Institute for Fiscal Studies

Identifying sibling influence on teenage substance use

IFS Working Paper W13/04

Joseph G. Altonji
Sarah Cattan
Iain Ware

E·S·R·C
ECONOMIC & SOCIAL RESEARCH COUNCIL

Identifying Sibling Influence on Teenage Substance Use

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

February 19, 2013

Abstract

The large sibling correlations in risky behavior between siblings raise 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 causal. 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 strong 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. We find that smoking, drinking, and marijuana use are affected by the example of older siblings, but only a small fraction of the link between siblings is causal.

*[1]Yale University and NBER, [2]Institute for Fiscal Studies, [3] Bain Capital. We are grateful to Jerome Adda, Monica Deza, Greg Duncan, Patrick Kline, Amanda Kowalski, Costas Meghir, Robert T. Michael, and participants in seminars at UC Berkeley, Brigham Young University, University of Chicago, The European Institute, MIT, 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, the SOLE/EALE 2010 Meetings, Stanford University, the University of Toronto, and Yale University for valuable comments. We thank Stan Panis for advice on how to adapt his program aML to estimate the joint dynamic probit and dynamic ordered probit models presented in the paper. We are among a large set of researchers who have benefited from Panis' generous support. 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 crime, and engagement 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 unobserved shared influences using two related empirical strategies. Both require panel data on sibling pairs. Both exploit a basic fact, and both are based on a key maintained assumption. The fact is that future actions of a youth cannot causally influence his or her sibling's actions today. The assumption is that younger siblings do not influence older siblings. 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.

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

[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.

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.

While the CRE design is a natural place to start, 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 younger sibling's error component effect. This would bias the CRE estimate of the older sibling's influence, although the direction of the bias is not clear. Consequently, our main focus is a series of dynamic models of the behavior of the older and younger siblings. The models allow for state dependence and unobserved heterogeneity at the individual and sibling pair levels. They consist 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 use the model 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 indicate that smoking by the older sibling raises the probability that the younger sibling smokes by about 15.6% of the sample mean for smoking. In the case of drinking alcohol the effect is about 9.1% of the sample mean. The results using the dynamic model show positive effects on smoking, drinking, and marijuana use. Smoking among older siblings in the period before we first observe the younger sibling increases smoking among younger siblings by about 14.1% of the baseline value. The corresponding values are 24.3% for drinking and 25.4% for marijuana use. However, the effects are smaller in later periods. Ordered probit results indicate that the sibling effect increases with the frequency of the older sibling's substance use. We also obtain positive point estimates of sibling effects on use of hard drugs and on selling drugs, but the estimates fall short of statistical significance.

Overall, we conclude that there is a modest positive sibling effect on substance use. On the other hand, simulations from the dynamic probit model indicate that sibling effects account for only a small fraction of the strong sibling correlation in substance use, although point estimates are noisy.

The results suggest that parenting behaviors and anti-substance use programs aimed at an adolescent would have beneficial spillovers on younger siblings, although we provide no direct evidence for specific behaviors or programs. And 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.[3] 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 joint dynamic probit model in sections 7, including a version that allows for gateway drugs. In section 8, we consider 3 category and 5 category joint dynamic ordered probit models. 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. We close with conclusions and a research agenda.

[3]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

We begin with a brief survey of the literature on family influences on risky behaviors, particularly substance use drawing across the social sciences.[4]

Developmental psychologists and sociologists were first to investigate the importance of social environment on adolescent development and behavior. 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; Barnes, 1990).

Within the family, siblings occupy a particular social position, and the psychology literature suggests two main mechanisms through which siblings may influence each other. The first one is that a sibling, most likely the younger one, may see his older sibling as a role model to observe, imitate, and use in shaping his notions about what types of behaviors are suitable (Widmer, 1997; Buhrmester, 1992; Rodgers and Rowe, 1988; Bikchandani et al., 1992). In the context of risky behavior, Patterson (1984) argues that siblings are more likely to learn this type of behaviors from each other when they have conflict ridden and aggressive relationships, because these promote antisocial behavior.

The second mechanism, which we refer to as the "opportunity hypothesis", suggests that siblings influence each other's behaviors 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 better and warmer relationships, share 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).

Although the economics literature does not focus specifically on siblings interactions, it also offers some rationales for conformity in behavior, which can be applied to siblings. For example, 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

[4] Cawley and Ruhm (2011) provide a comprehensive survey of the economics literature on risky health behaviors.

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.

In addition to the theoretical work reviewed above, a large number of studies have investigated social influence on youth behavior empirically. However, 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 a 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.

Using the same data set, Slomkowski et al. (2005) find that both genetic and environmental factors contribute to similarities between siblings' smoking behavior. Accordingly, parental behavior has been shown to be a source of imitation, although studies, such as Conger and Reuter (1996) and Windle (2000), show that it is less potent than sibling influences.

Researchers have also used data about the quality of the relationship between siblings to study sibling influences. For example, 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, and when siblings have more mutual friends and are of the same gender. 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. Overall however, results based on this type of data are mixed and often contradictory.[5]

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), estimates

[5]See, for example, Slomkowski et al. (2001). Several papers have also looked at sibling influences on smoking patterns, although results for this activity are also mixed (Otten et al., 2007; Bricker et al., 2005; Slomkowski et al., 2005).

linear probability models of high school graduation of an older sibling on the probability that the younger sibling graduates and vice-versa on the NLSY97. He finds that a sibling influence runs mostly from the older to the younger sibling, but his identification strategy relies on exclusion restrictions that seem questionable. [6]

Ouyang (2004) develops a dynamic model of the older and younger siblings' behaviors, which allows 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 with our approach, however, she does not allow for individual specific unobserved heterogeneity and 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 by observing their siblings' decisions. They allow the decision not to smoke to have a different effect than the decision to smoke cigarettes. 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, sibling effects account for the entire sibling correlation in the smoking. 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.[7]

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

[6]Oettinger (2000) 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.

[7]Consider a family with two siblings. Their model contains exogenous variables and is nonlinear, but 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.

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. We focus on smoking cigarettes, using marijuana, drinking alcohol, using cocaine and/or other hard drugs, and selling or helping to sell drugs.[8] The main outcome we analyze is whether the individual reports having engaged at all in the 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 NLSY97 variables. The first and most important one is a dummy variable indicating whether the respondent has engaged in the 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 in Web Appendix A we report the exact name, reference numbers, and survey years of the variables we used. The substance use information is first available 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

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

is composed of 1650 pairs of siblings, for whom we have between 1 and 6 rounds of observations.

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. We also present results based on 5 consumption categories.

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 1,456 pairs come from two-sibling families, while 176, 12, and 6 come from three, four, and five sibling families respectively.[9] 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.[10]

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

[9] 403 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. 359 of the families had children who were younger than 12 at the end of 1996. 167 had children who were older than 16 and younger than 12. No data were collected on these children.

[10] 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 needed 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.[11] 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 provide further details about variable construction and sample selection in Web Appendix A. 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, 61% of the males and 58% of the females report drinking alcohol during the previous year. 25% of the males and 19% of the females report using marijuana. The figure is about 40% 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 Web Appendix Table 1). The fractions who used the substance one or more days in the past month are lower, not surprisingly. Web Appendix Table 2 shows that incidence of the behaviors tends to increase with age until about age 20. The fractions of older siblings who engage in the behavior in all years and who engage in the behavior in some years are .21 and .42 (respectively) for smoking, .33 and .54 for drinking, .06 and .43 for marijuana, .005 and .19 for hard drugs, and .007 and .17 for selling hard drugs. The "some years" group plays a key role in distinguishing between family correlations and sibling effects.

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

[11] 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.

of ρ from the regression

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

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 siblings' 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.

The results are striking. Consider smoking cigarettes. If the older sibling smoked, the probability shifts by .226, which is very large relative to the sample mean of about .4. With controls, the shift in the probability remains large at .17. 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 .157, which is very large relative to the sample mean of about .22. 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 sibling by .092, a shift that is *larger* than the unconditional mean of .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.[12]

We also present separate results for brother pairs and sister pairs. The relationship across siblings tends to be larger for sister pairs, with the exception of selling drugs, a behavior in which females engage infrequently. Later in the paper, we explore whether the size of the peer effect depends on the gender composition of the pair.

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.

[12]The substance use questions are administered directly by the respondent into a computer. The interviewer cannot observe the responses. However, sometimes siblings are interviewed on the same day. We regressed $y^2_{a,t}$ on $y^1_{a+j,t}$ adding an intercept shift for whether the reports of $y^2_{a,t}$ and $y^1_{a+j,t}$ for year t were obtained on the same day. The coefficients main effects and interaction terms are positive and significant for smoking. This does not affect reports underlying the regression in Table 2 are from different

5 Model and methods

5.1 A model of substance use and sibling influences

In this section we present the joint dynamic model of substance use that underlies much of the econometric analysis. Consider a sibling pair. We continue to use a for age and the superscripts 1 and 2 to refer to the older and younger sibling, respectively. We leave family subscripts implicit throughout the paper. We focus on the case in which y is a binary choice.

In period t, the older sibling chooses y_t^1 according to

$$y_t^1 = 1(\gamma^1 y_{t-1}^1 + X^1\beta^1 + AGE_t^1\Gamma^1 + \alpha^1\varepsilon + \delta^1 v^1 + u_t^1 > 0) \tag{2}$$

The expression to the left of the inequality is the difference between the perceived benefit and the cost of y_t^1, including the opportunity costs of foregoing other goods.[13] The benefit of y_t^1 depends on a set of covariates X^1, the vector AGE_t^1 of age dummies indicating whether the older sibling is aged a in year t, the family specific component ε and the person specific component for the older sibling, v^1. u_t^1 is a transitory error component for the older sibling in period t.

The benefit of y_t^1 also depends on y_{t-1}^1, the choice of the older sibling in the previous period. The dependence parameter γ^1 reflects two mechanisms. The first one is the effect of habit formation and informational effects. The second is the effect of the information the parent has about the children, as well as the positive or negative influence of the parents' reaction on the net benefit of y_t^1 to the older sibling. In principle, the benefit of y_t^1 could depend on y_{t-1}^2 through direct peer influence from the younger to the older sibling or because the parental response to the behavior of the younger child has an effect on older sibling. We assume that both effects are 0 and leave y_{t-1}^2 out of (2).

The younger sibling faces a similar problem to that of the older one, except that the net benefit of his behavior also depends directly on the action of his older sibling in $t-1$.

[13] The budget constraint, which we leave implicit, is static. The decision function implicitly 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 costs and benefits of choosing y in the future. We do not allow the 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.

He chooses y_t^2 according to

$$y_t^2 = 1(\gamma^2 y_{t-1}^2 + \lambda^2 y_{t-1}^1 + \theta^2 a_{t-1}^1 + X^2\beta^2 + AGE_t^2\Gamma^2 + \alpha^2\varepsilon + \delta^2 v^2 + u_t^2 > 0) \qquad (3)$$

where v^2 is a component specific to the younger sibling and u_t^2 is a transitory error component for the younger sibling at time t. a_{t-1}^1 is the older sibling's age in the previous period, X^2 is a set of observed covariates, and AGE_t^2 is a vector of age dummies for the younger sibling. The main parameter of interest, λ^2, captures the direct influence of the older sibling on the younger sibling. We do not attempt to identify the specific mechanisms that underlie it, such as information provision, shaping of preferences, etc. Without loss of generality, v^1 and the corresponding component v^2 for the younger sibling are uncorrelated. Below we place restrictions on the distributions of u^1 and u^2 over time and across siblings.

We assume that substance use is 0 for all people if $a_t \leq A_0$ where A_0 is the age of initiation. Thus $y_{t-(a_t^2 - A_0)}^2 = 0$ and $y_{t-(a_t^1 - A_0)}^1 = 0$.

5.2 Using Correlated Random Effects (CRE) regression to estimate the direct sibling effect

The CRE estimates of the direct effect of older siblings on younger siblings' behavior are based on the following linear least squares projection equation:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + error\ . \qquad (4)$$

In Web Appendix B, we discuss in detail the assumptions on the above model that are required for β_2 to equal the sibling influence parameter λ^2, in the case in which y is continuous. They include: (A1) no state dependence; (A2) covariance stationarity of u_t^1; and (A3) the symmetry restriction $cov(u_t^2,\, u_{t-1}^1) = cov(u_t^2, u_{t+1}^1)$. Furthermore, β_2 will not equal λ^2 if one generalizes the above model to allow the effects of ε or v^1 on y_t^1 to vary with the age of older sibling a_t^1 or if one allows the influence of ε on y_t^2 to vary with the age of the younger sibling a_t^2. The basic argument carries over to the case in which y is a binary choice.

The assumption of no state dependence is problematic, given the evidence that the risky behaviors of interest are addictive. Furthermore, the presence of the terms $AGE_t^1\Gamma^1$ and $AGE_t^2\Gamma^2$ alone are enough to lead to age dependence in the influence of the error components on y_t^1 and y_t^2 in a nonlinear binary choice model such as (2) and (3). Following Chamberlain (1984), one could relax the assumption of no age dependence to some extent by replacing $\beta_1(y_{t-1}^1 + y_{t+1}^1)$ with $\beta_{1,a_{t-1}^1}y_{t-1}^1 + \beta_{1,a_{t+1}^1}y_{t+1}^1$ and still identify β_2, but we stick with the simpler CRE specification and use it in conjunction with the joint dynamic probit model that we discuss in the next section.

We also consider the case in which both contemporaneous and lagged behaviors of the older sibling influence the younger child with coefficients λ^{20} and λ^2, respectively. Consider the following projection equation:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_t^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + \beta_3 y_t^1 + error \tag{5}$$

In this case, we need to make two additional assumptions, in addition to (A1) through (A3), for β_2 and β_3 to capture λ^2 and λ^{20}. These assumptions are: (A4) u_t^2 and $u_{t'}^1$ are independent across siblings at all leads and lags and (A5) u_t^1 is serially uncorrelated. Under assumptions (A1) through (A5), one can identify the direct sibling effects. However, if any of these fail to hold, then in general $\beta_2 \neq \lambda^2$ and $\beta_3 \neq \lambda_0^{20}$ in (5).

If only (A5) fails, one can still estimate an average of λ^2 and λ^{20} by using the regression:

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

to test for sibling effects, as we do below. 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 and thus violate (A4). Consequently, we place less weight on specification (6). Finally, if one uses (4) when (5) is correct, then the coefficient on y_{t-1}^1 will pick up part of the effect of y_t^1, but one will still detect sibling influences.

Use of 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. While state dependence and nonstationarity will lead to inconsistency in the estimates of λ^2, the CRE approach has the advantage of simplicity and is a natural place to start the search for evidence of a causal effect of sibling behavior on substance use.

5.3 The econometric specification of the joint dynamic discrete choice model of sibling behavior

5.3.1 The joint dynamic probit model

Since behavior is dynamic and we do not observe behavior at the age of initiation, we augment equations (2) and (3) for y_t^1 and y_t^2 with equations for the initial condition of the older and younger siblings. The choice of y_t^1 in year t_{min}^1, the first year we observe the older sibling, is determined by:

$$y_t^1 = 1(X^1\beta_0^1 + AGE_t^1\Gamma_0^1 + \alpha_0^1\varepsilon + \delta_0^1 v^1 + u_t^1 > 0),\ t = t_{\min}^1 \tag{7}$$

The corresponding initial condition for the younger sibling is:

$$y_t^2 = 1(\lambda_0^2 y_{t-1}^1 + \theta_0^2 a_{t-1}^1 + X^2\beta_0^2 + AGE_t^2\Gamma_0^2 + \alpha_0^2\varepsilon + \delta_0^2 v^2 + u_t^2 > 0),\ t = t_{\min}^2 \tag{8}$$

where $t_{\min}^2$ is the first year we observe the behavior of the younger sibling.[14] From now on, the subscript 0 will be used to refer to parameters in the initial conditions.

