Identifying Sibling Influence on Teenage Substance Use

Joseph G. Altonji[1], Sarah Cattan[2] and Iain Ware[3] *

September 12, 2010

Abstract

A number of studies have found substantial correlations in risky behavior between siblings, raising the possibility that adolescents may directly influence the actions of their brothers or sisters. We assess the extent to which correlations in substance use and selling drugs are due to causal effects. Our identification strategy relies on panel data, the fact that the future does not cause the past, and the assumption that the direction of influence is from older siblings to younger siblings. Under this assumption along with other restrictions on dynamics, one can identify the causal effect from a regression of the behavior of the younger sibling on the past behavior and the future behavior of the older sibling. We also estimate a joint dynamic model of the behavior of older and younger siblings that allows for family specific effects, individual specific heterogeneity, and state dependence. We use the model to simulate the dynamic response of substance use to the behavior of the older sibling. Our results suggest that smoking, drinking, and marijuana use are affected by the example of older siblings, but most of the link between siblings arises from common influences.

---

*[1]Yale University and NBER, [2]University of Chicago, [3] Stanford University. We are grateful to Greg Duncan, Amanda Kowalski, Robert T. Michael and participants in seminars at Yale University, University of Chicago, New York University, the NLSY97 conference at the Bureau of Labor Statistics (May 2008), NBER Health 2009 Summer Institute, the Institute for Fiscal Studies, UC San Diego, and the SOLE/EALE 2010 Meetings for valuable comments. Our research has been supported by the Economic Growth Center, Yale University. We are responsible for the remaining shortcomings of the paper.

1 Introduction

Teenage smoking, substance abuse, involvement in property and violent crime, and involvement in risky sexual activity fluctuate, but remain at high levels.[1] Understanding the factors that lead adolescents to engage in these behaviors is a high research priority.

This paper examines whether substance use of one child directly influences the behavior of a younger sibling. Several studies have found significant correlations between risky behavioral patterns among siblings.[2] In keeping with this literature, in Table 2 below, we show that the probability an adolescent has smoked, used alcohol, smoked marijuana, used hard drugs, or sold drugs in the past year is dramatically higher if an older sibling engaged in the corresponding behavior when at the same age, even after one includes a basic set of control variables. Findings of this nature are consistent with the possibility that substance use and other risky behaviors are contagious among siblings in a household. However, siblings share many influences, including common family backgrounds, neighborhoods, schools, and genes. These common influences could potentially account for most or even all of the correlations. It is difficult to successfully control for the range of shared characteristics that affect siblings. As a result, there are very few convincing attempts to distinguish direct sibling influences from the plethora of unobserved factors that might contribute to the high correlation in delinquent behavior among siblings.

We address the problem of shared influences using two related empirical strategies. Both exploit a basic fact, and both are based on a key maintained assumption. The fact is that only actions of a youth that occur at or before a point in time can causally influence his or her sibling's action at that time. The assumption is that older siblings influence younger siblings, but younger siblings do not influence older siblings.

The first of our empirical strategies uses a correlated random effects (CRE) design

[1]See Levitt and Lochner (2001) on teenage homicide, Gruber and Zinman (2001) on smoking, Pacula et al. (2001) on marijuana usage, and Grossman et al. (2004) on teenage sex.

[2]For example,Amuedo-Derantes and Mach (2002) find that having a sibling who abuses illegal drugs significantly increases the likelihood that an adolescent will also take drugs. Duncan et al. (2005) compare correlations of various measures of achievement and delinquency across siblings, peers, neighbors, and schoolmates and find that these correlations are substantially stronger among siblings than among other groups.

in the spirit of Mundlak (1978) and Chamberlain (1984). We estimate models relating the behavior of the younger sibling at time t to the behavior of the older sibling before that date using the sum of the older sibling's behaviors before and after time t as a control variable. Our estimate of the sibling influence is the coefficient on the early behavior. The coefficient on the sum of the past and future behaviors identifies the part of the link in the behavior of siblings that is due to common unobserved influences. Our interpretation of the CRE evidence is conditional on the assumption that the direction of the influence between siblings is from the older sibling to the younger one. Several studies in the psychology literature support this assumption as a first approximation, including Buhrmester (1992) and Rodgers and Rowe (1988). To the extent that it is false, our estimates are likely to understate the influence of the older sibling on the younger one.

Both state dependence (e.g., habit formation) and nonstationarity with respect to age could lead the past behavior and the future behavior of the older sibling to have different relationships with the sibling fixed effect. This would bias the CRE estimate of the older sibling's influence, although the direction of the bias is not clear.[3] In part for this reason, we also develop and estimate a dynamic model of the behavior of the older and younger siblings. The model allows for state dependence and unobserved heterogeneity at the individual and sibling pair levels. It consists of a dynamic system of discrete choice equations in which the behavior of each sibling depends on exogenous variables, past behavior, and a person specific error component. The behavior of the younger sibling also depends on the past behavior of the older sibling. We estimate the system by maximum likelihood and use our estimates to simulate the dynamic response of the behavior of the younger sibling to the behavior of his or her older sibling.

Our results using the CRE approach point to positive effects of the behavior of the older sibling on smoking cigarettes and drinking alcohol of the younger sibling. The results using the dynamic model show positive effects on smoking, drinking, and marijuana use. We also obtain positive point estimates for using hard drugs and selling drugs, but the

[3]Below we also report regression estimates controlling for sibling fixed effects, which are subject to a similar problem.

estimates fall short of statistical significance. Overall, we conclude that there is a modest positive sibling effect on substance use. However, most of the very large correlation in sibling behavior is due to common influences rather than a peer effect.

While our focus is on sibling influences, the qualitative findings may be of some interest to the rapidly growing literature on peer influences among adolescents. Estimates of peer effects may be biased upward by the fact that adolescents select friends who share similar interests, while children cannot choose their siblings. On the other hand, the problem of common genes and family factors is less severe for friends and acquaintances than for siblings. Furthermore, some of the strategies that have been employed recently in studies of peer effects, such as variation arising from quasi-random assignment of roommates, are not feasible for siblings.[4] Perhaps for this reason, there is little quantitative evidence on peer influences among siblings. This knowledge gap provides the motivation for our study, despite the limitations of our identification strategies.

The paper continues in section 2, which provides a brief review of the existing economic and psychology literature on social influences on adolescent substance use, with a focus on sibling effects. In sections 3 and 4, we discuss the NLSY97 data and document the strong correlation in substance use across siblings. In section 5, we present a simple model of sibling links in behavior that underlies our econometric analysis. We explain the CRE strategy and present the joint dynamic probit model of substance use. We present our results using the CRE approach in section 6 and those using the dynamic probit and the dynamic ordered probit models in sections 7 and 8, respectively. In section 9, we explore the extent to which the link between siblings depends on the gender match, the age gap, and family process variables. Unfortunately, in most cases, our estimates of interaction effects are not sufficiently precise to support strong conclusions. We close with conclusions and a research agenda.

[4] See Sacerdote (2001), Marmaros and Sacerdote (2002), Duncan et al. (2005), and Stinebrickner and Stinebrickner (2006). One could examine whether the sibling influence is larger for siblings who share a bedroom. With data on the number bedrooms and the number of male and female children by age, one could create a proxy even if information on sharing a bedroom is unavailable. We do not have the necessary data to perform this analysis.

2 Literature Review

Developmental psychologists and sociologists were first to investigate the importance of social environment on adolescent development and behavior. Peers, parents, and siblings are widely viewed as the most significant domains of influence on adolescents, but their relative importance remains controversial. While some perceive peer group influence as the single most important factor shaping a child's behavior (Harris (1998)), a number of psychologists continue to emphasize the primacy of the family in shaping a child's attitudes and behaviors (Jessor and Jessor (1977), Kandel (1980)). With regard to alcohol for example, Barnes (1990) considers the diverse channels through which the family potentially exerts influence on behavior. Not only does a child often get his first exposure to and experience with alcohol within the family, but family members are powerful agents of socialization. A number of mechanisms, including imitation, role modeling, and parental tolerance toward certain behaviors, help shape an adolescent's attitudes and values and hence influence his or her future behavior.

Within the family, siblings occupy a particular social position. An extensive literature in psychology analyzes the different mechanisms through which siblings are hypothesized to influence each other's behaviors. One can distinguish two main hypotheses: the "role model" and the "opportunity" hypotheses. First, a sibling, most likely the younger one, may see his older sibling as a role model to observe, directly imitate, and use in shaping his notions about what types of behaviors are suitable (Widmer (1997), Buhrmester (1992), and Rodgers and Rowe (1988)). Patterson (1984) proposes a variant of this hypothesis in his "theory of siblings as key pathogens" to explain how siblings can encourage each other to have antisocial tendencies, such as delinquent behavior. He argues that siblings provide learning or training models to develop this behavior. This is particularly likely to happen between siblings who have conflict ridden and aggressive relationships, because these promote antisocial behavior.

The second mechanism through which siblings influence each other's behaviors is by providing opportunities (friends and settings) for substance use and sexual intercourse.

In contrast to Patterson's hypothesis, this mechanism is more likely to occur with siblings who have a better and warmer relationships, have common friends, and hence engage in risky behavior together. For the purpose of our study, it is important to note that most of the literature surveyed here argues that the pattern of influence runs from the older to the younger child (Buhrmester (1992), Rodgers and Rowe (1988)).

In the economics literature, the models of Akerlof (1997) and Bernheim (1994) concerning social conformity are also relevant to the present study of sibling influences. In his social distance model, Akerlof (1997) represents social interactions as a mutually beneficial trade between agents. Agents occupy a location on the social space, which is partly inherited. The model creates incentives for agents to interact with those that are close in the social space, thus possibly explaining their tendency to conform to the behavioral norms of those who share their inherited social location. Bikchandani et al. (1992) argue that it may be optimal for individuals to base their decisions on observed behavior of others because they lack complete information about the consequences of a specific action. These ideas can be applied to sibling interactions as well as to other groupings.

In addition to the theoretical work reviewed above, there are a number of empirical studies of social influence on youth behavior. While many focus on peer group effects only (e.g., Furman and Buhrmester (1992), Berndt (1992), Steinberg and Silverberg (1986)), some recent papers have incorporated parents and siblings into the analysis. Most of these papers provide evidence of large correlations between siblings in a variety of behaviors, without necessarily devising a strategy for distinguishing causality from the effect of common unobserved factors. For example, Duncan et al. (2001) examine sibling correlations in measures of delinquency for sample of adolescents in grades 7 through 12 (Add Health data). The sample includes genetically differentiated siblings within a family, peers, grade mates, and neighbors, thus allowing the authors to compare correlations in the same behavior across different types of relationships. The correlations are highest for siblings, especially for twins, thus suggesting a large scope for family influences.

Investigating further the role of genes in the same data set, Slomkowski et al. (2005) finds that both genetic and environmental factors contribute to similarities between siblings' smoking behavior. Data on the quality of the relationship between siblings reveal that a better sibling relationship magnifies the importance of shared environmental effects. Conger and Reuter (1996) stress the importance of both siblings' and parents' drinking behavior and find evidence that a sibling's drinking exacerbates an adolescent's tendency to drink both directly, through imitation or increased exposure, and indirectly, through the selection of friends who drink. Windle (2000) finds that peer and siblings' substance use more strongly predict adolescent's substance use than parents'. Furthermore, siblings' substance use is a strong predictor of coping motives for drinking, thus indicating that imitation and role modeling in stressful situations might be important channels of influence among siblings. Several papers have also looked at sibling influences on smoking patterns, although results for this activity are mixed (Otten et al. (2007), Bricker et al. (2005), and Slomkowski et al. (2005)).

Using the Arizona Sibling Study, Rowe and Gulley (1992) find that correlations in substance use and delinquent behavior are higher when interactions are warmer, less conflict ridden and more frequent, when siblings have more mutual friends, and for same-sex pairs of siblings. Although these results do not directly test for the presence of a direct sibling influence, they are consistent with one, as suggested by the opportunity hypothesis described above. In contrast, using sibling pair data from the longitudinal Iowa Youth and Family project, Slomkowski et al. (2001) obtains mixed results on how the level of support and hostility between siblings influences the strength of the relationship between sibling behavior.

While some of the studies mentioned above control for a large array of family and parental characteristics and some find interactions that are consistent with a sibling effect, the sibling effects they estimate could reflect the impact of unobserved common factors. A few studies, mostly in economics, attempt to identify a sibling causal effect by using instrumental variables strategies. One of them, Oettinger (2000), examines high school graduation by age 19 using the NLSY 1979. He estimates linear probability models of high school graduation of an older sibling on the probability that the younger sibling graduates and vice-versa. To address the endogeneity of the sibling's achievement, he uses the gender of the older sibling, measures of the family's "intactness" during his or her childhood, and local and national unemployment rates at age 18 as instrumental variables. He obtains a significant positive effect of the older sibling's graduation status on the younger sibling's graduation, but no effect of the younger sibling's graduation on the older sibling. This would suggest that a sibling influence runs mostly from the older to the younger sibling, although the exclusion restrictions seem questionable.

Ouyang (2004) develops a dynamic model of the older and younger siblings' behaviors, which allow for state dependence and for the older sibling's behavior to contemporaneously affect that of the younger sibling. She estimates the model with NLSY97 data on cigarettes, marijuana, and alcohol consumption and finds strong evidence of a sibling effect. In contrast to our approach, however, she does not allow for individual specific unobserved heterogeneity, and she proxies family specific heterogeneity with the older sibling's smoking history.

Finally, Harris and Lopez-Valcarel (2008) propose an interesting theoretical model in which siblings learn about whether smoking is desirable or not by observing their siblings' decisions. They allow the decision not to smoke to have a different effect than the decision to smoke. Using data on smoking behavior of family members from supplements to the CPS, they estimate a multivariate probit model in which the number of one's siblings who smoke appears on the right hand side. They find a powerful sibling influence as well as some evidence that the positive effect of smoking is stronger than the deterrent effect of not smoking. However, their estimates imply that the variance of the unobservable that affects the behavior of all siblings is zero. That is, conditional on a limited set of observables, they find that the entire correlation in the behavior of siblings is due to common influences. Consider a family with two siblings. In a simple regression model of the older sibling's behavior on the younger sibling's behavior, one cannot separately identify the causal effect of the younger sibling's behavior from the correlation in error components that determine the two. Although Harris and Lopez-Valcarel (2008)'s model contains exogenous variables and is nonlinear, we suspect that their finding of powerful sibling effects may be due in part to problems in separately identifying the common factors that influence smoking, from the sibling influence.

In sum, there are good theoretical reasons for believing that substance use and other behaviors of adolescents are causally influenced by siblings. However, the strong similarity in the behavior of siblings may be due to genes, shared environments, as well as a direct influence of one sibling on another. To date, little is known about the relative contribution of these mechanisms, let alone the precise nature of sibling interactions.

3 The NLSY97 Data

The empirical analysis uses the first eight rounds of the National Longitudinal Survey of Youth 1997 (NLSY97), which is a panel study of men and women who were between 12 and 16 years of age at the end of 1996. In the first round, the NLSY surveyed 8,984 individuals originating from 6,819 households in the United States. Because the sample design selected all household residents in the appropriate age range, the NLSY97 original cohort includes 1,892 households with more than one respondent. Using information about the relationship between the different respondents of the same household, we created a sample of pairs of biological siblings.

For every year since 1997, the NLSY97 contains extensive information about a wide range of risky behaviors. In this paper, we focus on smoking cigarettes, using marijuana, drinking alcohol, using cocaine and/or other hard drugs, and selling or helping to sell drugs.[5] The main outcome we analyze is whether the individual reports having engaged at all in the particular behavior since the last interview date. We construct this binary variable using answers to questions that were introduced in the survey in 1998 (1999 for cocaine and hard drug use), and we select those observations that are part of uninterrupted sequences of non-missing answers. Because individuals do not answer questions about all behaviors in every round, the analysis sample is slightly different for each behavior. In the case of cigarette smoking for example, the analysis sample is composed of 1646 pairs of siblings, for whom we have between 1 and 6 rounds of observations.

[5]In preliminary work, we also examined gang membership and sexual behavior. We did not find strong evidence of a sibling effect for these variables.

We also estimate models that use reports of the number of days the person engaged in the behavior in the previous month to construct an indicator for high consumption and an indicator for low consumption. We chose 7, 7, and 4 as the maximum number of days for the low consumption category for cigarettes, drinking, and marijuana use, respectively. These cutoffs insure that reasonable fractions of the observations fall in both the high and the low categories. Our results are fairly robust to the choice of cutoffs.

The younger siblings are between 15 to 19 years old when they enter our analysis sample, while the older siblings are between 16 and 20. The average age of the younger sibling is 16.04, while the average age of the older sibling is 18.06. We use all pairs with adjacent birth orders (i.e., the first born with the second born, and the second born with the third born if we have the three oldest siblings in our sample). A total of 1,453 pairs come from two-sibling families, while 375, 12, and 6 come from three, four, and five sibling families respectively.[6] Our sample is 24% Black and 23% Hispanic. The high minority proportions stem from the fact that we use supplemental and military samples along with the cross sectional sample. Unless we indicate otherwise, descriptive statistics and multivariate analyses we report are unweighted, and we do not account for nonrandom attrition.[7]

In all of our empirical work, we control for a set of individual and environmental characteristics. These consist of race, gender, AFQT percentile score, education completed by age 19, number of siblings, birth order dummies, mother's education, and a dummy for whether the child lived with both biological parents at age 12. We also include three dummy variables describing aspects of the individual's environment up to age 12. These consist of an indicator for whether the respondent ever heard gun shots or saw someone

[6]314 of the families who contribute sibling pairs have children who were excluded from NLSY97 because they were older than 16 at the end of 1996. 353 of the families had children who were younger than 12 at the end of 1996. 142 had children who were older than 16 and younger than 12. No data were collected on these children.

[7]One could use inverse probability weighting to account for effects of attrition at the sibling pair level in the correlated random effects analysis, but we are not entirely clear about how to construct the attrition weights for sibling pairs. One possibility would be to estimate the probability that data for a given observation on a sibling pair are available conditional on the age of the youngest sibling in the base year, the age gap, and base year characteristics. We are not sure how to correct for attrition when estimating the joint dynamic discrete choice model given that our models use data from multiple waves of the survey and that the data need depends on the equation of the model.

get shot at with a gun, an indicator for whether her house was broken into, and a third indicator for whether she ever was a frequent victim of bullying.[8] As a sensitivity check, we experimented with using the child's report of the percentage of his peers who engage in the behavior as an additional control, although the behavior of the child may influence his choice of peers. In some models, we use variables that characterize parenting styles and the degree to which the child is influenced by parents and siblings both as controls and as determinants of the strength of the direct sibling influence.

We wish to stress that the identification strategies underlying both the CRE approach and the joint dynamic probit approach relies on panel data, the fact that the future cannot cause the past, and the assumption that substance use by younger siblings does not causally influence that of older siblings. They are specifically designed to be valid in the presence of important omitted variables that influence the substance use of both siblings.[9]

We provide further details about variable construction and sample selection in the Data Appendix. Appendix Table 1 reports the age distribution of the sample. Appendix Table 2 reports unweighted and weighted descriptive statistics for the explanatory variables used in our analysis.

4 Sibling Correlations in Substance Use

To set the stage, we document the strong relationship in substance use among siblings. Table 1 reports the mean values of the substance use measures for males, females, and the combined sample. The values are high for many of the behaviors. For example, 62% of the males and 59% of the females report drinking alcohol during the previous year. 26% of the males and 19.7% of the females report using marijuana. The figure is about 40%

[8]Since the bullying measure reflects a possibly traumatic childhood experience, we think of it as measuring, albeit very imperfectly, aspects of the individual's mental health and social adjustment.

