

Peer effects and selection effects on smoking among Canadian youth

Brian Krauth
Simon Fraser University*

January 10, 2004

Abstract

A number of studies have indicated that peer smoking is a highly influential factor in a youth's decision to smoke. However, these results are suspect because the studies often fail to account for selection and simultaneity bias. This paper develops an econometric model of youth smoking which incorporates both peer effects and selection effects. Identification is achieved by using the degree of selection on observables as a proxy for the degree of selection on unobservables. The results indicate that peers have some influence on a young person's decision to smoke but that influence is much weaker than is suggested by reduced form models.

1 Introduction

Youth smoking is a major public health concern in much of the world. The World Health Organization (Mackay and Eriksen 2002, p. 36) estimates that 4.2 million individuals die every year from smoking-related conditions. Health Canada (2002) estimates that over 45,000 Canadians die from smoking every year, and that the social cost to Canadians from tobacco use is over \$15 billion per year. As a result, governments often spend large sums on programs to reduce tobacco use. For example, in 2001 the federal government of Canada instituted the Federal Tobacco Control Strategy, a program with a budget of \$500 million over its first five years. Because tobacco is highly addictive and most smokers begin when they are teenagers (Mackay and Eriksen 2002, p.28), such tobacco control efforts are primarily focused on discouraging young people from starting to smoke. The research literature on tobacco control has a similar focus on youth. Research using behavioral data has generally found that peers are highly influential in the decision of a young person to smoke (Mackay and Eriksen 2002, Tyas and Pederson 1998); as a result attempts to alter the social context of youth smoking represent a "major component of the Canadian government's efforts to reduce tobacco use" (Mintz, Layne, Ladouceur, Hazel and Desrosiers 1997). Because they represent an externality, peer effects may

*Email: bkrauth@sfu.ca, Web: <http://www.sfu.ca/~bkrauth>. This paper has benefited by comments from Stephen Coate, Phil Oreopoulos, Kevin Milligan, Lisa Powell, and Anindya Sen, as well as seminar participants at SFU, the 2003 CEA meetings, the 2003 Canadian Public Economics Group meetings, the 2003 Society for Computational Economics meetings, and the 2003 North American Summer meetings of the Econometric Society. All errors are mine.

also be used to justify increased tobacco control efforts. For example, Kenkel, Reed, and Wang (2002) calibrate an rational addiction model of lifetime tobacco use and find that the optimal tobacco tax increases by over 50% when peer effects are included. In spite of the apparent consensus, the influence of peers in youth smoking is far from well-established. Most empirical studies in the literature fail to account for both endogenous peer selection and the simultaneity of choice among peers, both of which may lead a researcher to dramatically overestimate the strength of peer influence (Manski 1993).

This paper uses data from Health Canada's 1994 Youth Smoking Survey (YSS) (Stephens and Morin 1996) to estimate the strength of peer influence in smoking among Canadian youth. I use a simulation-based structural estimation method developed in Krauth (2002) which explicitly allows for both endogenous peer selection and simultaneity of choice. This approach is complementary to recent attempts to estimate peer effects in youth smoking using instrumental variables (Gaviria and Raphael 2001, Norton, Lindrooth and Ennett 1998) or panel data methods (Engels, Knibbe, Drop and de Haan 1997, Wang, Eddy and Fitzhugh 2000), and has some significant advantages over these other approaches. The structural estimator addresses simultaneity by treating the group outcome as an endogenous variable, and addresses selection by explicitly allowing correlation in unobservables between peers. The model is identified by assumptions on the nature of this correlation which are about as strong as those needed for identification under IV methods. A key advantage of the structural approach is that it can be used to generate informative interval estimates of the peer effect under much weaker assumptions.

The empirical results reported in this paper include both point estimates under a strong but plausible restriction on the correlation in unobservables (that it is equal to the correlation in observable characteristics) and interval estimates consistent with much weaker restrictions. These results indicate that peers are influential but not as influential as would be suggested by a "naive" probit model that treats peer smoking as an exogenous regressor. The naive estimates indicate a very strong peer effect: the probability of smoking for a representative individual with five close friends increases by 15 percentage points in response to one friend becoming a smoker. Point estimates for the structural model imply a much lower but still positive peer effect; the representative individual's probability of smoking increases by five percentage points in response to one friend becoming a smoker. The interval estimates reported in the paper indicate that this finding is not highly sensitive to moderate violations of the identifying assumption. Accounting for peer selection and simultaneity thus reduces but does not eliminate the apparent influence of peers on youth smoking.

1.1 Related literature

Because of the addictive nature of smoking, much of the research into the determinants of smoking behavior deals with youth smoking. Among adults, the strongest predictor of current smoking is smoking as a young person. For example, Gruber and Zinman (2001) find that 75% of adult current or former smokers began before their 19th birthday. The literature on youth smoking has found a number of variables that are closely associated with a young person's smoking behavior in multiple surveys. These variables include parental or sibling smoking, performance in school, race and ethnicity, and prices¹ (Gruber and Zinman 2001). Peer smoking is also, in the words of a recent review article, "consistently found to be related to smoking initiation, maintenance, and intentions" (Tyas and Pederson 1998).