In the above equation, λ_0^2 is the sibling influence parameter in the initial condition ($t = t_{\min}^2$). We assume $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. The person specific error components $v^1 \sim N(0, \sigma_{v^1}^2)$ and $v^2 \sim N(0, \sigma_{v^2}^2)$ are independent across siblings. The error components $u_t^1 \sim N(0, 1)$ and $u_t^2 \sim N(0, 1)$ are independent across siblings and years. The coefficients on the age dummies and X differ between the initial conditions and the equations for the later periods. They also differ between the older and younger siblings.

We use two alternative specifications of the error structure. We refer to our baseline specification as error specification A. It restricts the factor loadings α on the family effect ϵ and the factor loadings δ on the individual effects v^1 and v^2 to equal 1 in all equations. Note, however, that we allow the variance of $v^1$and v^2 to differ. In error specification B, the family effect ε enters (7), (8), (2) and (3) with the factor loadings $1, \alpha^1$, α_0^2, and α^2

[14]The value of $t_{\min}^1$ varies from 1998 to 2000 in (7) while $a_{t_{\min}^1}^1$ ranges from 15 to 20. The value of $t_{\min}^2$ varies from 1999 to 2001 while $a_{t_{\min}^2}^2$ varies from 15 to 19.

respectively, v^1 enters (7) and (2) with the factor loadings 1 and δ^1 respectively, and v^2 enters (8) and (3) with the factor loadings 1 and δ^2 respectively. We restrict the variance of v_1 and v_2 to be the same across siblings.[15] For some outcomes, we have difficulty identifying the separate roles of family heterogeneity and individual heterogeneity when we use the less restricted version, and the estimates of the sibling influence parameters λ^2 and λ_0^2 tend to be noisier.

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 (7) and (2) and linear interactions between X^2 and a_t^2 in (8) and (3).[16] For the most part, the state dependence parameters and sibling effects parameters are not very sensitive to the addition of the interaction terms, and so we present the models without the interaction terms.[17] We estimate the models by maximum likelihood.[18]

5.3.2 A joint 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 also estimate a joint dynamic ordered probit model. Consider cigarettes. Let $y_{L,t}^1$ equal to 1 if the older sibling smoked between one and 7 days during the last month and let $y_{H,t}^1$ to be 1 if he or she smoked more than 7 days. The corresponding threshold values are 7 days for alcohol and 4 days for marijuana. The indicators are determined according to $y_{L,t}^1 = 1(q_L \leq y_t^{1*} < q_H)$ and $y_{H,t}^1 = 1(y_t^{1*} \geq q_H)$ where q_H and q_L are threshold parameters

[15]Note that we restrict the variance of the idiosyncratic error components to be 1 in both the initial condition and the later years for both the younger and older siblings equations. 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.

[16]One would expect age interactions to be particularly important in the initial condition equations.

[17]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.

[18]For computational ease, each pair coming from the same household is assumed to receive an independent draw of the common component ε. Thus 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 and 3 and Web Appendix Tables 3 and 4 are clustered at the household level.

and y_t^{1*} is the latent index given by:

$$y_t^{1*} = \gamma_H^1 y_{H,t-1}^1 + \gamma_L^1 y_{L,t-1}^1 + X^1\beta^1 + AGE_t^1\Gamma^1 + \alpha^1\varepsilon + \delta^1 v^1 + u_t^1, \quad t > t_{\min}^1.$$

The values of $y_{L,t}^1$ and $y_{H,t}^1$ for $t = t_{\min}$ are determined by an ordered probit model with the latent index equation

$$y_t^{1*} = X^1\beta_0^1 + AGE_t^1\Gamma_0^1 + \alpha_0^1\varepsilon + \delta_0^1 v^1 + u_t^1, \quad t = t_{\min}^1.$$

We expect $\gamma_H^1 > \gamma_L^1$, 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 $y_{L,t}^2 = 1(q_L \leq y_t^{2*} < q_H)$ and $y_{H,t}^2 = 1(y_t^{2*} \geq q_H)$, where

$$\begin{aligned} y_t^{2*} &= \lambda_{H,0}^2 y_{H,t-1}^1 + \lambda_{L,0}^2 y_{L,t-1}^1 + X^2\beta_0^2 + \theta_0^2 a_{t-1}^1 + AGE_t^2\Gamma_0^2 + \alpha_0^2\varepsilon + \delta_0^2 v^2 + u_t^2, \quad t = t_{\min}^2 \\ y_t^{2*} &= \gamma_H^2 y_{H,t-1}^2 + \gamma_L^2 y_{L,t-1}^2 + \lambda_H^2 y_{H,t-1}^1 + \lambda_L^2 y_{L,t-1}^1 + X^2\beta^2 + \theta^2 a_{t-1}^1 + AGE_t^2\Gamma^2 \\ &\quad + \alpha^2\varepsilon + \delta^2 v^2 + u_t^2, \quad t > t_{\min}^2. \end{aligned}$$

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 $\lambda_{H,0}^2 > \lambda_{L,0}^2 > 0$ and $\lambda_H^2 > \lambda_L^2 > 0$. In our main specification, we stop at three categories (zero, low, and high consumption) because of sample size considerations. However, in Section 8.2, we use a larger number of categories while restricting the state dependence parameters to lie on a piecewise linear spline. Error specifications A and B are the same as in the binary probit case.

6 Sibling effect estimates based on the CRE approach

Table 3 presents 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 presents estimates of our main specification, which we refer to as Model 1. Model 1 is a variant of (4) for the case in which y_t^2 is binary and the control variables X and age dummies are added:

$$y_t^2 = 1(\beta_0 + \beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + X^2\beta_3 + AGE_t^1\Gamma_1 + AGE_t^2\Gamma_2 + error > 0) \quad (9)$$

In the middle panel, we replace $\beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1$ in the equation above with $\beta_1(y_{t-2}^1 + y_{t-1}^1 + y_{t+1}^1 + y_{t+2}^1) + \beta_2(y_{t-2}^1 + y_{t-1}^1)$. We refer to this specification as 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, the peer influence coefficient on $(y_{t-1}^1 + y_t^1)$ in Model 3 is likely to be positively biased if transitory environmental factors are correlated across siblings. It may also be positively biased as the result of an interview effect in cases in which the interviews occur on the same day.[19]

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 .062 (.026). This estimate is statistically significant and is equal to 15.6% of the mean probability. The marginal effect of $(y_{t-1}^1 + y_{t+1}^1)$ is .085 (.018), 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

[19]The substance use questions are administered directly by the respondent into a computer, with computer turned away from the interviewer. Only the respondent can observe the responses. However, sometimes siblings are interviewed on the same day. To investigate whether this affects the response pattern, we estimated the linear probability model $y_{a,t}^2 = b_0 + b_1 y_{a+j,t}^1 + b_2 y_{a+j,t}^1 \times SAMEDAY_t + b_3 SAMEDAY_t + error_t$, where $SAMEDAY_t$ is 1 if the siblings were interviewed on the same day in year t and is 0 otherwise. The values of b_1 and b_2 are 0.192(0.019) and 0.085(0.022) for smoking, 0.266(0.018) and 0.013(0.023) for drinking, 0.163(0.018) and 0.058(0.024) for marijuana, 0.065(0.025) and 0.092(0.036) for hard drugs, and 0.050(0.020) and 0.008(0.030) for selling drugs. The results are mixed, but overall the evidence indicates reports are more strongly linked when the interview occurs on the same day. The estimates of b_3 are negative and significant in all cases.

[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 Web Appendix Table 3, 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.

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^1_{t-1} raises the probability of drinking by about .054 (.024), which is 9.1% of the mean probability. The link due to common influences is .118 (.018). 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 of the older sibling on the younger one.

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^1_{t-1} is .011 (.019) 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.[21]

7 Results for the joint dynamic probit model

7.1 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

[21] We also tried a fixed effect approach. Specifically, we estimated

$$y^2_t = \lambda^2 y^1_{t-1} + AGE^1_t \Gamma_1 + AGE^2_t \Gamma_2 + \varepsilon + v^2 + u^2_t, \tag{10}$$

treating $\varepsilon + v^2$ as a fixed effect. The advantage of the fixed effect estimator is that it requires assumptions (A1) and (A2), but not (A3). On the other hand, it requires (A5), while u^2_t and $u^1_{t'}$ 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. The estimates of the coefficient on y^1_{t-1} are .028 (.014) for smoking and .045 (.014) for drinking (see Web Appendix Table 4). Both coefficients are significant at the .05 level, but are smaller than the estimates based on (9). 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 (9), 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 (10) to match the sample for (9) and to using a linear probability specification for the CRE model in place of the probit specification. Keep in mind that in Table 3, we report marginal effects on the probability of substance use rather than probit coefficients.

smoking cigarettes. The estimates of the state dependence parameters are .947 (.068) for the younger sibling and .906 (.062) 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 .508 (.051) next year and by .040 (.010) two years out relative to the baseline.

The value of $\hat{\sigma}_\varepsilon$ is .746 (0.051). 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 1.034 and .837, respectively. 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_0^2 and λ^2, which are the coefficients on y_{t-1}^1 in the equations for y_t^2 in the initial condition and the subsequent periods respectively. A priori, we would expect both to be positive. We also would expect λ_0^2 to exceed λ^2 because we do not condition on y_{t-1}^2 in the initial condition. $\hat{\lambda}_0^2$ is .213 (0.102), which is significant at the 5% 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 $\hat{\lambda}^2$ for subsequent years is .054 (0.069), which is positive but not significant.

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 reported in Web Appendix Table 5 indicate that state dependence is indeed a bit weaker for drinking.

The sibling influence parameter λ_0^2 is .405 and is highly significant. The estimate of λ^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 expected less of a difference between λ_0^2 and λ^2. The results for the other error structures are basically similar.

Column 3 reports estimates for marijuana, which are very similar to the results for drinking. We find strong evidence for a sibling effect that operates primarily through the initial condition. However, the point estimate of λ^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 but they are not statistically significant. In the case of selling drugs (column 5), 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.[22]

Table 4b reports estimates using error specification B, which allows the factor loadings associated with ε, $v^1$and v^2 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. 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. The point estimates are positive for hard drug use and selling drugs but are not significant.

7.2 The dynamic response to the older sibling's substance use

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

[22]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.

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.[23]

Figure 1a presents the results for smoking using the model in the first column of Table 4a. The vertical axis measures the change in behavior relative to the baseline probability. The horizontal axis measures the time period relative to $(t^2_{\min} - 1)$, so 0 corresponds to $(t^2_{\min} - 1)$. Web Appendix Table 5 reports point estimates and standard errors, which are based on a parametric bootstrap method.[24]

We begin with the older sibling's response. The solid line in the 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_t, relative to the baseline average for y^1_t.[25] 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 one year later by .208, which is the fraction .508 of the baseline value (.409). The effect is 13.6% 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_t, relative to the baseline average of y^2_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 among older siblings increases smoking among

[23]In all but ten cases, $t^1_{\min} = (t^2_{\min} - 1)$, so we use the actual age of the older sibling in creating the age dummies AGE^1_t. In the 10 cases where $t^1_{\min} \neq (t^2_{\min} - 1)$, we set the age of the older sibling in year $t^1_{\min}$ to the actual age plus 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 specification A, we first simulate y^1_t from $(t^2_{\min} - 1)$ to $(t^2_{\min} + 4)$ using (7) and (2). 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 (8) and (3). 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 $y^1_{t^2_{\min}-1}$ set to 0 for all pairs rather than the value implied by (7) and a simulation with $y^1_{t^2_{\min}-1}$ set to 1 for all pairs. For each sibling pair i, we performed each of the three simulations 100 times. We then averaged over the 100 simulations for all the pairs.

[24]We draw 150 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 100 simulations and take the average, as described in the previous footnote. The standard errors are the standard deviations across the 150 averages. The 90% confidence bands are computed from the point estimate and standard error estimates under a normality assumption.

[25]To be more specific, for each older sibling, we first set y^1_{t-1} to 1 in $(t^2_{\min} - 1)$, simulate forward and take the average of y^1_t for the values of t reported at the top of each column of Web Appendix Table 5 and on the horizontal axis of Figure 1a. We repeat the procedure with y^1_{t-1} set to 0 in $(t^2_{\min} - 1)$, take the difference in the two averages for each value of t, and then divide by baseline value in the top row of Web Appendix Table 5.

younger siblings in $t^2_{\min}$ by .141 (.066) of the baseline value. This is 28 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 .036 (.017) in the second period. The effect on the probability that the younger sibling smokes relative to baseline is essentially zero after three years.[26]

Figure 1b displays simulations for drinking. For the older sibling, drinking last year period raises the probability of drinking this year by .302 (.031) of the baseline value, which is about .569. After three periods, the effect is only .010 (.003) 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 .243 (.058) of the baseline probability (.504). The effect on the younger sibling is essentially zero after three periods.

Figure 1c shows that marijuana use by the older sibling in $(t^2_{\min}-1)$ increases the probability that the older sibling uses marijuana one year later by .642 (.071) of the baseline probability of .248. The effect on the older sibling's behavior is .032 (.009) three 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 .254 of the baseline value. The effect on the younger sibling is under .011 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 Web Appendix Table 6). However, the effect of smoking by the older sibling on the younger sibling is essentially zero, although the standard error is large.[27]

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

[26]In Web Appendix Table 5, we report the baseline simulation for y^2_t. In the rows for the younger sibling labelled *W/ Feedback* we report the path of the difference in y^2_t relative to the baseline simulation for younger siblings when y^1_{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_{t-1} is allowed to affect future values of y^1_{t-1} in accordance with the model. The effect of the shift on y^2_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 sibling's behavior, which are reported in the rows *W/out feedback* and graphed in Figure 1.

[27]We would not make too much of this. It reflects the fact that in the case of smoking, $\hat{\lambda}^2_0$ fall to essentially 0 under error specification B. This decline is offset by an increase in the sibling effect in later periods, $\hat{\lambda}^2$, but $\hat{\lambda}^2$ does not matter for the sibling response to $y^1_{t^2_{\min}-1}$.

use in all periods, we find only modest effects on the behavior of the younger sibling for $t > t^2_{\min} + 2$ (not reported).

7.3 The relative contribution of sibling effects and common influences to sibling correlations in substance use

The fact that our estimates imply that younger siblings' behavior is relatively insensitive to whether or not the older sibling consumes the substance in all periods suggests that only a small part of the large sibling links in substance use reported in Table 2 is causal. To quantify this, we simulated data from our model using the parameter estimates corresponding to error specification A and estimated the parameter ρ in the descriptive regression (1) relating the behavior of the younger sibling at age a to the behavior of the older sibling at the same age. Next, we performed a similar simulation, this time setting the peer influence parameters λ^2_0 and λ^2 to zero and all the other parameters to their estimated values.

Results of this exercise are reported in Table 5. For ease of comparison, we present $\hat{\rho}$ based on the actual data in column 1. Columns 2 and 3 report estimates of ρ based on data from a simulation in which the peer influence parameters are set to their estimated values and to zero, respectively. Column 4 reports the difference between columns 2 and 3 divided by column 2. This is the fraction of the sibling link ρ that is due to peer influence. In the case of smoking, the point estimate is 0.083 (0.070). The corresponding fractions are 0.046 (.036) for alcohol and -0.010 (.087) for marijuana. The relatively large standard error estimates reflect the difficulty of estimating a ratio, particularly when ρ is small.[28] We obtain similar estimates using error specification B.

7.4 Robustness checks

7.4.1 Bias if the younger sibling influences the older sibling

If the younger sibling positively influences the behavior of the older sibling, then we are likely to underestimate the sibling effect. To see why, first consider the static CRE model

[28]The corresponding fractions are 0.129 (0.238) for using hard drugs, and .014 (0.295) for selling drugs. We focus on the results that exclude controls from (1) since the correlation among the observed characteristics of siblings is part of the common influence in sibling behavior. The bottom panel of the table reports results with controls included. The part of $\hat{\rho}$ due to peer effects (column 2 - column 3) is similar to the values in the top panel, but this difference is a larger fraction of the value of column 2.

and reparameterize equation (4) 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 this will lead to underestimate the direct sibling influence in econometric models that assume that the sibling influence goes in only one direction. Simulations support this intuition.[29]

7.4.2 Bias from treatment of the initial conditions

The fact we typically find a stronger sibling effect in the initial condition than in the equation for subsequent periods could reflect the fact that λ_0^2 captures influence over more than one 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_0^2 and biases downward λ^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_0^2 set to 0.213 and λ^2 set to 0.054. The estimates of λ_0^2 and λ^2 using the simulated data are 0.163 (0.109) and 0.057 (0.073) (table omitted). These results suggest that there is little bias in λ^2 and that, if anything, we are underestimating λ_0^2.[30]

[29] 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 of 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_0^2 and λ^2 in the dynamic probit model presented above. As expected, the estimates of λ_0^2 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^2 using the CRE specification(4). 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, thus confirming our conjecture.