[9]Strictly speaking, the fact that we exclude the older sibling's individual characteristics from the younger siblings's substance use equations, and vice versa, also contributes to identification of the joint dynamic probit and ordered probit models. We impose these questionable exclusion restrictions to reduce the computational burden. They probably play only a minor role in identification of the sibling effect.

for cigarette smoking. About 6% of the sample reports having used hard drugs in the previous year. The unweighted means are similar to the weighted means (see Appendix Table 3). Panel B of the Table 1 reports probabilities of engaging in the activity in at least one year between age 15 and age 20. The probability of ever using hard drugs is .147. The probability of ever selling drugs is .219 for males and .100 for females.[10] The fractions who used the substance one or more days in the past month are lower, not surprisingly. Appendix Table 4 shows that incidence of the behaviors tends to increase with age until about age 20.

In Table 2, we use a regression to summarize the relationship between substance use of the sibling pairs when they were at the same age. Specifically, we report OLS estimates of γ from the regression:

$$y^2_{a,t} = \beta_0 + \gamma y^1_{a,t-j} + X^2\beta_1 + AGE^2_t\Gamma + u^2_{at}$$

where $y^2_{a,t}$ and $y^1_{a,t-j}$ are the behaviors of the younger and older siblings at age a, respectively, j is the sibling age gap, AGE^2_t is a set of age dummies for the younger sibling, X^2 is a vector of controls that refer to the younger sibling and that are listed in Section 2. Throughout the paper, the superscripts 1 and 2 indicate whether a variable refers to the older sibling or the younger sibling, respectively. We also report estimates with controls excluded.[11]

The results are striking. Consider smoking cigarettes. If the older sibling smoked, the probability shifts by .239, which is very large relative to the sample mean of about .4. With controls, the shift in the probability remains large at .18. In the case of marijuana, if the older sibling smoked at a given age, the probability that the younger sibling uses marijuana at that age increases by .162, which is very large relative to the sample mean of about .23. Adding controls leads to only a modest reduction in this figure.

Having an older sibling who uses hard drugs shifts the probability for the younger

[10] These are estimated using sample members who are observed every year between the age of 15 and 20.

[11] The controls are listed in the table footnote.

sibling by .102, a shift that is *larger* than the unconditional mean of about .06. The mean shift for selling drugs is also extremely large relative to the sample mean. In all cases, adding control variables weakens the relationship to some degree, but a strong relationship remains.

We also present separate results for brother pairs and sister pairs. The relationship across siblings tend to be larger for sister pairs, with the exception of selling drugs, a behavior in which females engage infrequently. Later in the paper, we experiment with whether the size of the peer effect depends on the gender match between the older sibling and the younger sibling.

In the remainder of the paper, we address the key but difficult question of whether the sibling correlations are due, at least in part, to a causal effect of the older sibling's behavior.

5 A Model of Substance Use and Sibling Influences

We present a simple model of substance use with the purpose of motivating the econometric strategies used in the paper and to help interpret the parameters. In particular, we treat the sibling influence parameter as a reduced form parameter and do not attempt to identify the specific mechanisms that underlie it, such as information provision, shaping of preferences, etc.[12] As we shall see, even with a very simple formulation, the CRE estimation strategy does not work if there is state dependence from sources such as habit formation or information effects or because of parental reactions to past behavior of children.

Consider a set of families, each with two children. We continue to use a for age and to refer to the older and younger sibling as 1 and 2, respectively. Normalize a so that substance use is 0 for all people when $a < 1$. Without loss of generality, assume that the older sibling is aged 1 in year $t-1$ and the younger sibling is aged 1 in t. Let $y^1_{t+a^1-2}$, denote the behavior of sibling 1 in year $t + a^1 - 2$ when sibling 1 is age a^1. Let $y^2_{t+a^2-1}$,

[12]See the papers cited in the literature survey for a discussion of theories of sibling influence on substance use.

denote the behavior of the younger sibling in year $t+a^2-1$ when he is age a^2. We usually suppress the age subscripts. We also suppress the control variables that appear in our empirical models of y because they are not essential to our identification strategy. We leave the subscript for the family implicit throughout the paper. For present purposes, it is convenient to treat y as a continuous variable and ignore corner solutions at $y = 0$, although we work with a discrete indicator in the empirical work.

5.1 Choices of the older sibling

In every period t, the older sibling chooses y_t^1 to solve the objection function

$$\max_{y_t^1}[\varepsilon + v^1 + u_t^1 + (g+h)y_{t-1}^1 + hy_{t-1}^2]y_t^1 - \frac{1}{2}(y_t^1)^2, \tag{1}$$

where the above expression captures the difference between the perceived benefit and the cost of consuming y_1^1 (including the opportunity costs of foregoing other goods).[13] In the above equation, the term between brackets represents the marginal benefit of an additional unit of y, where ϵ is a family specific component, v^1 is the value of the person specific component for the older sibling, which, without loss of generality, is not correlated with the corresponding component v^2 of the younger sibling. The error component u_t^1 is a transitory error component u for the older sibling in period t. Below we place restrictions on the distributions of u^1 and u^2 over time and across siblings. Additionally, the marginal benefit of an additional unit of y depends on the actions of both siblings in the previous period, through two mechanisms. The first one, captured by the parameter g, is the effect of habit formation and informational effects. The second one, captured by the parameter h, is the effect of the information the parent has about the children, as well as the positive

[13]The above objective function allows for the possibility that agents account for the action's costs and benefits that play out over time. They may also consider the effects of their actions on the utility of others, including parents and siblings. The costs include punishment by the parents, school authorities, criminal sanctions, etc. However, we assume that agents are myopic in the sense that they do not account for the effects of the choice of y today on the marginal costs and benefits of choosing y in the future. The budget constraint, which we leave implicit, is static. Costs of substance use include opportunity cost of consuming other goods. We do not allow the marginal benefit to the older child of an action to depend upon the characteristics or choices of the younger child. Furthermore, older siblings do not consider the influence of their behavior on the younger sibling's choice.

or negative influence of the parents' reaction on the marginal net benefit of y to the older sibling. Although there is no direct peer influence from the younger to the older sibling, the parental response creates an indirect dependence of the older child's behavior on the behavior of the younger child. To keep the notation simple, we assume that the parental reaction affects the marginal benefits of future choices of y by the same amount for all siblings. One could instead assume that parents' reactions are birth order specific or that they only influence the behavior of the particular child. In the latter case, hy^2_{t-1} would drop out of the older sibling's objective function.

The younger sibling faces a similar problem to that of the older one, except that the marginal benefit of his behavior also depends directly on the action of his older sibling in $t-1$. He chooses y^2_t to solve

$$\max_{y^2_t}[\varepsilon + \upsilon^2 + u^2_t + (g+h)y^2_{t-1} + (\lambda_1 + h)y^1_{t-1}]y^2_t - \frac{1}{2}(y^2_t)^2 \ ,$$

where ϵ is the family specific component introduced in (1), υ^2 is a component specific to the younger sibling, and u^2_t is a transitory error component for the younger sibling at time t. The parameter λ_1 captures the direct influence of the older sibling on the younger sibling.[14]

Since the older sibling is age 1 and the younger sib is age 0 in $t-1$, $y^1_{t-2} = 0$ and $y^2_{t-1} = 0$. For periods $t-1$, t, $t+1$, the behavior of the older sibling can be expressed as

$$\begin{aligned}
y^1_{t-1} &= \varepsilon + \upsilon^1 + u^1_{t-1} \\
y^1_t &= \varepsilon + \upsilon^1 + u^1_t + (g+h)y^1_{t-1} \\
y^1_{t+1} &= \varepsilon + \upsilon^1 + u^1_{t+1} + (g+h)y^1_t + hy^2_t
\end{aligned}$$

Notice that the behavior of the younger sibling only starts affecting that of the older sibling at $t+1$ because y^2 is 0 until t.

[14] The parental reaction parameter h and state dependence parameter g could be different for the older and younger children. In the joint dynamic probit model we allow state dependence to differ but do not allow for unobserved heterogeneity in the state dependence or sibling influence parameters.

Similarly, the optimal choices for the younger sibling at t and $t+1$ are

$$\begin{aligned} y_t^2 &= \varepsilon + v^2 + u_t^2 + (\lambda_1 + h) y_{t-1}^1 \\ y_{t+1}^2 &= \varepsilon + v^2 + u_{t+1}^2 + (g+h) y_t^2 + (\lambda_1 + h) y_t^1 \end{aligned}$$

5.2 Using Correlated Random Effects Regression to Estimate the Direct Sibling Effect

Consider the linear least squares projection:

$$y_t^2 = \beta_0 + \beta_1 (y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + error \ . \tag{2}$$

Assume:

A1. $h = 0$, i.e. no parental response

A2. $g = 0$, i.e. no state dependence

A3. The distribution of $u_{a,t}^1$ is covariance stationary over a and t, with $var(u_{a,t}^1) = \sigma_{u^1}^2$. $u_{a,t}^1$ may be serially dependent.

A4. $cov(u_t^2,\, u_{t-1}^1) = cov(u_t^2, u_{t+1}^1)$.

In this case,

$$\begin{aligned} y_{t-1}^1 &= \varepsilon + v^1 + u_{t-1}^1 \\ y_{t+1}^1 &= \varepsilon + v^1 + u_{t+1}^1 \end{aligned}$$

Using the above equations and A3 and A4, it is easy to show that the coefficients of the projection of $\varepsilon + v^2 + u_t^2$ onto y_{t-1}^1 and y_{t+1}^1 both equal $[var(\varepsilon) + cov(u_t^2, u_{t-1}^1)]/[2var(\varepsilon) +$

$var(v^1) + \sigma^2_{u^1} + cov(u^1_{t-1}, u^1_{t+1})]$. Consequently, β_1 and β_2 in (2) are given by

$$\begin{aligned}
\beta_1 &= [var(\varepsilon) + cov(u^2_t, u^1_{t-1})]/[2var(\varepsilon) + var(v^1) + \sigma^2_{u^1} + cov(u^1_{t-1}, u^1_{t+1})], \\
\beta_2 &= \lambda_1.
\end{aligned}$$

In this case, β_2 is λ_1, the direct sibling effect. The basic argument carries over to the case in which y is a binary variable determined according to:

$$y^1_{t-1} = 1(\varepsilon + v^1 + u^1_{t-1}) > 0) \tag{3}$$

$$y^1_{t+1} = 1(\varepsilon + v^1 + u^1_{t+1} > 0) \tag{4}$$

$$y^2_t = 1(\varepsilon + v^2 + u^2_t + \lambda_1 y^1_{t-1} > 0), \tag{5}$$

although one must replace A3 with the assumption that the $u^1_{a,t+a-1}$ are identically distributed. However, if any of the four assumptions above are false, then $\beta_2 \neq \lambda_1$ in (2), except in special cases. Indeed, if any of the assumptions fail, then the coefficients of the projection of $\varepsilon + v^2 + u^2_{1t}$ on y^1_{t-1} and y^1_{t+1} will differ, and the difference will be reflected in β_2. For the same reason, if the effects of ε or v^1 on $y^1_{a,t+a-1}$ vary with a, as would be the case if preferences and costs are such that

$$y^1_{a,t+a-1} = f(a) + g_a\varepsilon + \pi_a v^1 + u^1_{a,t+a-1},$$

where g_a and π_a are age dependent coefficients, then the equality restriction on the coefficients of the projection of $\varepsilon + v^2 + u^2_t$ on y^1_{t-1} and y^1_{t+1} will fail. The function $f(a)$ is not a problem if the model is additively separable in age, provided that one also controls for the age of each of the siblings in year t. However, in a nonlinear setting such as (4) the presence of $f(a)$ is enough to invalidate the restriction on the projection coefficients, even if g_a and π_a do not depend on age.

5.2.1 Bias If the Younger Sibling Influences the Older Sibling

If the assumption that younger siblings do not influence older siblings fails and the younger sibling positively influences the behavior of the older sibling, then we are likely to underestimate the sibling effect λ_1. To see why, maintain assumptions 1-4, but now allow the behavior of the older sibling in $t+1$ to depend upon y_t^2 with coefficient γ_1. Then the model becomes

$$\begin{aligned} y_{t-1}^1 &= \varepsilon + v^1 + u_{t-1}^1 \\ y_{t+1}^1 &= \gamma_1 y_t^2 + \varepsilon + v^1 + u_{t+1}^1 \\ y_t^2 &= \lambda_1 y_{t-1}^1 + \varepsilon + v^2 + u_t^2. \end{aligned}$$

Reparameterize (2) as:

$$y_t^2 = \beta_0 + \beta_1 y_{t+1}^1 + (\beta_2 + \beta_1) y_{t-1}^1 + error.$$

The dependence of y_{t+1}^1 on y_t^2 will raise the coefficient on y_{t+1}^1 relative to the coefficient on y_{t-1}^1. This will reduce the estimate of the causal effect of y_{t-1}^1 since the estimate is the difference in the coefficients on y_{t+1}^1 and y_t^2. In the presence of state dependence, the implications of reverse causality are less transparent. However, it will tend to increase the strength of the link between future values of y^1 and past values of y^2. Intuitively, we expect that this will lead to underestimation of the direct sibling influence in econometric models that assume that the sibling influence goes in only one direction. Note 22 below summarizes a simulation experiment that supports this intuition.

5.2.2 Contemporaneous Sibling Effects

Suppose both contemporaneous and lagged behaviors of the older sibling influence the younger child with coefficients λ_0 and λ_1, respectively. Consider the projection equation:

$$y_t^2 = \beta_0 + \beta_1 (y_{t-1}^1 + y_t^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + \beta_3 y_t^1 + error \tag{6}$$

In addition to assumptions A1-A4 above, assume:

A5. The idiosyncratic error components u_t^2 and $u_{t'}^1$ are independent across siblings at all leads and lags.

A6. u_t^1 is serially uncorrelated.

Then,

$$\beta_1 = var(\varepsilon)/[3var(\varepsilon) + var(v^1) + \sigma^2_{u^1}]$$

and

$$\beta_2 = \lambda_1 \;\text{ and }\; \beta_3 = \lambda_0$$

where λ_0 is the direct effect of y_t^1 on y_t^2. Consequently, under the six assumptions, one can identify the direct sibling effects.

However, if any of the assumptions A1 through A6 fails, then in general $\beta_2 \neq \lambda_1$ and $\beta_3 \neq \lambda_0$ in (6). (Nonseparable forms of age dependence will also pose problems in this case.) If only A6 fails, one can still estimate an average of λ_0 and λ_1 and test, as we do below, for sibling effects using the regression:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_t^1 + y_{t+1}^1 + y_{t+2}^1) + \beta_2(y_{t-1}^1 + y_t^1) + error. \tag{7}$$

We are particularly concerned that temporal variation in factors such as stresses within the family (e.g., parental unemployment, marital conflict, parental substance abuse) or variation in access to drugs or alcohol in a neighborhood or in a school will lead u_t^2 and u_t^1 to co-vary. Consequently, we place less weight on specification (7). If one uses (2) when (6) is correct, then the coefficient on y_{t-1}^1 will pick up part of the effect of y_t^1, but we will still detect sibling influences.

The idea of using the difference between the effect of the past or contemporaneous value and the future value of a treatment variable to identify the causal effect of the treatment is, of course, a standard approach in the program evaluation literature. However, it is sometimes forgotten that one requires strong assumptions about how time or age interacts with the error components that influence both the treatment and the outcome. One also needs strong restrictions on dynamics. An alternative approach is to estimate a joint dynamic model of the outcome of interest and the "treatment", which, in our case, is the past behavior of the older sibling. In Section 5.3, we provide such a model, which is close in spirit to the model above. Nevertheless, the CRE approach has the advantage of simplicity. While state dependence and nonstationarity will lead to inconsistency in the estimates of λ_1, it is a natural place to start the search for evidence of a causal effect of sibling behavior on substance use.

5.2.3 Fixed Effects Estimation

We also estimate the sibling influence parameter in

$$y_t^2 = \lambda_1 y_{t-1}^1 + \varepsilon + v^2 + u_t^2, \tag{8}$$

treating $\varepsilon + v^2$ as a fixed effect. The advantage of the fixed effects estimator is that it requires assumptions A1 and A2, but not A3. On the other hand, it requires A5, while u_t^2 and $u_{t'}^1$ may be correlated in the case of the CRE procedure subject to A1-A4. This is a substantial disadvantage. A second disadvantage is that the fixed effect estimator requires multiple observations on the younger sibling, which reduces power. When we include fixed effects we use a linear probability model rather than a probit specification.

5.3 A Joint Dynamic Model of Sibling Behavior

We work with two specifications. The first treats substance use as a binary choice. The second distinguishes the level of consumption. We start with the binary choice model, which has four equations. Since behavior is dynamic and we do not observe behavior at the age of initiation, we include an equation for the initial condition of the older sibling. This equation refers to the choice in year $t_{\min}^1$, the first year we observe the older sibling. The second equation refers to the choice of the older sibling in year t, given the choice in the previous year. The last two equations are the initial conditions for the younger sibling and the younger sibling's choice in year t given the younger sibling's choice in $t-1$. The initial condition refers to behavior in $t^2_{\min}$, the first year we observe the behavior of the younger sibling.[15]

Older sibling's choice at $t = t^1_{\min}$ (initial condition):

$$y^1_t = \mathbf{1}(X^1\beta^1_1 + AGE^1_t\Gamma^1_1 + \varepsilon + v^1 + u^1_t > 0) \tag{9}$$

Older sibling's choices at $t > t^1_{\min}$:

$$y^1_t = \mathbf{1}(\gamma^1 y^1_{t-1} + X^1\beta^1_2 + AGE^1_t\Gamma^1_2 + \delta^1_{2\varepsilon}\varepsilon + \delta^1_{2v}v^1 + u^1_t > 0) \tag{10}$$

Younger sibling's choice at $t = t^2_{\min}$ (initial condition):

$$y^2_t = \mathbf{1}(\lambda^2_1 y^1_{t-1} + X^2\beta^2_1 + \theta^2_1 a^1_{t-1} + AGE^2_t\Gamma^2_1 + \delta^2_{1\varepsilon}\varepsilon + v^2 + u^2_t > 0) \tag{11}$$

Younger sibling's choices at $t > t^2_{\min}$:

$$y^2_t = \mathbf{1}(\gamma^2 y^2_{t-1} + \lambda^2_2 y^1_{t-1} + X^2\beta^2_2 + \theta^2_2 a^1_{t-1} + AGE^2_t\Gamma^2_2 + \delta^2_{2\varepsilon}\varepsilon + \delta^2_{2v}v^2 + u^2_t > 0) \tag{12}$$

where:

- y^1_t and y^2_t are the behaviors of the older and younger siblings in year t
- X^1 and X^2 are vectors of control variables for the older and the younger siblings
- a^1_{t-1} is the age of the older sibling in year $t-1$
- AGE^1_t and AGE^2_t are vectors of age dummies indicating whether the sibling is aged a in year t
- $\varepsilon \sim N(0, \sigma^2_\varepsilon)$ is a sibling pair specific error component
- $v^1 \sim N(0, \sigma^2_{v^1})$ and $v^2 \sim N(0, \sigma^2_{v^2})$ are independent person specific error components
- $u^1_t \sim N(0,1)$ and $u^2_t \sim N(0,1)$ are person/time specific error components that are independent across siblings and years

[15]The value of $t^1_{\min}$ varies from 1998 to 2000 in (9) while $a^1_{t^1_{\min}}$ ranges from 15 to 20. The value of $t^2_{\min}$ varies from 1999 to 2001 while $a^2_{t^2_{\min}}$ varies from 15 to 19.

- λ_1^2 is the sibling influence parameter in the initial condition ($t = t_{\min}^2$)

and

- λ_2^2 is the sibling influence parameter for $t > t_{\min}^2$.