However, the finding of a close statistical association between peer smoking and a respondent's own smoking does not necessarily imply a causal relationship. Indeed, selection and simultaneity imply that the statistical association may dramatically overstate the strength of any causal relationship. While this issue has been known for some time, it is only recently that empirical studies of youth smoking that attempt to account for selection and simultaneity have appeared. These studies have used either instrumental variables or panel data methods. In the economics literature, the dominant approach uses group characteristics as instrumental variables for group behavior. When these group characteristics are simply the group averages of the individual-level regressors, as in Gaviria and Raphael (2001), this approach can lead to consistent estimates under some strong but plausible assumptions. However, other researchers (Norton et al. 1998, for example) have used other group characteristics (for example, population density of the neighborhood the group lives in) which cannot in principle be valid instruments - if they have a direct impact on the group's average behavior, they must have a corresponding direct impact on the behavior of individual group members. This difficulty in selecting truly exogenous instruments provides some motivation for the use of alternative approaches to identification.

In the public health literature, concern with selection issues has been addressed mostly using panel data. These papers (Engels et al. 1997, Wang et al. 2000, for example) do not usually use regression-based econometric analysis, but rather ask whether smoking tends to precede or follow membership in a peer group of smokers. If membership in a peer group of smokers predicts transition into smoking, this is interpreted as evidence for a peer effect. If smoking predicts transition into a higher-smoking peer group, this is interpreted as evidence for a selection effect. They note that initial smoking appears to be associated with subsequent moves into a higher-smoking peer group, but initial membership in a higher-smoking peer group is not strongly associated with initial nonsmokers becoming smokers. This

¹DeCicca, Kenkel, and Mathios (2002) argue that the common finding of a high price elasticity of youth smoking is an artifact of correlation between prices and unobserved local factors.

result is generally taken to mean that the correlation in smoking among peers is caused primarily by selection effects and not peer effects. However, these results are somewhat weakened by the lack of a formal model of youth decision making. For example, if peer effects operate without a significant lag, current behavior of group members may not be strongly associated with future changes in group behavior even if peers are highly influential. While the results from these longitudinal studies provide some evidence for selection, they may actually have little to say about the strength of peer influence.

As discussed earlier, this paper pursues a structural approach. Although this approach is fairly novel in the analysis of youth smoking, there is a recently-growing literature in econometrics on empirical models of binary games. Kooreman (1994) and Tamer (2003) develop econometric models of interdependent binary choice for the special case with two group members and no selection effect. Krauth (2002) and Kooreman and Soetevent (2002) extend this approach to an arbitrary group size and incorporate inter-peer correlation in characteristics (i.e., peer selection). The structural approach has several advantages over the IV and panel methods. First, it allows for more flexibility in the choice of identifying assumptions. While point identification requires assumptions as strong as those required by the IV methodology, it is only under the structural method used here that informative interval estimates can be obtained with weaker assumptions. Second, the results can be much more easily used to generate quantitative behavioral predictions than is possible with the panel data approach in public health.

2 Data

The primary data source is the 1994 Youth Smoking Survey (YSS), a national survey of Canadian youth aged 10-19. Summary statistics for the data, with cleaning and imputations as described below, are reported in Table 1. In addition to data on a respondent's smoking behavior and that of his or her peers, the YSS has data on other potentially relevant demographic and environmental factors, including province of residence. Information on province of residence is critical because many important policies such as tax rates and restrictions on access or usage vary across provinces. The YSS is divided into two components, a school-based sample of 10 to 14 year olds, and a household-based sample of 15 to 19 year olds. I restrict attention to the household sample, which has 9,491 observations. Of these, 281 observations are dropped because the respondent reports having no friends, or fails to report his or her own smoking behavior or friends' smoking behavior, leaving 9,210 observations to be used in estimation. For all other variables, missing values are replaced with the sample mean.

The endogenous variables in the regressions are a binary variable indicating whether the respondent

is a current smoker, and a discrete variable indicating the fraction of the respondent’s close friends that smoke. In keeping with the conventional definition in the smoking literature, a “current smoker” is defined as an individual who reports smoking at least one cigarette in the past 30 days and over 100 cigarettes in his or her lifetime. In addition to being asked about his or her own smoking, each respondent is asked both how many close friends he or she has, and how many of them smoke. Respondents can claim up to 95 close friends. To keep the computational cost of the estimator reasonable, those respondents who claim more than 6 close friends are coded as having 6 close friends and the fraction of their close friends who smoke is rounded to the nearest sixth. For example, if a person identifies 10 close friends, 7 of whom smoke, this is recoded as having 6 close friends and $(6 * \frac{7}{10}) = 4.2 \approx 4$ close friends who smoke. This recoding affects 3,405 observations.

One serious issue with this and similar data sets is that adolescent self-reports of their smoking and that of their peers are known to be biased. Norton, Lindrooth, and Ennett (2003) provide a recent discussion of these issues. First, despite efforts by survey collectors to emphasize and ensure confidentiality, young people tend to systematically underreport their own smoking. Audit studies compare self-reported smoking with breath or saliva tests and usually find substantial underreporting among youth (Wagenknecht et al., 1992). Second, young people also tend to overestimate the extent to which their peers smoke. Smokers have a particularly high propensity to overestimate the fraction of their peers that smoke, an example of what psychologists call the ‘false consensus’ effect. The YSS data exhibit clear signs of some combination of underreported own smoking and overreported peer smoking: 23.1% of respondents are current smokers, whereas 42.5% of their friends are smokers. As the respondents and their friends are drawn from the same population, and should thus have similar smoking rates, biased reporting is a clear issue. Section 3.4 outlines how biased reporting is incorporated in the econometric model.