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

7.4.3 Robustness to allowing for gateway drugs

There is considerable policy interest in the idea that cigarettes may be a gateway to marijuana, alcohol a gateway to marijuana, marijuana a gateway to hard drugs, and so on. Policies to control marijuana are justified in part by a concern that it leads to hard drug use. The idea of gateway drugs seems plausible, although the patterns of causal influence substances are not well established.[31]

We experimented with extended versions of the joint dynamic probit model that jointly model pairs of drugs. Our motivation was primarily to check on the robustness of our findings rather than to test the gateway drug hypothesis. Consider the case in which cigarettes are a gateway drug for marijuana. The model consists of equations for cigarette smoking that have the same form as the model above. After the initial period, marijuana use depends on lagged cigarette use as well as lagged marijuana use. An implication of the model is that cigarette smoking by the older sibling can influence marijuana use by the younger sibling through its effect on marijuana use by the older sibling. We estimated models with smoking as the gateway drug for marijuana, drinking for marijuana, smoking for hard drugs, drinking for hard drugs, and marijuana for hard drugs. We worked with two different error specifications.

The results are exploratory but may be summarized as follows. First, the joint models indicate state dependence for both the gateway substance and the paired substance that is fully consistent with the results when we examine each substance in isolation. Second, family heterogeneity and individual heterogeneity are important and contribute to the correlation in substance use across siblings and across substances for each sibling. Third, the effect of the gateway drug on future consumption of the paired drug is never significantly positive, although we could not rule out modest positive effects. Fourth, and most importantly for present purposes, the estimates of the sibling influence parameters are generally a bit larger than those obtained when we model the substances separately. We present the model in Web Appendix C and model estimates in Web Appendix Tables 7 and 8.

[31] See Deza (2011) for references and evidence of a modest causal effect of prior use of consumption of soft drugs on consumption of harder drugs using methods somewhat similar to ours to account for unobserved heterogeneity. She does not model siblings' behavior.

8 Results for the joint ordered probit model

8.1 The three consumption category case

We now turn to the estimates of the joint dynamic ordered probit models using error specification A. These are reported in Table 6.[32] We limit the analysis to smoking cigarettes, drinking alcohol, and smoking marijuana because these behaviors are more common in the sample and the quantities are more varied. The estimates of the sibling influence parameters in the initial conditions, $\lambda^2_{H,0}$ and $\lambda^2_{L,0}$, are both positive for all three outcomes. In the case of cigarettes and drinking, $\lambda^2_{H,0}$ is larger than $\lambda^2_{L,0}$, 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_H and λ^2_L, the sibling influence parameters for the periods $t > t^2_{\min}$, are small and not always positive. In the case of alcohol, λ^2_H 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.[33] 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 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 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 2b graph simulations based on 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 based on the dynamic ordered probit model with error specifications A and B are in Web Appendix Table 10 and 11, respectively). One period later, the shift raises the low consumption probability for

[32] The results based upon error specification B are similar. See Web Appendix Table 9 for parameter estimates and Web Appendix Table 11 for corresponding simulations.

[33] 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 cut-off 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 cut-off between y_L and y_H.

the older sibling by .262 and the high consumption probability by .551 relative to the baseline averages. The effects become very close to 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 .392 relative to baseline and also boosts the probability of low consumption. The effects are very small after two periods. In the case of alcohol (Figures 2c and d) and marijuana (Figures 2e and f), the dynamic effect on the behavior of the younger sibling is similar to smoking, but slightly smaller.

8.2 A joint dynamic ordered probit model with many categories

In this section, we propose a simple way to allow for flexible forms of nonlinearity in the dynamic behavior of substance use and sibling influence while continuing to allow for sibling pair effects and individual effects. Expanding to an arbitrary number of categories, the equations for the latent variables y_t^{1*} and y_t^{2*} become

$$\begin{aligned} y_t^{1*} &= \sum_{m=2}^{M} \gamma_m^1 y_{m,t-1}^1 + X^1\beta^1 + AGE_t^1\Gamma^1 + \alpha^1\epsilon + \delta^1 v^1 + u_t^1, \qquad t > t_{\min}^1 \\ y_t^{2*} &= \sum_{m=2}^{M} \gamma_m^2 y_{m,t-1}^2 + \sum_{m=2}^{M} \lambda_m^2 y_{m,t-1}^1 + X^2\beta^2 + \theta^2 a_{t-1}^1 + AGE_t^2\Gamma^2 + \alpha^2\epsilon + \delta^2 v^2 + u_t^2, \\ & \qquad t > t_{\min}^2 \end{aligned}$$

with thresholds $q_1, q_2, ..., q_{M-1}$, where M is the total number of categories and consumption of zero corresponds to category $m = 1$.[34]

The equations for the initial conditions y_t^{2*} and y_t^{1*} at $t = t_{min}^1$ and $t = t_{min}^2$, respectively, take the same form but exclude own lags. By including a large number of groups, one can accommodate an arbitrary nonlinear relationship between substance use today and own past substance use as well as past substance use by the older sibling. Of course, with a large number of categories, freely estimating the thresholds q, as well as the γ and λ parameters would be hopeless without a very large sample. However, one can restrict these parameters to lie on a flexible but relatively parsimonious function, such as a linear spline with a number of break points less than M.

We have estimated models for days of use of cigarettes, alcohol, and marijuana last

[34] $y_{1,t}^1$ equals 1 if $y_t^{1*} < q_1$ and 0 otherwise, $y_{m,t}^1$ equals 1 if $q_{m-1} \leq y_t^{1*} < q_m$ and is 0 otherwise for $m = 2, \ldots, M-1$, and $y_{M,t}^1$ equals 1 if $q_{M-1}^1 \leq y_t^{1*}$ and is 0 otherwise. y_t^2 is determined by y_t^{2*} in a similar fashion.

month using 5 categories (M=5). For all three substances, the categories are 0, 1-7, 8-14, 15-22 and 23-30 days of use. We restrict the γ and λ parameters to lie on a linear spline with two breakpoints in slope but leave the thresholds q_m $(m = 1, \ldots, M-1)$ free. The changes in slope occur at 8 and 18 days. We provide more detail about the model in Web Appendix D.[35]

The parameter estimates in Web Appendix Table 12. These estimates are used to construct Table 7, which presents the implied values of the state dependence parameters and sibling influence parameters for $m = 2$ to 5. The state dependence parameters are all positive, and they are increasing in m in all cases, with the exception that $\hat{\gamma}_5^2$ is less than $\hat{\gamma}_4^2$ for the younger sibling for drinking. This could easily be due to sampling error. We also find that the sibling influence parameters $\lambda_{m,0}^2$ are all positive in the initial condition equations for all three substances. The size of $\lambda_{m,0}^2$ increase in m through category 4 in the cases of smoking and drinking and through category 3 in the case of marijuana, but the effects are smaller for the highest category for all three outcomes. However, the point estimates are relatively imprecise. The sibling effect in later periods is smaller, and the point estimates are negative in several cases.

Web Appendix Figure 1 reports the effect of exogenously shifting the older sibling's substance use in $t_{\min}^2 - 1$ from 0 to one of the two highest categories (with equal probability) on substance use in subsequent periods, relative to the baseline probabilities. The effects on relative probabilities are larger for the high consumption categories, which makes sense. The effects decay to essentially 0 after 3 periods.

Overall, the results for the five-category ordered probit model correspond fairly closely to those for the three-category model. The size of the sibling effect increases with the level of consumption.

[35] M must be less than or equal to the number of categories distinguished in the data and could be as large as 31 in our case. We chose 5 because of sample size limitations and the likelihood of diminishing returns. We did not experiment with this number. As is common in surveys, responses tend to cluster at 5, 10, 15, 20, etc. If one were to use a large M, one would want to extend the model to account for this, because the tendency to cluster would account for a larger fraction of the difference in the response probabilities across clusters. One could also restrict the threshold parameters q_m $(m = 1, \ldots, M-1)$ to lie on a flexible function, but we did not need to do so with $M = 5$.

9 Determinants of the strength of the sibling effect

Next we examine a number of possible determinants of the sibling effect, including the gender mix, the age gap, and a variety 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.[36] In Web Appendix Table 14, we report estimates for a specification that replaces $\lambda_0^2 y_{t-1}^1$ in (8) and $\lambda^2 y_{t-1}^1$ in (3) respectively with:

$$[\lambda_{mm,0}^2(M^2 \times M^1) + \lambda_{ff,0}^2(F^2 \times F^1) + \lambda_{mf,0}^2(M^2 \times F^1 + M^1 \times F^2)]y_{t-1}^1$$

and

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

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.[37]

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 $\lambda_0^2 y_{t-1}^1$ and $\lambda^2 y_{t-1}^1$ in (8) and (3) with $[\lambda_0^2 + \lambda_{2+,0}^2 1(a_t^1 - a_t^2 > 2)]y_{t-1}^1$ and $[\lambda^2 + \lambda_{2+}^2 1(a_t^1 - a_t^2 > 2)]y_{t-1}^1$, 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

[36] 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 NLSY97 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.

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

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.[38] The point estimates of $\lambda^2_{2+,0}$ and λ^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 Web Appendix Table 15). We simply do not have enough data to draw strong conclusions about how the age gap between siblings influences the sibling effect.[39]

9.2 Family process interactions

In view of the child psychology literature's emphasis on the importance of family process variables for child outcomes, one might expect the nature of the child's relationship with his parents and siblings to affect the size of the peer effect. We investigate this issue in the NLSY97 with data about the child's relationship to family members. In particular, we use measures of parental supportiveness, parental monitoring, and parenting style (uninvolved, permissive, authoritarian, or authoritative). We also use a variable indicating whether a sibling is the first person the youth turns to for advice and another one indicating whether the youth turns to someone other than the parents for advice.[40] We incorporated these 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).[41] One might expect the sibling effect to be larger for adolescents who get advice from siblings. For smoking and drinking, the marginal effects of the interaction of whether the youth gets

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

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

[40] Details about the construction of these measures for the analysis are available in Web Appendix A.

[41] In keeping with discussion in the literature, we expected negative effects of parents being more supportive and involved through authority and monitoring. We did not have a clear prior about the sign of the main effect of turning to a sibling for advice.

advice from a sibling with the older sibling's lagged behavior are .061 (.079) and .182 (.076), respectively. However, it is -.029 (.060) 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 .064 (.055) in the case of smoking, by .068 (.051) in the case of drinking, and by .046 (.045) in the case of 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 for children who do not live with both biological parents. However, 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. We start with a correlated random effects (CRE) design, 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. The CRE approach has the advantage of simplicity but rules out 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 rely primarily on a joint dynamic probit model and a joint dynamic ordered probit model that allow for state dependence and nonstationarity. A secondary contribution of the paper is to propose the use of single equation or multiple equation dynamic ordered response models with large numbers of categories but restrictions on the category specific model parameters as a way to allow for nonlinear state dependence in the presence of unobserved heterogeneity.

The CRE results indicate that smoking, drinking, and, more tentatively, marijuana use by the older sibling increase the probability that the younger sibling engages in these behaviors. The 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. Simulations using the dynamic probit model indicate that only a small fraction of the large sibling correlation in substance use is causal, although the estimates of the fractions are noisy.

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 some previous studies of sibling links in risky behavior, is a natural possibility. The availability of genetic markers that influence substance use could be incorporated, building on some of the work discussed in Fletcher and Lehrer (2011). However, the time between interviews makes Add Health less than ideal. Second, other behaviors, including positive behaviors such as volunteering and study time, could be examined. Third, as evidence accumulates on the dynamic interrelationship among the use of different substances, it would be desirable to revisit our analysis of sibling effects in models with multiple substances. The question of "gateway" drugs is salient in policy discussions of drug law reform, but given limited information in the data it is 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. Statistical power is a problem, at least in the NLSY97, but 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.

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.

Cawley, J. and C. Ruhm (2011). Chapter three - the economics of risky health behaviors. In T. M. Mark Pauly and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, pp. 95 – 199. Amsterdam, Holland: Elsevier.

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.

Fletcher, J. and S. Lehrer (2011). Genetic lotteries within families. *Journal of Health Economics 30*(4), 647–659.

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 parentsï¿œ current and former smoking, and family structure. *Journal of Behavioral Medicine 30*(2), 143 – 154.

Ouyang, L. (2004). 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.

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.

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.

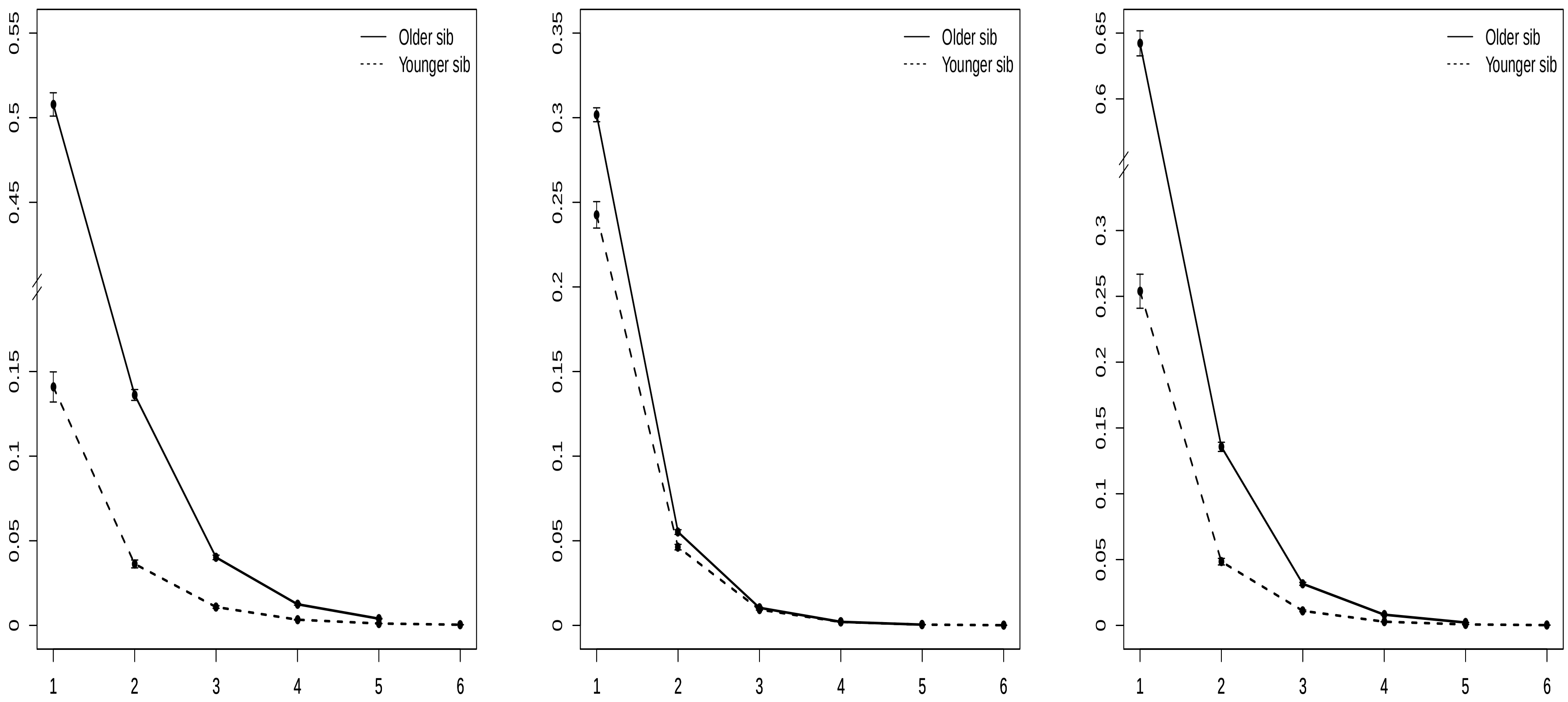
Widmer, E. (1997). Influence of older siblings on initiation of sexual intercourse. *Journal of Marriage and the Family 59*(4), 928 – 938.