We allow the coefficients on X to be different in the initial conditions and in the equations for the later periods. We also allow them to differ between the older and younger siblings. We experiment with two specifications for the error structure. Error specification A restricts the factor loadings on the family effect ε to be 1 in all equations. It also restricts the factor loadings on the individual effects v^1 and v^2 to equal 1 in all equations. That is, $\delta_{2\varepsilon}^1 = \delta_{2\varepsilon}^2 = \delta_{1\varepsilon}^2 = 1$ and $\delta_{2v}^1 = \delta_{2v}^2 = 1$. Note, however, that we allow the variance of the individual person specific effect v to differ between the older sibling and the younger sibling. In error specification B, we allow all the factor loadings on the family effect ($\delta_{2\varepsilon}^1$, $\delta_{2\varepsilon}^2$, and $\delta_{1\varepsilon}^2$) and the factor loadings on the individual personal specific effect in the later equations (δ_{2v}^1 and δ_{2v}^2) to be freely estimated. We restrict the variance of v to be the same across siblings.[16] For some outcomes, we have difficulty identifying the separate roles of family heterogeneity and individual heterogeneity when we use the less restricted version.

We also experimented with a more general version of the above model in which we use error specification A but allow linear interactions between the elements of X^1 and a_t^1 in (9) and (10) and linear interactions between X^2 and a_t^2 in (11) and (12).[17] For the most part, the state dependence parameters and sibling effects parameters are not very sensitive to addition of the interaction terms, and so we present the models without the

[16]Note that we restrict the variance of the idiosyncratic error components to be 1 in both the initial condition and the later years. This is implicitly a normalization, because we allow the coefficients of all variables to differ across these equations for both the older and younger siblings.

[17]One would expect age interactions to be particularly important in the initial conditions equation. We do not allow the state dependence effects or the sibling influence effects to interact with the age of the older sibling or the age of the younger sibling in the specification of the index for the latent variable that determines y_t^1 and y_t^2.

interaction terms.[18] We estimate the models by maximum likelihood.[19]

5.3.1 A Dynamic Ordered Probit Model

The degree of state dependence and the strength of the peer influence are likely to depend on the amount of substance use. To investigate this parsimoniously, we supplement our main analysis by estimating a joint dynamic ordered probit model. Consider cigarettes. We define y^1_{Lt} equal to 1 if the person smoked between one and 7 days during the last month. We define y^1_{Ht} to be 1 if the older sibling smoked more than 7 days in the previous month. The corresponding threshold values are 7 days for alcohol and 4 days for marijuana. The indicators are determined according to $y^1_{L,t} = 1(q_H \geq y^{1*}_t > q_L)$ and $y^1_{H,t} = 1(y^{1*}_t \geq q_H)$ where q_H and q_L are threshold parameters and y^{1*}_t is the latent index given by:

$$y^{1*}_t = \gamma^1_H y^1_{H,t-1} + \gamma^1_L y^1_{L,t-1} + X^1\beta^1_2 + AGE^1_t\Gamma^1_2 + \delta^1_{2\varepsilon}\varepsilon + \delta^1_{2v}v^1 + u^1_t,\ t > t^1_{\min}.$$

The initial condition for y^{1*}_t for $t = t_{\min}$ is determined by an ordered probit model that is an obvious generalization of (9). We expect $\gamma^1_H > \gamma^1_L$, since the positive influence of habit, social connections, and information on the propensity to engage in substance use is likely to be increasing in the quantity consumed in the previous period.

Similarly, the younger sibling's choice is summarized by $y^2_{L,t} = 1(q_H \geq y^{*2}_t > q_L)$ and

[18]For both error specifications A and B, the state dependence parameter for the younger sibling is lower for all five behaviors when age interaction terms are added. In the case of error specification B, the sibling influence parameters are higher for all of the behaviors except smoking, although the coefficients are also less precisely estimated. Some of the factor loadings change, but there is no clear pattern.

[19]For computational ease, each pair coming from the same household is assumed to receive an independent draw of the common component ε. This means that we are implicitly allowing for the possibility that the common household environment is sibling pair specific. Our reported standard errors for the joint dynamic probit and ordered probit models (see below) do not account for the possible error correlation across pairs that come from the same household. Relatively few households supply more than one pair of observations, so any bias in the standard errors is likely to be small (see Section 3). Standard errors for the regression and probit results in Tables 2, 3, A5, and A6 are clustered at the household level.

$y_{H,t}^2 = 1(y_t^{*2} \geq q_H)$, where

$$\begin{aligned} y_t^{2*} &= \lambda_{1H}^2 y_{H,t-1}^1 + \lambda_{1L}^2 y_{L,t-1}^1 + X^2\beta_1^2 + \theta_1^2 a_{t-1}^1 + AGE_t^2\Gamma_1^2 + \delta_{1\varepsilon}^2\varepsilon + v^2 + u_t^2,\ t = t_{\min}^2 \\ y_t^{2*} &= \gamma_H^2 y_{H,t-1}^2 + \gamma_L^2 y_{L,t-1}^2 + \lambda_{2H}^2 y_{H,t-1}^1 + \lambda_{2L}^2 y_{L,t-1}^1 + X^2\beta_2^2 + \theta_2^2 a_{t-1}^1 \\ &\quad + AGE_t^2\Gamma_2^2 + \delta_{2\varepsilon}^2\varepsilon + \delta_{2v}^2 v^2 + u_t^2,\ t\ > t_{\min}^2\ . \end{aligned}$$

We expect the state dependence parameters to obey $\gamma_H^2 > \gamma_L^2 > 0$. If sibling influences are positive and increasing in the intensity of the older sibling's behavior, then $\lambda_{1H}^2 > \lambda_{1L}^2 > 0$ and $\lambda_{2H}^2 > \lambda_{2L}^2 > 0$. One can easily generalize the model to allow for additional positive categories. We stop at two because of sample size considerations. Error specification A and error specification B are the same as in the binary probit case.

6 Sibling Effect Estimates Based on the CRE Model

In Table 3, we present estimates of sibling effects using the correlated random effect model discussed in section 5.2. Each column refers to a different outcome. The top panel of Table 3 presents estimates of our main specification, which we refer to as Model 1. Model 1 is a variant of (2) for the case in which y_t^2 is binary and control variables and age dummies are added:

$$y_t^2 = \mathbf{1}(\beta_0 + \beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + X^2\beta_3 + AGE_t^1\Gamma_1 + AGE_t^2\Gamma_2 + e > 0). \qquad (13)$$

In the middle panel, we replace $\beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1$ with $\beta_1(y_{t-2}^1 + y_{t-1}^1 + y_{t+1}^1 + y_{t+2}^1) + \beta_2(y_{t-2}^1 + y_{t-1}^1)$ (Model 2). If the sibling influence operates over two or more periods, adding the additional lead and lag might increase power, but it comes at a substantial cost in sample size. In the bottom panel, we allow for the possibility of a contemporaneous influence. We replace $\beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1$ with $\beta_1(y_{t-1}^1 + y_t^1 + y_{t+1}^1 + y_{t+2}^1) + \beta_2(y_{t-1}^1 + y_t^1)$ (Model 3). As we discussed in Section 5.2.2, the peer influence coefficient on $(y_{t-1}^1 + y_t^1)$ in Model 3 are likely to be positively biased if transitory environmental factors are correlated across siblings.

We report marginal effects of the raw variables on the probability that $y_t^2 = 1$ based on MLE probit estimates of β_1, β_2 and the other parameters in the model. Standard errors are clustered at the household level.[20]

Column 1 refers to smoking. The results for Model 1 indicate that y_{t-1}^1 raises the smoking probability by .063 (.027). This estimate is statistically significant and substantial relative to the mean probability. The marginal effect of $(y_{t-1}^1 + y_{t+1}^1)$ is .086, so about 3/5th of the link between the older sibling's past smoking and the younger sibling's current smoking is due to common influences and 2/5th is due to the sibling effect. The results for Model 2 and Model 3 suggest an even stronger causal sibling effect on smoking.

For drinking, the estimates of Model 1 indicate that y_{t-1}^1 raises the probability of drinking by about .060 (.025), while the link due to common influences is .116. The evidence for a causal effect in the case of marijuana is weak. The estimates are positive, but are statistically significant only in the case of Model 3, which allows for a contemporaneous influence.

The point estimate for use of hard drugs and selling drugs are positive and substantial relative to the sample mean. For example, in the case of hard drugs, the marginal effect of y_{t-1}^1 is .0147 (0.0204) for Model 1 while the sample mean is .062. However, the effect is not statistically significant. We obtain even larger estimates using Model 2 and Model 3. In the case of selling drugs, we obtain a large, positive, and statistically significant estimate using Model 2. Overall, the results for hard drugs and selling drugs suggest a positive causal effect but are too noisy to support strong conclusions.

[20]The sample sizes differ substantially across models due to the requirement for additional leads and lags in the case of Model 2 and, to a minor extent, the loss of observations due to missing data on y_t^1 in the case of Model 3. In Appendix Table 5, we report the marginal effects of the control variables for Model 1. The estimates for variables that are correlated across siblings are reduced by about 10% in absolute value by the presence of y_{t-1}^1 , y_{t+1}^1 and the age dummies for the older siblings. We also experimented with a number of additional controls, including self reports of the percentage of peers who engage in the behavior. These did not have much effect on the correlated random effects estimates or the joint dynamic probit estimates of the sibling influence parameters.

6.1 Fixed Effects Estimates

In Appendix Table 6, we report estimates of (8) using dummy variables for each younger sibling, thus treating $\varepsilon + v^2$ as a fixed effect for every sibling pair. The estimates of the coefficient on y^1_{t-1} are .029 (.014) for smoking and .043 (.015) for drinking. Both coefficients are significant at the .05 level, but are smaller than the estimates based on (13). We also obtain a small positive coefficient for marijuana that is larger than the CRE estimate, but is significant at only the 0.25 level. The coefficients for use of hard drugs and selling drugs are also positive and close to the CRE values but not statistically significant. Thus the results are qualitatively consistent with our findings based upon (13), but the point estimates tend to be smaller. We do not know why this is the case, although the nature of the variation in the behavior of the older sibling that the two estimators use to identify the sibling effect is different. The difference in the magnitude across estimation strategies is robust to selecting the sample for (8) to match the sample for (13) and to using a linear probability specification for the CRE model in place of the probit specification.[21]

7 Estimates of the Joint Dynamic Probit Model

We now turn to estimates of the joint dynamic probit model. Table 4A presents the results for error specification A, our basic specification. The first column reports the results for smoking cigarettes. The estimates of the state dependence parameter λ are .947 (.068) for the younger sibling and .899 (.064) for the older sibling. Thus, lagged behavior matters. Dynamic simulations reported below indicate that smoking today raises the probability that the older sibling smokes by .196 (.021) next year and by .051 (.010) two years out.

The value of $\hat{\sigma}_{\varepsilon}$ is .710. This confirms the CRE result that there is a substantial common error component that drives the smoking behavior of siblings. We also find an important individual specific error component: $\hat{\sigma}_{v^1}$ and $\hat{\sigma}_{v^2}$ are 1.05 and .815, respec-

[21] Keep in mind that in Table 3, we report marginal effects on the probability of substance use rather than the probit coefficients.

tively. Consequently, temporal correlation in cigarette smoking comes from the influence of the family specific and individual specific error components, as well as from true state dependence.

Next we turn to the sibling influence parameters λ_1^2 and λ_2^2, which are the coefficients on y_{t-1}^1 in the equations for y_t^2. A priori, we would expect both to be positive. We also would expect λ_1^2 to exceed λ_2^2 because we do not condition on y_{t-1}^2 in the initial condition. $\hat{\lambda}_1^2$ is .215 (0.101), which is significant at the 1% level. Comparing this value to the state dependence term indicates that having the older sibling smoke shifts the latent variable for smoking by about one fourth the amount that smoking in the past does. The coefficient $\hat{\lambda}_2^2$ for subsequent years is .049 (0.069), which is positive but not significant. [22]

Column 2 reports results for drinking. We find strong evidence of state dependence, although the lag coefficient is somewhat smaller than for cigarette smoking. One must keep in mind that the coefficients on the lagged dependent variables should be judged relative to the standard deviation of the composite error, which is smaller for drinking than for smoking. Nevertheless, the dynamic simulations in Table 5 indicate that state dependence is indeed a bit weaker for drinking.

The sibling influence parameter λ_1^2 is .411 and is highly significant. The estimate of λ_2^2 is close to zero and insignificant. The results suggest that siblings have a substantial influence at early ages but not later, which makes some intuitive sense, but we would have expected less of a difference between λ_1^2 and λ_2^2. The results for the other error structures are basically similar.

Column 3 reports estimates for marijuana. The results are very similar to the results for drinking. We find strong evidence for a sibling effect that operates primarily

[22]In Section 5.3.1 we argued that our estimates of λ_1 and λ_2 will be biased downward if younger siblings influence older siblings. To investigate this, we simulated data from the joint dynamic probit model after adding a term that allows the younger sibling to positively influence the older sibling. We set the coefficient to a positive value. All other parameters were set to the estimates for the dynamic probit model for smoking reported in Table 4A. We then used the simulated data to estimate the model with the parameter governing influence of the younger sibling on the older sibling set to 0, and examined the effect on sibling influence parameters λ_1 and λ_2 in the dynamic probit model presented below, confirming our conjecture. As expected, the estimates of λ_1 and λ_2 decline when the data come from a model in which younger sibling influences the other sibling. We also used the simulated data to examine the behavior of the estimates of λ_1 using the CRE specification(2). Increasing the size of the effect of the younger sibling on the older sibling leads to a reduction in the coefficient on y_{t-1}^1.

through the initial condition. The point estimate of λ_2^2 is actually negative, although it is not significant. Overall, the evidence from the dynamic model for a sibling effect on marijuana use is substantially stronger than the evidence from the CRE model. We also find substantial state dependence and an important role for both family and individual heterogeneity.

Column 4 reports results for the use of hard drugs. Qualitatively, the results are similar to the results for drinking and marijuana use. The point estimates suggest a considerable sibling influence. However, the estimates are not statistically significant. In the case of selling drugs (column 5), we find that family heterogeneity is less important relative to individual heterogeneity. State dependence in this behavior is substantial. The point estimates of the peer influence terms are large in magnitude and substantial relative to the state dependence term, but they are not statistically significant. We do not have enough power to determine whether there is an important sibling influence on selling drugs.[23]

Table 4B reports estimates using error restriction B, which allows the factor loadings in ε to differ between the younger and older siblings and to differ between the initial condition and the subsequent periods. The results for alcohol and marijuana are similar to those in Table 4A and show strong evidence of a sibling influence. Sibling coefficients for hard drugs and selling drugs rise, but are imprecisely estimated. In the case of cigarettes, the sibling coefficient in the initial condition falls while the sibling coefficient for subsequent periods rises, although neither is statistically significant.

Overall, the evidence from the joint dynamic probit model points to a positive sibling effect on substance use. The evidence is strongest for smoking, drinking and marijuana use, although the point estimates are positive for hard drug use and selling drugs.

Investigating Possible Bias from Treatment of Initial Conditions The fact we typically find a stronger sibling effect in the initial condition than in the equation for subsequent periods could reflect that fact that λ_1^2 captures influence over more than one

[23]We noted earlier that selling drugs is more a male than a female activity. Our model includes a gender dummy but does not allow the factor loading on the family error component to depend upon gender. This may have the effect of increasing the importance of the individual specific error component. Below we discuss models that allow the sibling influence to depend upon the gender pairing.

period but also raises questions. We conducted a simulation exercise to investigate the possibility that misspecification of the initial condition biases upward the estimate of λ_1^2 and biases downward λ_2^2. We generated data from our model for smoking from age 13 forward using the estimated parameter values (Table 4A, column 1). We then estimated the model using the simulated data corresponding to the ages that we see in the NLSY97. The data were generated with λ_1^2 set to 0.215 and λ_2^2 set to 0.0488. The estimates of λ_1^2 and λ_2^2 using the simulated data are 0.1330 (0.0239) and 0.0486 (0.0157) (table omitted). These results suggest that there is little bias in λ_2^2 and that, if anything, we are underestimating λ_2^2.[24]

7.1 The Dynamic Response to the Older Sibling's Substance Use Behavior

The estimates of the parameters of the dynamic probit model refer to effects on the latent variable index rather than to effects on the probability of substance use. Furthermore, they do not provide a quantitative sense of how persistent the effects are. To address these issues, we simulate the effect of an exogenous switch in the behavior of the older sibling from 0 to 1 in period $(t^2_{\min} - 1)$ on the time paths of substance use of both the older and younger siblings.[25]

Figure 1a presents the results for smoking using the model reported in the first two columns of Table 4A. The vertical axis measure the change in behavior relative to the

[24]Interestingly, the state dependence parameters for the younger sibling seem to be underestimated and the variance of person specific error component for the younger sib (v^2) seems to be overestimated.

[25]In all but four cases, $t^1_{\min} = (t^2_{\min} - 1)$, and so we use the actual age of the older sibling in creating the D^1_{at}. In the 4 cases, we set the age of the older sibling in year $t^1_{\min}$ to the actual age minus the value of $(t^2_{\min} - t^1_{\min} - 1)$ for the pair and construct dummies for subsequent years accordingly.

We obtain the mean baseline path as follows. Using the sample distribution of X^1 and estimated parameters based on error structure 1, we first simulate y^1_t from $(t^2_{\min} - 1)$ to $(t^2_{\min} + 4)$ using (9) and (10). With simulated values of y^1_t and the estimated model parameters for the younger siblings, we simulate y^2_t from $t^2_{\min}$ to $(t^2_{\min} + 5)$ using (11) and (12). All error terms are drawn from the distributions implied by the model estimates. We obtain the effect of an exogenous shift in behavior of the older sibling from 0 to 1 in period $(t^2_{\min} - 1)$ by conducting a similar simulation with $y^1_{t^2_{\min}-1}$ to 0 for all pairs rather than the value implied by (9) and a simulation with $y^1_{t^2_{\min}-1}$ to 1 for all pairs. For each sibling pair i, we performed each of the three simulations 20 times. We then average over the 20 simulations for all the pairs.

baseline probability. The horizontal axis measure the time period relative to $(t^2_{\min} - 1)$, so 0 corresponds to $(t^2_{\min} - 1)$. Appendix Table 7A reports point estimates and standard errors, which are based on a parametric bootstrap method.[26]

We begin with the older sibling's response. The solid line graph reports the effect of exogenously switching y^1_{t-1} from 0 to 1 on the time path of the average value of $y^1_{a,t}$, relative to the baseline average for $y^1_{a,t}$.[27] The vertical bars represent 90% confidence bands. One can see that the exogenous change in smoking behavior from 0 to 1 in $(t^2_{\min} - 1)$ raises the probability of smoking 1 year later by .467 of the baseline value (.419). The effect is .119 of the baseline value 2 years later and essentially dies out after 4 periods.

The broken line in the graph displays the effect on the time path of $y^2_{a,t}$, relative to the baseline average of $y^2_{a,t}$, of a one-time exogenous shift in the smoking behavior of the older sibling from 0 to 1 in $(t^2_{\min} - 1)$, with the distribution of the future behavior of the older siblings unaffected. Smoking by the older sibling increases smoking among younger siblings in $t^2_{\min}$ by .143 (.087) of the baseline value. This is 31 percent of the effect of the older sibling's behavior in $(t^2_{\min} - 1)$ on his own behavior in the next period. The value is .034 (.021) in the second period. The effect on the probability of the younger sibling smokes relative to baseline is essentially zero after three years.[28]

Figure 1b displays simulations for drinking. For the older sibling, drinking last year period raises the probability of drinking this year .292 (.042) of the baseline value. The

[26]We draw 75 values of the parameter vector for the joint dynamic probit model from a multivariate normal distribution with mean and variance matrix set to the point estimates of mean and variance of the parameter vector. For each draw of the parameter vector we perform 20 simulations and take the average, as described in the previous footnote. The standard errors are the standard deviations across the 75 averages. The 90% confidence bands are computed from the point estimate and standard error estimates under a normality assumption.