3 Model

The econometric model is similar in spirit to the standard model of discrete choice with social interaction effects due to Brock and Durlauf (2001). Both the model and estimation method are described in greater detail in Krauth (2002).

3.1 Preferences and peer groups

In the model, each individual is a member of a peer group. Peer groups are indexed by $g \in Z+$, and group g has $n_g \geq 2$ members, where n_g is exogenous and may vary across groups. Within each peer

group, individuals are indexed by i , so that the pair i, g identifies an individual. Each member of a peer group is influenced symmetrically by each other member, and there are no cross-group influences.

Individuals choose either to smoke ($s_{i,g} = 1$) or not ($s_{i,g} = 0$). An individual's utility function $u_{i,g}(s_{i,g})$ satisfies the following:

$$u_{i,g}(1) - u_{i,g}(0) = \alpha + \beta x_{i,g} + \lambda z_g + \gamma \bar{s}_{i,g} + \epsilon_{i,g} \quad (1)$$

where $x_{i,g}$ is a vector of individual-level exogenous variables (e.g., ethnicity, sex, parental and sibling behavior, disposable income), z_g is a vector of group-level exogenous variables (e.g., prices, school characteristics, or aggregate fixed effects), $\bar{s}_{i,g}$ is the fraction of the other group members that smoke:

$$\bar{s}_{i,g} = \frac{1}{n_g - 1} \sum_{j \neq i} s_{j,g} \quad (2)$$

and $\epsilon_{i,g}$ is an unobserved individual-level term. The model thus incorporates an endogenous² peer effect ($\gamma \bar{s}_{i,g}$) into an otherwise standard discrete choice model of smoking. An individual will prefer to smoke ($s_{i,g} = 1$) if and only if his or her incremental utility from doing so is positive. However, because this incremental utility depends in part on the average choice of peers ($\bar{s}_{i,g}$), predictions on the behavior of individuals cannot be derived in isolation. Instead, it is necessary to analyze behavior in equilibrium.

3.2 Equilibrium

The equilibrium concept used is Nash, supplemented with a selection rule to deal with multiplicity of Nash equilibria. Given the exogenous variables, the endogenous variables are given by a Nash equilibrium in pure strategies of a complete information simultaneous move game, where player i, g 's strategy is given by $s_{i,g}$ and his or her payoff function is described by equation (1). Because there may be multiple Nash equilibria for some draws of the exogenous variables, the model is supplemented with an equilibrium selection rule: the endogenous variables are assumed to take on the values associated with the lowest-activity Nash equilibrium for the given exogenous variables. This selection rule is chosen because it corresponds to the steady state of the myopic best-reply dynamics reached when

²The model does not include what are known as “contextual effects”, in which the characteristics (as opposed to smoking behavior) of one's peers have a direct effect on the relative utility of smoking. Since Manski (1993) pointed out the difficulties in doing so, empirical researchers rarely attempt to simultaneously estimate endogenous effects and contextual effects. It is more common (Gaviria and Raphael 2001, Hoxby 2000, Sacerdote 2001, for example) to simply ignore one or the other and note that the estimated endogenous (contextual) effect could be reinterpreted as contextual (endogenous). That caveat applies to this article as well.

agents start as nonsmokers. Given that young people are born as nonsmokers and that teen smokers rarely quit while young, such a dynamic process may be a reasonable model of behavior. Monte Carlo results reported in Krauth (2002) imply that for low or moderate peer effects, misspecification in the equilibrium selection rule does not have a large impact on the resulting estimates.

3.3 Peer selection and correlation in characteristics

As indicated in the introduction, the structural model estimated in this paper allows for the likelihood that peer choice is not entirely random. This is done by incorporating a reduced-form correlation in observable characteristics and in unobservable characteristics across members of the same peer group. That is, the exogenous variables are identically distributed across all individuals in the population, and are independent across members of different peer groups but are not independent across members of the same peer group. Although it might be desirable to formally model the process of selecting one’s peer group, estimating such a model would require substantially more detail on group composition than is available.

The details are as follows. The joint distribution of the individual-level exogenous variables is assumed to be multivariate normal across the group members. For a group of size $n_g = 3$:

$$\begin{bmatrix} \beta x_{1,g} \\ \beta x_{2,g} \\ \beta x_{3,g} \\ \epsilon_{1,g} \\ \epsilon_{2,g} \\ \epsilon_{3,g} \end{bmatrix} \sim N \left(\begin{bmatrix} \mu \\ \mu \\ \mu \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho_x \sigma^2 & \rho_x \sigma^2 & 0 & 0 & 0 \\ \rho_x \sigma^2 & \sigma^2 & \rho_x \sigma^2 & 0 & 0 & 0 \\ \rho_x \sigma^2 & \rho_x \sigma^2 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \rho_\epsilon & \rho_\epsilon \\ 0 & 0 & 0 & \rho_\epsilon & 1 & \rho_\epsilon \\ 0 & 0 & 0 & \rho_\epsilon & \rho_\epsilon & 1 \end{bmatrix} \right) \quad (3)$$

with the distribution being defined analogously for other values of n_g .