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

FIGURE 1

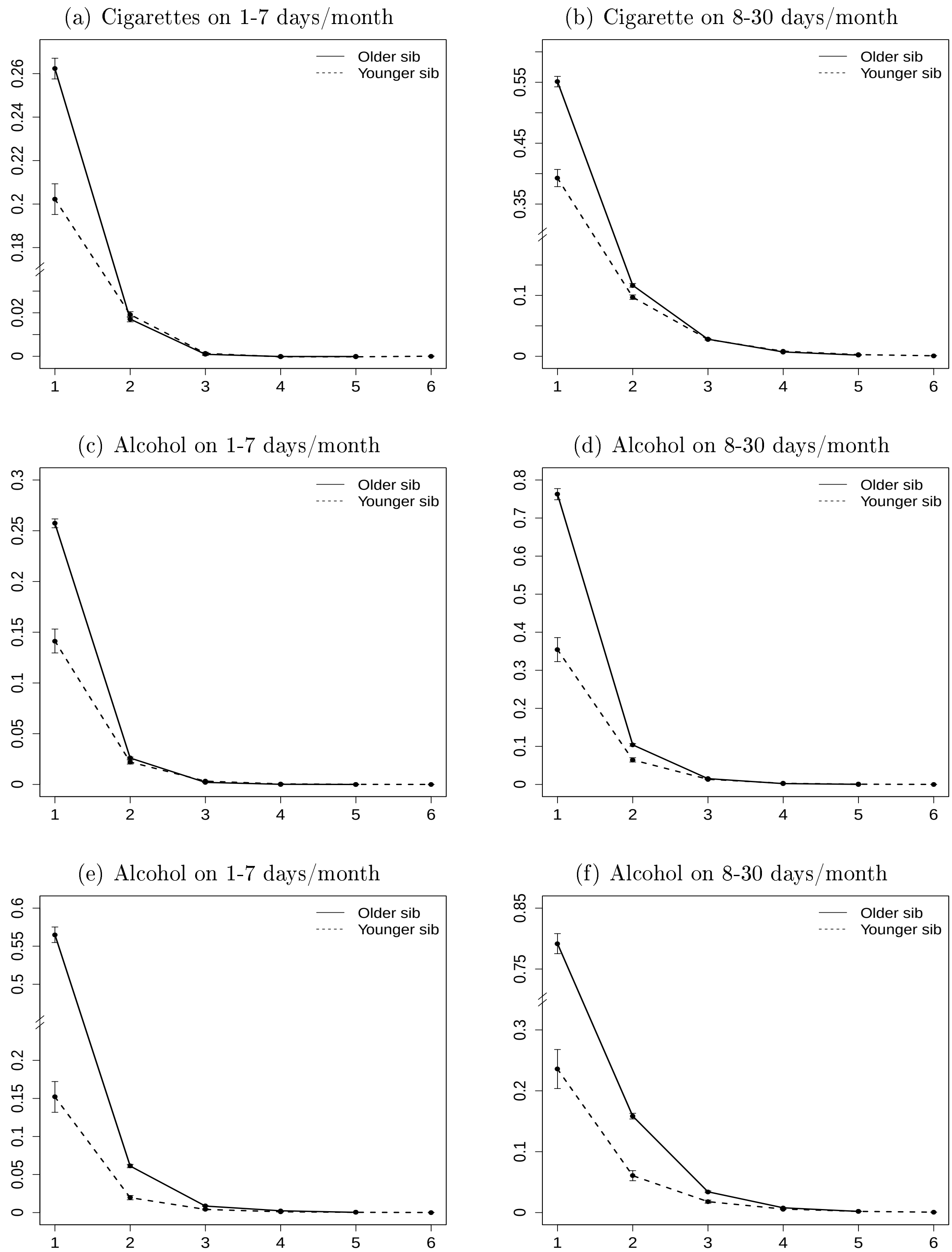
Effect of shifting the older sibling's probability of substance use from zero to one on the older and younger siblings' probabilities of behavior, relative to baseline (based on dynamic probit model, error A)

(a) Any cigarettes (b) Any alcohol (c) Any marijuana



Note: The solid line and the broken line represent the effects on the probabilities of behavior, relative to baseline, of the older sibling and of the young 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: 0.3992(0.0114) and 0.4266(0.0106) for the older sibling, and 0.3376 (0.0119) and 0.3870 (0.0314) for the younger sibling. For drinking, they are 0.5559 (0.0115) and 0.7237 (0.0108) for the older sibling and 0.4610 (0.0120) and 0.5867 (0.0300) for the younger sibling. For smoking marijuana, they are 0.2529 (0.0106) and 0.1972 (0.0084) for the older sibling and 0.2087 (0.0097) and 0.2781 (0.0360) for the younger sibling.

FIGURE 2
Effect of shifting the older sibling's probability of substance use from zero to the highest category on the older and younger siblings' probabilities of consumption relative to baseline (based on ordered probit model, 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 of the young 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 being in the low smoking category in the first and last period displayed on the graphs for the older and younger siblings, respectively are: 0.0947(0.0044), 0.0902(0.0041), 0.0956(0.0040), 0.0916(0.0058). For being in the highs smoking category, they are: 0.2096(0.0091), 0.2910(0.0096), 0.1714(0.0083), 0.2839(0.0380). For being in the low drinking category, they are: 0.3174(0.0089), 0.4068(0.0083), 0.2984(0.0078), 0.3906(0.0140). For being in the high drinking category, they are: 0.0713(0.0053), 0.2135(0.0078), 0.0568(0.0045), 0.1916(0.0304). For being in the low marijuana category, they are: 0.0724(0.0045), 0.0673(0.0036), 0.0735(0.0041), 0.0805(0.0118). For being in the high marijuana category, they are: 0.0761(0.0057), 0.0872(0.0064), 0.0685(0.0050), 0.1247(0.0432).

TABLE 1
Sample Means for Substance Use Behaviors

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Drugs
	A - Probability of engaging in behavior last year				
Full sample	0.398	0.596	0.220	0.060	0.054
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Males	0.418	0.613	0.251	0.064	0.074
	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)
Females	0.377	0.579	0.187	0.055	0.033
	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)
	B - Probability of engaging in behavior between age 15 and 20				
Full sample	0.387	0.561	0.441	0.147	0.162
	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)
Males	0.407	0.810	0.481	0.159	0.219
	(0.006)	(0.009)	(0.012)	(0.009)	(0.010)
Females	0.366	0.779	0.399	0.134	0.100
	(0.006)	(0.010)	(0.012)	(0.009)	(0.008)
	C - Number of days engaged in behavior in the last month				
Full Sample	6.59	2.98	1.78	-	-
	(0.080)	(0.037)	(0.041)		
Males	6.95	3.64	2.37	-	-
	(0.113)	(0.059)	(0.067)		
Females	6.21	2.29	1.16	-	-
	(0.112)	(0.043)	(0.046)		
	D - Distribution of sample in each consumption category (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 21,293 to 21,460 for full sample, from 10,893 to 10,998 for males, and from 10,397 to 10,462 for females. Sample sizes in Panel B vary from 3,297 to 3,300 for full panel, 1,712 to 1,715 for males, and equal 1,585 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	0.226***	0.259***	0.157***	0.092***	0.049***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)
Full set of controls	0.173***	0.173***	0.135***	0.076***	0.035***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)
	Panel B: Brothers				
No controls	0.215***	0.234***	0.151***	0.070*	0.076***
	(0.032)	(0.031)	(0.030)	(0.038)	(0.029)
Full set of controls	0.168***	0.157***	0.133***	0.048	0.058**
	(0.031)	(0.033)	(0.030)	(0.033)	(0.026)
	Panel C: Sisters				
No controls	0.336***	0.325***	0.228***	0.213***	0.006
	(0.038)	(0.030)	(0.035)	(0.073)	(0.024)
Full set of controls	0.226***	0.203***	0.177***	0.164**	-0.007
	(0.038)	(0.033)	(0.032)	(0.065)	(0.022)

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 7,122 to 7,213 for full sample, 1,869 to 1,929 for brothers, 1,752 to 1,769 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		
$y_{t-1}^1 \quad (\beta_2)$	0.062**	0.054**	0.008	0.011	0.005
	(0.026)	(0.024)	(0.022)	(0.019)	(0.018)
$y_{t-1}^1 + y_{t+1}^1 \quad (\beta_1)$	0.085***	0.118***	0.089***	0.034**	0.016
	(0.018)	(0.018)	(0.015)	(0.014)	(0.013)
			Model 2		
$y_{t-2}^1 + y_{t-1}^1 \quad (\beta_2)$	0.092***	0.022	0.031	0.024	0.041**
	(0.027)	(0.024)	(0.021)	(0.018)	(0.017)
$y_{t-2}^1 + y_{t-1}^1 + y_{t+1}^1 + y_{t+2}^1 \quad (\beta_1)$	0.024	0.080***	0.041***	0.017	-0.013
	(0.015)	(0.016)	(0.013)	(0.013)	(0.013)
			Model 3		
$y_{t-1}^1 + y_t^1 \quad (\beta_2)$	0.096***	0.046**	0.052***	0.025*	0.013
	(0.024)	(0.021)	(0.018)	(0.015)	(0.015)
$y_{t-1}^1 + y_t^1 + y_{t+1}^1 + y_{t+2}^1 \quad (\beta_1)$	0.020	0.070***	0.031***	0.013	0.003
	(0.014)	(0.014)	(0.011)	(0.011)	(0.011)

Note: The table reports marginal effects 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 5,383 and 6,975 for Model 1, 2,507 and 3,916 for Model 2, and 3,841 and 5,309 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 parameters					
Old sibling (γ^1)	0.906 ***	0.632 ***	0.688 ***	0.501 ***	0.575 ***
	(0.062)	(0.054)	(0.061)	(0.132)	(0.131)
Young sibling (γ^2)	0.947 ***	0.668 ***	0.737 ***	0.735 ***	0.510 ***
	(0.068)	(0.056)	(0.065)	(0.144)	(0.128)
Sibling's influence parameters					
Initial condition (λ_0^2)	0.213 **	0.405 ***	0.290 ***	0.329	0.083
	(0.102)	(0.086)	(0.107)	(0.284)	(0.224)
Later periods (λ^2)	0.054	0.007	-0.052	0.043	-0.006
	(0.069)	(0.057)	(0.067)	(0.216)	(0.177)
Standard deviation of error term specific to:					
Family (σ_ε)	0.746 ***	0.615 ***	0.627 ***	0.559 ***	0.484 ***
	(0.051)	(0.033)	(0.042)	(0.090)	(0.087)
Older sibling (σ_{v^1})	1.034 ***	0.599 ***	0.677 ***	0.767 ***	0.740 ***
	(0.079)	(0.061)	(0.067)	(0.131)	(0.104)
Younger sibling (σ_{v^2})	0.837 ***	0.617 ***	0.706 ***	0.843 ***	0.734 ***
	(0.081)	(0.063)	(0.069)	(0.152)	(0.117)
Log likelihood value	-7494.55	-8373.25	-6935.76	-2622.23	-3059.78

Note: The table reports probit model parameters rather than marginal effects. Standard errors in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary from 1,286 to 1,661 for the older siblings' models and from 1,079 to 1,661 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. All factor loadings are normalized to 1 in this specification.

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

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Drugs
State dependence parameters					
Older sibling (γ^1)	0.833 ***	0.636 ***	0.632 ***	0.418 ***	0.540 ***
	(0.067)	(0.056)	(0.066)	(0.151)	(0.134)
Younger sibling (γ^2)	0.911 ***	0.612 ***	0.713 ***	0.724 ***	0.473 ***
	(0.072)	(0.059)	(0.067)	(0.152)	(0.137)
Sibling's influence parameters					
Initial condition (λ_0^2)	0.008	0.273 **	0.238 *	0.231	-1.555
	(0.178)	(0.129)	(0.143)	(0.432)	(2.985)
Later periods (λ^2)	0.120	0.037	-0.035	0.092	0.104
	(0.077)	(0.062)	(0.074)	(0.224)	(0.177)
Family-specific error term					
Standard deviation (σ_ε)	0.823 ***	0.760 ***	0.530 ***	1.404	1.403
	(0.124)	(0.108)	(0.102)	(1.122)	(2.217)
Factor loadings:					
Older sibling, later periods (α^1)	1.110 ***	0.865 ***	1.254 ***	0.755	0.294
	(0.188)	(0.145)	(0.242)	(0.658)	(0.277)
Younger sib, initial period (α_0^2)	1.038 ***	0.808 ***	1.295 ***	0.271	1.182
	(0.285)	(0.186)	(0.351)	(0.219)	(4.760)
Younger sib, later periods (α^2)	0.753 ***	0.754 ***	1.214 ***	0.205	0.241
	(0.174)	(0.156)	(0.346)	(0.188)	(0.520)
Individual-specific error term					
Standard deviation (σ_v)	0.754 ***	0.455 ***	0.552 ***	0.945 ***	0.636 *
	(0.071)	(0.066)	(0.076)	(0.192)	(0.351)
Factor loadings:					
Older sib, later periods (δ^1)	1.460 ***	1.198 ***	1.406 ***	-0.163	1.314 **
	(0.177)	(0.269)	(0.260)	(0.609)	(0.559)
Younger sib, later periods (δ^2)	1.343 ***	1.726 ***	1.363 ***	1.039 ***	1.399
	(0.158)	(0.274)	(0.256)	(0.268)	(0.862)
Log likelihood value	-7486.50	-8367.19	-6931.05	-2620.88	-3066.55

Note: See Table 4a. In this specification, the factor loadings $\alpha_0^1, \delta_0^1, \delta_0^2$ are normalized to 1.

TABLE 5

Proportion of the Correlation Between Siblings' Behavior Explained by the Direct Effect (Based on the Joint Dynamic Probit Model with Error A)

	Estimates of the effect of old sibling's behavior on younger sibling's behavior at the same age			Proportion of the coefficient due to sibling effect
	NLSY 97 dataset	Simulated data all estimates	Simulated data zero sibling effect	
	(1)	(2)	(3)	(4)
	No controls			
Smoking cigarettes	0.226***	0.208	0.191	0.083
	(0.011)	(0.012)	(0.018)	(0.070)
Drinking alcohol	0.259***	0.257	0.245	0.046
	(0.012)	(0.011)	(0.013)	(0.036)
Smoking marijuana	0.157***	0.146	0.148	-0.010
	(0.011)	(0.015)	(0.018)	(0.087)
Using hard drugs	0.092***	0.076	0.066	0.129
	(0.012)	(0.020)	(0.022)	(0.238)
Selling drugs	0.049***	0.051	0.051	0.014
	(0.011)	(0.017)	(0.016)	(0.295)
	With controls			
Smoking cigarettes	0.173***	0.170	0.153	0.102
	(0.011)	(0.013)	(0.018)	(0.084)
Drinking alcohol	0.173***	0.171	0.156	0.092
	(0.012)	(0.011)	(0.013)	(0.054)
Smoking marijuana	0.135***	0.131	0.132	-0.009
	(0.011)	(0.015)	(0.019)	(0.096)
Using hard drugs	0.076***	0.066	0.056	0.149
	(0.012)	(0.020)	(0.022)	(0.270)
Selling drugs	0.035***	0.042	0.041	0.017
	(0.011)	(0.017)	(0.016)	(0.428)

Note: In column 1, standard errors are clustered at the household level. Parametric boostrap standard errors in columns (2), (3), and (4) based on 150 replications. The controls consist of male, 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. The numbers in the fourth column are obtained by taking the difference between numbers in column (2) and column (3) and dividing by column (2).