[27]To be more specific, for each older sibling we first set $y^1_{a,t-1}$ to 1 in $(t^2_{\min} - 1)$, simulate forward, and take the average of $y^1_{a,t-1}$ for the values of $t - 1$ reported in the column. We repeat the procedure with $y^1_{a,t-1}$ set to 0 in $(t^2_{\min} - 1)$, take the difference in the two averages for each value of $t - 1$, and then divide by baseline value in the top row of Appendix Table 7A.

[28]In Appendix Table 7A, we report the baseline simulation for $y^2_{a,t}$. In the rows for the younger sibling labelled "W/ Feedback" we report the path of the difference in $y^2_{a,t}$ relative to the baseline simulation for younger siblings when $y^1_{a,t-1}$ is set to 1 in $(t^2_{\min} - 1)$ and when it is set to 0 in $(t^2_{\min} - 1)$, respectively, and the shift in $y^1_{a,t-1}$ is allowed to affect future values of $y^1_{a,t-1}$ in accordance with the model. The effect of the shift on $y^2_{a,t}$ is the same in $t^2_{\min}$ (by construction). It is a bit larger in subsequent periods, because of the persistence in the behavior of the older sibling when we allow for feedback. However, the values are pretty similar to the effect of a one time shift in the older sib's behavior, which are reported in the rows W/ out feedback and graphed in Figure 1.

baseline value is about .56. After three periods the effect is only .01 (.004) of the baseline value. An exogenous change in the drinking behavior of the older sibling in $(t^2_{\min} - 1)$ increases drinking among younger siblings by .240 (.071) of the baseline probability (.47). The effect on the younger sibling is essentially zero after three periods.

Figure 1c graphs changes relative to baseline for marijuana. Marijuana use by the older sibling in $(t^2_{\min} - 1)$ increases the probability that the older sibling uses marijuana 1 year later by 0.58 (.08) of the baseline probability (.27). The effect on the older sibling's behavior is .0229 (.009) 3 years later and close to 0 after that. A one-time exogenous shift in the smoking behavior of the older sibling from 0 to 1 in $(t^2_{\min} - 1)$ increases the probability that the younger sibling uses marijuana in $t^2_{\min}$ by .239 of the baseline value. The effect on the younger sibling is under .01 after two periods.

When we use the model parameters for error specification B to perform the simulations, we obtain similar results to those in the figure in the case of marijuana and drinking (see Appendix Table 7B). However, the effect of smoking by the older sibling on the younger sibling is essentially zero, although the standard error is large.

Overall, the effects of substance use by the older sibling in one period on the younger sibling are substantial, but die out fairly quickly. It is important to note that most of our parameter estimates indicate that the peer influence is biggest in the initial condition for the younger sibling. For this reason, when we simulate the average effect of exogenously shifting the behavior of the older sibling from no substance use in all periods to substance use in all periods, we find only modest effects on the behavior of the younger sibling for $t > t^2_{\min} + 2$ (not reported). The fact that our estimates imply the younger siblings' behavior is relatively insensitive to whether the older sibling consumes the substance in all periods versus not at all indicates that only a small part of the strong substance use correlations reported in Table 2 are causal.

8 Ordered Probit Results

We now turn to the estimates of the joint dynamic ordered probit models using error specification A, which are reported in Table 5.[29] We limit the analysis to smoking cigarettes, drinking alcohol, and smoking marijuana because these behaviors are more common in the sample. The estimates of the sibling influence parameters in the initial conditions, λ^2_{1H} and λ^2_{1L}, are both positive for all three outcomes. In the case of cigarettes and drinking, λ^2_{1H} is larger than λ^2_{1L}, which accords with our expectation, and is statistically significant. The opposite is true in the case of marijuana, but the standard errors on these estimates are substantial. The estimates of λ^2_{2H} and λ^2_{2L}, the sibling influence parameters for the periods $t > t^2_{\min}$, are small and not always positive. In the case of alcohol, λ^2_{2H} is actually negative and statistically significant at the 5% level. This runs counter to our expectations and is troubling. However, we are looking at results for multiple parameters so sampling error might be the explanation.[30] In keeping with the binary probit results, we find that both family heterogeneity and individual heterogeneity are important for all three outcomes. We also find evidence of substantial state dependence for both the older sibling and the younger sibling. As expected, γ^1_H, the coefficient on the indicator $y^1_{H,t-1}$for the high consumption level, is substantially larger than γ^1_L, the coefficient on the indicator for the low consumption level. The same is true of the state dependence parameters for the younger sibling.

Figures 2a and b graph simulation based estimates of the effects of an exogenous shift in the behavior of the older sibling from no smoking to the highest consumption category in $t^2_{\min} - 1$. (Point estimates and standard errors are in Appendix Table 9A.) One period

[29]The results based upon error specification B are similar. See Appendix Table 8 and corresponding simulations in Appendix Table 9B

[30]We examined the sensitivity of our results to the specific categories we chose for the cut off values of y_H and y_L. In the case of smoking and drinking, we estimated the models using all possible partitions between 0/1-3/4-30 days in the last month to 0/1-20/21-30 days (the partition we actually use is 0/1-7/8-30 days). In the case of marijuana, we tried all partitions ranging from 0/1-2/3-30 days to 0/1-14/15-30 days. The state dependence coefficients on y_L and y_h tend to rise a bit as we increase the cutoff between y_L and y_H. The sibling effect parameters do not vary much relative to standard errors, although in the case of marijuana the sibling effect parameters tend to be a bit larger for partitions in the range of 0/1-4/5-30, which is the one we report results for, than when we choose a high cutoff between y_L and y_H.

later, the shift raises the low consumption probability for the older sibling by .19 and the high consumption probability by .47 relative to the baseline averages. The effects become very close 0 after four periods. The shift in the older sibling's behavior increases the probability that the younger sibling is in the high consumption category one year later by .316 relative to baseline and also boosts the probability of low consumption. The effects are very small after two periods. In the case of marijuana (Figures 3a and b) and alcohol (Figures 4a and b), the dynamic effect on the behavior of the younger sibling is similar to smoking, but smaller.

9 Determinants of the Strength of the Sibling Effect

In the section, we examine a number of possible determinants of the sibling effect, including the gender mix, the age gap, and a number of family process and relationship variables. For simplicity, and because power is limited, we restrict the analysis to the binary probit specification.

9.1 Gender Mix Interactions and Age Difference Interactions

As we noted earlier, the psychology literature (and common sense) might lead one to expect the strength of the peer influence to depend on the gender mix of the siblings.[31] In Appendix Table 10, we report estimates for a specification that replaces $\lambda_1^2 y_{t-1}^1$ and

[31] We would like to control for siblings' co-residence and examine whether the sibling influence varies with co-residence, as one would expect it would. Unfortunately, it is impossible to infer this information from the NLSY97. Data on co-residence is contained in the household roster, where respondents are indexed by an identification number that is different from their identification number in the youth questionnaire, which we use for the rest of the analysis. The NLSY does not provide a direct way to match these two identification numbers. One could match respondents based on their characteristics, but this method only allows one to match about half of the sample. We thank Steven McClasky for helpful consultations on this point.

$\lambda_2^2 y_{t-1}^1$ with

$$[\lambda_{1mm}^2(M^2 \times M^1) + \lambda_{1ff}^2(F^2 \times F^1) + \lambda_{1mf}^2(M^2 \times F^1 + M^1 \times F^2)]y_{t-1}^1$$

and

$$[\lambda_{2mm}^2(M^2 \times M^1) + \lambda_{2ff}^2(F^2 \times F^1) + \lambda_{2mf}^2(M^2 \times F^1 + M^1 \times F^2)]y_{t-1}^1,$$

where M^1 and M^2 (F^1 and F^2) are dummies that equal 1 if the older and the younger siblings are males (females), respectively. For smoking and marijuana use, the sibling influence parameters are substantially larger for sister pairs. However, the standard errors are relatively large (see Appendix Table 10).[32]

We also estimated models in which we allow the sibling influence to depend upon whether the siblings were more than two years apart by replacing the terms $\lambda_1^2 y_{t-1}^1$ and $\lambda_2^2 y_{t-1}^1$ with $[\lambda_1^2 + \lambda_{1,2+}^2 \mathbf{1}(a_t^1 - a_t^2 > 2)]y_{t-1}^1$ and $[\lambda_2^2 + \lambda_{2,2+}^2 \mathbf{1}(a_t^1 - a_t^2 > 2)]y_{t-1}^1$, respectively. On the one hand, siblings who are close in age may spend more time together and have a closer bond. On the other hand, the difference between the younger and the older siblings in the degree of access to alcohol, marijuana, and other drugs may increase with the age gap, thus increasing the impact of the older sibling even when the age of the older sibling and the younger sibling are held constant. Furthermore, with a wider age gap, the assumption that older siblings influence younger siblings, but not vice versa is more likely to be true.[33] The point estimates of $\lambda_{1,2+}^2$ and $\lambda_{2,2+}^2$ are positive in the models for cigarette smoking, drinking, and marijuana use, but have large standard errors and are never statistically significant (see Appendix Table 11). We simply do not have enough data to draw strong conclusions about how the age gap between siblings influences the sibling effect.[34]

[32] We examined whether the effects for mixed pairs depend on whether female is oldest but the estimates are imprecise.

[33] This discussion mirrors the different predictions of opportunity versus role model views of sibling influence that we touched upon in the literature review.

[34] Coefficients on gender mix interactions and age gap interactions in the correlated random effects models are also imprecise (not reported).

9.2 Family Process Interactions

The child psychology literature stresses the importance of family process variables for child outcomes. It would be interesting to know how parenting styles and the nature of the child's relationship with his parents and with his sibling influences the size of the peer effect. The NLSY97 contains a rich set of measures, and we use several indices that were constructed from them. The variable *Youth-Parent Relationship* is an index that measures the supportiveness of the parent-child relationship on a scale of 0 to 32. The index *Parental Monitoring* measures how much the parents know about their child's environment, friends, and school on a 0 to 16 scale. The *Parenting Style* index is defined to be 1 if uninvolved, 2 if permissive, 3 if authoritarian, and 4 if authoritative. The variable *Sib_Advice* is 1 if the first person the youth turns to for advice is his or her "brother or sister"(the question does not identify which sibling). The variable *NoParent_advice* is 1 if the youth turns to someone other than the parents for advice.[35] We incorporated the additional variables one at a time into our CRE specification by adding the interaction between the family process variable and the older sibling's lagged behavior as well as with the sum of the sibling's lagged and lead behaviors. We also included the family variable itself as a control variable.

The coefficients on the interaction terms with y^1_{t-1} often have the sign that we expected, but they are usually not statistically significant (not reported).[36] We lack sufficient power

[35]The questions underlying the first index (youth-parent relationship) are administered in the first three rounds of the survey to kids between 12 and 14 years old by 12/31/96. The second one (parental monitoring) are administered in the first 4 rounds of the survey to kids between 12 and 14 years old by 12/31/96. The third one (parenting style) is administered in the first four rounds to all respondents of the survey. Given the requirement for lagged values of the behavior of the older sibling, our sample for the random effects models start with the 1999 survey. Consequently, to increase sample size, we use the average of the available responses for the younger sibling rather than the value corresponding to year t. This reduces possibility of reverse causality from child's choice of y to the family process variables. The question *Sib_advice* and *NoParent_advice* are based on the response to the question "If you had an emotional problem or personal relationship problem, who would you first turn to for help?". It is asked of kids aged 12-14 as of December 31 1996, in years 1997, 1998, 1999, 2000, 2002, 2004. We use average of the 2000 and 2002 responses as the value for 2001 and the average of 2002 and 2004 as the value for 2003.

[36]In keeping with discussion in the literature, we expected negative effects for parent relationship, parental monitoring, parenting style and *NoParent_advice*. We do not have a clear prior about the sign of the main effect of *Sib_Advice*.

to detect small or modest interaction effects. One might expect the sibling effect to be larger for adolescents who get advice from siblings. The marginal effect for *Sib_advice* $\times$ y^1_{t-1} is .116 (.084) for smoking and .136 (.108) for drinking. However, the estimate is -.030 (.064) for marijuana.

We also estimated CRE models with interactions between a dummy for whether the child lives with both of their biological parents and y^1_{t-1} and $y^1_{t-1}+ y^1_{t+1}$. On the one hand, one might speculate that adolescents living with their biological mother and father are subject to more influence from parents and less influence from older siblings. Alternatively, the presence of both biological parents might strengthen the family in general, making both parental influences and sibling influences more important relative to outside influences, particularly peers. As it turns out, living with one's biological mother and father at age 12 boosts the effect of y^1_{t-1} by .040 (.056) in the case of smoking, by .075 (.049) in the case of drinking, and by .021 (.045) for marijuana (table not reported).

The main effects of several of the variables are significant. The point estimates indicate that smoking, drinking, and marijuana use are more likely for children who have unsupportive parents, uninformed parents, and uninvolved parents. They are also much more likely for children who receive advice from people other than their parents and who do not live with both biological parents. However, all the estimates of the main effects should be taken with a grain of salt because of the possibility of reverse causality and omitted variable bias.

10 Conclusion

Parents frequently implore their older children to set a good example for younger brothers and sisters. Social scientists, particularly psychologists, have long been interested in the influences that siblings have on each other. Many studies, including ours, have found strong sibling correlations in a variety of behaviors, including substance use, that are robust to the inclusion of a rich set of controls. The difficult question is whether these correlations reflect causal influences or result from shared genes and environment. To identify causal effects, we use the fact that the future cannot cause the past and make the key assumption that older siblings influence younger siblings, but not vice versa. Our first empirical strategy is based upon a CRE model, in which we regress the younger sibling's behavior on the lagged behavior of the older sibling and the sum of the lagged and future behaviors of the older sibling. As we point out, the CRE design is only valid in the absence of state dependence and requires strong stationarity assumptions that are unlikely to hold for behaviors that gradually emerge during adolescence. Furthermore, the estimates do not provide much information about how the effect of sibling behavior plays out over time. Consequently, we also make use of a joint dynamic probit model and a joint dynamic ordered probit model that allow for state dependence and nonstationarity.

Using a CRE design, we find that smoking, drinking, and, more tentatively, marijuana use by the older sibling increases the probability that the younger sibling engages in these behaviors. Our sibling influence estimates are too imprecise in the case of hard drugs and selling drugs to draw strong conclusions, although the point estimates suggest a positive effect. Using the dynamic probit models, we find a positive and significant sibling effect for cigarettes, drinking, and marijuana use. We also find a positive effect for hard drugs and selling drugs, but the coefficients are not statistically significant. For the most part, the effects are largest in the equation for the initial condition for the younger sibling. Although we find large and significant effects of past behavior on the latent variable that determines substance use, the effect on the younger sibling of a one-time shift in the behavior of the older sibling dies out quickly.

There is a substantial research agenda. First, the analysis should be repeated with additional data sets containing panel data on substance for large samples. These are steep data requirements. Add Health, which has been used in a number of previous studies of sibling links in risky behavior, is a natural possibility, but the time between interviews makes it less than ideal. Second, other behaviors, including positive behaviors such as volunteering and study time, could be examined. Third, one might model the dynamic interrelationship among the use of the substances rather than consider them one at a time, as we have done in this paper. Such a model would permit one to consider the degree to which smoking, say, leads to marijuana use. The question of "gateway" drugs is salient in policy discussions of drug law reform, but given limited information in the data it would be hard to quantify the linkages without strong a priori information about which linkages are most likely. Finally, one could also examine how family process determines the strength and direction of the sibling effect. We are pessimistic, however, that much more can be done on this question with the NLSY97 because of power considerations. A more structured approach in which the researcher constrains the way in which home environment measures alter the strength of the sibling effect on a multiple set of behaviors is worth trying.

References

Akerlof, G. (1997). Social distance and social decisions. *Econometrica 65*(5), 1005 – 1027.

Amuedo-Derantes, C. and T. Mach (2002). Impact of families on juvenile substance abuse. *Journal of Bioeconomics 4*, 269 – 281.

Barnes, G. (1990). Impact of the family on adolescent drinking patterns. In R. L. Collins, K. E. Leonard, and J. S. Searles (Eds.), *Alcohol and the family: research and clinical perspectives*. Guilford Press.

Berndt, T. (1992). Friendship and friends' influence in adolescence. *Current Directions in Psychological Science 1*(5), 158 – 159.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 841 – 877.

Bikchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy 100*(5), 992 – 997; 1010 – 1014.

Bricker, J., A. Peterson, B. Leroux, M. Andersen, K. Rajan, and I. Sarason (2005). Prospective prediction of children's smoking transitions: role of parents' and older siblings' smoking. *Addiction 101*(1), 128 – 136.

Buhrmester, D. (1992). The developmental courses of sibling and peer relationships. In F. Boer and J. Dunn (Eds.), *Children's Sibling Relationships: Developmental and Clinical Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chamberlain, G. (1984). Panel data. In Griliches and Intrilogator (Eds.), *Handbook of Econometrics, Volume 2*. Amsterdam, Holland: Elsevier.

Conger, R. and M. Reuter (1996). Siblings, parents, and peers: a longitudinal study of social influences in adolescent risk for alcohol use and abuse. In G. Brody (Ed.), *Sibling relationships: their causes and consequences*. Normood, NJ: Ablex.

Duncan, G., J. Boisjoly, and K. Harris (2001). Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography 38*, 437 – 447.

Duncan, G., J. Boisjoly, M. Kramer, D. Levy, and J. Eccles (2005). Peer effects in drug use and sex among college students. *Journal of Abnormal Child Psychology 33*(3), 375 – 385.

Furman, W. and D. Buhrmester (1992). Age and sex differences in perceptions of networks of personal relationships. *Child Development 63*(1), 103.

Grossman, M., R. Kaestner, and S. Markowitz (2004). Get high and get stupid: The effect of alcohol and marijuana use on teen sexual. *Review of Economics of Household 2*(4), 413 – 441.

Gruber, J. and J. Zinman (2001). Youth smoking in the United States. In J. Gruber (Ed.), *Risky Behavior Among Youths*. Chicago: University of Chicago Press.

Harris, J. (1998). *The Nurture Assumption*. New York: Touchstone.

Harris, J. and B. Lopez-Valcarel (2008). Asymmetric peer effects in the analysis of cigarette smoking among young people in the United States, 1992-1999. *Journal of Health Economics 27*, 249 – 264.

Jessor, R. and S. Jessor (1977). *Problem behavior and psychosocial development: a longitudinal study of youth*. New York: Academic Press.

Kandel, B. (1980). Drug and drinking behavior among youth. *Annual Review of Sociology 6*, 235 – 285.

Levitt, S. and L. Lochner (2001). The determinants of juvenile crime. In J. Gruber (Ed.), *Risky Behavior Among Youths*. Chicago: University of Chicago Press.

Marmaros, D. and B. Sacerdote (2002). Peer effects in occupational choice for Dartmouth students. *European Economic Review 46*.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*, 69 – 85.

Oettinger, G. (2000). Sibling similarity in high school graduation outcomes: Causal interdependency or unobserved heterogeneity? *Southern Economic Journal 66*(3), 631 – 648.

Otten, R., R. Engels, M. van de Ven, and J. Bricker (2007). Parental smoking and adolescent smoking stages: the role of parents current and former smoking, and family structure. *Journal of Behavioral Medicine 30*(2), 143 – 154.

Ouyang, L. (2004, November). Sibling effects on teen risky behaviors. Unpublished paper, Department of Economics, Duke University.

Pacula, R., M. Grossman, and F. Chaloupka (2001). Marijuana and youth. In J. Gruber (Ed.), *Risky Behavior Among Youths*. Chicago: University of Chicago Press.