Equation (3) defines a probability distribution with several useful characteristics. First, with the exception of the dependence across peers, the distribution corresponds to that in a standard probit model. In particular, for all individuals i, g the conditional distribution of unobservables has $\epsilon_{i,g} | x_{i,g} \sim N(0, 1)$. Unlike a standard probit model, this model allows for correlation in both observable characteristics (ρ_x) and unobservable characteristics (ρ_ϵ) across group members, as would occur when young people are prone to selecting friends who are similar to themselves. The “naive” analysis pursued in much of the literature on peer effects implicitly assumes that peer selection is random ($\rho_\epsilon = \rho_x = 0$).

In order to obtain point estimates of model parameters, it is necessary to impose an additional restriction on ρ_ϵ , the between-peer correlation in unobservables. The primary restriction used in this

paper is that the correlation is the same as the correlation in observables, i.e. $\rho_\epsilon = \rho_x$. The idea of using the degree of selection on observables as a proxy for the degree of selection on unobservables was first proposed by Altonji, Elder, and Taber (2000) to correct for selection effects in measuring the effect of attending a Catholic school. These authors demonstrate that equality in these two correlations will hold (in expectation) if the observables are a random subset of a large set of relevant variables. Alternatively, if the observed variables are more highly correlated between peers than the unobserved variables, the equal-correlation point estimate of the peer effect will be biased downwards. This is a distinct possibility, as personal information that is particularly easily gathered in surveys (race, sex, age) may also be more easily observed by potential friends. In any case, the model can also be estimated under alternative restrictions on ρ_ϵ , including interval restrictions. As a result, the results reported in this paper allow readers with different beliefs about the selection process to construct their own range of estimates consistent with both the data and their prior beliefs.

3.4 Actual smoking and reported smoking

As discussed in Section 2, the data exhibit signs of either overreporting of peer smoking, underreporting of own smoking, or some combination of the two. In order to close the model and have a well-defined likelihood function it is necessary to explicitly model the connection between actual behavior and reported behavior. Unfortunately, the data provide limited information on this connection.

In light of these limitations, the model of reporting behavior used here is quite simple. First, it is assumed that peer smoking is reported truthfully. Second, it is assumed that smokers falsely report as nonsmokers with a fixed probability p_r ³. Let $r_{i,g}$ indicate whether a person self-reports as a smoker. Then:

$$r_{i,g} = \begin{cases} s_{i,g} & \text{with probability } p_r \\ 0 & \text{with probability } 1 - p_r \end{cases} \quad (4)$$

where p_r is a parameter. Note that conditional on $s_{i,g}$, $r_{i,g}$ is independent of all other variables, and that

$$p_r = \frac{E(r_{i,g})}{E(\bar{s}_{i,g})} \quad (5)$$

The parameter p_r can thus be estimated using the sample analogue to this expression.

³The model was also estimated (not reported, available on request) allowing p_r to vary across provinces, with almost no effect on results.

4 Estimation

Given the model outlined in Section 3, this section describes how the model is estimated using the YSS data. Let N be the number of observations in the data. Because the YSS is a household-based random sample from a large population, it can be assumed that each observation represents a member of a different peer group. Without loss of generality, let the g th observation in the data set be identified as describing person 1 of group g .

4.1 Naive estimator

In order to facilitate comparison with previous work, I first apply a naive probit estimator to the data. The naive estimator is simply the standard maximum likelihood probit estimator which treats average peer behavior \bar{s} as an exogenous variable. In other words, I use maximum likelihood to estimate a model in which

$$\Pr(r_{1,g} = 1|x_{1,g}, z_g) = \Phi\left(\tilde{\alpha} + \tilde{\beta}x_{1,g} + \tilde{\lambda}z_g + \tilde{\gamma}\bar{s}_{1,g}\right) \quad (6)$$

where Φ is the CDF for the standard normal distribution. In addition I estimate a second naive model which corrects for reporting bias as described in Section 3.4. In this case:

$$\Pr(r_{1,g} = 1|x_{1,g}, z_g) = p_r \Phi\left(\tilde{\alpha} + \tilde{\beta}x_{1,g} + \tilde{\lambda}z_g + \tilde{\gamma}\bar{s}_{1,g}\right) \quad (7)$$

where $p_r = \frac{\sum_{g=1}^N r_{1,g}}{\sum_{g=1}^N \bar{s}_{1,g}}$. Again, the remaining parameters are estimated by maximum likelihood. Because both these naive estimators treat peer behavior as exogenous, the maximum likelihood estimators of $\tilde{\gamma}$ and $\tilde{\lambda}$ are not consistent estimators of the true peer effect γ . The estimates are reported for comparison with the structural estimates, and to demonstrate the cost of ignoring selection and simultaneity.

