ (eq. 21)                                       |                   | -.0629<br>(.0181) |                     |
| $\alpha$ (eq. 20) or $\alpha_0$ (eq. 21)                  | -.1501<br>(.0103) | -.1135<br>(.0240) |                     |
| $\alpha_1$ (eq. 21)                                       |                   | -.0131<br>(.0120) |                     |
| Correlation Coefficient of Within-Household Disturbances* | 0                 | 0                 | .5222<br>(.0125)    |

¶ Deflated to constant 1992 dollars based on the Consumer Price Index.

§ For respondents who were  $\geq 18$  years old in the 1995-96 Wave, this variable

is computed as  $\left( \frac{p_{t-3} - p_t}{p_{t-3}} \right)$ , where  $p_t$  is the current real price and  $p_{t-3}$  is the

3-year lagged real price. For all other respondents, the variable equals 0.

\* For equations (20) and (21), the correlation coefficient was set equal to zero, based on separate maximum likelihood estimates for 2- and 3-sibing households.

## REFERENCES

- ACCRA, *ACCRA Cost of Living Index Manual*(Arlington VA: ACCRA, 2000).
- Alexander, C., M. Piazza, D. Mekos, and T. Valente, "Peers, schools, and adolescent cigarette smoking," *J Adolesc Health*, 29 (2001), 22-30.
- Anderson, C., and D.M. Burns, "Patterns of Adolescent Smoking Initiation Rates by Ethnicity and Sex," *Tobacco Control*, 9 (2000), 114-118.
- Arcidiacono, P., H. Sieg, and F. Sloan. 2001. Living Rationally under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking. Cambridge MA: National Bureau of Economic Research Working Paper No. 8602.
- Arday, D.R., S.L. Tomar, D.E. Nelson, R.K. Merritt, M.W. Schooley, and P. Mowery, "State Smoking Prevalence Estimates: A Comparison of the Behavioral Risk Factor Surveillance System and Current Population Surveys," *American Journal of Public Health*, 87 (1997), 1665-1669.
- Baltagi, B.H., and D. Levin, "Estimating Dynamic Demand for Cigarettes Using Panel Data: The Effects of Bootlegging, Taxation, and Advertising Reconsidered," *Review of Economics and Statistics*, 68 (1986), 148-155.
- Becker, Gary S., Michael Grossman, and Kevin M. Murphy, "An empirical analysis of cigarette addiction," *American Economic Review*, 84 (1994), 396-418.

Becker, Gary S., and Kevin M. Murphy, "A theory of rational addiction.," *Journal of Political Economy*, 96 (1988), 675-700.

Berry, Steven T., "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, LX (1992), 889-917.

Bikhchandani , Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, C (1992), 992-1026.

Bikhchandani , Sushil, David Hirshleifer, and Ivo Welch, "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, 12 (1998), 151-170.

Bjorn, P.A., and Q.H. Vuong, "Modeles d 'equations simultanees pour variables endogenes fictives: Une formulation par la theorie des jeux avec application a la participation au marche du travail.," *L'Actualite Economique*, 73 (1997), 161-205.

Bresnahan, Timothy F., and Peter C. Reiss, "Entry in Monopoly Markets," *Review of Economic Studies*, LVII (1990), 531-553.

Brock, William A., and Steven N. Durlauf, "Discrete Choice With Social Interactions," *Review of Economic Studies*, LXVIII (2001), 235-260.

Centers for Disease Control, "Changes in the Cigarette Brand Preferences of Adolescent Smokers - United States, 1989-1993," *Morbidity and Mortality Weekly Report*, 43 (1994), 577-581.

Centers for Disease Control, "Youth Tobacco Surveillance, 1998-1999," *Morbidity and Mortality Weekly Report*, 49 (SS-10) (2000), 12-13.

Chaloupka, F.J., "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 94 (1991), 722-742.

Chaloupka, F.J., and R.L. Pacula, "Sex and Race Differences in Young People's Responsiveness to Price and Tobacco Control Policies," *Tobacco Control*, 8 (1999), 373-377.

Chaloupka, F.J., and H Wechsler, *Price, Tobacco Control Policies and Smoking Among Young Adults*(Cambridge MA: National Bureau of Economic Research Working Paper No. 5012, February, 1995).

Chao, Li-Wei, "A comparison of consensus and nonconsensus approaches to modeling contraceptive choice behavior," *Health Economics*, 11 (2002), 599 – 622.

Clark, Andrew E., and Andrew J. Oswald, "Satisfaction and comparison income.," *Journal of Public Economics*, 61 (1996), 359-381.

Cummings, K.M., and D. Shah, "Trends in Smoking Initiation among Adolescents and Young Adults - United States, 1980-1989," *Morbidity and Mortality Weekly Report*, 44 (1995), 521-525.

DeCicca, Philip, Donald Kenkel, and Donald Mathios, "Racial Difference in the Determinants of Smoking Onset," *Journal of Risk and Uncertainty*, 21 (2000), 311-340.