Patterson, G. (1984). Siblings: fellow travelers in coercive family process. In R. Blanchard and D. Blanchard (Eds.), *Advances in the Study of Aggression, Vol. I.* Orlando, Fla.: Academic Press.

Rodgers, J. and D. Rowe (1988). Influence of siblings on adolescent sexual behavior. *Developmental Psychology 24*(5), 722 – 728.

Rowe, D. and B. Gulley (1992). Sibling effects on substance use and deliquency. *Criminology 30*, 217 – 233.

Sacerdote, B. (2001). Peer effects with random assignment: results for Dartmouth roommates. *Quarterly Journal of Economics 166*.

Slomkowski, C., R. Rende, K. Conger, J. Simons, and R. Conger (2001). Sisters, brothers, and delinquency: evaluating social influence during early and middle adolescence. *Child Development 72*(1), 271 – 283.

Slomkowski, C., R. Rende, S. Novak, E. Lloyd-Richardson, and R. Niaura (2005). Sibling effects on smoking in adolescence: evidence for social influence from a genetically informative design. *Addiction 100*, 430 – 438.

Steinberg, L. and S. Silverberg (1986). The vicissitudes of autonomy in early adolescence. *Child Development 57*(4), 841– 846.

Stinebrickner, R. and T. Stinebrickner (2006). What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics 90*(8-9), 1435 – 1454.

Widmer, E. (1997). Influence of older siblings on initiation of sexual intercourse. *Journal of Marriage and the Family 59*(4), 928 – 938.

Windle, M. (2000). Parental, sibling, and peer influences on adolescent substance use and alcohol problems. *Applied Developmental Science 4*(2), 98 – 110.

Data Appendix

The paper uses data from the first eight rounds of the NLSY97. In the following paragraphs, we explain how we constructed the variables used in the analysis and list the question names and reference numbers (in parentheses) of the NLSY97 variables we used to construct our dataset.

Sibling pairs

The NLSY original cohort includes 1,892 households with more than one NLSY97 respondent. In order to link respondents to their siblings, we used the variables: YOUTH_SIBID01.01 (R1308300), YOUTH_SIBID02.01 (R1308400), YOUTH_SIBID03.01 (R1308500), YOUTH_SIBID04.01 (R1308600). For each respondent, these variables return the identification number of up to four other respondents from the same household. Then, we used the variable HHI2_RELY.01 (R1309100, R1309200, R1309300, R1309400), which characterizes the type of relationship between these respondents. For siblings, the NLSY97 distinguishes between full (biological), half, step, foster, and adoptive siblings. The analysis presented in the paper is conducted on a sample of full siblings only. In preliminary work, we estimated many of the models using pairs of full, half, and step siblings, and obtained results similar to those reported in the paper. Finally, as mentioned in the paper, in households supplying more than one sibling pair, we only included pairs with adjacent birth order. To select these pairs, we used the variable CV_AGE_12/31/96, which gives the age of each respondent as of December 31, 1996.

Control Variables

Our set of controls includes several individual, familial and environmental characteristics. Below, we describe each of them and list the raw variables we used to construct them.

- Age is computed using the variable named *CV_AGE*_12/31/96 (R1194000), which measures the respondent's age as of December 31st 1996

- A male dummy, which equals 1 if the respondent is a male, was created using the variable

$KEY!SEX$(R0536300).

- Two separate dummy variables for race were created for the categories of black and Hispanic, using the variable $KEY!RACE\_ETHNICITY$ (R1482600). Each category is mutually exclusive, and white is the reference group.

- Education is measured as the respondent's highest grade completed by age 19, and the grade is normalized by subtracting 12 from it. This variable is constructed by combining the age of the respondent and the yearly variables returning the respondent's highest grade completed by each survey round CV_HGC_EVER (R1204400, R2563100, R3884700, R5463900, R7227600, S1541500, S2011300, S3812200).

- Mother's education is measured as the biological mother's highest grade completed, as reported by the respondent in 1997. Her grade is also normalized by subtracting 12 from it. This variable was constructed from the variable $CV\_HGC\_BIO\_MOM$ (R1302500).

- AFQT score is measured in percentile and standardized by the age of the respondent at the time of the test. From the summer of 1997 through the spring of 1998, most NLSY97 round 1 respondents took the computer-adaptive form of the Armed Services Vocational Aptitude Battery (CAT-ASVAB). The results of the different math and verbal tests were combined and weighted by the NLS Program staff to produce the percentile score recorded under the variable $ASVAB\_MATH\_VERBAL\_SCORE\_PCT$ (R9829600), which is similar to the AFQT score. This variable assumes three decimal places, so we constructed our variable by simply dividing the score by 1000.

- Family structure is measured by a dummy for whether the individual lived with both biological parents at age 12. In 1997, the question $CV\_YTH\_REL\_HH\_AGE\_12$ (R1205000) asks respondents about their relationship to the parent figure or guardian in the household at age 12. If the individual replied that the parent figure was both the biological mother and the biological father, we set our dummy variable to 1, otherwise to 0.

- We created three binary variables, describing aspects of the individuals' environment up to age 12. We build these directly from three NLSY questions about particularly violent or traumatizing childhood experiences. The first one is the variable $YSAQ-517$ (R0443900), which records whether the respondent ever had her house or apartment broken into before turning 12 years old. The second one is the variable $YSAQ-519$ (R0444100), which records whether the respondent ever saw anyone get shot or shot at with a gun before turning 12. The third one is the variable $YSAQ-518$ (R0444000), which records whether the respondent was ever the victim of repeated bullying before turning 12. Since the bullying measure reflects a possibly traumatic childhood experience, it may be thought of as measuring, albeit very imperfectly, some aspect of the individual's mental state and social adjustment.

- We created birth order dummies and a variable measuring the number of full siblings who live in the household, using the household roster data. In particular, we used the variable "YOUTH_ID.01" (R0533400), which gives the respondent's ID number in the household roster, and the variables describing the relationship between household members. These variables have names of the form "HHI2_RELX.0Y", where X is the respondent's roster ID and Y is the ID of the other household respondents.

Substance Use Measures

In most of our analysis, the main dependent variable is a dummy indicating whether the respondent reports having engaged at least once in a particular behavior since the last interview date. For example, for smoking, the variable takes the value 1 if the respondent reports having smoked since the last interview, and 0 otherwise. For each behavior, we construct this variable from two NLSY variables. The first and most important one is a dummy variable indicating whether the respondent has engaged in this type of behavior since the last date of interview. When it is available (i.e. for the first survey rounds in general), we use a second dummy variable, which indicates whether the respondent has ever engaged in this type of behavior. This second variable allows checking the consistency of some of the answers in the first question, as well as filling in some of the missing observations. These questions were not asked in every year, and we report below the exact name, reference numbers (in parentheses), and years of the variables we used.

Smoking, Drinking, Marijuana, and Selling drugs

For smoking, drinking, marijuana smoking, and selling drugs, the first question (about the respondent's activity last year) was not asked in the first survey round (1997). As a result, we only use data starting in 1998, when respondents are aged 14 through 18. The NLSY variables used to form the dependent variables are:

-Smoking: $YSAQ359$ (R2189400, R3508500, R4906600, R6534100, S0921600, S2988300, S4682900) for 1998 through 2004, and $YSAQ360C$ (R0357900, R2189100, R3508200, R4906400) for 1997 through 2000.

- Drinking: $YSAQ364D$ (R2190200, R3509300, R4907400, R6534700, S0922200, S2988900, S4683700) from 1998 through 2004, and $YSAQ363$ (R0358300, R2189900, R3509000, R4907100) from 1997 through 2000.

- Marijuana: $YSAQ370C$ (R2191200, R3510300, R4908400, R6535600, R6535600, S0923200, S2989700) from 1998 through 2004, and $YSAQ369$ (R0358900, R2190900, R3510000, R4908100) from 1997 through 2000.

- Selling or helping to sell drugs: $YSAQ394B$ (R2196400, R3516000, R4914000, R6540500, S0928000, S2994000) for 1998 through 2004, and $YSAQ430$ (R0365000, R2199300, R3518900, R4916900, R6543400, S0930900) for 1997 through 20002.

Cocaine and other hard drugs use

The NLSY97 asked respondents about cocaine and other hard drugs use starting in the second survey round (1998). In 1998, the survey asked whether the respondent had ever used these types of drugs, and it is only in 1999 that it started asking whether the respondent had used hard drugs since the last interview. As a result, we restricted our analysis to the last six rounds (1999 to 2004) for this behavior, starting when respondents are between 15 and 19. We used the following variables: $YSAQ372CC$ (R3511100, R4909200, R6536400, S0924000, S2990300, S4685500) for 1999 through 2004, and $YSAQ372B$ (R2191500, R3510800, R4908900, R6536100, S0923700) for 1998 through 2002.

Finally, to estimate the dynamic ordered probit models, we created indicators of zero, low, and high consumption of cigarettes, alcohol, and marijuana. These indicators were constructed using NLSY97 questions about how many days the respondent engaged in the behavior in the previous month. Respectively, these refer to the NLSY97 questions $YSAQ361$ (R035810, R2189500, R3508600, R4906700, R6534200, S0921700, S2988400, S4683000) for smoking cigarettes, $YSAQ365$ (R0358500, R2190300, R3509400, R4907500, R6534800, S0922300, S2989000, S4683800) for drinking alcohol, and $YSAQ371$ (R0359100, R2191300, R3510400, R4908500, R6535700, S0923300, S2989800, S4684800) for smoking marijuana. Note that all of these questions were asked to all respondents from 1997 through 2004. However, since the rest of the analysis is conducted on data from 1998 onwards, we did not use the first round of the survey when using these variables.

Family processes and parenting variables

In several rounds of the survey, the NLSY asks respondents about their relationship to their residential and non-residential parents. Based on these questions, Child Trends, Inc. created a number of scales that measure different aspects of this relationship. In the paper, we used three of these scales for both residential mother and residential father. The first one is an index from 0 to 32, measuring how supportive the youth reports her parents to be (a high score indicates a more supportive relationship). The second one is an index from 0 to 16, measuring the youth's perception of her parents' degree of monitoring (a high score indicates greater monitoring). Results for this index were very noisy and are not discussed in the paper. The third index is a four-category variable describing the youth's perception of her parents' parenting style; this variable equals 1 if the parents are uninvolved, 2 if permissive, 3 if authoritarian, and 4 if authoritative. The corresponding NLSY variables are: FP_YMSUPP (R1485200, R2600700, R3924100) and FP_YFSUPP (R1485300, R2600800, R3924200) for the first index, $FP\_YMMONIT$ (R1485700, R2601000, R3924400, R5510900) and $FP\_YFMONIT$ (R1485800, R2601100, R3924500, R5511000) for the second index, and $FP\_YMPSTYL$ (R1486500, R2601400, R3924800, R5511100) and $FP\_YFPSTYL$ (R1486600, R2601500, R392490, R5511200) for the third index. Note that questions used to create the first and second indexes were only asked to respondents aged 12 to 14 as of December 31, 1996, while questions underlying the third index were asked to the entire cohort. These NLSY variables are available for 1997 through 1999 for the first index, and for 1997 through 2000 for the other two. In our analysis, the variable we use is the index mean over the years with available data. If the respondent's answers were missing for one residential parent, we used the mean for the residential parent that had non-missing values. If the respondent had a non-missing value for both residential parents, we averaged the answers across the parents and used that value in our regressions. Finally, we constructed a dummy that equals 1 if the first person the youth turns to for advice is his or her "brother or sister" and another dummy that equals 1 if the youth turns to someone other than the parents for advice. To build these variables, we used a variable reporting who the youth turns to for help if he or she has an emotional problem or personal relationship problem. This variable is named YSAQ-351A(R0357300, R2176000, R3493900, R4892300, S0919200, S4681600).

Treatment of Missing Data

With the exception of the black, Hispanic, and male dummies, the other variables used in the analysis contain a small number of missing values. We dropped the few observations for which we were missing household roster data and were not able to determine the number of siblings and birth order. In the case of highest grade completed, AFQT, mother's education, family structure, and the three childhood environment dummies, we imputed missing values using predicted values from a regression of the variables on all other six variables. For the substance use measures of the younger sibling and older sibling, we dropped cases involving missing values for current values, leads, or lags of y^2 or y^1 that appear in a particular model as well as cases for subsequent years even if the necessary data are available. For example, if an individual has non-missing answers from 1998 to 2000, a missing one in 2001, and a non-missing one in 2002, we only included his answers for 1998 through 2000. We made this choice because we wanted to estimate each of the equations of the dynamic model in section 7 on a sample that is fairly stable across the years. We estimated both the correlated random effect models on the same sample as the one for the joint dynamic model, so the same observation selection rules apply.

TABLE 1
Sample Means for Substance Use Behaviors

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Drugs
	Panel A: Probability of engaging in behavior last year				
Full sample	0.402	0.603	0.229	0.062	0.058
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Males	0.424	0.619	0.261	0.067	0.081
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)
Females	0.380	0.586	0.197	0.057	0.034
	(0.005)	(0.005)	(0.004)	(0.003)	(0.002)
	Panel B: Probability of engaging in behavior between age 15 and 20				
Full sample	0.573	0.795	0.441	0.147	0.162
	(0.009)	(0.007)	(0.009)	(0.006)	(0.006)
Males	0.602	0.810	0.481	0.159	0.219
	(0.012)	(0.009)	(0.012)	(0.009)	(0.010)
Females	0.542	0.779	0.399	0.134	0.100
	(0.013)	(0.010)	(0.012)	(0.009)	(0.008)
	Panel C: Number of days engaged in behavior in the last month				
Full Sample	6.73	3.00	1.87	-	-
	(0.091)	(0.059)	(0.069)		
Males	7.08	3.66	2.48	-	-
	(0.092)	(0.067)	(0.080)		
Females	6.37	2.32	1.24	-	-
	(0.090)	(0.048)	(0.054)		
	Panel D: Distribution of sample in each category consumption (last month)				
1-4 days last month	0.067	0.269	0.075	-	-
	(0.002)	(0.003)	(0.002)		
5-7 days last month	0.022	0.091	0.016	-	-
	(0.001)	(0.003)	(0.001)		
8 + days last month	0.245	0.130	0.074	-	-
	(0.003)	(0.003)	(0.002)		

Note: Standard errors of the sample means in parentheses. Sample sizes in Panels A, C, and D vary from 16379 to 16456 for full sample, from 8320 to 8339 for males, and from 8057 to 8062 for females. Sample sizes in Panel B vary from 3172 to 3313 for full panel, 1647 to 1720 for males, and from 1525 to 1593 for females.

TABLE 2

Estimates of the Coefficient on the Older Sibling's Behavior in a Linear Probability Model of Younger Sibling's Behavior at the Same Age

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Hard Drugs
	Panel A: All siblings				
No controls	0.239***	0.259***	0.162***	0.102***	0.057***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)
Full set of controls	0.185***	0.172***	0.139***	0.085***	0.041***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)
	Panel B: Brothers				
No controls	0.238***	0.243***	0.151***	0.081*	0.083***
	(0.033)	(0.033)	(0.032)	(0.042)	(0.031)
Full set of controls	0.192***	0.172***	0.129***	0.051	0.059**
	(0.032)	(0.035)	(0.031)	(0.037)	(0.028)
	Panel C: Sisters				
No controls	0.339***	0.332***	0.240***	0.244***	0.014
	(0.039)	(0.032)	(0.035)	(0.079)	(0.028)
Full set of controls	0.229***	0.205***	0.191***	0.189***	-0.002
	(0.039)	(0.034)	(0.033)	(0.071)	(0.026)

Note: Standard errors clustered at the household level in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary from 6504 to 6551 for full sample, 1705 to 1723 for brothers, 1629 to 1637 for sisters. The controls consist of male (Panel A only), black and hispanic dummies, younger sibling's age dummies, highest grade completed by age 19, mother's highest grade completed, AFQT percentile score, number of siblings, birth order dummies, whether the respondent reported that her house had been broken in by age 12, whether she reported that she had been a victim of bullying by age 12, whether she reported having witnessed a shooting by age 12, and whether she lived with both biological parents at age 12.

TABLE 3

Marginal Effects of the Older Sibling's Behavior on the Younger Sibling's Behavior (Correlated Random Effects Model)

	Younger Sibling's Behavior y_t^2				
	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Hard Drugs
Older Sibling's Behavior:	Model 1				
y_{t-1}^1	0.063**	0.060**	0.011	0.015	0.009
	(0.027)	(0.025)	(0.022)	(0.020)	(0.019)
$y_{t-1}^1 + y_{t+1}^1$	0.086***	0.116***	0.087***	0.032**	0.016
	(0.018)	(0.018)	(0.015)	(0.015)	(0.013)
	Model 2				
$y_{t-2}^1 + y_{t-1}^1$	0.092***	0.023	0.035	0.025	0.043**
	(0.027)	(0.025)	(0.022)	(0.018)	(0.018)
$y_{t-2}^1 + y_{t-1}^1 + y_{t+1}^1 + y_{t+2}^1$	0.025	0.080***	0.040***	0.016	-0.013
	(0.016)	(0.016)	(0.013)	(0.013)	(0.013)
	Model 3				
$y_{t-1}^1 + y_t^1$	0.093***	0.044**	0.054***	0.025*	0.015
	(0.024)	(0.021)	(0.018)	(0.015)	(0.015)
$y_{t-1}^1 + y_t^1 + y_{t+1}^1 + y_{t+2}^1$	0.023*	0.071***	0.031***	0.013	0.002
	(0.014)	(0.014)	(0.011)	(0.011)	(0.011)

Note: Marginal effects are based on probit estimates. Standard errors clustered at the household level in parentheses.* denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary between 6132 and 6596 for Model 1, 2435 and 3776 for Model 2, and 3721 and 5108 for Model 3. All models include the set of controls listed in the footnote to Table 2, as well as older sibling's age dummies.