4.2 Structural estimator: Point estimates

Next, the model is estimated by the simulated maximum likelihood (SML) method developed in Krauth (2002). The vector of parameters to be estimated is $\theta \equiv (p_r, \alpha, \beta, \lambda, \gamma, \mu, \sigma, \rho_x, \rho_\epsilon)$. Observations are indexed by g ; observation g in the data set is treated as describing person 1 in group g . For each observation, the GHK simulator (Hajivassiliou, McFadden and Ruud 1996) is used to estimate the conditional likelihood $\Pr(r_{1,g}, \bar{s}_{1,g}|x_{1,g}, z_g; \theta)$. This approximate conditional likelihood is then logged

and added up across all observations to give

$$\hat{L}(\theta) \equiv \sum_{g=1}^n \ln \hat{\Pr}(r_{1,g}, \bar{s}_{1,g} | x_{1,g}, z_g; \theta)$$

where $\hat{\Pr}$ is the simulation-based estimate of the true probability \Pr , and $\hat{L}(\theta)$ is thus a simulation-based estimate of the true log-likelihood function $L(\theta)$. The parameter vector θ is then chosen to maximize the conditional log-likelihood subject to the restrictions:

$$\rho_\epsilon, \rho_x \in \left(\frac{-1}{\max_g(n_g)}, 1 \right) \tag{8}$$

$$\gamma \geq 0 \tag{9}$$

$$\rho_\epsilon = \rho_x \tag{10}$$

Restriction (8) implies that the peer group covariance matrix is positive definite. The nonnegativity constraint on the peer effect in restriction (9) is necessary for computational purposes; this constraint does not bind for this particular data set. However, it should be noted that if γ is exactly zero, the bootstrap method used to estimate the covariance matrix of estimates is inconsistent (Andrews 2000), and the t-statistic for γ has a nonstandard asymptotic distribution. As a result, significance tests for γ are not performed in this paper. Finally equation (10) gives the equal-correlation restriction discussed in Section 3.3.

Krauth (2002) describes the estimation method and its properties in detail, and reports the outcome of various Monte Carlo experiments. Results in that paper include:

1. If the model is correctly specified, the structural estimator of the peer effect is consistent and has much lower bias than the naive probit estimator.
2. Deviations from normality in the explanatory variables do not have a substantial effect on parameter estimates.
3. The equilibrium selection rule has no effect on parameter estimates when the peer effect is zero; in that case equilibrium is always unique. The potential bias due to an incorrectly specified equilibrium selection rule increases gradually in the strength of the peer effect.
4. The variance of the estimator depends strongly and negatively on the explanatory power of the individual-level observables. When these variables have strong explanatory power, the peer effect will be more precisely estimated.

4.3 Structural estimator: Interval estimates

As the equal-correlation assumption (10) reflects a strong and possibly incorrect restriction on the model, one might also want to explore the effect of alternative or weaker restrictions. While some sort of restriction on ρ_ϵ is needed to achieve identification, the structural model allows one to obtain interval estimates of the peer effect under interval restrictions on ρ_ϵ .

Interval estimates are composed as follows. Choose some arbitrary $P \in \left(\frac{-1}{\max_g(n_g)}, 1\right)$. Estimate the model, replacing the equal-correlation restriction (10) with the restriction:

$$\rho_\epsilon = P \tag{11}$$

Define the function $\hat{\gamma}(P)$ as the point estimate of the peer effect γ under the restriction ($\rho_\epsilon = P$). This function can then be used to construct interval estimates. Let $[P_0, P_1]$ be an arbitrary closed subinterval of $\left(\frac{-1}{\max_g(n_g)}, 1\right)$. Then:

$$\rho_\epsilon \in [P_0, P_1] \Rightarrow \hat{\gamma} \in \left[\min_{P \in [P_0, P_1]} \hat{\gamma}(P), \max_{P \in [P_0, P_1]} \hat{\gamma}(P) \right] \tag{12}$$

The interval estimates defined in such a way are consistent, efficient, and sharp (Krauth 2002). Because the selection of a plausible interval restriction on ρ_ϵ is a subjective matter, in reporting results I report values⁴ for the function $\hat{\gamma}(\rho_\epsilon)$ and leave the reader to apply the formula (12).

Another alternative to interval estimation is to estimate both ρ_x and ρ_ϵ under the restriction:

$$\gamma = 0 \tag{13}$$

in place of (10). The resulting estimates answer the question “How high does the within-group correlation in unobservables need to be to explain the observed within-group correlation in behavior without peer effects?”

⁴It should be noted that it is only possible to calculate a values of $\hat{\gamma}(P)$ at a finite number of points in the interval, implying that the true minimum or maximum may not be found. However, it can be proved that $\hat{\gamma}(P)$ is continuous, implying that checking a sufficiently fine grid of points in the interval will yield a reasonable approximation to the true minimum and maximum. In addition, experience with the estimator suggests that $\hat{\gamma}(P)$ is strictly decreasing, though that has yet to be proved. If true, this would imply one only needs to calculate $\hat{\gamma}(P_0)$ and $\hat{\gamma}(P_1)$.

5 Results

5.1 Overview of figures and tables

Table 2 reports point estimates using both naive and structural estimators. Estimated standard errors are calculated for the naive estimators using the standard asymptotic approximation, and calculated for the structural model using the simple bootstrap with 100 replications. Province level fixed effects are also included in the regressions, but are not reported. These fixed effects control for variations in cigarette prices, taxes, and availability across jurisdictions, as well as possible cultural or environmental differences between provinces. The fixed effects are treated as aggregate variables (z_g in the model) while all other characteristics are treated as individual-level variables ($x_{i,g}$ in the model).