- Dendukuri, N., and L. Joseph, "Bayesian Approaches to Modeling the Conditional Dependence Between Multiple Diagnostic Tests," *Biometrics*, 57 (2001), 158-167.
- Douglas, S., and G. Hariharan, "The Hazard of Starting Smoking: Estimates from a Split Population Duration Model," *Journal of Health Economics*, 13 (1994), 213-230.
- Emery, S., M.M. White, and John P. Pierce, "Does Cigarette Price Influence Adolescent Experimentation?," *Journal of Health Economics*, 20 (2001), 261-270.
- Engels, Rutger C. M. E., Ronald A. Knibbe, Maria J. Drop, and Ypie T. de Haan, "Homogeneity of Cigarette Smoking Within Peer Groups: Influence or Selection?," *Health Education and Behavior*, 24 (1997), 781-811.
- Farkas, Arthur J., Janet M. Distefan, Won S. Choi, Elizabeth A. Gilpin, and John P. Pierce, "Does Parental Smoking Cessation Discourage Adolescent Smoking?," *Preventive Medicine*, 28 (1999), 213-218.
- Farkas, Arthur J., Elizabeth A. Gilpin, M.M. White, and John P. Pierce, "Association Between Household and Workplace Smoking Restrictions and Adolescent Smoking," *Journal of the American Medical Association*, 284 (2000), 717-722.
- Flay, Brian R., Frank B. Hu, Ohidul Siddiqui, L. Edward Day, Donald Hedeker, John Petraitis, Jean Richardson, and Steve Sussman, "Differential Influence of Parental Smoking and Friends' Smoking on Adolescent Initiation and

- Escalation of Smoking," *Journal of Health and Social Behavior*, 35 (1994), 248-265.
- Gaviria, Alejandro, and Steven Raphael, "School-Based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics*, LXXXIII (2001), 257-268.
- Gerlach, K.K., D.R. Shopland, A.M. Hartman, J.T. Gibson, and T.F. Pechacek, "Workplace Smoking Policies in the United States: Results from a National Survey of More than 100,000 Workers," *Tobacco Control*, 6 (1997), 199-206.
- Gilpin, Elizabeth A., Won S. Choi, Steven T. Berry, and John P. Pierce, "How Many Adolescents Start Smoking Each Day in the United States?," *Journal of Adolescent Health*, 25 (1999), 248-255.
- Gilpin, Elizabeth A., F.A. Stillman, A.M. Hartman, J.T. Gibson, and John P. Pierce, "Index for State Tobacco Control Initial Outcomes," *American Journal of Epidemiology*, 152 (2000), 727-738.
- Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics*, CXI (1996), 507-548.
- Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman, "The Social Multiplier," HIER Discussion Paper 1968, Harvard University, 2002.
- Gruber, J. 2001. Risky Behavior Among Youths: An Economic Analysis, Introduction. In *Risky Behavior Among Youths: An Economic Analysis*, edited by J. Gruber. Chicago: University of Chicago Press.
- Harris, J. E., and S. W. Chan, "The continuum-of-addiction: cigarette smoking in relation to price among Americans aged 15-29," *Health Econ*, 8 (1999), 81-6.

- Hersch, J., "Gender, Income Levels, and the Demand for Cigarettes," *Journal of Risk and Uncertainty*, 21 (2000), 263-282.
- Hiedemann, B., "Stackelberg model of social security acceptance decisions in dual career households," *Journal of Economic Behavior and Organization*, 34 (1998), 263 -278.
- Hoxby, Caroline, "The Power of Peers: How Does the Makeup of a Classroom Influence Achievement," *Education Next*, 2 (2002), 56-63.
- Jones, Andrew M., "A Double-Hurdle Model of Cigarette Consumption," *Journal of Applied Econometrics*, 4 (1989), 23-39.
- Jones, Andrew M., "Health, addiction, social interaction and the decision to quit smoking," *Journal of Health Economics*, 13 (1994), 93-110.
- Kindleberger, Charles P., *Manias, Panics, and Crashes: A History of Financial Crises*(London: MacMillan, 1989).
- Kooreman, Peter, "Estimation of Econometric Models of Some Discrete Games," *Journal of Applied Econometrics*, 9 (1994), 255-268.
- Krauth, B. 2001. Simulation-Based Estimation of Peer Effects: Working Paper, Simon Fraser University, November 27.
- Lahiri, K., and J.G. Song, "The Effect of Smoking on Health Using a Sequential Self-selection Model," *Health Economics*, 9 (2000), 491-511.
- Lewit, E. M., and D. Coate, "The Potential for Using Excise Taxes to Reduce Smoking," *Journal of Health Economics*, 2 (1982), 121-45.

Lewit, E. M., D. Coate, and M. Grossman, "The Effects of Government Regulation on Teenage Smoking," *Journal of Law and Economics*, XXIV (1981), 545-69.

Lindbeck, Assar, Sten Nyberg, and Jorgen Weibull, "Social Norms and Economic Incentives in the Welfare State," *Quarterly Journal of Economics*, CXIV (1999), 1-35.