TABLE 4A

Estimates of Dynamic Probit Model (Error A)

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Drugs
State dependence					
γ^1 Old sibling	0.899***	0.632***	0.680***	0.564***	0.570***
	(0.064)	(0.057)	(0.063)	(0.138)	(0.136)
γ^2 Young sibling	0.947***	0.666***	0.734***	0.754***	0.502***
	(0.068)	(0.056)	(0.065)	(0.144)	(0.131)
Sibling's influence					
λ_1^2 $t_{\min}$	0.215**	0.411***	0.312***	0.379	0.128
	(0.101)	(0.086)	(0.107)	(0.288)	(0.220)
λ_2^2 $t > t_{\min}$	0.049	0.006	-0.053	0.071	0.006
	(0.069)	0.057	0.067	0.218	0.177
Standard deviation of error term specific to:					
σ_{ε} Family	0.771***	0.629***	0.640***	0.535***	0.472***
	(0.052)	(0.034)	(0.043)	(0.093)	(0.089)
σ_{υ^1} Older sibling	1.055***	0.609***	0.690***	0.758***	0.742***
	(0.082)	(0.064)	(0.069)	(0.135)	(0.107)
σ_{υ^2} Younger sibling	0.815***	0.610***	0.693***	0.848***	0.750***
	(0.082)	(0.064)	(0.069)	(0.149)	(0.118)
Log likelihood value	-7248.88	-8094.33	-6691.13	-2503.31	-2955.64

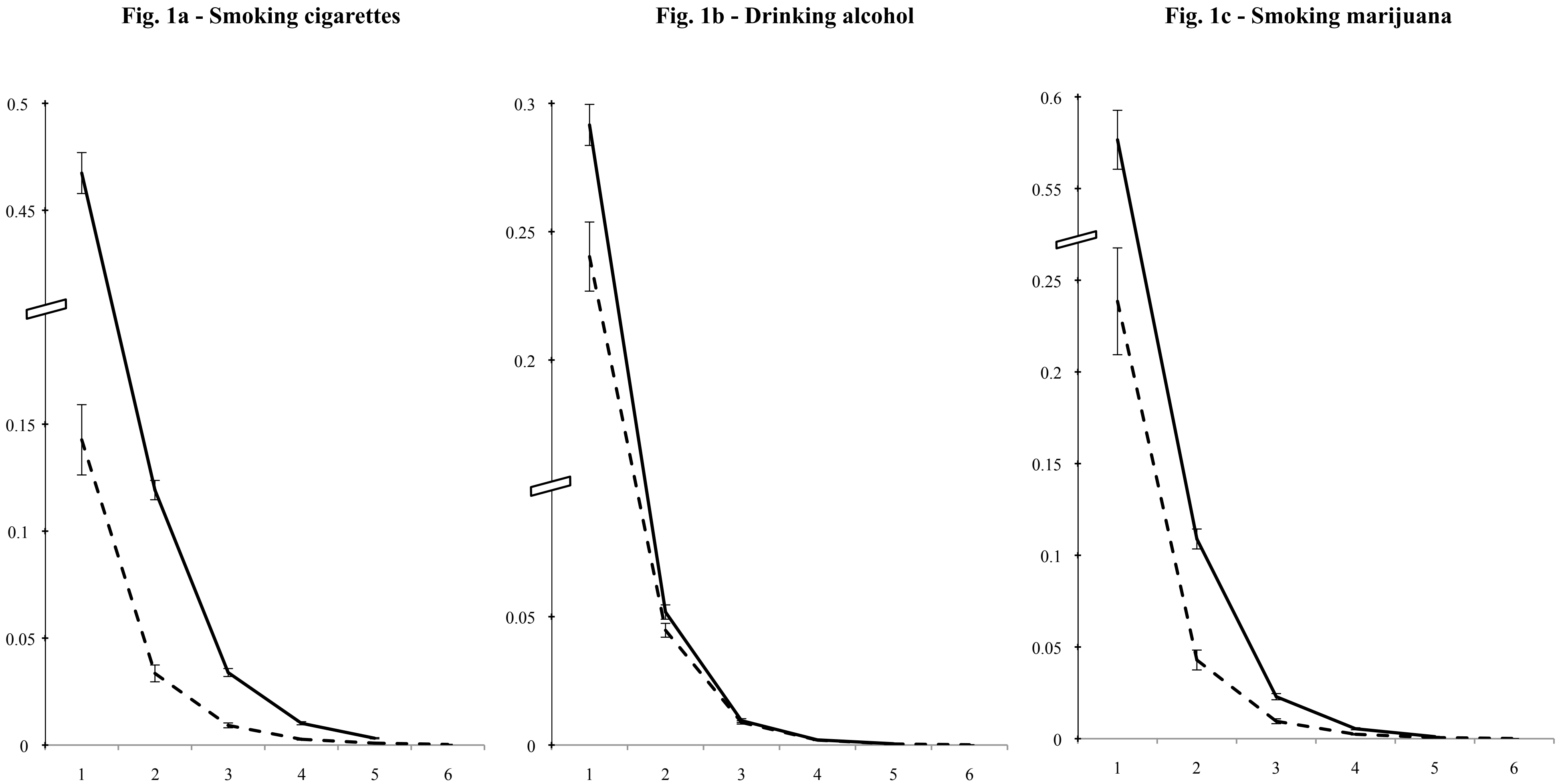
Notes: The table reports probit model parameters rather than marginal effects. Standard errors in parentheseses.
* denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary from 1173 to 1586 for the older siblings' models and from 1079 to 1658 for the younger siblings' models. All models include the set of controls listed in the footnote to Table 2, as well as older sibling's age dummies.

TABLES 4B
Estimates of Dynamic Probit Model (Error B)

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Doing Hard Drugs	Selling Drugs
State dependence					
γ^1 Older sibling	0.824***	0.632***	0.622***	0.524***	0.536***
	(0.069)	(0.058)	(0.068)	(0.144)	(0.142)
γ^2 Younger sibling	0.912***	0.611***	0.709***	0.767***	0.461***
	(0.073)	(0.059)	(0.068)	(0.154)	(0.140)
Sibling's Influence					
λ_1^2 $t_{\min}$	0.008	0.278**	0.264*	0.362	0.229
	(0.173)	(0.124)	(0.138)	(0.394)	(0.226)
λ_2^2 $t > t_{\min}$	0.116	0.039	-0.036	0.032	0.030
	(0.076)	(0.062)	(0.074)	(0.223)	(0.182)
Family-specific error term					
σ_ε Standard deviation	0.862***	0.754***	0.547***	0.364	0.345
	(0.126)	(0.105)	(0.101)	(0.238)	(0.257)
Factor loadings:					
$\delta_{2\varepsilon}^1$ Older sibling $(t > t_{\min})$	1.111***	0.918***	1.289***	1.266*	0.777*
	(0.183)	(0.152)	(0.243)	(0.728)	(0.440)
$\delta_{1\varepsilon}^2$ Younger sibling $(t_{\min})$	1.003***	0.840***	1.239***	2.099	1.243
	(0.264)	(0.195)	(0.337)	(1.513)	(1.952)
$\delta_{2\varepsilon}^2$ Younger sibling $(t > t_{\min})$	0.728***	0.755***	1.166***	2.020	2.181
	(0.162)	(0.153)	(0.325)	(2.089)	(2.754)
Individual- specific error term					
σ_υ Standard deviation	0.747***	0.451***	0.558***	0.696***	0.733***
	(0.071)	(0.067)	(0.078)	(0.178)	(0.144)
Factor loadings					
$\delta_{2\upsilon}^1$ Older sibling $(t > t_{\min})$	1.485***	1.191***	1.381***	1.273***	1.207***
	(0.184)	(0.298)	(0.276)	(0.378)	(0.272)
$\delta_{2\upsilon}^2$ Younger sibling $(t > t_{\min})$	1.349***	1.756***	1.355***	0.954	0.798
	(0.160)	(0.276)	(0.248)	(0.692)	(0.903)
Log likelihood value	-7241.23	-8088.18	-6686.51	-2502.85	-2954.12

Note: See Table 4A.

Effect of Shifting the Older Sibling's Probability of Behavior from 0 to 1 on the Older and Young Sibling Probabilities of Behavior, Relative to Baseline (Error A)



Note: The solid line and the broken line represent the effects on the probabilities of behavior, relative to baseline, of the older sibling and on the younger sibling, respectively. Error bars show the 90% confidence intervals. The x-axis measures the number of periods after the exogenous change in the older sibling's behavior. Baseline probabilities for smoking in the first and last period displayed on the graphs are, respectively: 0.4188 (0.0433) and 0.4346 (0.0592) for the older sibling and 0.3322 (0.0795) and 0.3796 (0.0926) for the younger sibling. For drinking, they are 0.5573 (0.0527) and 0.6926 (0.0428) for the older sibling and 0.4664 (0.0803) and 0.5754 (0.0829) for the younger sibling. For smoking marijuana, they are 0.2511 (0.0609) and 0.2889 (0.0489) for the older sibling and 0.2291 (0.0682) and 0.3088 (0.0883) for the younger sibling.

TABLE 5
Estimates of Joint Dynamic Ordered Probit Model (Error A)

			Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence					
Older sibling	γ_L^1	Low consumption	0.328***	0.341***	0.321***
			(0.080)	(0.044)	(0.082)
	γ_H^1	High consumption	0.797***	0.602***	0.700***
			(0.066)	(0.065)	(0.079)
Younger sibling	γ_L^2	Low consumption	0.587***	0.541***	0.611***
			(0.077)	(0.045)	(0.080)
	γ_H^2	High consumption	1.023***	0.848***	1.035***
			(0.071)	(0.068)	(0.087)
Sibling's influence					
$t_{\min}$	λ_{1L}^2	Low consumption	0.103	0.081	0.340*
			(0.183)	(0.080)	(0.177)
	λ_{1H}^2	High consumption	0.462***	0.217*	0.218
			(0.121)	(0.131)	(0.170)
$t > t_{\min}$	λ_{2L}^2	Low consumption	0.060	-0.020	0.073
			(0.089)	(0.044)	(0.089)
	λ_{2H}^2	High consumption	0.085	-0.147**	-0.009
			(0.077)	(0.065)	(0.095)
Standard deviation of error term					
specific to:	σ_ε	Family	0.796***	0.553***	0.652***
			(0.058)	(0.027)	(0.052)
	σ_{υ^1}	Older sibling	1.308***	0.521***	0.864***
			(0.079)	(0.043)	(0.073)
	σ_{υ^2}	Younger sibling	0.888***	0.464***	0.600***
			(0.077)	(0.046)	(0.081)
Low consumption threshold			-0.069	-0.824	0.851
			(0.697)	(0.537)	(0.994)
High consumption threshold			0.520	0.799	1.470
			(0.698)	(0.538)	(0.994)
Log likelihood value			-9180.28	-12850.29	-7126.49

Note: Standard errors in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary from 1165 to 1555 for the older siblings' models and from 1054 to 1639 for the younger sibling's models. All models include the set of controls listed in the footnote to Table 2, as well as older sibling's age dummies.

Effect of Shifting the Older Sibling's Behavior from Zero to High Consumption on the Older and Younger's Probabillities of Consumption, Relative to Baseline (Error B)

Fig. 2a - Low consumption of cigarettes

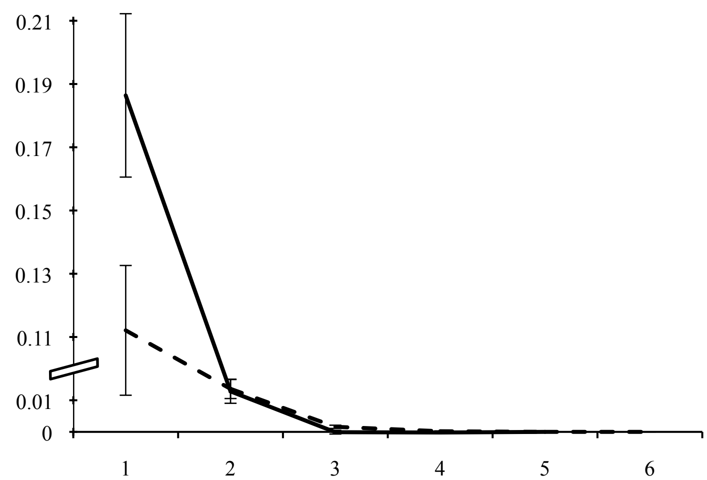


Fig. 2b - High consumption of cigarettes

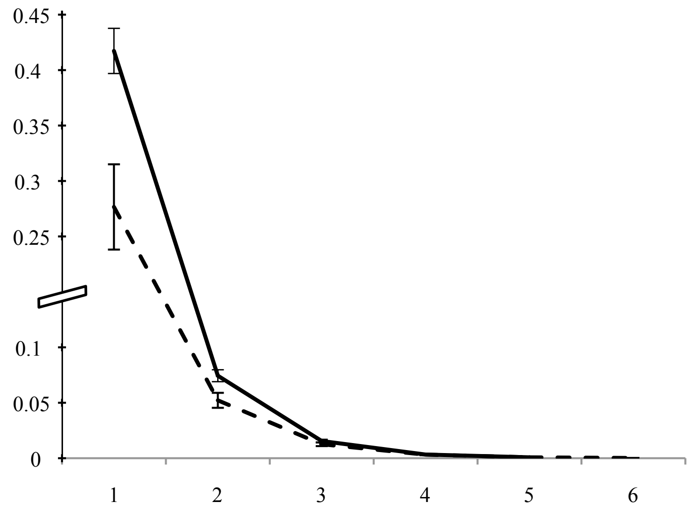


Fig. 3a - Low consumption of alcohol

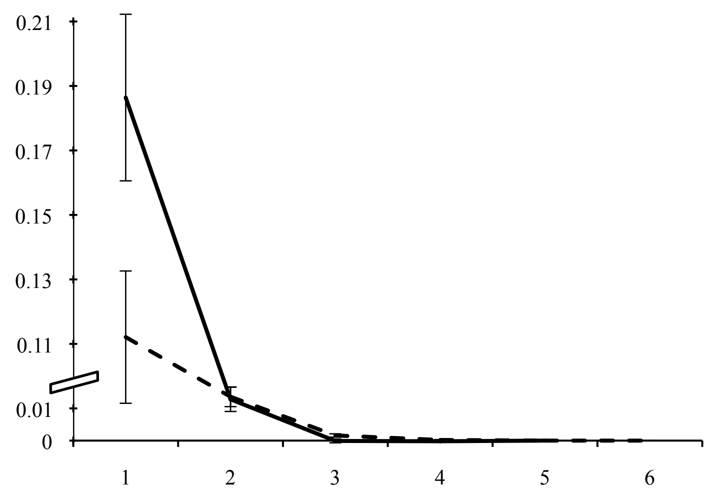


Fig. 3b - High consumption of alcohol

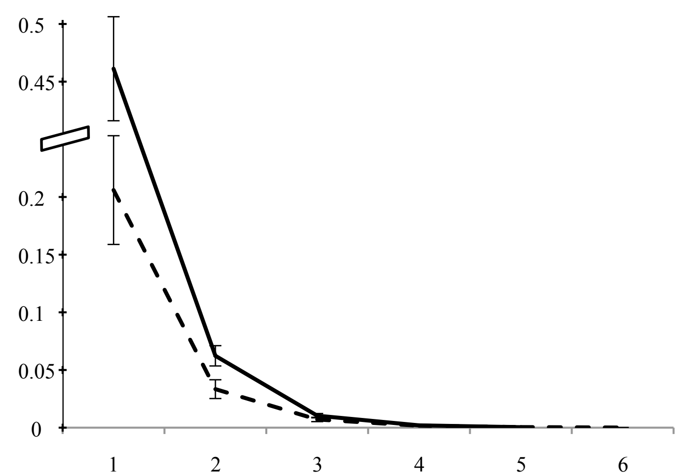


Fig. 4a - Low consumption of marijuana

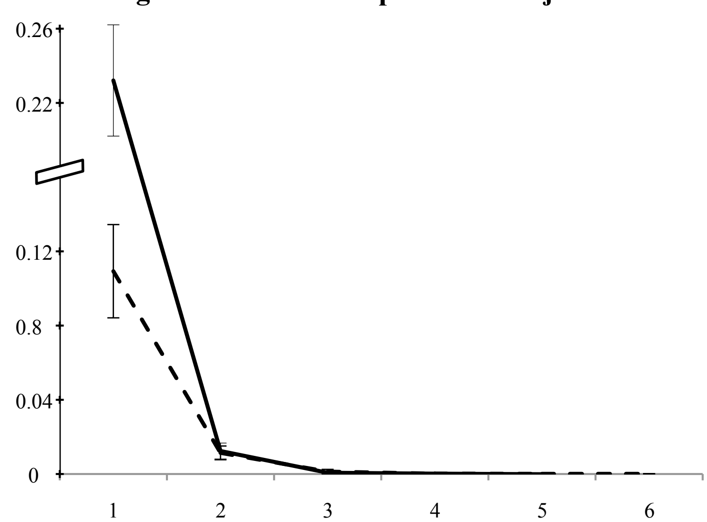
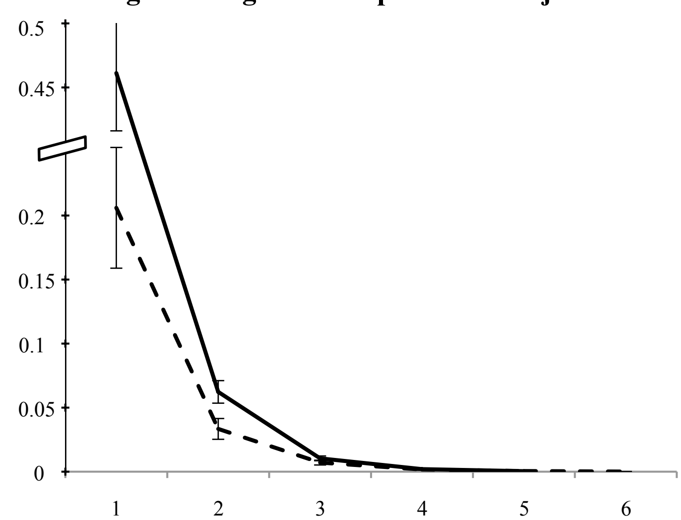


Fig. 4b - High consumption of marijuana



Note: The solid line and the broken line represent the effect for the older siblings and the younger siblings, respectively. Error bars show the 90% confidence intervals. The x-axis measures the number of periods after the exogenous change in the older sibling's behavior. Baseline probabilities for being in the low smoking category in the first and last period displayed on the graphs for the older and younger siblings, respectively are: 0.0818 (0.0154), 0.0857 (0.0115), 0.0907 (0.0254), and 0.0875 (0.0137). For being in the high smoking category, they are: 0.2781 (0.1233), 0.3275 (0.1137), 0.2072 (0.1080), and 0.3075 (0.1124). For being in the low drinking category, they are: 0.3257 (0.0724), 0.3802 (0.0347), 0.2644 (0.0785), and 0.3692 (0.0366). For being in the high drinking category, they are: 0.1532 (0.0947), 0.2895 (0.1113), 0.0955 (0.0606), and 0.2629 (0.0975). For being in the low marijuana category, they are: 0.0797 (0.0351), 0.0866 (0.0284), 0.084 (0.0465), and 0.0879 (0.0267). For being in the high marijuana category, they are: 0.1849 (0.1429), 0.2081 (0.1170), 0.1415 (0.1263), and 0.1889 (0.1178).

APPENDIX TABLE 1

Age Distribution by Birth Order

Age	Younger sibling	Older siblings
15	36.45	
16	31.71	6.56
17	23.88	21.63
18	7.59	34.39
19	0.36	34.87
20		1.94
21		0.61

Note: This distribution refers to the age distribution of the younger and older siblings in the year used as the initial condition for the younger siblings' dynamic probit model.

APPENDIX TABLE 2

Summary Statistics of Sample Characteristics

Variable	Unweighted sample	Weighted sample
Male	0.509	0.512
Black	0.240	0.135
Hispanic	0.232	0.140
Lived with biological parents at 12	0.519	0.575
House broken in by age 12	0.158	0.147
Witness of gun shooting by age 12	0.102	0.080
Victim of bullying by age 12	0.179	0.179
Highest grade completed by 19	11.837	11.970
	(1.173)	(1.095)
AFQT percentile score	44.402	50.930
	(28.515)	(28.211)
Mother's highest grade completed	12.243	12.784
	(2.933)	(2.733)
Number of (full) siblings	2.116	1.992
	(1.290)	(1.225)
Age of younger sibling	16.037	15.983
	(0.969)	(0.952)
Age of older sibling	18.058	18.048
	(0.978)	(0.977)
Age gap	2.021	2.065
	(0.877)	(0.879)

Note: Based on the sample used for estimation of the dynamic smoking model (N=16456). Weighted statistics are computed using a set of cross-sectional weights for the survey round in which the respondent is 19 years old. Ages of the younger and older siblings refer to the age in the first year the behavior of younger siblings is observed.

APPENDIX TABLE 3

Means of Risky Behaviors in Weighted Sample

	Full Sample	Male Sample	Female Sample
Smoking cigarettes last year	0.438	0.443	0.434
	(0.004)	(0.005)	(0.006)
Drinking alcohol last year	0.652	0.654	0.650
	(0.004)	(0.005)	(0.005)
Smoking marijuana last year	0.248	0.270	0.226
	(0.003)	(0.005)	(0.005)
Using hard drugs last year	0.071	0.071	0.070
	(0.002)	(0.003)	(0.003)
Selling drugs last year	0.062	0.082	0.041
	(0.002)	(0.003)	(0.002)
Days smoked cigarettes last month	7.726	7.759	7.692
	(0.096)	(0.135)	(0.138)
Days drank last month	3.374	3.999	2.716
	(0.044)	(0.069)	(0.051)
Days smoked marijuana last month	2.003	2.542	1.434
	(0.049)	(0.079)	(0.057)

Note: Standard errors of sample means in parentheses. Means are computed using a set of cross-sectional weights for each survey round in which the data are available. Sample sizes vary from 16379 to 16456 for full sample, from 8320 to 8339 for males, and from 8057 to 8062 for females.