To give the coefficients a more intuitive interpretation which is comparable across models, Table 3 reports the results from a simple thought experiment. Consider a representative individual who has five close friends (the median in the YSS data), all nonsmokers, and has a set of observed characteristics (including average peer choice) such that the model predicts his or her probability of self-reporting as a smoker to be exactly 23.1%, the mean in the YSS data. Suppose that 20% of the relevant peer group (e.g., one of the five friends) suddenly switches from a non-smoker to a smoker. By how much would the model predict the representative individual’s probability of smoking increases? This quantity, for both the models estimated here and for a few other recent studies, is reported in Table 3.

Table 4 displays estimates of $(\gamma, \rho_x, \rho_\epsilon)$ from the structural model with the equal correlation restriction (10) replaced by alternative restrictions, as described in Section 4.3. This table can be used, in combination with the formula in (12), to construct interval estimates of the peer effect under user-specified restrictions on the selection effect. For example, suppose that $\rho_\epsilon \in [0.0, 0.5]$, i.e., the correlation in unobservables is positive, and not much more than twice the estimated correlation in observables (from the point estimates). In that case the interval estimate of the peer effect would be $[0.578, 1.745]$, with a pointwise 95% confidence interval of $[0.388, 1.863]$.

Figure 1 summarizes the results of the paper graphically by plotting the estimated peer effect (γ) against the assumed or estimated selection effect (ρ_ϵ). The two points on the left side of the graph represent the naive estimates (which implicitly assume no selection effect). The point in the center represents the point estimate from the structural model, and the dotted ellipse surrounding it is a 95% joint confidence region. The solid line represents the estimated $\hat{\gamma}(\rho_\epsilon)$ function, and the dotted lines represent a pointwise 95% confidence band.

5.2 Interpretation

Taken as a whole, the tables and figure reveal several key results.

1. The estimated impact of the exogenous variables on current smoking is consistent with previous results in the youth smoking literature (Gruber and Zinman 2001). Age, absence from school, and exposure to smokers at home are all contributing factors to the decision to smoke, and students who are doing well in school tend to smoke less than students who are doing poorly. Marginal effects of the exogenous variables are roughly comparable in magnitude across the three estimators.⁵
2. The naive estimates suggest extremely strong peer influence; the hypothetical effect of one additional close friend smoking on the probability the representative individual smokes is over 15 percentage points. This quantitative finding is within the range set by previous studies using the naive method.
3. The estimated γ under the restriction that $\rho_\epsilon = 0$ essentially corrects for the simultaneity effect without correcting for any selection effect. Because it is so much lower than the naive estimate (which ignores both selection and simultaneity), it is clear that the naive estimator suffers from large simultaneity bias.
4. The point estimate from the structural model is well below both the naive estimates and the structural estimate under the no-selection-effect restriction, implying that the naive estimates also suffer from substantial bias due to endogenous peer selection.
5. Even after accounting for selection and simultaneity, the point estimate from the structural model indicates peers exhibit a nontrivial influence on youth smoking. The hypothetical effect of one additional close friend smoking on the probability the representative individual smokes is over 5 percentage points.
6. Even if one is concerned with the equal-correlation restriction, a positive and nontrivial degree of peer influence is consistent with a fairly wide range of values for the selection effect. For example, if $\rho_\epsilon \in [0.0, 0.5]$, then $\hat{\gamma}(\rho_\epsilon) \in [0.578, 1.745]$.
7. All three of the studies referenced in Table 3 that use instrumental variables to identify peer effects have found little difference between naive and IV estimates. This finding is something of a puzzle, given that longitudinal studies have found convincing evidence for peer selection and the structural estimates reported here also suggest peer selection. This suggests one of

⁵The underreporting correction, used in columns 2 and 3, will tend to increase the magnitude of reported coefficients over the standard probit model, but leave the associated marginal effects about the same.

two possibilities: either the naive approach was acceptable all along, there is no peer selection on unobservables, and the other (panel and structural) studies are in some way flawed, or the similar instruments in these IV studies are not exogenous.

6 Conclusion

It has long been known that naive estimation leads to upwardly biased measures of peer effects. For years, empirical researchers have used naive estimation with a few caveats in the absence of credible ways of compensating for simultaneity and peer selection. Recent years have seen an increased awareness of selection bias, and a number of valuable papers have used instrumental variables or panel methods to evaluate the strength of peer effects in youth smoking and other contexts.

This paper has used an alternative but complementary approach, the estimation of an equilibrium model of discrete choice which explicitly incorporates the possibility of relative homogeneity within peer groups. As with structural estimation in general, this approach has the advantages of theoretical clarity and the ability to make predictions on the equilibrium response of behavior to changes in incentives. In addition the method is flexible enough to yield interval estimates under weaker interval restrictions, providing a valuable means of performing sensitivity analysis. In this case, there is also the advantage that estimation does not require complex data sets or natural experiments, only standard survey data. While natural experiments have proved a valuable tool in the analysis of social interactions, they are not available in all questions of interest. Natural experiments will be particularly hard to find if the relevant peer group is formed informally (like groups of close friends) rather than by some central authority (like classrooms). In the absence of natural experiments, applied researchers will continue to investigate the question of peer effects in youth smoking using the tools available. For example, at least two recently published papers by prominent public health researchers (Alexander et al. (2001) and Norton et al. (2003)) estimate peer effects without any attempt to deal with simultaneity or selection.