TABLE 6

Estimates of Joint Dynamic Ordered Probit Model (Error A)

		Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence parameters:				
Older sibling	Low consumption (γ_L^1)	0.345 ***	0.339 ***	0.334 ***
		(0.077)	(0.043)	(0.079)
	High consumption (γ_H^1)	0.797 ***	0.577 ***	0.686 ***
		(0.064)	(0.062)	(0.076)
Younger sibling	Low consumption (γ_L^2)	0.588 ***	0.544 ***	0.619 ***
		(0.077)	(0.045)	(0.080)
	High consumption (γ_H^2)	1.022 ***	0.855 ***	1.049 ***
		(0.070)	(0.068)	(0.087)
Sibling's influence parameters:				
Initial condition	Low consumption ($\lambda_{L,0}^2$)	0.108	0.072	0.324 *
		(0.181)	(0.080)	(0.177)
	High consumption ($\lambda_{H,0}^2$)	0.441 ***	0.206*	0.155
		(0.123)	(0.130)	(0.166)
Later periods	Low consumption (λ_L^2)	0.057	-0.016	0.069
		(0.089)	(0.044)	(0.089)
	High consumption (λ_H^2)	0.078	-0.132 **	-0.009
		(0.078)	(0.066)	(0.096)
Standard deviation of error term specific to:				
Family (σ_ε)		0.775 ***	0.538 ***	0.645 ***
		(0.057)	(0.026)	(0.051)
Older sibling (σ_{v^1})		1.247 ***	0.527 ***	0.867 ***
		(0.075)	(0.041)	(0.071)
Younger sibling (σ_{v^2})		0.894 ***	0.467 ***	0.598 ***
		(0.076)	(0.045)	(0.081)
Thresholds:				
Low consumption (q_L)		-0.107	-0.373	0.888 *
		(0.472)	(0.354)	(0.516)
High consumption (q_H)		0.474	1.240 ***	1.505 ***
		(0.471)	(0.355)	(0.517)
Log likelihood value		-9549.89	-13337.27	-7374.18

Note: Standard errors in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary from 1,278 to 1,645 for the older siblings' models and from 1,054 to 1,645 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. All factor loadings are normalized to 1 in this specification. In this specification, all factor loadings are normalized to 1.

TABLE 7

State Dependence and Sibling Effect Parameters Implied from the Estimates of the Joint Dynamic Ordered Probit Model with Five Consumption Categories

		Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana
State dependence parameters:				
Older sibling	1-7 days (γ_2^1)	0.301	0.329	0.322
		(0.029)	(0.016)	(0.026)
	8-14 days (γ_3^1)	0.500	0.484	0.614
		(0.046)	(0.028)	(0.048)
	15-22 days (γ_4^1)	0.672	0.596	0.822
		(0.058)	(0.039)	(0.065)
	23-30 days (γ_5^1)	0.808	0.676	0.910
		(0.066)	(0.048)	(0.078)
Younger sibling	1-7 days (γ_2^2)	0.532	0.532	0.595
		(0.029)	(0.016)	(0.029)
	8-14 days (γ_3^2)	0.728	0.812	0.842
		(0.048)	(0.030)	(0.054)
	15-22 days (γ_4^2)	0.900	0.852	0.982
		(0.062)	(0.043)	(0.071)
	23-30 days (γ_5^2)	1.084	0.660	1.046
		(0.070)	(0.055)	(0.083)
Sibling's influence parameters:				
Initial condition	1-7 days ($\lambda_{2,0}^2$)	0.098	0.077	0.273
		(0.066)	(0.029)	(0.061)
	8-14 days ($\lambda_{3,0}^2$)	0.376	0.178	0.522
		(0.119)	(0.062)	(0.118)
	15-22 days ($\lambda_{4,0}^2$)	0.496	0.218	0.438
		(0.154)	(0.113)	(0.161)
	23-30 days ($\lambda_{5,0}^2$)	0.384	0.178	-0.010
		(0.176)	(0.164)	(0.194)
Later periods	1-7 days (λ_2^2)	0.084	-0.021	0.084
		(0.032)	(0.016)	(0.032)
	8-14 days (λ_3^2)	0.066	-0.102	-0.114
		(0.056)	(0.028)	(0.060)
	15-22 days (λ_4^2)	0.058	-0.134	-0.182
		(0.072)	(0.041)	(0.080)
	23-30 days (λ_5^2)	0.082	-0.094	-0.038
		(0.082)	(0.053)	(0.095)

Note: Coefficients and standard errors are calculated based on the estimates of the joint dynamic ordered probit model with five consumption categories (error specification A) reported in Web Appendix Table 12.

APPENDIX TABLE 1

Age Distribution by Birth Order

Age	Younger sibling	Older siblings
15	36.42	
16	31.64	6.48
17	23.76	21.58
18	7.7	34.3
19	0.48	34.85
20		2.18
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 smoking probit model.

APPENDIX TABLE 2

Summary Statistics of Sample Characteristics

Variable	Unweighted sample	Weighted sample
Male	0.522	0.522
Black	0.243	0.144
Hispanic	0.237	0.143
Lived with biological parents at 12	0.522	0.573
House broken in by age 12	0.159	0.152
Witness of gun shooting by age 12	0.104	0.086
Victim of bullying by age 12	0.169	0.176
Highest grade completed by 19	11.780	11.895
	(0.030)	(0.030)
AFQT percentile score	41.738	47.264
	(0.681)	(0.717)
Mother's highest grade completed	12.224	12.746
	(0.072)	(0.071)
Number of (full) siblings	2.124	2.006
	(0.032)	(0.032)
Age of younger sibling	16.042	15.985
	(0.024)	(0.024)
Age of older sibling	18.065	18.044
	(0.024)	(0.025)
Age gap	2.023	2.060
	(0.022)	(0.022)

Note: Standard errors of the sample means in parentheses. Statistics are based on the sample used for estimation of the dynamic smoking model including 1650 pairs of siblings. 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.

Web Appendix to

Identifying Sibling Influence on Teenage Substance Use

Joseph G. Altonji, Sarah Cattan, and Iain Ware

February 19, 2013

A Data

The paper uses data from the first eight rounds of the National Longitudinal Survey of Youth 1997 (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.

A.1 Sibling pairs

The NLSY97 original cohort includes 1,892 households with more than one 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) to characterize 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.

A.2 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 Black and Hispanic categories, 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 in 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_MO* (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 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 and to 0 otherwise.

- 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 and the variables returning the ages of the other 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, and *HHI2_AGE.0Y* where Y is the ID of the other household respondents.

A.3 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 the 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 2000.

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.

Cigarette, alcohol and marijuana consumption level 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 only used these variables from the second round of the survey onwards.

A.4 Family processes and parenting variables

In several rounds, the NLSY97 asked respondents about their relationship with their residential and non-residential parents. Based on these questions, Child Trends, Inc. created a number of scales measuring different aspects of the 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 *P_ 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 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. In the NLSY97, this variable's name is *YSAQ-351A* (R0357300, R2176000, R3493900, R4892300, S0919200, S4681600).

A.5 Treatment of Missing Data

With the exception of the race and gender 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 substance use measures, 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 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 for both strategies.

B Discussion of the Correlated Random Effects (CRE) approach

We reproduce equations (2) and (3) from section 5.1 of the paper:

$$\begin{aligned} y_t^1 &= 1(\gamma^1 y_{t-1}^1 + X^1\beta^1 + AGE_t^1\Gamma^1 + \alpha^1\epsilon + \lambda^1 v^1 + u_t^1 > 0) \\ y_t^2 &= 1(\gamma^2 y_{t-1}^2 + \lambda^2 y_{t-1}^1 + X^2\beta^2 + AGE_t^2\Gamma^2 + \alpha^2\epsilon + \lambda^2 v^2 + u_t^2 > 0) \end{aligned}$$

The first equation already assumes that y_t^1 does not depend directly on y_{t-1}^2 and any parent's response to y_{t-1}^2 does not influence the older sibling's behavior. For simplicity, assume that the outcome y is a continuous variable, the factor loadings are all equal to 1, and that $\beta^1, \beta^2, \Gamma^1$,and Γ^2 are 0. Under these assumptions, the choices of y_t^1 and y_t^2 are determined by:

$$\begin{aligned} y_t^1 &= \gamma^1 y_{t-1}^1 + \varepsilon + v^1 + u_t^1 \\ y_t^2 &= \gamma^2 y_{t-1}^2 + \lambda^2 y_{t-1}^1 + \varepsilon + v^2 + u_t^2 \end{aligned}$$

Consider the linear least squares projection:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + error \tag{11}$$

Keep in mind that the error components u_t^1 and u_t^2 are person specific, although we have suppressed person subscripts throughout the paper. Assume the following:

(A1) $\gamma^1 = \gamma^2 = 0$, i.e. there is no state dependence from any source, including parental response.

(A2) The distribution of u_t^1 is covariance stationary over t and the age of the older sibling at t with variance $var(u_t^1)$. u_t^1 may be serially dependent.

(A3) $cov(u_t^2, u_{t-1}^1) = cov(u_t^2, u_{t+1}^1)$.

Under assumption (A1), we obtain:

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

Using the above equations and assuming that (A2) and (A3) hold, some straightforward algebra establishes that the coefficients of the projection of $\epsilon + v^2 + u_t^2$ onto y_{t-1}^1 and y_{t+1}^1 both equal to $[var(\varepsilon) + cov(u_t^2, u_{t-1}^1)]/[var(\varepsilon) + var(v^1) + var(u_t^1) + cov(u_{t-1}^1, u_{t+1}^1)]$. Consequently, β_1 and β_2 in (11) are given by:

$$\begin{aligned} \beta_1 &= \frac{var(\varepsilon) + cov(u_t^2, u_{t-1}^1)}{2var(\varepsilon) + var(v^1) + var(u_t^1) + cov(u_{t-1}^1, u_{t+1}^1)} \\ \beta_2 &= \lambda^2 \end{aligned}$$

Thus, under assumptions (A1), (A2) and (A3), β_2 identifies λ^2, the direct sibling effect.

The basic argument carries over to the case in which y is a binary variable determined according to:

$$\begin{aligned} y_{t-1}^1 &= 1(\epsilon + v^1 + u_{t-1}^1) > 0) \\ y_{t+1}^1 &= 1(\epsilon + v^1 + u_{t+1}^1 > 0) \\ y_t^2 &= 1(\lambda^2 y_{t-1}^1 + \epsilon + v^2 + u_t^2 > 0), \end{aligned} \tag{12}$$

although one must replace (A2) with the assumption that the $u_{a,t+a-1}^1$ are identically distributed. However, if any of the three assumptions above are false, then $\beta_2 \neq \lambda^2$ in (11), except in special cases. Indeed, if any of the assumptions fail, then the coefficients of the projection of $\varepsilon + v^2 + u_t^2$ on y_{t-1}^1 and y_{t+1}^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_{t-1}^1 vary with age a in period t, as would be the case if preferences and costs are such that:

$$y_t^1 = f(a) + \alpha_a^1 \varepsilon + \delta_a^1 v^1 + u_t^1,$$

where α_a^1 and δ_a^1 are age dependent coefficients, then the equality restriction on the coefficients of the projection of $\varepsilon + v^2 + u_t^2$ on y_{t-1}^1 and y_{t+1}^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 (12), the presence of $f(a)$ is enough to invalidate the restriction on the projection coefficients, even if α_a^1 and δ_a^1 do not depend on age.

Following Chamberlain (1984), one could generalize the approach by imposing the assumption that u_t^1 and u_t^2 are uncorrelated at all leads and lags, but allowing the coefficients of the projection of $\epsilon + v^2$ on leads and lags of y_t^1 to depend on a_t^2 and a_t^1. We do not pursue this.

Contemporaneous sibling effects Suppose both contemporaneous and lagged behaviors of the older sibling influence the younger child with coefficients λ^{20} and λ^2, respectively. We reproduce the projection equation (5) from section 5.1 of the paper:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_t^1 + y_{t+1}^1) + \beta_2 y_{t-1}^1 + \beta_3 y_t^1 + error \tag{13}$$

In addition to assumptions (A1)-(A3) above, assume:

(A4) The idiosyncratic error components u_t^2 and $u_{t'}^1$ are independent across siblings at all leads and lags.

(A5) u_t^1 is serially uncorrelated.

Then,

$$\beta_1 = \frac{var(\varepsilon)}{3var(\varepsilon) + var(v^1) + var(u_t^1)}$$

and

$$\beta_2 = \lambda^2 \text{ and } \beta_3 = \lambda^{20}$$

where λ^{20} is the contemporaneous effect of y_t^1 on y_t^2. Consequently, under the five assumptions, one can identify the contemporaneous and lagged direct sibling effects.

However, if any of the assumptions (A1) through (A5) fails, then in general $\beta_2 \neq \lambda^2$ and $\beta_3 \neq \lambda^{20}$ in (13). Non-separable forms of age dependence will also pose problems in this case. If only (A6) fails, one can still estimate an average of λ^{20} and λ^2 and test, as we do in the paper, for sibling effects using the regression:

$$y_t^2 = \beta_0 + \beta_1(y_{t-1}^1 + y_t^1 + y_{t+1}^1 + y_{t+2}^1) + \beta_2(y_{t-1}^1 + y_t^1) + error. \tag{14}$$

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 (14). If one uses (11) when (13) is correct, then the coefficient on y^1_{t-1} will pick up part of the effect of y^1_t, but we will still detect sibling influences.

C A joint dynamic model of substance use with gateway drugs

When estimating the joint dynamic model of siblings behavior, we explored in preliminary analysis the idea that some drugs may serve as gateway to others. In the model, this idea is captured by letting an individual's substance use be affected by past use of that particular substance, but also by past use of another substance. In order to estimate the direct effect of the gateway drug on the paired substance, we add a set of equations to the system we previously estimated in order to model the dynamic use of the gateway drug.

Denote g^1_t and g^2_t the older and younger siblings' use of the gateway drug in period t. The model with gateway drugs includes the following equations for the older sibling for all $t > t^1_{\min}$:

$$
\begin{aligned}
y^1_t &= 1(\gamma^1 y^1_{t-1} + \eta^1 g^1_{t-1} + X^1\beta^1 + AGE^1_t\Gamma^1 + \alpha^1\varepsilon + \delta^1 v^1 + u^1_t > 0)\\
g^1_t &= 1(\gamma^{1g} g^1_{t-1} + X^1\beta^{1g} + AGE^1_t\Gamma^{1g} + \alpha^{1g}\epsilon + \delta^{1g} v^1 + u^{1g}_t > 0)
\end{aligned}
$$

and for the younger sibling for all $t > t^2_{\min}$:

$$
\begin{aligned}
y^2_t &= 1(\lambda^2 y^1_{t-1} + \gamma^2 y^2_{t-1} + \eta^2 g^2_{t-1} + \theta^2 a^1_{t-1} + X^2\beta^2_1 + AGE^2_t\Gamma^2 + \alpha^2\varepsilon + \delta^2 v^2 + u^2_t > 0)\\
g^2_t &= 1(\lambda^{2g} g^1_{t-1} + \gamma^{2g} g^2_{t-1} + \theta^{2g} a^1_{t-1} + X^2\beta^{2g} + AGE^2_t\Gamma^{2g} + \alpha^{2g}\varepsilon + \delta^{2g} v^2 + u^{2g}_t > 0)
\end{aligned}
$$

Initial conditions similar to equations (8) and (7) specified in section 5.3 of the paper are also included in the model for $t = t^1_{min}$and $t = t^2_{min}$ for both drugs and siblings.

Web Appendix Table 7 reports results for models in which smoking cigarettes and drinking alcohol are considered as gateways to marijuana use and models in which cigarettes, alcohol and marijuana are gateways to hard drug use. The results reported in this table correspond to error specification A, in which we allow v^1 and v^2 to have different variances, normalize all the factor loadings in the model for the outcome drug to 1, and freely estimate the factor loadings on the family and individual specific components in the gateway equations.