APPENDIX TABLE 4

Risky Behaviors by Age

Age	Smoking cigarettes last year	Drinking alcohol last year	Smoking marijuana last year	Using hard drugs last year	Selling drugs last year	Days smoked cigarettes last month	Days drank alcohol last month	Days smoked marijuana last month
15	0.300	0.405	0.164	0.059	0.061	3.033	1.154	0.792
	(0.458)	(0.491)	(0.370)	(0.237)	(0.239)	(0.109)	(0.042)	(0.051)
16	0.337	0.450	0.216	0.059	0.066	4.342	1.416	1.211
	(0.473)	(0.498)	(0.412)	(0.236)	(0.249)	(0.088)	(0.031)	(0.043)
17	0.362	0.523	0.243	0.067	0.078	5.533	1.808	1.725
	(0.481)	(0.499)	(0.429)	(0.249)	(0.268)	(0.079)	(0.029)	(0.043)
18	0.416	0.579	0.244	0.073	0.067	6.678	2.684	1.805
	(0.493)	(0.494)	(0.429)	(0.260)	(0.249)	(0.077)	(0.034)	(0.039)
19	0.423	0.627	0.248	0.060	0.061	7.120	3.145	2.167
	(0.494)	(0.484)	(0.432)	(0.238)	(0.240)	(0.081)	(0.037)	(0.045)
20	0.417	0.653	0.239	0.065	0.049	7.770	3.434	2.266
	(0.493)	(0.476)	(0.427)	(0.246)	(0.215)	(0.086)	(0.040)	(0.048)
21	0.438	0.721	0.229	0.052	0.046	8.112	4.629	2.226
	(0.496)	(0.449)	(0.420)	(0.222)	(0.209)	(0.100)	(0.052)	(0.054)
22	0.440	0.729	0.182	0.051	0.027	8.224	4.559	1.782
	(0.496)	(0.444)	(0.386)	(0.220)	(0.161)	(0.127)	(0.064)	(0.063)
23	0.450	0.724	0.174	0.055	0.028	9.059	5.012	1.651
	(0.498)	(0.447)	(0.380)	(0.227)	(0.166)	(0.206)	(0.115)	(0.092)

Note: Standard errors of sample means in parentheses. Based on the sample used for the estimation of the dynamic smoking model (N=16456).

APPENDIX TABLE 5

Estimated Marginal Effect of Control Variables for the CRE Model

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Hard Drugs
Male	0.028	-0.010	0.039**	0.004	0.033***
	(0.022)	(0.020)	(0.016)	(0.008)	(0.008)
Black	-0.214***	-0.175***	-0.065***	-0.043***	-0.031***
	(0.026)	(0.028)	(0.021)	(0.009)	(0.009)
Hispanic	-0.149***	-0.019	-0.017	-0.008	-0.008
	(0.027)	(0.028)	(0.022)	(0.011)	(0.010)
Highest grade completed by 19	-0.079***	-0.022**	-0.038***	-0.011***	-0.015***
	(0.011)	(0.010)	(0.007)	(0.004)	(0.003)
AFQT percentile score	0.0002	0.002***	0.001***	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's highest grade completed	0.002	0.005	0.001	0.004**	-0.001
	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)
Lived with biological parents at 12	-0.052**	0.001	-0.025	-0.017	-0.011
	(0.026)	(0.023)	(0.020)	(0.011)	(0.010)
Number of (full) siblings	-0.027**	-0.038***	-0.030***	-0.006	-0.011**
	(0.011)	(0.009)	(0.008)	(0.004)	(0.004)
2nd born	-0.093*	-0.145***	-0.113***	-0.038*	-0.009
	(0.051)	(0.043)	(0.038)	(0.023)	(0.020)
3rd born	-0.016	-0.080*	-0.054*	-0.019	0.009
	(0.047)	(0.043)	(0.030)	(0.015)	(0.020)
House broken in by 12	0.089***	0.004	0.030	0.002	0.009
	(0.029)	(0.028)	(0.020)	(0.011)	(0.010)
Witness of gun shooting by 12	0.087**	0.103***	0.073***	0.038***	0.050***
	(0.037)	(0.035)	(0.025)	(0.014)	(0.011)
Victim of bullying by 12	0.044	0.049*	0.048**	-0.006	0.013
	(0.030)	(0.028)	(0.021)	(0.011)	(0.009)

Young sib age 16	0.020	0.034	0.046*	-	-0.014
	(0.027)	(0.028)	(0.026)	-	(0.010)
Young sib age 17	0.0360	0.111***	0.059*	-0.025***	-0.005
	(0.035)	(0.033)	(0.031)	(0.008)	(0.014)
Young sib age 18	0.111**	0.187***	0.052	-0.010	-0.028**
	(0.044)	(0.039)	(0.037)	(0.012)	(0.013)
Young sib age 19	0.138**	0.237***	0.060	-0.034***	-0.033**
	(0.055)	(0.045)	(0.045)	(0.013)	(0.015)
Young sib age 20	0.189***	0.262***	0.029	-0.036***	-0.042***
	(0.067)	(0.050)	(0.053)	(0.013)	(0.012)
Young sib age 21	0.245***	0.317***	0.062	-0.037***	-0.038***
	(0.080)	(0.046)	(0.068)	(0.013)	(0.014)
Young sib age 22	0.352***	0.341***	0.062	-0.044***	-0.050***
	(0.095)	(0.049)	(0.097)	(0.012)	(0.008)
Old sib age 17	0.086	0.038	0.016	-	0.024
	(0.055)	(0.058)	(0.051)	-	(0.036)
Old sib age 18	0.089	0.042	0.032	0.242*	0.036
	(0.062)	(0.062)	(0.054)	(0.127)	(0.037)
Old sib age 19	0.092	0.027	0.062	0.171*	0.053
	(0.067)	(0.068)	(0.060)	(0.100)	(0.040)
Old sib age 20	0.068	-0.011	0.071	0.186**	0.071
	(0.072)	(0.075)	(0.065)	(0.095)	(0.048)
Old sib age 21	0.060	-0.036	0.069	0.182*	0.083
	(0.080)	(0.081)	(0.070)	(0.101)	(0.055)
Old sib age 22	0.019	-0.038	0.089	0.258**	0.095
	(0.087)	(0.089)	(0.079)	(0.129)	(0.065)
Old sib age 23	-0.040	-0.058	0.053	0.307**	0.146
	(0.093)	(0.099)	(0.086)	(0.156)	(0.092)

Note: These estimates refer to the coefficients on control variables in Model 1 of Table 3. Standard errors clustered at household level in parentheses.* denotes significant at 10% level, ** at 5% level, and *** at 1% level. The sample sizes vary between 6132 and 6596. For all behaviors but doing hard drugs, the reference category for dummies is age 15 for the younder siblings and age 16 for the older siblings. For doing hard drugs, the reference cateogory is taken to be one year later since there are no data available on hard drug use for younger siblings at 15 and older siblings at 16.

APPENDIX TABLE 6

Linear Probability Model of Young Sibling's Behavior with Fixed Effects

	Smoking cigarettes last year	Drinking alcohol last year	Smoking marijuana last year	Using hard drugs last year	Selling drugs last year	Days smoked cigarettes last month	Days drank alcohol last month	Days smoked marijuana last month
y^1_{t-1}	0.029**	0.043***	0.017	0.017	0.002	0.024*	0.022	-0.005
	(0.014)	(0.015)	(0.015)	(0.016)	(0.015)	(0.014)	-(0.015)	(0.017)
Younger sibling's age dummies:								
16	0.019	0.031	0.062***	-	-0.003	0.597	0.357*	0.896***
	(0.020)	(0.022)	(0.018)	-	(0.012)	(0.367)	(0.207)	(0.242)
17	0.016	0.099***	0.089***	-0.010	0.013	1.733***	0.823***	1.614***
	(0.023)	(0.025)	(0.021)	(0.013)	(0.014)	(0.454)	(0.259)	(0.338)
18	0.059**	0.150***	0.095***	0.015	-0.007	2.397***	1.555***	1.412***
	(0.025)	(0.027)	(0.024)	(0.015)	(0.014)	(0.567)	(0.313)	(0.376)
19	0.064**	0.193***	0.119***	0.003	-0.010	2.787***	1.587***	1.863***
	(0.026)	(0.031)	(0.027)	(0.016)	(0.016)	(0.637)	(0.344)	(0.398)
20	0.076***	0.201***	0.107***	0.007	-0.029*	3.339***	1.530***	1.770***
	(0.026)	(0.032)	(0.028)	(0.016)	(0.015)	(0.659)	(0.379)	(0.385)
21	0.089**	0.229***	0.108***	-0.001	-0.033*	3.250***	2.453***	1.811***
	(0.028)	(0.034)	(0.030)	(0.016)	(0.017)	(0.682)	(0.430)	(0.391)
22	0.101***	0.211***	0.025	-0.001	-0.044**	3.406***	1.609***	1.248***
	(0.031)	(0.038)	(0.031)	(0.017)	(0.019)	(0.733)	(0.506)	(0.373)
Older sibling's age dummies:								
17	-0.017	-0.068*	0.019	-	0.020	-1.569**	-0.893**	-0.080
	(0.030)	(0.037)	(0.030)	-	(0.019)	(0.685)	(0.408)	(0.327)
18	0.011	-0.053	0.022	0.062***	0.019	-0.967	-1.142***	-0.493
	(0.030)	(0.035)	(0.028)	(0.020)	(0.017)	(0.669)	(0.391)	(0.364)
19	0.020	-0.054	0.015	0.032*	0.023	-0.709	-1.346***	-0.207
	(0.029)	(0.033)	(0.027)	(0.016)	(0.016)	(0.649)	(0.391)	(0.397)
20	0.020	-0.068**	0.003	0.023	0.023	-0.791	-1.298***	-0.394
	(0.028)	(0.032)	(0.026)	(0.016)	(0.015)	(0.618)	(0.371)	(0.412)
21	0.026	-0.070**	-0.017	0.008	0.025*	-0.637	-1.032***	-0.304
	(0.025)	(0.029)	(0.025)	(0.016)	(0.014)	(0.593)	(0.358)	(0.398)
22	0.019	-0.058**	-0.016	0.018	0.020	-0.310	-0.454	-0.215
	(0.022)	(0.025)	(0.023)	(0.014)	(0.013)	(0.506)	(0.332)	(0.346)
23	-0.014	-0.051**	-0.038**	0.009	0.022*	-0.058	-0.261	-0.223
	(0.018)	(0.021)	(0.019)	(0.012)	(0.012)	(0.407)	(0.306)	(0.288)

Note: Standard errors clustered at household level in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Ssample sizes vary between 6733 and 8733. For all behaviors but doing hard drugs, the reference category for dummies is age 15 for the younder siblings and age 16 for the older siblings. For doing hard drugs, the reference cateogory is taken to be one year later since there are no data available on hard drug use behavior for younger siblings at 15 and older siblings at 16.

APPENDIX TABLE 7A

Effect of Shifting the Older Sibling's Probability of Behavior from 0 to 1 in $t^2_{\min}-1$ on the Older and Younger Sibling's Probabilities of Behavior Relative to Baseline (Error A)

Smoking cigarettes

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline	0.4027	0.4188	0.4275	0.4339	0.4364	0.4346	
	(0.0637)	(0.0433)	(0.0477)	(0.0545)	(0.0578)	(0.0592)	
W/ feedback	2.483	0.4674	0.1192	0.0339	0.0102	0.0032	
	(0.0000)	(0.0504)	(0.0238)	(0.0098)	(0.0038)	(0.0015)	
				Younger Siblings			
Baseline		0.3322	0.3676	0.3786	0.3913	0.3849	0.3796
		(0.0795)	(0.0597)	(0.0724)	(0.0787)	(0.0865)	(0.0928)
W/ feedback		0.1427	0.0374	0.0113	0.0037	0.0013	0.0004
		(0.0865)	(0.0273)	(0.0102)	(0.0040)	(0.0016)	(0.0006)
W/out feedback		0.1427	0.0335	0.0092	0.0027	0.0009	0.0003
		(0.0865)	(0.0207)	(0.0060)	(0.0020)	(0.0008)	(0.0003)

Drinking Alcohol

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline	0.5502	0.5573	0.6031	0.6558	0.6861	0.6926	
	(0.0773)	(0.0527)	(0.0519)	(0.0458)	(0.0415)	(0.0428)	
W/ feedback	1.8175	0.2916	0.0518	0.0095	0.002	0.0005	
	(0.0000)	(0.0421)	(0.0147)	(0.0040)	(0.0011)	(0.0004)	
				Younger Siblings			
Baseline		0.4664	0.5041	0.5428	0.5657	0.5767	0.5754
		(0.0803)	(0.0536)	(0.0617)	(0.0673)	(0.0759)	(0.0829)
W/ feedback		0.2403	0.0448	0.0089	0.002	0.0004	0.0001
		(0.0709)	(0.0158)	(0.0042)	(0.0012)	(0.0004)	(0.0001)
W/out feedback		0.2403	0.0447	0.0088	0.0019	0.0004	0.0001
		(0.0709)	(0.0142)	(0.0034)	(0.0009)	(0.0003)	(0.0001)

Smoking Marijuana

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
	Older Siblings						
Baseline	0.2511	0.2663	0.2832	0.2939	0.2912	0.2889	
	(0.0609)	(0.0496)	(0.0558)	(0.0544)	(0.0524)	(0.0489)	
W/ feedback	3.9824	0.5766	0.1089	0.0229	0.0054	0.0011	
	(0.0000)	(0.0846)	(0.0286)	(0.0089)	(0.0029)	(0.0009)	
	Younger Siblings						
Baseline		0.2291	0.2643	0.2804	0.2938	0.3044	0.3088
		(0.0682)	(0.0530)	(0.0614)	(0.0701)	(0.0806)	(0.0883)
W/ feedback		0.2385	0.0375	0.0072	0.0015	0.0003	0
		(0.1533)	(0.0341)	(0.0099)	(0.0032)	(0.0010)	(0.0003)
W/out feedback		0.2385	0.0429	0.0095	0.0024	0.0006	0.0001
		(0.1533)	(0.0285)	(0.0070)	(0.0020)	(0.0005)	(0.0001)

Note:"Baseline" corresponds to probabilities of simulated behaviors using the dynamic probit model. "W/ feedback" corresponds to an exogenous shift of the older sibling's probability of behavior from 0 to 1 in the first period, allowing for the effect of this shift on the older sibling's behavior in the later periods. "W/out feedback" corresponds to an exogenous shift of the older sibling's probability of behavior from 0 to 1 in the first period, setting the older sibling's behavior in the later periods to its baseline level. The numbers recorded in the rows labeled "W/out feedback" and "W/ feedback" refer to the average change in said probabilities due to the corresponding exogenous switches in older siblings' behavior, divided by the baseline probability of these behaviors. Bootstrapped standard errors are in parentheses.

APPENDIX TABLE 7B

Effect of Shifting the Older Sibling's Probability of Behavior from 0 to 1 in $t^2_{\min}-1$ on the Older and Younger Siblings' Probabilities of Behavior Relative to Baseline (Error B)

Smoking cigarettes

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline	0.4114	0.4062	0.4129	0.4213	0.4243	0.4223	
	(0.0667)	(0.0541)	(0.0604)	(0.0592)	(0.0591)	(0.0610)	
W/ feedback	2.4309	0.4169	0.1003	0.0272	0.0079	0.0024	
	(0.0000)	(0.0686)	(0.0271)	(0.0103)	(0.0040)	(0.0015)	
				Younger Siblings			
Baseline		0.3504	0.3747	0.3809	0.3918	0.3818	0.3718
		(0.0712)	(0.0535)	(0.0686)	(0.0787)	(0.0883)	(0.0955)
W/ feedback		0.0000	0.0586	0.0709	0.0713	0.0738	0.0743
		(0.1519)	(0.0676)	(0.0697)	(0.0689)	(0.0720)	(0.0727)
W/out feedback		0.0000	0.0117	0.0061	0.0024	0.001	0.0003
		(0.1519)	(0.0359)	(0.0110)	(0.0036)	(0.0014)	(0.0005)

Drinking Alcohol

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline	0.5435	0.555	0.6003	0.6455	0.6803	0.6956	
	(0.0616)	(0.0417)	(0.0497)	(0.0497)	(0.0490)	(0.0496)	
W/ feedback	1.8398	0.2871	0.0504	0.009	0.0017	0.0003	
	(0.0000)	(0.0451)	(0.0135)	(0.0034)	(0.0009)	(0.0002)	
				Younger Siblings			
Baseline		0.4546	0.5082	0.5531	0.5793	0.5897	0.5919
		(0.0706)	(0.0695)	(0.0820)	(0.0871)	(0.0951)	(0.1050)
W/ feedback		0.1577	0.0444	0.0262	0.0215	0.0196	0.0189
		(0.1054)	(0.0397)	(0.0395)	(0.0371)	(0.0353)	(0.0339)
W/out feedback		0.1577	0.0274	0.0054	0.0012	0.0003	0.0001
		(0.1054)	(0.0170)	(0.0039)	(0.0010)	(0.0003)	(0.0001)

Smoking Marijuana

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
	Older Siblings						
Baseline	0.2488	0.2602	0.2781	0.2922	0.2966	0.2929	
	(0.0573)	(0.0486)	(0.0550)	(0.0568)	(0.0572)	(0.0584)	
W/ feedback	4.0187	0.4898	0.081	0.0146	0.0028	0.0006	
	(0.0000)	(0.0932)	(0.0259)	(0.0065)	(0.0017)	(0.0005)	
	Younger Siblings						
Baseline		0.2155	0.261	0.2865	0.3026	0.3109	0.312
		(0.0589)	(0.0637)	(0.0815)	(0.0889)	(0.1012)	(0.1102)
W/ feedback		0.2145	0.0311	0.0060	0.0011	0.0003	0.0000
		(0.1757)	(0.0289)	(0.0071)	(0.0019)	(0.0006)	(0.0002)
W/out feedback		0.2145	0.0346	0.0071	0.0015	0.0004	0.0001
		(0.1757)	(0.0286)	(0.0062)	(0.0014)	(0.0004)	(0.0001)

Note: See Appendix Table 7A.