Lux, Thomas, "The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions," *Journal of Economic Behavior & Organization*, 33 (1998), 143-165.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, LX (1993), 531-542.

Marcus, A.C., D.R. Shopland, L.A. Crane, and W.R. Lynn, "Prevalence of Cigarette Smoking in the United States: Estimates from the 1985 Current Population Survey," *Journal of the National Cancer Institute*, 81 (1989), 409-414.

Norton, E. C., R. C. Lindrooth, and S. T. Ennett, "How measures of perception from survey data lead to inconsistent regression results: evidence from adolescent and peer substance use," *Health Econ*, 12 (2003), 139-48.

Norton, Edward C., Richard C. Lindrooth, and Susan T. Ennett, "Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco use," *Health Economics*, 7 (1998), 439 - 453.



Ohsfeldt, R.L., R.G. Boyle, and E.I. Capilouto, *Tobacco Taxes, Smoking Restrictions, and Tobacco Use*(Cambridge MA: National Bureau of Economic Research Working Paper No. 6486, March, 1998).

Orzechowski, and Walker, *The Tax Burden on Tobacco, Historical Compilation, Volume 36, 2001*(Arlington VA, 2002).

Pacula, R.L., and F. Chaloupka, *The Effects of Macro-level Interventions on Addictive Behavior*(Chicago: University of Chicago at Illinois, Working Paper, December, 1999).

Proescholdbell, R.J., L. Chassin, and D.P. MacKinnon, "Home Smoking Restrictions and Adolescent Smoking," *Nicotine and Tobacco Research*, 2 (2000), 159-167.

Qu, Y., and A. Hadgu, "A model for evaluating sensitivity and specificity for correlated diagnostic tests in efficacy studies with an imperfect reference test," *Journal of the American Statistical Association*, 93 (1998), 920-928.

Sacerdote, Bruce, "Peer Effects With Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, CXVI (2001), 681-704.

Shopland, D.R., K.K. Gerlach, D.M. Burns, A.M. Hartman, and J.T. Gibson, "State-Specific Trends in Smoke-Free Workplace Policy Coverage: The Current Population Survey Tobacco Use Supplement, 1993 to 1999," *Journal of Occupational and Environmental Medicine*, 43 (2001), 680-686.

Shopland, D.R., A.M. Hartman, J.T. Gibson, M.D. Mueller, L.G. Kessler, and W.R. Lynn, "Cigarette Smoking among U.S. Adults by State and Region:

- Estimates from the Current Population Survey," *Journal of the National Cancer Institute*, 88 (1996), 1748-1758.
- Simons-Morton, B., D. L. Haynie, A. D. Crump, S. P. Eitel, and K. E. Saylor, "Peer and parent influences on smoking and drinking among early adolescents," *Health Educ Behav*, 28 (2001), 95-107.
- Sweeney, C.T., D.R. Shopland, A.M. Hartman, J.T. Gibson, C.M. Anderson, K.B. Gower, and D.M. Burns, "Sex Differences in Workplace Smoking Policies: Results from the Current Population Survey," *Journal of the American Medical Womens Association*, 55 (2000), 311-315.
- Tamer, Elie, "Empirical Strategies for Estimating Discrete Games with Multiple Equilibria," mimeo, Princeton University, 2002.
- Tamer, Elie, "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, LXX (2003), 147-166.
- Torrance-Rynard, V. L., and S. D. Walter, "Effects of dependent errors in the assessment of diagnostic test performance," *Stat Med*, 16 (1997), 2157-75.
- U.S. Bureau of the Census, *Current Population Survey, January 1996: Tobacco Use Supplement Technical Documentation*(Washington DC: Bureau of the Census, 1996).
- U.S. Bureau of the Census, *Current Population Survey, September 1998: Tobacco Use Supplement Technical Documentation*(Washington DC: Bureau of the Census, 1998).

U.S. Department of Health and Human Services, *Preventing Tobacco Use among Young People: A Report of the Surgeon General*(Atlanta GA: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, Office on Smoking and Health, 1994).

Wang, M. Q., E. C. Fitzhugh, R. C. Westerfield, and J. M. Eddy, "Family and peer influences on smoking behavior among American adolescents: an age trend," *J Adolesc Health*, 16 (1995), 200-3.

Wang, Min Qi, James M. Eddy, and Eugene C. Fitzhugh, "Smoking Acquisition: Peer Influence and Self-selection," *Psychological Reports*, 86 (2000), 1241-1246.

Wasserman, J.W., W.G. Manning, J.P. Newhouse, and J.D. Winkler, "The Effects of Excise Taxes and Regulations on Cigarette Smoking," *Journal of Health Economics*, 10 (1991), 43-64.

Winston, Gordon C., and David J. Zimmerman. 2003. Peer Effects in Higher Education. In *College Decisions: How Students Actually Make Them and How They Could*, edited by C. Hoxby. Chicago: University of Chicago Press & National Bureau of Economic Research, in press.