Web Appendix Table 8 reports results from error specification B, in which we restrict $v^1$and v^2 to have the same variance, normalize the factor loadings α_0^1, δ_0^2 and δ_0^1 to one and freely estimate all the other factor loadings. Note that these are broad generalizations of the error specifications A and B we imposed for the models without gateway drugs.

D A joint dynamic ordered probit model with many categories

This appendix describes the joint dynamic ordered probit model with parameter restrictions in more detail. Section 8.2 presented a simple way to allow for flexible forms of nonlinearity in the dynamic behavior of substance use and sibling influence while continuing to allow for sibling pair and individual effects. In particular, we generalized the three category ordered probit model to the M category case by re-writing equations for the latent variables y_t^{1*} and y_t^{2*} as:

$$\begin{aligned}
y_t^{1*} &= X^1\beta_0^1 + AGE_t^1\Gamma_0^1 + \alpha_0^1\epsilon + \delta_0^1 v^1 + u_t^1, \quad t = t_{\min}^1 \\
y_t^{1*} &= \sum_{m=2}^{M} \gamma_m^1 y_{m,t-1}^1 + X^1\beta^1 + AGE_t^1\Gamma^1 + \alpha^1\epsilon + \delta^1 v^1 + u_t^1, \quad t > t_{\min}^1 \\
y_t^{2*} &= \sum_{m=2}^{M} \lambda_{m,0}^2 y_{m,t-1}^1 + X^2\beta_0^2 + \theta_0^2 a_{t-1}^1 + AGE_t^2\Gamma_0^2 + \alpha_0^2\epsilon + \delta_0^2 v^2 + u_t^2, \quad t = t_{\min}^2 \\
y_t^{2*} &= \sum_{m=2}^{M} \gamma_m^2 y_{m,t-1}^2 + \sum_{m=2}^{M} \lambda_m^2 y_{m,t-1}^1 + X^2\beta^2 + \theta^2 a_{t-1}^1 + AGE_t^2\Gamma^2 + \alpha^2\epsilon + \delta^2 v^2 + u_t^2, \quad t > t_{\min}^2
\end{aligned}$$

with thresholds $q_1, q_2, \ldots, q_{M-1}$, where M is the number of categories and category $m = 1$ corresponds to zero consumption. The values of $y_{m,t}^1$ are determined by the indicator function $1(y_t^{1*} < q_1)$ for $m = 1$, by $1(q_{m-1} \leq y_t^{1*} < q_m)$ for $2 \leq m \leq M-1$, and $1(q_{M-1} \leq y_t^{1*})$ for $m = M$. $y_{m,t}^2$ is determined by $y_{m,t}^{2*}$ in the same fashion. By including a large number of groups, one can accommodate an arbitrary non-linear relationship between own current substance use and own past use as well as past use by the older sibling.

With a large number of categories, freely estimating the γ and λ parameters would be hopeless without a very large sample. The solution we propose is to restrict these parameters to lie on a flexible but relatively parsimonious function. A linear spline with a number of break points less than M is a convenient choice. To illustrate the idea, suppose that splines with break points at $m = c_1$ and $m = c_2$ provide a good approximation. Then the M-1 γ_m^1 parameters can be written as a function of 3 parameters g_j^1 (j=1,2,3):

$$\gamma_m^1 = g_1^1 \min(m, c_1) + g_2^1 1(j \geq c_1). \min(m - c_1, c_2 - c_1) + g_3^1 1(m \geq c_2).(m - c_2)$$

Similarly, the other γ and λ parameters can be written as:

$$\begin{aligned}
\gamma_m^2 &= g_1^2 \min(m, c_1) + g_2^2 1(j \geq c_1). \min(m - c_1, c_2 - c_1) + g_3^2 1(m \geq c_2).(m - c_2) \\
\lambda_{m,0}^2 &= l_{1,0}^2 \min(m, c_1) + l_{2,0}^2 1(m \geq c_1). \min(m - c_1, c_2 - c_1) + l_{3,0}^2 1(m \geq c_2).(m - c_2) \\
\lambda_m^2 &= l_1^2 \min(m, c_1) + l_2^2 1(m \geq c_1). \min(m - c_1, c_2 - c_1) + l_3^2 1(m \geq c_2).(m - c_2)
\end{aligned}$$

One could also reduce the number of parameters by constraining the q_m thresholds to lie on a spline or an alternative functional form that is strictly increasing in m. Since we consider a model with five categories, we do not need to do so.

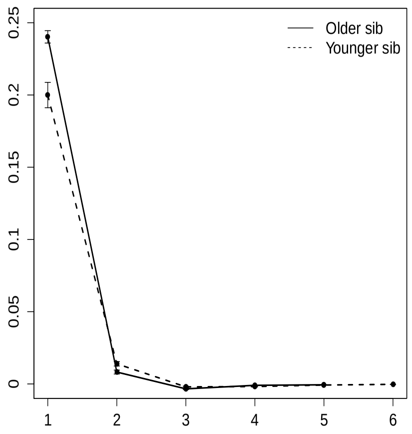
In Web Appendix Table 12, we report estimates of models where we define five categories of substance use (0; 1-7; 8-14; 15-22; 23-30 days) and use splines with break points at c_1=8 days and c_2=18 days. The values of c_1 and c_2 are chosen by assumption and parameters g_j^1, g_j^2, $l_{j,0}^2$ and l_j^2 $(j = 1, 2, 3)$ are freely estimated. The estimates of the γ and λ parameters in Table 7 are computed by substituting the estimates of the spline parameters reported in Web Appendix Table 12 into the above equations.

Note that Fortin and Lemieux (1998) proposed the use of an ordered probit model with a large number of categories as a way to allow for arbitrary non-linearity in the link between a continuous variable (the wage in their case) and a latent variable determined by an index of regressors and an additive error term. Essentially, the ordered probit model provides a model of the probability that the continuous variable falls in a particular interval. Our model is very different, in that it involves a system of equations with dynamics and unobserved heterogeneity. But basically we are combining the idea of using the ordered categorical response model with a large number of categories as the specification for the link function relating y to the observed and unobserved variables that determine the latent variable y^* with the idea of specifying the category specific parameters parsimoniously. Our approach seems well suited to the estimation of nonlinear dynamic panel data models with unobserved heterogeneity.

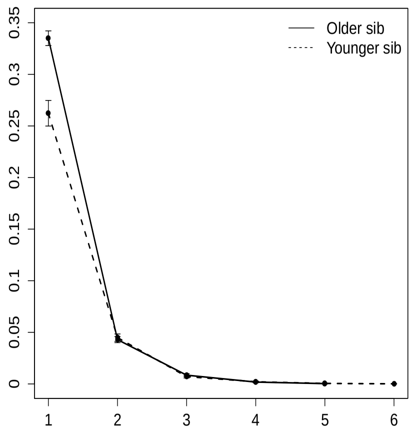
WEB APPENDIX FIGURE 1

Effect of shifting the older sibling's probability of substance use from zero to one of the two highest categories (with equal probability) on the older and younger siblings' probabilities of consumption, relative to baseline (based on ordered probit model with 5 categories, error A)

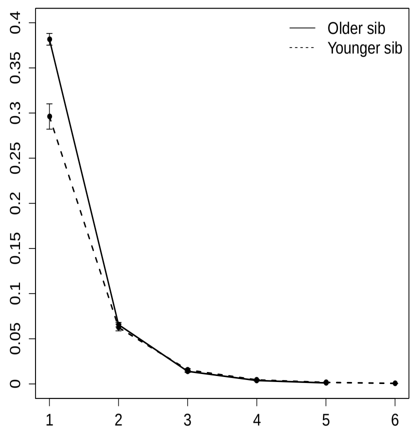
(a) Cigarettes on 1-7 days/month



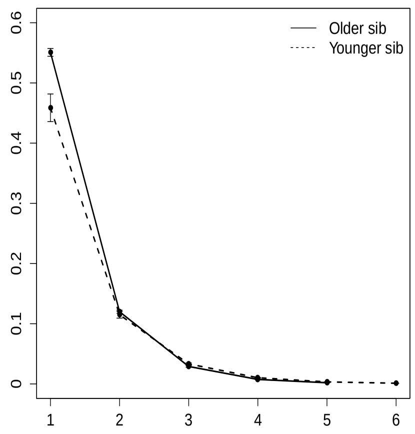
(b) Cigarettes on 8-14 days/month



(c) Cigarettes on 15-22 days/month

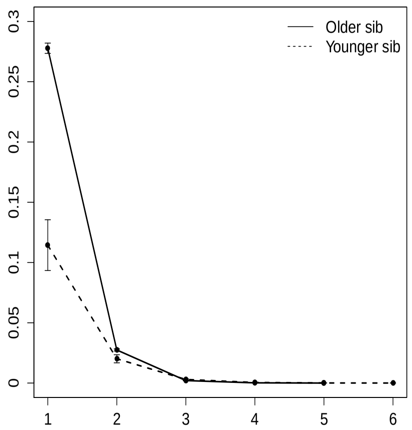


(d) Cigarettes on 23-30 days/month

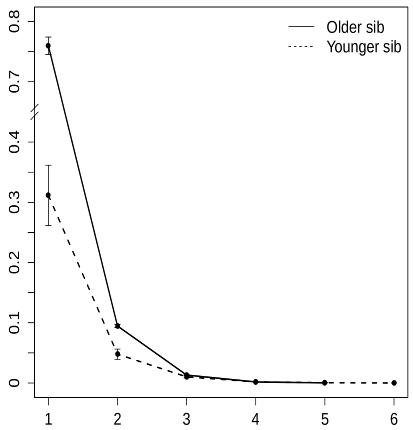


Note: The solid line and the broken line represent the effects on the probabilities of behavior, relative to baseline, of the older sibling and of the young sibling, respectively. Error bars show the 90% confidence intervals. The x-axis measures the number of periods after the exogenous change in the older sibling's behavior. Baseline probabilities for smoking on 1-7 days a month in the first and last period shown in the graphs are: 0.0952 (0.004)and 0.0897(0.004) for the older sibling and 0.0957(0.005) and 0.0928(0.005) for the younger sibling. In the same order, for smoking on 8-14 days: 0.0221(0.001), 0.0217(0.001), 0.0211(0.001) and 0.0225(0.002). For smoking on 15-22 days: 0.0344(0.003),0.0351(0.003),0.0318(0.003),0.0364(0.004). For smoking on 23-30 days: 0.1623 (0.009),0.2361(0.011),0.1213(0.007),0.2448(0.036). For drinking on 1-7 days: 0.3201 (0.011),0.4115(0.009),0.3016(0.011),0.3910(0.017). For drinking on 8-14 days: 0.0431 (0.003),0.0985 (0.004),0.0362(0.002),0.0906(0.011). For drinking on 15-22 days: 0.0216 (0.002),0.0706(0.005),0.0168(0.002),0.0644(0.012). For drinking on 23-30 days: 0.0070 (0.001),0.0408(0.004),0.0050(0.001),0.0355(0.011). For marijuana on 1-7 days: 0.0881 (0.006),0.0805(0.006),0.0874(0.005),0.1040(0.013). For marijuana on 8-14 days: 0.0129(0.002),0.0127(0.002),0.0126(0.002),0.0179(0.005). For marijuana on 15-22 days: 0.0189(0.002),0.0197(0.002),0.0180(0.002),0.0288(0.007). For marijuana on 23-30 days: 0.0283(0.003),0.0408(0.004),0.0259(0.003),0.0688(0.029).

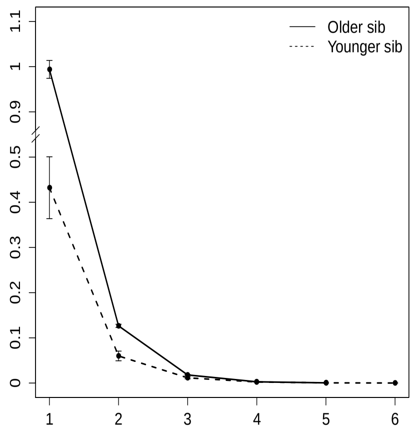
(e) Alcohol on 1-7 days/month



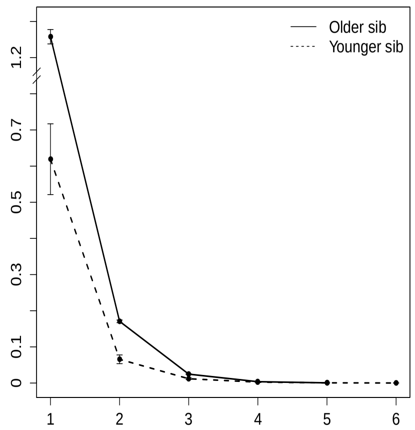
(f) Alcohol on 8-14 days/month



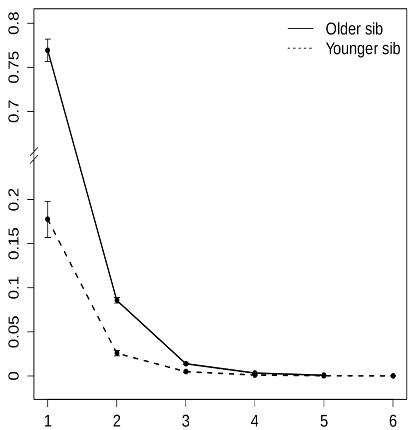
(g) Alcohol on 15-22 days/month



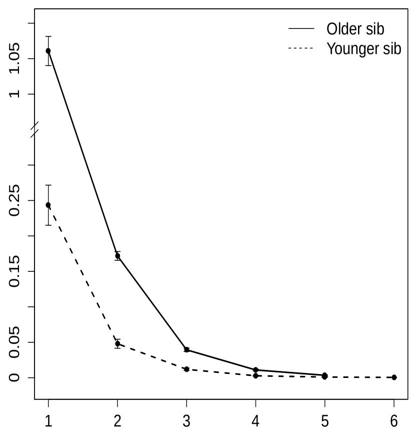
(h) Alcohol on 23-30 days/month



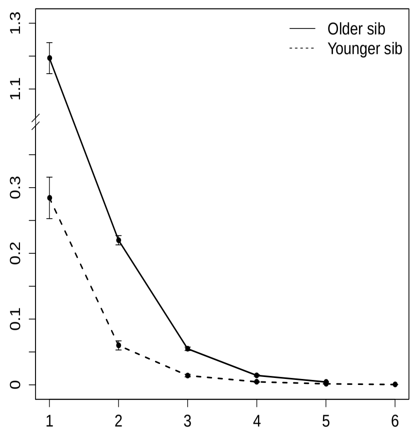
(i) Marijuana on 1-7 days/month



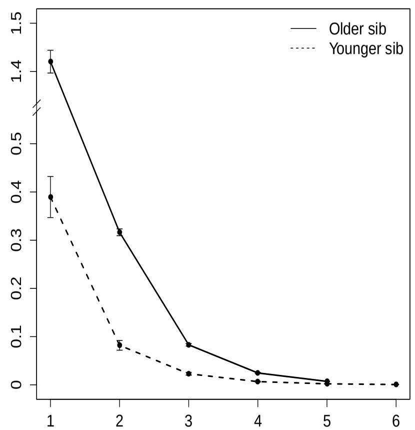
(j) Marijuana on 8-14 days/month



(k) Marijuana on 15-22 days/month



(l) Marijuana on 23-30 days/month



WEB APPENDIX TABLE 1

Weighted Means of Substance Use Measures

	Full Sample	Male Sample	Female Sample
Smoking cigarettes last year	0.433	0.445	0.419
	(0.003)	(0.005)	(0.005)
Drinking alcohol last year	0.646	0.652	0.639
	(0.003)	(0.005)	(0.005)
Smoking marijuana last year	0.236	0.248	0.223
	(0.003)	(0.004)	(0.004)
Using hard drugs last year	0.067	0.070	0.064
	(0.002)	(0.002)	(0.002)
Selling drugs last year	0.057	0.068	0.045
	(0.002)	(0.002)	(0.002)
Days smoked cigarettes last month	7.579	7.748	7.393
	(0.084)	(0.117)	(0.120)
Days drank last month	3.371	3.629	3.089
	(0.039)	(0.056)	(0.053)
Days smoked marijuana last month	1.902	2.167	1.611
	(0.042)	(0.064)	(0.055)

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 21,293 to 21,460 for full sample, from 11,043 to 11,153 for males, and from 10,250 to 10,307 for females.