APPENDIX TABLE 8

Estimates of Joint Dynamic Ordered Probit Model (Error B)

			Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence					
Older sibling	γ_L^1	Low consumption	0.300***	0.348***	0.298***
			(0.082)	(0.045)	(0.083)
	γ_H^1	High consumption	0.755***	0.616***	0.646***
			(0.071)	(0.068)	(0.084)
Younger sibling	γ_L^2	Low consumption	0.560***	0.531***	0.563***
			(0.080)	(0.045)	(0.083)
	γ_H^2	High consumption	0.971***	0.821***	0.952***
			(0.077)	(0.070)	(0.095)
Sibling's influence					
t_min	λ_{1L}^2	Low consumption	0.057	0.093	0.351*
			(0.201)	(0.112)	(0.185)
	λ_{1H}^2	High consumption	0.392**	0.255	0.247
			(0.199)	(0.186)	(0.188)
t>t_min	λ_{2L}^2	Low consumption	0.070	-0.022	0.074
			(0.091)	(0.045)	(0.095)
	λ_{2H}^2	High consumption	0.098	-0.155**	-0.011
			(0.086)	(0.069)	(0.104)
Family-specific error term					
Stand. deviation	σ_ε		1.158***	0.454***	0.881***
			(0.155)	(0.054)	(0.155)
Factor loadings	$\delta_{2\varepsilon}^1$	Older sibling $(t > t_{\min})$	0.920***	0.922***	1.282***
			(0.128)	(0.110)	(0.206)
	$\delta_{1\varepsilon}^2$	Younger sibling $(t_{\min})$	0.525***	1.654***	0.418***
			(0.147)	(0.298)	(0.133)
	$\delta_{2\varepsilon}^2$	Younger sibling $(t > t_{\min})$	0.543***	1.679***	0.493***
			(0.125)	(0.219)	(0.124)
Individual- specific error term					
Stand. deviation	σ_υ		0.934***	0.626***	0.661***
			(0.077)	(0.060)	-(0.086)
Factor loadings	$\delta_{2\upsilon}^1$	Older sibling $(t > t_{\min})$	1.302***	1.006***	0.365
			(0.130)	(0.106)	(0.287)
	$\delta_{2\upsilon}^2$	Younger sibling $(t > t_{\min})$	1.173***	-0.156	1.345***
			(0.134)	(0.153)	(0.199)
Low consumption threshold			-0.098	-0.841	0.876
			(0.702)	(0.538)	(1.013)
High consumption threshold			0.496	0.811	1.508
			(0.702)	(0.539)	(1.014)
Log likelihood value			-9181.54	-12852.93	-7121.72

Note: See Table 5.

APPENDIX TABLE 9A

Effect of Shifting the Older Sibling's Behavior from Zero to High Consumption in $t_{\min}^{2}-1$ on the Older and Younger Siblings' Probabilities of Behavior Relative to the Baseline (Error A)

Smoking cigarettes

	$t_{\min}^{2}-1$	$t_{\min}^{2}$	$t_{\min}^{2}+1$	$t_{\min}^{2}+2$	$t_{\min}^{2}+3$	$t_{\min}^{2}+4$	$t_{\min}^{2}+5$
				Older Siblings			
Baseline							
Low consumption	0.0878	0.0835	0.0837	0.0845	0.0848	0.084	
	(0.0192)	(0.0147)	(0.0138)	(0.0124)	(0.0114)	(0.0105)	
High consumption	0.2343	0.2677	0.2903	0.3036	0.3067	0.3075	
	(0.1048)	(0.1056)	(0.1147)	(0.1170)	(0.1139)	(0.1135)	
W/ feedback							
Low consumption	0.0000	0.1882	0.0113	0.0003	0.0004	0.0000	
	(0.0000)	(0.1281)	(0.0263)	(0.0099)	(0.0040)	(0.0019)	
High consumption	4.2681	0.4665	0.088	0.0192	0.0045	0.0011	
	(0.0000)	(0.1018)	(0.0271)	(0.0081)	(0.0025)	(0.0008)	
				Younger Siblings			
Baseline							
Low consumption		0.0946	0.0879	0.0876	0.089	0.088	0.088
		(0.0258)	(0.0190)	(0.0158)	(0.0147)	(0.0142)	(0.0147)
High consumption		0.2175	0.2438	$t_{\min}^{2}-1$552	0.2853	0.2872	0.2932
		(0.1034)	(0.1019)	(0.1133)	(0.1172)	(0.1193)	(0.1183)
W/ feedback							
Low consumption		0.1384	0.019	0.0019	-0.0002	0.0003	0.0002
		(0.0877)	(0.0323)	(0.0104)	(0.0039)	(0.0019)	(0.0013)
High consumption		0.3164	0.0767	0.0203	0.0058	0.0017	0.0005
		(0.1689)	(0.0386)	(0.0107)	(0.0037)	(0.0013)	(0.0005)
W/out feedback							
Low consumption		0.1384	0.0158	0.0018	-0.0002	0.0001	0.0001
		(0.0877)	(0.0306)	(0.0093)	(0.0036)	(0.0016)	(0.0009)
High consumption		0.3164	0.0693	0.0171	0.0046	0.0014	0.0004
		(0.1689)	(0.0363)	(0.0094)	(0.0030)	(0.0011)	(0.0004)

Drinking alcohol

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
	Older Siblings						
Baseline							
Low consumption	0.2905	0.3127	0.3327	0.3541	0.3629	0.3684	
	(0.0895)	(0.0791)	(0.0677)	(0.0555)	(0.0454)	(0.0400)	
High consumption	0.0922	0.1467	0.1838	0.2212	0.2583	0.2754	
	(0.0682)	(0.0935)	(0.0983)	(0.1035)	(0.1096)	(0.1076)	
W/ feedback							
Low consumption	0	0.1969	0.0166	0.0007	0.0000	-0.0001	
	(0.0000)	(0.1352)	(0.0138)	(0.0025)	(0.0006)	(0.0002)	
High consumption	10.8517	0.6148	0.0791	0.0113	0.0016	0.0003	
	(0.0000)	(0.2543)	(0.0308)	(0.0050)	(0.0008)	(0.0002)	
	Younger Siblings						
Baseline							
Low consumption		0.2758	0.2939	0.324	0.3424	0.3527	0.3639
		(0.0996)	(0.0884)	(0.0764)	(0.0655)	(0.0575)	(0.0493)
High consumption		0.0816	0.1189	0.1613	0.1989	0.2252	0.2499
		(0.0690)	(0.0815)	(0.0994)	(0.1062)	(0.1081)	(0.1031)
W/ feedback							
Low consumption		0.0811	0.0063	0.0002	0.0001	0.0001	0.0000
		(0.0985)	(0.0168)	(0.0030)	(0.0006)	(0.0004)	(0.0002)
High consumption		0.2335	0.0169	0.0006	-0.0004	-0.0002	0.0000
		(0.3081)	(0.0474)	(0.0099)	(0.0024)	(0.0006)	(0.0002)
W/out feedback							
Low consumption		0.0811	0.0114	0.0013	0.0002	0.0000	0.0000
		(0.0985)	(0.0161)	(0.0028)	(0.0005)	(0.0002)	(0.0001)
High consumption		0.2335	0.0359	0.0066	0.0013	0.0002	0.0001
		(0.3081)	(0.0468)	(0.0087)	(0.0018)	(0.0004)	(0.0001)

Smoking marijuana

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
	Older Siblings						
Baseline							
Low consumption	0.0712	0.0773	0.0805	0.0821	0.0816	0.0814	
	(0.0398)	(0.0390)	(0.0363)	(0.0341)	(0.0328)	(0.0324)	
High consumption	0.1279	0.1652	0.1842	0.1899	0.1882	0.1863	
	(0.1278)	(0.1448)	(0.1470)	(0.1408)	(0.1370)	(0.1318)	
W/ feedback							
Low consumption	0	0.2648	0.023	0.003	0.0012	0.0002	
	(0.0000)	(0.1643)	(0.0268)	(0.0059)	(0.0024)	(0.0009)	
High consumption	7.8174	0.5241	0.0724	0.0122	0.0023	0.0005	
	(0.0000)	(0.2880)	(0.0491)	(0.0099)	(0.0021)	(0.0006)	
	Younger Siblings						
Baseline							
Low consumption		0.0755	0.0828	0.0817	0.0815	0.081	0.0826
		(0.0428)	(0.0400)	(0.0355)	(0.0330)	(0.0319)	(0.0290)
High consumption		0.1382	0.1782	0.1814	0.1821	0.1712	0.1707
		(0.1323)	(0.1633)	(0.1654)	(0.1601)	(0.1460)	(0.1328)
W/ feedback							
Low consumption		0.1297	0.0108	0.0026	0.0003	0.0003	0.0001
		(0.1474)	(0.0258)	(0.0077)	(0.0029)	(0.0014)	(0.0007)
High consumption		0.1855	0.0324	0.0076	0.0021	0.0006	0.0002
		(0.2333)	(0.0482)	(0.0132)	(0.0040)	(0.0013)	(0.0005)
W/out feedback							
Low consumption		0.1297	0.0097	0.0018	0.0004	0.0001	0.0000
		(0.1474)	(0.0244)	(0.0072)	(0.0019)	(0.0012)	(0.0005)
High consumption		0.1855	0.0341	0.0086	0.0024	0.0007	0.0002
		(0.2333)	(0.0458)	(0.0122)	(0.0037)	(0.0012)	(0.0004)

Note: "Baseline" corresponds to probabilities of simulated behaviors using the dynamic ordered probit model. "W/ feedback" corresponds to an exogenous shift of the older sibling's behavior from zero to high consumption in the first period, allowing for the effect of this shift on the older sibling's behavior in the later periods. "W/out feedback" corresponds to an exogenous shift of the older sibling's behavior from zero to high consumption in the first period, setting the older sibling's behavior in the later periods to the baseline level. The numbers recorded in the rows labeled "W/out feedback" and "W/ feedback" refer to the average change in said probabilities due to the corresponding exogenous switches in older siblings' behavior, divided by the baseline probability of these behaviors. Bootstrapped standard errors are in parentheses.

APPENDIX TABLE 9B

Effect of Shifting the Older Sibling's Behavior from Zero to High Consumption in $t^2_{\min}-1$ on the Older and Younger Siblings' Probabilities of Behavior Relative to the Baseline (Error B)

Smoking cigarettes

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline							
Low consumption	0.0916	0.0818	0.0828	0.0852	0.0855	0.0857	
	(0.0190)	(0.0154)	(0.0142)	(0.0134)	(0.0125)	(0.0115)	
High consumption	0.2491	0.2781	0.3044	0.3209	0.3242	0.3275	
	(0.1203)	(0.1233)	(0.1268)	(0.1214)	(0.1160)	(0.1137)	
W/ feedback							
Low consumption	0	0.1737	0.0045	-0.0014	0.0001	-0.0001	
	(0.0000)	(0.1325)	(0.0325)	(0.0092)	(0.0032)	(0.0012)	
High consumption	4.0147	0.4173	0.0744	0.0155	0.0033	0.0008	
	(0.0000)	(0.1072)	(0.0283)	(0.0080)	(0.0022)	(0.0007)	
				Younger Siblings			
Baseline							
Low consumption		0.0907	0.0865	0.0856	0.088	0.0881	0.0875
		(0.0254)	(0.0176)	(0.0155)	(0.0139)	(0.0136)	(0.0137)
High consumption		0.2072	0.2501	0.2712	0.293	0.2987	0.3075
		(0.1080)	(0.1076)	(0.1126)	(0.1099)	(0.1073)	(0.1124)
W/ feedback							
Low consumption		0.1352	0.0139	0.0023	0.0000	0.0002	0.0000
		(0.1126)	(0.0277)	(0.0082)	(0.0040)	(0.0020)	(0.0011)
High consumption		0.2766	0.0604	0.0158	0.0044	0.0012	0.0004
		(0.2024)	(0.0350)	(0.0091)	(0.0028)	(0.0009)	(0.0003)
W/out feedback							
Low consumption		0.1352	0.0110	0.0011	0.0000	0.0001	-0.0001
		(0.1126)	(0.0261)	(0.0067)	(0.0029)	(0.0013)	(0.0008)
High consumption		0.2766	0.0522	0.0126	0.0032	0.0008	0.0002
		(0.2024)	(0.0356)	(0.0089)	(0.0026)	(0.0008)	(0.0003)

Drinking alcohol

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
	Older Siblings						
Baseline							
Low consumption	0.3096	0.3257	0.3483	0.367	0.3771	0.3802	
	(0.0826)	(0.0724)	(0.0612)	(0.0477)	(0.0397)	(0.0347)	
High consumption	0.1079	0.1532	0.1988	0.244	0.272	0.2895	
	(0.0787)	(0.0947)	(0.1071)	(0.1093)	(0.1090)	(0.1113)	
W/ feedback							
Low consumption	0	0.1864	0.0128	-0.0001	-0.0002	0	
	(0.0000)	(0.1360)	(0.0198)	(0.0031)	(0.0009)	(0.0002)	
High consumption	9.2711	0.6314	0.0847	0.0127	0.002	0.0003	
	(0.0000)	(0.2708)	(0.0414)	(0.0072)	(0.0016)	(0.0003)	
	Younger Siblings						
Baseline							
Low consumption		0.2644	0.3113	0.3295	0.3513	0.3611	0.3692
		(0.0785)	(0.0704)	(0.0578)	(0.0461)	(0.0391)	(0.0366)
High consumption		0.0955	0.1264	0.1715	0.2055	0.2329	0.2629
		(0.0606)	(0.0702)	(0.0866)	(0.0892)	(0.0896)	(0.0975)
W/ feedback							
Low consumption		0.1121	0.0082	0.0005	0.0000	-0.0001	0.0000
		(0.1080)	(0.0161)	(0.0028)	(0.0007)	(0.0003)	(0.0001)
High consumption		0.2877	0.0244	0.0013	-0.0003	-0.0001	0.0000
		(0.2931)	(0.0422)	(0.0076)	(0.0018)	(0.0005)	(0.0001)
W/out feedback							
Low consumption		0.1121	0.0136	0.0016	0.0002	0.0000	0.0000
		(0.1080)	(0.0161)	(0.0027)	(0.0006)	(0.0002)	(0.0001)
High consumption		0.2877	0.0458	0.008	0.0015	0.0003	0.0001
		(0.2931)	(0.0447)	(0.0078)	(0.0015)	(0.0004)	(0.0001)

Smoking marijuana

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
				Older Siblings			
Baseline							
Low consumption	0.0779	0.0797	0.0847	0.0878	0.0868	0.0866	
	(0.0407)	(0.0351)	(0.0334)	(0.0301)	(0.0283)	(0.0284)	
High consumption	0.1511	0.1849	0.2093	0.2184	0.2104	0.2081	
	(0.1427)	(0.1429)	(0.1456)	(0.1351)	(0.1232)	(0.1170)	
W/ feedback							
Low consumption	0	0.2321	0.0124	0.0007	0.0002	-0.0001	
	(0.0000)	(0.1579)	(0.0225)	(0.0052)	(0.0021)	(0.0007)	
High consumption	6.6167	0.4612	0.0623	0.0104	0.002	0.0004	
	(0.0000)	(0.2374)	(0.0463)	(0.0098)	(0.0028)	(0.0007)	
				Younger Siblings			
Baseline							
Low consumption		0.084	0.0857	0.0836	0.0855	0.087	0.0879
		(0.0465)	(0.0380)	(0.0329)	(0.0303)	(0.0286)	(0.0267)
High consumption		0.1415	0.1711	0.1836	0.187	0.1851	0.1889
		(0.1263)	(0.1392)	(0.1468)	(0.1310)	(0.1195)	(0.1178)
W/ feedback							
Low consumption		0.1093	0.0100	0.0006	-0.0001	0.0003	0.0000
		(0.1323)	(0.0204)	(0.0076)	(0.0024)	(0.0014)	(0.0008)
High consumption		0.2060	0.0325	0.0066	0.0015	0.0003	0.0001
		(0.2481)	(0.0411)	(0.0093)	(0.0023)	(0.0008)	(0.0003)
W/out feedback							
Low consumption		0.1093	0.0114	0.0013	0.0002	0.0001	0.0001
		(0.1323)	(0.0195)	(0.0063)	(0.0015)	(0.0009)	(0.0005)
High consumption		0.206	0.0334	0.0071	0.0017	0.0005	0.0001
		(0.2481)	(0.0428)	(0.0098)	(0.0027)	(0.0008)	(0.0003)

Note: See Appendix Table 9A.

APPENDIX TABLE 10

Estimates of Dynamic Probit Model Allowing for Dependence of Sibling's Influence on Gender Mix

			Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence	γ^1	Older Sibling	0.824***	0.632***	0.622***
			(0.069)	(0.059)	(0.068)
	γ^2	Younger sibling	0.914***	0.612***	0.710***
			(0.073)	(0.059)	(0.068)
Sibling's influence					
	λ^2_{1mm}	Brothers	-0.151	0.273	0.193
			(0.231)	(0.185)	(0.215)
	λ^2_{1ff}	Sisters	0.195	0.234	0.492**
			(0.252)	(0.192)	(0.241)
	λ^2_{1mf}	Mixed Pair	0.004	0.283*	0.233
			(0.202)	(0.150)	(0.170)
	λ^2_{2mm}	Brothers	0.026	0.022	-0.179
			(0.120)	(0.105)	(0.127)
	λ^2_{2ff}	Sisters	0.323**	0.114	0.141
			(0.137)	(0.109)	(0.150)
	λ^2_{2mf}	Mixed Pair	0.088	0.018	-0.012
			(0.097)	(0.080)	(0.092)
Family-specific error term					
Stand. Deviation	σ_ε		0.863***	0.758***	0.542***
			(0.127)	(0.106)	(0.102)
Factor Loadings	$\delta^1_{2\varepsilon}$	Older sibling $(t > t_{\min})$	1.109***	0.911***	1.304***
			(0.184)	(0.152)	(0.250)
	$\delta^2_{1\varepsilon}$	Younger sibling $(t_{\min})$	1.014***	0.842***	1.253***
			(0.274)	(0.196)	(0.349)
	$\delta^2_{2\varepsilon}$	Younger sibling $(t > t_{\min})$	0.716***	0.748***	1.165***
			(0.160)	(0.153)	(0.329)
Individual-specific error term					
Stand. Deviation	σ_υ		0.748***	0.452***	0.562***
			(0.072)	(0.068)	(0.078)
Factor loadings	$\delta^1_{2\upsilon}$	Older sibling $(t > t_{\min})$	1.487***	1.196***	1.371***
			(0.185)	(0.297)	(0.275)
	$\delta^2_{2\upsilon}$	Younger sibling $(t > t_{\min})$	1.346***	1.747***	1.342***
			(0.161)	(0.277)	(0.247)
Log likelihood value			-7241.73	-8087.47	-6684.64

Note: See Table 4A.

APPENDIX TABLE 11

Estimates of Dynamic Probit Model Allowing for Dependence of Sibling's Influence on Age Gap

			Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence	γ^1	Older Sibling	0.834***	0.632***	0.622***
			(0.069)	(0.059)	(0.068)
	γ^2	Younger sibling	0.910***	0.612***	0.708***
			(0.073)	(0.059)	(0.068)
Sibling's influence					
$t_{\min}$	λ_1^2	Brothers	0.134	0.245*	0.233
			(0.131)	(0.133)	(0.154)
	$\lambda_{1,2+}^2$	Sisters	0.078	0.133	0.118
			(0.207)	(0.185)	(0.217)
$t > t_{\min}$	λ_2^2	Brothers	0.079	0.026	-0.072
			(0.084)	(0.067)	(0.080)
	$\lambda_{1,2+}^2$	Sisters	0.109	0.058	0.146
			(0.132)	(0.124)	(0.141)
Family-specific error term					
Stand. Deviation	σ_ε		1.160***	0.757***	0.540***
			(0.173)	(0.105)	(0.102)
Factor Loadings	$\delta_{2\varepsilon}^1$	Older sibling $(t > t_{\min})$	1.219***	0.916***	1.304***
			(0.187)	(0.152)	(0.250)
	$\delta_{1\varepsilon}^2$	Younger sibling $(t_{\min})$	0.445***	0.834***	1.253***
			(0.099)	(0.194)	(0.349)
	$\delta_{2\varepsilon}^2$	Younger sibling $(t > t_{\min})$	0.379***	0.749***	1.165***
			(0.090)	(0.152)	(0.329)
Individual-specific error term					
Stand. Deviation	σ_υ		0.910***	0.453***	0.562***
			(0.087)	(0.067)	(0.079)
Factor loadings	$\delta_{2\upsilon}^1$	Older sibling $(t > t_{\min})$	0.339*	1.187***	1.369***
			(0.196)	(0.298)	(0.279)
	$\delta_{2\upsilon}^2$	Younger sibling $(t > t_{\min})$	1.223***	1.744***	1.351***
			(0.139)	(0.274)	(0.245)
Log likelihood value			-7241.73	-8087.47	-6684.64

Note: See Table 4A.