Under the fairly strong assumption of equal correlation in observables and unobservables, I find that the estimated peer effect is substantially lower than the naive estimate, but is still quite large. Less restrictive assumptions lead to interval estimates, but the intervals associated with relatively weak restrictions are also well below the naive estimate and well above zero. These empirical results suggest a middle ground between extreme skepticism and credulity on peer effects in youth smoking: peers are moderately influential in a young person's decision to smoke.

This research can be extended in a number of useful ways. The recent work of Tamer (2003) on estimation in the presence of multiple equilibria can be applied to investigate the role of equilibrium

selection. This methodology can also be extended to work with newly available school-based data sets used in Alexander et al. (2001) or Norton et al. (2003) which sample all students in a given school and ask students to identify their friends directly. Not only will such data provide a more consistent measure of peer choice because all behavior is self-reported, but it will also provide more detailed information on how peer influence works across complex social networks.

References

- Alexander, Cheryl, Marina Piazza, Debra Mekos, and Thomas Valente**, “Peers, schools, and adolescent cigarette smoking,” *Journal of Adolescent Health*, 2001, 29, 22–30.
- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber**, “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” Working Paper 7831, NBER 2000.
- Andrews, Donald W. K.**, “Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space,” *Econometrica*, 2000, 68 (2), 399–406.
- Brock, William A. and Steven N. Durlauf**, “Discrete choice with social interactions,” *Review of Economic Studies*, 2001, 68 (2), 235–260.
- DeCicca, Philip, Donald Kenkel, and Alan Mathios**, “Putting out the fires: Will higher taxes reduce the onset of youth smoking?,” *Journal of Political Economy*, 2002, 110 (1), 144–169.
- 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*, 1997, 24, 81–811.
- Gaviria, Alejandro and Steven Raphael**, “School-based peer effects and juvenile behavior,” *Review of Economics and Statistics*, 2001, 83 (2), 257–268.
- Gruber, Jonathan and Jonathan Zinman**, “Youth smoking in the U.S.: Evidence and implications,” in Jonathan Gruber, ed., *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, 2001.
- Hajivassiliou, Vassilis, Daniel McFadden, and Paul Ruud**, “Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results,” *Journal of Econometrics*, 1996, 72 (1), 85–134.
- Health Canada**, *Report on Tobacco Control: An Update*, Health Canada, 2002.
- Hoxby, Caroline M.**, “Peer effects in the classroom: Learning from gender and race variation,” Working Paper 7867, NBER 2000.
- Kenkel, Donald S., Robert R. Reed III, and Ping Wang**, “Rational addiction, peer externalities, and long run effects of public policy,” Working Paper 9249, NBER 2002.
- Kooreman, Peter**, “Estimation of econometric models of some discrete games,” *Journal of Applied Econometrics*, 1994, 9, 255–268.
- and **Adriaan Soetevent**, “A discrete choice model with social interactions: An analysis of high school teen behavior,” Working Paper, University of Groningen 2002.
- Krauth, Brian V.**, “Simulation-based estimation of peer effects,” Working Paper, Simon Fraser University 2002.
- Mackay, Judith and Michael Eriksen**, *The Tobacco Atlas*, World Health Organization, 2002.
- Manski, Charles F.**, “Identification of endogenous social effects: The reflection problem,” *Review of Economic Studies*, 1993, 60 (3), 531–542.
- Mintz, James H., Neville Layne, Rachel Ladouceur, Jane Hazel, and Monique Desrosiers**, “Social Advertising and Tobacco Demand Reduction in Canada,” in M.E. Goldberg, M. Fishbein, and S.E. Middlestadt, eds., *Social Marketing: Theoretical & Practical Perspectives*, Erlbaum Associates, 1997.
- 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*, 1998, 7, 439–453.
- , —, and —, “How measures of perception from survey data lead to inconsistent regression results: Evidence from adolescent and peer substance use,” *Health Economics*, 2003, 12, 139–148.
- Powell, Lisa M., John A. Tauras, and Hana Ross**, “Peer Effects, Tobacco Control Policies, and Youth Smoking Behavior,” Working Paper, University of Illinois - Chicago 2003.
- Sacerdote, Bruce**, “Peer effects with random assignment: Results for Dartmouth roommates,” *Quarterly Journal of Economics*, May 2001, 116 (2), 681–704.
- Stephens, T. and M. Morin**, “Youth Smoking Survey, 1994: Technical Report,” Working Paper, Health Canada 1996.
- Tamer, Elie T.**, “Incomplete simultaneous discrete response model with multiple equilibria,” *Review of Economic Studies*, 2003, 70 (1), 147–167.
- Tyas, Suzanne L. and Linda L. Pederson**, “Psychosocial factors related to adolescent smoking: A critical review of the literature,” *Tobacco Control*, 1998, 7, 409–420.
- Wagenknecht, K.E., G.L. Burke, L.L. Perkins, N.J. Haley, and G.D. Friedman**, “Misclassification of smoking status in the CARDIA study: a comparison of self-report with serum cotinine levels,” *American Journal of Public Health*, 1992, 82, 33–36.
- Wang, Min Qi, James M. Eddy, and Eugene C. Fitzhugh**, “Smoking acquisition: Peer influence and self-selection,” *Psychological Reports*, 2000, 86, 1241–1246.