WEB APPENDIX TABLE 2

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	0.296	0.420	0.170	0.054	0.062	3.148	1.152	0.802
	(0.013)	(0.014)	(0.011)	(0.006)	(0.007)	(0.239)	(0.085)	(0.108)
16	0.341	0.448	0.218	0.058	0.068	4.290	1.376	1.249
	(0.011)	(0.011)	(0.009)	(0.005)	(0.006)	(0.219)	(0.076)	(0.110)
17	0.363	0.520	0.242	0.066	0.079	5.610	1.779	1.732
	(0.009)	(0.010)	(0.008)	(0.005)	(0.005)	(0.214)	(0.077)	(0.116)
18	0.411	0.578	0.246	0.072	0.066	6.607	2.698	1.825
	(0.009)	(0.009)	(0.008)	(0.005)	(0.004)	(0.208)	(0.092)	(0.107)
19	0.423	0.625	0.243	0.059	0.058	7.163	3.143	2.096
	(0.009)	(0.009)	(0.008)	(0.004)	(0.004)	(0.218)	(0.097)	(0.118)
20	0.414	0.647	0.237	0.063	0.049	7.663	3.401	2.254
	(0.009)	(0.009)	(0.008)	(0.004)	(0.004)	(0.225)	(0.105)	(0.127)
21	0.432	0.712	0.215	0.050	0.043	7.948	4.593	2.104
	(0.010)	(0.009)	(0.008)	(0.004)	(0.004)	(0.253)	(0.134)	(0.136)
22	0.435	0.712	0.186	0.055	0.026	8.167	4.497	1.782
	(0.012)	(0.011)	(0.009)	(0.005)	(0.004)	(0.294)	(0.151)	(0.146)
23	0.445	0.731	0.171	0.049	0.019	8.320	4.641	1.596
	(0.015)	(0.013)	(0.011)	(0.006)	(0.004)	(0.379)	(0.200)	(0.177)

Note: Standard errors of sample means in parentheses. Based on the sample used for the estimation of the dynamic smoking model (N=21,398).

WEB APPENDIX TABLE 3

Estimated Marginal Effect of Control Variables in the CRE Model

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Hard Drugs
Male	0.027	-0.006	0.040**	0.003	0.033***
	(0.022)	(0.019)	(0.016)	(0.008)	(0.008)
Black	-0.211***	-0.176***	-0.066***	-0.043***	-0.031***
	(0.026)	(0.027)	(0.020)	(0.009)	(0.009)
Hispanic	-0.145***	-0.017	-0.014	-0.009	-0.008
	(0.027)	(0.027)	(0.022)	(0.011)	(0.010)
Highest grade completed by 19	-0.079***	-0.021**	-0.036***	-0.010***	-0.014***
	(0.011)	(0.010)	(0.007)	(0.004)	(0.003)
AFQT percentile score	0.000	0.002***	0.001***	0.000	0.000*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Mother's grade	0.002	0.005	0.001	0.004**	-0.001
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Lived w/ bio parents at 12	-0.052**	0.003	-0.026	-0.015	-0.012
	(0.025)	(0.023)	(0.019)	(0.011)	(0.010)
Number of (full) siblings	-0.025**	-0.038***	-0.028***	-0.006	-0.010**
	(0.010)	(0.009)	(0.008)	(0.004)	(0.004)
2nd born	-0.085*	-0.138***	-0.103***	-0.036	-0.006
	(0.050)	(0.042)	(0.038)	(0.022)	(0.020)
3rd born	-0.012	-0.074*	-0.049	-0.017	0.010
	(0.047)	(0.042)	(0.030)	(0.015)	(0.019)
House broken in by 12	0.088***	0.006	0.032	0.002	0.009
	(0.028)	(0.027)	(0.020)	(0.011)	(0.009)
Witness of gun shooting by 12	0.089**	0.102***	0.071***	0.037***	0.048***
	(0.036)	(0.033)	(0.025)	(0.014)	(0.011)
Victim of bullying by 12	0.050*	0.052*	0.054***	-0.004	0.014
	(0.029)	(0.027)	(0.020)	(0.011)	(0.009)
Young sibling age 16	0.021	0.036	0.048*		-0.013
	(0.027)	(0.028)	(0.026)		(0.010)
Young sibling age 17	0.038	0.113***	0.061*	-0.025***	-0.004
	(0.035)	(0.033)	(0.031)	(0.008)	(0.014)

WEB APPENDIX TABLE 3 (cont.)

	Smoking Cigarettes	Drinking Alcohol	Smoking Marijuana	Using Hard Drugs	Selling Hard Drugs
Young sibling age 18	0.113***	0.188***	0.056	-0.010	-0.028**
	(0.044)	(0.039)	(0.037)	(0.012)	(0.013)
Young sibling age 19	0.142***	0.237***	0.060	-0.034***	-0.033**
	(0.054)	(0.045)	(0.045)	(0.013)	(0.015)
Young sibling age 20	0.195***	0.262***	0.031	-0.036***	-0.041***
	(0.066)	(0.049)	(0.052)	(0.013)	(0.013)
Young sibling age 21	0.250***	0.323***	0.048	-0.038***	-0.038***
	(0.077)	(0.044)	(0.065)	(0.013)	(0.013)
Young sibling age 22	0.390***	0.327***	0.044	-0.044***	-0.046***
	(0.083)	(0.051)	(0.089)	(0.011)	(0.010)
Old sibling age 17	0.086	0.037	0.015		0.023
	(0.055)	(0.058)	(0.050)		(0.035)
Old sibling age 18	0.088	0.041	0.030	0.240*	0.035
	(0.062)	(0.062)	(0.054)	(0.127)	(0.036)
Old sibling age 19	0.091	0.025	0.058	0.169*	0.053
	(0.067)	(0.068)	(0.059)	(0.100)	(0.040)
Old sibling age 20	0.067	-0.016	0.068	0.183*	0.068
	(0.072)	(0.075)	(0.064)	(0.095)	(0.047)
Old sibling age 21	0.058	-0.033	0.064	0.179*	0.081
	(0.079)	(0.081)	(0.069)	(0.100)	(0.054)
Old sibling age 22	0.004	-0.046	0.080	0.247**	0.087
	(0.085)	(0.088)	(0.077)	(0.125)	(0.062)
Old sibling age 23	-0.036	-0.054	0.058	0.297*	0.126
	(0.092)	(0.097)	(0.085)	(0.153)	(0.085)

Note: See Table 3

WEB APPENDIX TABLE 4

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

	Smoked last year	Drank last year	Marijuana last year	Used Hard drugs last year	Sold hard drugs last year	Days smoked last month	Days drank last month	Day used marijuana last month
y^1_{t-1}	0.028**	0.045***	0.017	0.017	0.004	0.022	0.012	0.005
	(0.014)	(0.015)	(0.014)	(0.016)	(0.014)	(0.014)	(0.014)	(0.016)
Younger sibling's age dummies:								
16	0.017	0.027	0.064***		-0.003	0.648*	0.291	0.890***
	(0.019)	(0.022)	(0.018)		(0.012)	(0.359)	(0.205)	(0.237)
17	0.012	0.091***	0.088***	-0.008	0.012	1.668***	0.667***	1.575***
	(0.022)	(0.024)	(0.020)	(0.013)	(0.014)	(0.434)	(0.246)	(0.325)
18	0.056**	0.140***	0.095***	0.018	-0.008	2.341***	1.443***	1.416***
	(0.024)	(0.026)	(0.023)	(0.015)	(0.014)	(0.535)	(0.298)	(0.354)
19	0.060**	0.179***	0.115***	0.002	-0.010	2.679***	1.421***	1.817***
	(0.025)	(0.029)	(0.025)	(0.015)	(0.015)	(0.595)	(0.327)	(0.380)
20	0.067***	0.181***	0.105***	0.005	-0.027*	3.224***	1.325***	1.724***
	(0.025)	(0.030)	(0.027)	(0.015)	(0.015)	(0.610)	(0.360)	(0.368)
21	0.081***	0.216***	0.096***	-0.006	-0.028*	3.067***	2.323***	1.601***
	(0.027)	(0.032)	(0.028)	(0.016)	(0.016)	(0.636)	(0.410)	(0.371)
22	0.101***	0.189***	0.030	-0.006	-0.037**	3.606***	1.558***	1.033***
	(0.029)	(0.035)	(0.029)	(0.016)	(0.017)	(0.724)	(0.484)	(0.359)

WEB APPENDIX TABLE 4 (cont.)

	Smoked last year	Drank last year	Marijuana last year	Used Hard drugs last year	Sold hard drugs last year	Days smoked last month	Days drank last month	Day used marijuana last month
Older sibling's age dummies:								
17	-0.0150	-0.0613*	0.0174		0.0212	-1.424**	-1.007**	-0.102
	(0.0291)	(0.0355)	(0.0287)		(0.0184)	(0.662)	(0.400)	(0.328)
18	0.0131	-0.0533	0.0239	0.0599***	0.0232	-0.774	-1.183***	-0.408
	(0.0288)	(0.0336)	(0.0271)	(0.0197)	(0.0161)	(0.643)	(0.379)	(0.356)
19	0.0203	-0.0544*	0.0142	0.0258	0.0256*	-0.588	-1.381***	-0.152
	(0.0277)	(0.0318)	(0.0259)	(0.0160)	(0.0150)	(0.617)	(0.371)	(0.379)
20	0.0198	-0.0670**	0.00423	0.0175	0.0236	-0.583	-1.281***	-0.360
	(0.0260)	(0.0298)	(0.0243)	(0.0154)	(0.0145)	(0.581)	(0.350)	(0.387)
21	0.0280	-0.0616**	-0.0156	0.00409	0.0274**	-0.429	-0.976***	-0.323
	(0.0230)	(0.0267)	(0.0238)	(0.0151)	(0.0134)	(0.545)	(0.335)	(0.379)
22	0.0180	-0.0509**	-0.0142	0.0112	0.0183	-0.0979	-0.407	-0.209
	(0.0203)	(0.0235)	(0.0212)	(0.0133)	(0.0122)	(0.469)	(0.309)	(0.331)
23	-0.00698	-0.0415**	-0.0292*	0.00817	0.0200*	-0.00515	-0.255	-0.0450
	(0.0169)	(0.0196)	(0.0171)	(0.0111)	(0.0112)	(0.387)	(0.284)	(0.267)

Note: Standard errors clustered at household level in parentheses. * denotes significant at 10% level, ** at 5% level, and *** at 1% level. Sample sizes vary between 7,056 and 8,698. 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.

WEB APPENDIX TABLE 5

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 (Based on Dynamic Probit Model, 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	0.3992	0.4092	0.4184	0.4261	0.4282	0.4266	
	(0.0114)	(0.0105)	(0.0101)	(0.0094)	(0.0099)	(0.0106)	
W/ feedback	2.5068	0.5078	0.1362	0.0402	0.0125	0.0040	
	(0.0716)	(0.0513)	(0.0239)	(0.0099)	(0.0041)	(0.0016)	
	Younger Siblings						
Baseline		0.3376	0.3715	0.3839	0.3973	0.3932	0.3870
		(0.0119)	(0.0113)	(0.0144)	(0.0182)	(0.0240)	(0.0314)
W/ feedback		0.1409	0.0423	0.0141	0.0048	0.0017	0.0006
		(0.0663)	(0.0215)	(0.0079)	(0.0030)	(0.0011)	(0.0005)
W/out feedback		0.1409	0.0363	0.0108	0.0034	0.0011	0.0004
		(0.0663)	(0.0174)	(0.0054)	(0.0018)	(0.0006)	(0.0002)

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	0.5559	0.5688	0.6171	0.6705	0.7056	0.7237	
	(0.0115)	(0.0102)	(0.0096)	(0.0091)	(0.0091)	(0.0108)	
W/ feedback	1.7995	0.3017	0.0551	0.0104	0.0021	0.0004	
	(0.0370)	(0.0306)	(0.0103)	(0.0028)	(0.0007)	(0.0002)	
	Younger Siblings						
Baseline		0.4610	0.5037	0.5461	0.5709	0.5874	0.5867
		(0.0120)	(0.0115)	(0.0136)	(0.0174)	(0.0226)	(0.0300)
W/ feedback		0.2426	0.0472	0.0097	0.0021	0.0005	0.0001
		(0.0581)	(0.0143)	(0.0038)	(0.0010)	(0.0003)	(0.0001)
W/out feedback		0.2426	0.0462	0.0094	0.0020	0.0004	0.0001
		(0.0581)	(0.0122)	(0.0029)	(0.0008)	(0.0002)	(0.0001)

WEB APPENDIX TABLE 5 (cont.)

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	0.2529	0.2478	0.2454	0.2367	0.2187	0.1972	
	(0.0106)	(0.0099)	(0.0090)	(0.0085)	(0.0085)	(0.0084)	
W/ feedback	3.9618	0.6422	0.1356	0.0316	0.0081	0.0022	
	(0.1667)	(0.0708)	(0.0260)	(0.0086)	(0.0029)	(0.0010)	
	Younger Siblings						
Baseline		0.2087	0.2453	0.2641	0.2745	0.2805	0.2781
		(0.0097)	(0.0105)	(0.0145)	(0.0200)	(0.0264)	(0.0360)
W/ feedback		0.2539	0.0410	0.0079	0.0017	0.0004	0.0001
		(0.0963)	(0.0226)	(0.0066)	(0.0020)	(0.0007)	(0.0002)
W/out feedback		0.2539	0.0484	0.0110	0.0028	0.0007	0.0002
		(0.0963)	(0.0187)	(0.0047)	(0.0013)	(0.0004)	(0.0002)

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. Parametric boostrap standard errors based on 150 replications in parentheses.

WEB APPENDIX TABLE 6

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 (Based on Dynamic Probit Model, 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	0.3992	0.4096	0.4185	0.4265	0.4272	0.4250	
	(0.0113)	(0.0105)	(0.0107)	(0.0100)	(0.0099)	(0.0110)	
W/ feedback	2.5069	0.4414	0.1097	0.0299	0.0087	0.0026	
	(0.0708)	(0.0509)	(0.0214)	(0.0082)	(0.0031)	(0.0011)	
	Younger Siblings						
Baseline		0.3388	0.3713	0.3826	0.3984	0.3947	0.3890
		(0.0117)	(0.0105)	(0.0131)	(0.0172)	(0.0234)	(0.0311)
W/ feedback		-0.0193	0.0061	0.0042	0.0018	0.0008	0.0003
		(0.1117)	(0.0274)	(0.0083)	(0.0027)	(0.0009)	(0.0004)
W/out feedback		-0.0193	-0.0050	-0.0015	-0.0005	-0.0002	-0.0001
		(0.1117)	(0.0275)	(0.0079)	(0.0024)	(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	0.5556	0.5681	0.6173	0.6718	0.7069	0.7248	
	(0.0120)	(0.0103)	(0.0095)	(0.0093)	(0.0095)	(0.0104)	
W/ feedback	1.8007	0.3056	0.0566	0.0105	0.0021	0.0005	
	(0.0384)	(0.0334)	(0.0111)	(0.0029)	(0.0008)	(0.0002)	
	Younger Siblings						
Baseline		0.4599	0.5028	0.5445	0.5694	0.5837	0.5834
		(0.0124)	(0.0112)	(0.0143)	(0.0200)	(0.0264)	(0.0350)
W/ feedback		0.1667	0.0309	0.0062	0.0013	0.0003	0.0001
		(0.0890)	(0.0152)	(0.0033)	(0.0008)	(0.0002)	(0.0001)
W/out feedback		0.1667	0.0271	0.0049	0.0010	0.0002	0.0000
		(0.0890)	(0.0150)	(0.0029)	(0.0006)	(0.0002)	(0.0000)

WEB APPENDIX TABLE 6 (cont.)