A Tables and Figures

Variable	Mean	Std. Dev.
Current smoker	0.231	0.421
Number of friends	6.140	2.804
Fraction of friends smoking	0.425	0.368
Age on 1/1/94	16.815	1.406
Attending school	0.851	0.356
Performs above average in school	0.287	0.453
Performs below average in school	0.032	0.177
Number of smokers in house	0.828	1.038
Has seen ads for tobacco-sponsored events	0.463	0.487

Table 1: Summary statistics for YSS data, 9,210 observations.

Variable Name	Naive Probit	Naive with underreporting	Structural Model
Selection effect (ρ_x, ρ_ϵ)	–	–	0.237 (0.071)
Peer effect (γ)	2.240 (0.055)	3.626 (0.119)	1.251 (0.270)
Age on 1/1/94	0.091 (0.014)	0.164 (0.025)	0.098 (0.032)
Attending school	-0.328 (0.051)	-0.517 (0.100)	-0.574 (0.114)
Above avg in school	-0.192 (0.046)	-0.297 (0.075)	-0.554 (0.066)
Below avg in school	0.227 (0.089)	0.458 (0.181)	0.566 (0.210)
# smokers in house	0.258 (0.016)	0.453 (0.036)	0.627 (0.050)
Seen ads for events sponsored by tobacco	0.242 (0.037)	0.367 (0.066)	0.296 (0.061)

Table 2: Point estimates for both naive and structural estimators. Province-level fixed effects also included, but not reported. Estimated standard errors in parentheses.

Source (peer group)	Method of Identification	Implied $\Delta \Pr(\text{smoke})$
Wang et al., 1995 (4 best friends)	Naive	37.3% - 53.1%
Norton et al., 1998 (neighborhood)	Naive IV	23.9% 24.8%
Gaviria and Raphael, 2001 (schools)	Naive IV	3.0% 3.0%
Norton et al., 2003 (3 best friends)	Naive	10.5%
(school)	Naive	7.9%
Powell et al., 2003 (school)	Naive IV	11.1% 10.8%
This paper (close friends)	Naive Naive w/underrep. Structural, equal corr.	15.6% 15.2% 5.4%

Table 3: Comparison of estimated peer effects in this paper and in other studies. To facilitate comparison across different models, results are stated in terms of the increase in probability that a representative individual will be a smoker in response to a 20% increase in peer smoking. See text for details.

Identifying Restriction	$\hat{\gamma}$	$\hat{\rho}_x$	$\hat{\rho}_\epsilon$
$\rho_\epsilon = \rho_x$	1.251 (0.270)	0.237 (0.071)	0.237 (0.071)
$\rho_\epsilon = 0.0$	1.745 (0.059)	0.201 (0.028)	–
$\rho_\epsilon = 0.1$	1.499 (0.072)	0.231 (0.028)	–
$\rho_\epsilon = 0.2$	1.260 (0.066)	0.261 (0.028)	–
$\rho_\epsilon = 0.3$	1.040 (0.068)	0.286 (0.032)	–
$\rho_\epsilon = 0.4$	0.795 (0.091)	0.318 (0.036)	–
$\rho_\epsilon = 0.5$	0.578 (0.095)	0.349 (0.044)	–
$\rho_\epsilon = 0.6$	0.400 (0.095)	0.353 (0.041)	–
$\rho_\epsilon = 0.7$	0.155 (0.119)	0.374 (0.041)	–
$\gamma = 0.0$	–	0.420 (0.033)	0.738 (0.035)

Table 4: Estimated peer/selection effect under alternative identifying restrictions. Estimated standard errors in parentheses.

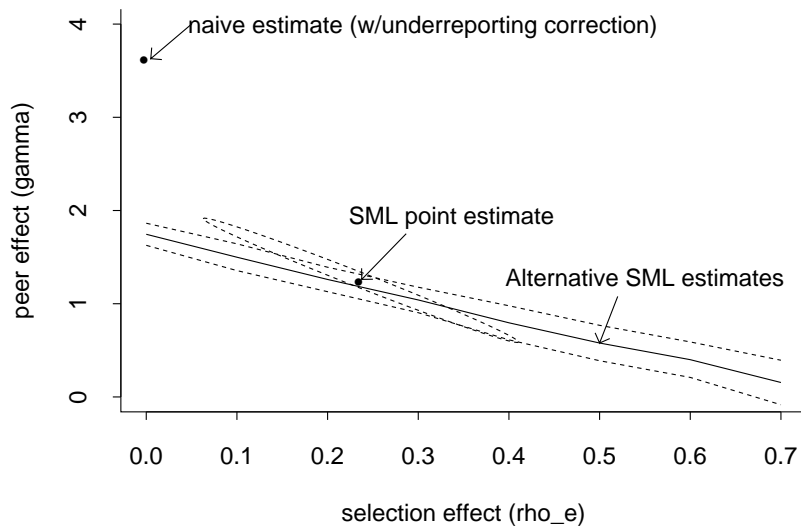


Figure 1: Estimated peer effect (γ) for alternative assumptions about selection effect (ρ_ϵ). Individual points represent point estimate from naive probit model or baseline structural model with equal-correlation assumption. Solid line represents alternative estimates of γ as a function of the assumed value of ρ_ϵ . Dashed lines depict 95% confidence intervals or ellipses.