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	0.2520	0.2472	0.2463	0.2363	0.2188	0.1962	
	(0.0106)	(0.0101)	(0.0099)	(0.0094)	(0.0089)	(0.0095)	
W/ feedback	3.9754	0.5668	0.1082	0.0237	0.0056	0.0013	
	(0.1677)	(0.0758)	(0.0253)	(0.0077)	(0.0023)	(0.0007)	
	Younger Siblings						
Baseline		0.2066	0.2434	0.2631	0.2735	0.2784	0.2764
		(0.0104)	(0.0103)	(0.0138)	(0.0200)	(0.0273)	(0.0372)
W/ feedback		0.2236	0.0369	0.0074	0.0016	0.0004	0.0001
		(0.1539)	(0.0286)	(0.0071)	(0.0020)	(0.0006)	(0.0002)
W/out feedback		0.2236	0.0406	0.0090	0.0022	0.0005	0.0001
		(0.1539)	(0.0289)	(0.0068)	(0.0018)	(0.0005)	(0.0002)

Note: See Web Appendix Table 5

WEB APPENDIX TABLE 7
Estimates of Dynamic Probit Model Allowing for the Lagged Effect of a Gateway Drug (Error A)

Outcome drug:	Marijuana	Marijuana	Hard drugs	Hard drugs	Hard drugs
Gateway drug:	Smoking	Drinking	Smoking	Drinking	Marijuana
State dependence for outcome drug					
Older sibling (γ^1)	0.743 ***	0.756 ***	0.581 ***	0.604 ***	0.616 ***
	(0.064)	(0.062)	(0.127)	(0.132)	(0.113)
Younger sibling (γ^2)	1.017 ***	0.805 ***	1.401 ***	1.163 ***	0.790 ***
	(0.054)	(0.062)	(0.092)	(0.120)	(0.124)
State dependence for gateway drug					
Older sibling (γ^{1g})	0.906 ***	0.725 ***	0.852 ***	0.632 ***	0.483 ***
	(0.064)	(0.049)	(0.085)	(0.073)	(0.087)
Younger sibling (γ^{2g})	1.007 ***	0.681 ***	0.941 ***	0.562 ***	0.631 ***
	(0.066)	(0.055)	(0.086)	(0.076)	(0.087)
Lagged effect of the gateway drug					
Older sibling (η^1)	-0.028	-0.127 **	0.158	0.072	-0.091
	(0.062)	(0.064)	(0.108)	(0.138)	(0.097)
Younger sibling (η^2)	-0.295 ***	-0.020	-0.087	-0.105	0.025
	(0.059)	(0.061)	(0.108)	(0.134)	(0.103)
Sibling's influence for outcome drug					
Initial condition (λ_0^2)	0.452 ***	0.387 ***	0.444 **	0.430 *	0.390
	(0.100)	(0.104)	(0.224)	(0.243)	(0.287)
Later periods (λ^2)	0.158 ***	0.092	0.182	0.225	0.080
	(0.053)	(0.060)	(0.151)	(0.157)	(0.172)
Sibling's influence for gateway drug					
Initial condition (λ_0^{2g})	0.208 *	0.352 ***	0.301 **	0.192	-0.149
	(0.116)	(0.114)	(0.145)	(0.124)	(0.264)
Later periods (λ^{2g})	0.132 **	0.079	0.143 *	-0.039	0.073
	(0.058)	(0.054)	(0.084)	(0.074)	(0.079)
Standard deviation of error terms specific to:					
Family (σ_ε)	0.300 ***	0.535 ***	0.140 ***	0.331 ***	0.560 ***
	(0.028)	(0.036)	(0.032)	(0.043)	(0.058)
Older sibling (σ_{v^1})	0.820 ***	0.730 ***	0.828 ***	0.806 ***	0.720 ***
	(0.054)	(0.057)	(0.096)	(0.111)	(0.088)
Younger sibling (σ_{v^2})	0.738 ***	0.720 ***	0.459 ***	0.568 ***	0.823 ***
	(0.045)	(0.059)	(0.062)	(0.072)	(0.110)

WEB APPENDIX TABLE 7 (cont.)

Outcome drug:	Marijuana	Marijuana	Hard drugs	Hard drugs	Hard drugs
Gateway drug:	Smoking	Drinking	Smoking	Drinking	Marijuana
Factor loadings of family specific error term in gateway drug model					
α_0^{1g}	2.757 ***	1.221 ***	8.141 ***	2.174 ***	1.584 ***
	(0.347)	(0.189)	(1.959)	(0.398)	(0.366)
α^{1g}	3.611 ***	0.993 ***	9.384 ***	2.482 ***	1.080 ***
	(0.384)	(0.112)	(2.225)	(0.401)	(0.166)
α_0^{2g}	1.796 ***	1.243 ***	3.443 ***	1.856 ***	2.024 ***
	(0.324)	(0.243)	(1.020)	(0.376)	(0.648)
α^{2g}	1.579 ***	1.085 ***	2.668 ***	1.746 ***	1.077 ***
	(0.248)	(0.152)	(0.667)	(0.320)	(0.205)
Factor loadings on older sibling error term in gateway drug model					
δ_0^{1g}	0.954 ***	0.882 ***	0.811 ***	0.705 ***	1.089 ***
	(0.126)	(0.148)	(0.157)	(0.191)	(0.177)
δ^{1g}	0.755 ***	0.668 ***	0.778 ***	0.416 ***	1.459 ***
	(0.082)	(0.080)	(0.140)	(0.109)	(0.216)
Factor loadings on younger sibling error term in gateway drug model					
δ_0^{2g}	1.268 ***	0.691 ***	2.709 ***	1.704 ***	1.161 ***
	(0.141)	(0.119)	(0.444)	(0.280)	(0.188)
δ^{2g}	1.262 ***	0.919 ***	2.371 ***	1.497 ***	1.032 ***
	(0.119)	(0.083)	(0.394)	(0.219)	(0.139)
Log likelihood value	-13943.69	-14787.61	-8640.72	-9261.26	-7877.30

Note: 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 1,278 to 1,640 for the older siblings' models and from 1,066 to 1640 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. In this specification, the factor loadings $\alpha_0^1, \alpha^1, \alpha_0^2, \alpha^2, \delta_0^1, \delta^1, \delta_0^2, \delta^2$ are normalized to 1.

WEB APPENDIX TABLE 8

Estimates of Dynamic Probit Model Allowing for the Lagged Effect of a Gateway Drug (Error B)

Outcome drug: Gateway drug:	Marijuana Smoking	Marijuana Drinking	Hard drugs Smoking	Hard drugs Drinking	Hard drugs Marijuana
State dependence for outcome drug					
Older sibling (γ^1)	1.024 ***	0.814 ***	0.866 ***	0.585 ***	0.604 ***
	(0.045)	(0.058)	(0.112)	(0.142)	(0.117)
Younger sibling (γ^2)	0.977 ***	0.791 ***	1.252 ***	1.144 ***	0.823 ***
	(0.058)	(0.062)	(0.116)	(0.124)	(0.127)
State dependence for gateway drug					
Older sibling (γ^{1g})	0.966 ***	0.707 ***	0.908 ***	0.640 ***	0.473 ***
	(0.063)	(0.050)	(0.080)	(0.073)	(0.085)
Younger sibling (γ^{2g})	1.098 ***	0.737 ***	0.915 ***	0.577 ***	0.628 ***
	(0.061)	(0.052)	(0.093)	(0.076)	(0.088)
Lagged effect of the gateway drug					
Older sibling (η^1)	-0.187 ***	-0.159 **	0.027	0.042	-0.151
	(0.064)	(0.066)	(0.127)	(0.144)	(0.109)
Younger sibling (η^2)	-0.259 ***	0.003	0.023	-0.114	0.050
	(0.061)	(0.061)	(0.124)	(0.150)	(0.115)
Sibling's influence for outcome drug					
Initial condition (λ_0^2)	0.418 ***	0.386 ***	0.268	0.437 *	0.118
	(0.104)	(0.112)	(0.241)	(0.251)	(0.337)
Later periods (λ^2)	0.127 **	0.036	-0.172	0.206	0.153
	(0.060)	(0.066)	(0.183)	(0.165)	(0.176)
Sibling's influence for gateway drug					
Initial condition (λ_0^{2g})	0.267 **	0.490 ***	0.510 ***	0.184	-0.528 **
	(0.105)	(0.089)	(0.123)	(0.126)	(0.260)
Later periods (λ^{2g})	0.168 ***	0.108 *	0.261 ***	-0.033	0.107
	(0.064)	(0.055)	(0.076)	(0.077)	(0.076)
Standard devition of error terms specific to:					
Family (σ_ε)	0.314 ***	0.459 ***	0.244 ***	0.352 ***	0.403 ***
	(0.042)	(0.045)	(0.053)	(0.065)	(0.070)
Individual (σ_v)	0.834 ***	0.840 ***	0.409 ***	0.639 ***	0.552 ***
	(0.057)	(0.056)	(0.075)	(0.088)	(0.078)

WEB APPENDIX TABLE 8 (cont.)

Outcome drug: Gateway drug:	Marijuana Smoking	Marijuana Drinking	Hard drugs Smoking	Hard drugs Drinking	Hard drugs Marijuana
Factor loadings on family specific error term in the outcome drug equations					
α^1	2.043 ***	1.963 ***	2.573 ***	1.219 ***	1.189 ***
	(0.299)	(0.228)	(0.693)	(0.396)	(0.288)
α_0^2	1.318 ***	1.058 ***	2.357 ***	0.846 ***	2.845 ***
	(0.271)	(0.173)	(0.798)	(0.322)	(0.683)
α^2	0.927 ***	0.951 ***	2.153 ***	0.947 ***	1.346 ***
	(0.180)	(0.131)	(0.661)	(0.272)	(0.328)
Factor loadings on family specific error term in the gateway drug equations					
α_0^{1g}	3.329 ***	3.089 ***	2.039 ***	2.123 ***	2.033 ***
	(0.672)	(0.697)	(0.614)	(0.492)	(0.489)
α^{1g}	3.486 ***	1.527 ***	1.904 ***	2.364 ***	1.375 ***
	(0.493)	(0.167)	(0.524)	(0.498)	(0.278)
α_0^{2g}	1.635 ***	0.989 ***	2.481 ***	1.680 ***	6.194 ***
	(0.314)	(0.164)	(0.730)	(0.401)	(2.014)
α^{2g}	1.164 ***	0.771 ***	1.918 ***	1.546 ***	1.601 ***
	(0.210)	(0.124)	(0.550)	(0.330)	(0.344)
Factor loadings on older sibling specific error term					
δ^1	0.370 ***	0.292 ***	0.882 ***	1.277 ***	1.579 ***
	(0.063)	(0.059)	(0.269)	(0.261)	(0.281)
δ_0^{1g}	2.196 ***	3.313 ***	2.988 ***	0.781 ***	1.519 ***
	(0.486)	(0.800)	(0.601)	(0.262)	(0.289)
δ^{1g}	0.573 ***	0.288 ***	3.097 ***	0.430 ***	1.984 ***
	(0.085)	(0.049)	(0.608)	(0.161)	(0.324)
Factor loadings on younger sibling specific error term					
δ^2	0.875 ***	0.953 ***	0.594 **	0.921 ***	1.514 ***
	(0.084)	(0.088)	(0.239)	(0.210)	(0.311)
δ_0^{2g}	1.081 ***	0.758 ***	2.897 ***	1.527 ***	1.621 ***
	(0.127)	(0.094)	(0.592)	(0.264)	(0.380)
δ^{2g}	0.995 ***	0.832 ***	2.592 ***	1.324 ***	1.521 ***
	(0.095)	(0.074)	(0.523)	(0.213)	(0.242)
Log likelihood value	-13927.22	-14754.10	-8641.74	-9261.98	-7861.85

Note: See Web Appendix Table 7. The factor loadings $\alpha_0^1, \delta_0^1, \delta_0^2$ are normalized to 1.

WEB APPENDIX TABLE 9

Estimates of Joint Dynamic Ordered Probit Model (Error B)

		Cigarettes	Alcohol	Marijuana
State dependence parameters				
Older sibling	Low consumption (γ_L^1)	0.325 ***	0.345 ***	0.315 ***
		(0.078)	(0.043)	(0.080)
	High consumption (γ_H^1)	0.759 ***	0.588 ***	0.645 ***
		(0.070)	(0.065)	(0.080)
Younger sibling	Low consumption (γ_L^2)	0.560 ***	0.531 ***	0.568 ***
		(0.080)	(0.045)	(0.083)
	High consumption (γ_H^2)	0.969 ***	0.824 ***	0.958 ***
		(0.076)	(0.070)	(0.095)
Sibling's influence parameters				
Initial condition	Low consumption ($\lambda_{L,0}^2$)	0.053	0.104	0.339 *
		(0.202)	(0.109)	(0.184)
	High consumption ($\lambda_{H,0}^2$)	0.351 *	0.281	0.196
		(0.207)	(0.181)	(0.185)
Later periods	Low consumption (λ_L^2)	0.069	-0.022	0.069
		(0.092)	(0.046)	(0.095)
	High consumption (λ_H^2)	0.093	-0.146 **	-0.011
		(0.087)	(0.070)	(0.104)
Family-specific error term				
Standard deviation (σ_ε)		1.060 ***	0.433 ***	0.886 ***
		(0.147)	(0.053)	(0.153)
FL Older sibling, later periods (α^1)		0.917 ***	0.929 ***	1.252 ***
		(0.129)	(0.113)	(0.193)
FL Younger sib, initial period (α_0^2)		0.603 ***	1.594 ***	0.407 ***
		(0.166)	(0.321)	(0.126)
FL Younger sib, later periods (α^2)		0.609 ***	1.751 ***	0.482 ***
		(0.143)	(0.232)	(0.120)
Individual- specific error term				
Standard deviation (σ_v)		0.909 ***	0.637 ***	0.663 ***
		(0.075)	(0.057)	(0.085)
FL Older sibling, later periods (δ^1)		1.310 ***	0.995 ***	0.380
		(0.127)	(0.099)	(0.286)
FL Younger siblings, later periods (δ^2)		1.185 ***	-0.081	1.346 ***
		(0.139)	(0.156)	(0.197)
Low consumption threshold (q_L)		-0.141	-0.462	0.889 *
		(0.465)	(0.372)	(0.512)
High consumption threshold (q_H)		0.444	1.177 ***	1.517 ***
		(0.465)	(0.372)	(0.512)
Log likelihood value		-9546.68	-13332.51	-7367.66

Note: See Table 6. In this specification, the factor loadings $\alpha_0^1, \delta_0^1, \delta_0^2$ are normalized to 1.

WEB APPENDIX TABLE 10

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 (Based on Dynamic Ordered Probit Model, 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	0.0947	0.0865	0.0877	0.0889	0.0899	0.0902	
	(0.0044)	(0.0040)	(0.0038)	(0.0041)	(0.0041)	(0.0041)	
High consumption	0.2096	0.2389	0.2619	0.2767	0.2856	0.2910	
	(0.0091)	(0.0087)	(0.0089)	(0.0095)	(0.0092)	(0.0096)	
W/ feedback							
Low consumption	0.0000	0.2623	0.0171	0.0009	-0.0001	-0.0001	
	(0.0000)	(0.0354)	(0.0094)	(0.0040)	(0.0019)	(0.0008)	
High consumption	4.7792	0.5511	0.1166	0.0281	0.0071	0.0019	
	(0.2079)	(0.0640)	(0.0230)	(0.0078)	(0.0027)	(0.0009)	
	Younger Siblings						
Baseline							
Low consumption		0.0956	0.0875	0.0887	0.0914	0.0913	0.0916
		(0.0040)	(0.0040)	(0.0045)	(0.0047)	(0.0052)	(0.0058)
High consumption		0.1714	0.2157	0.2449	0.2684	0.2788	0.2839
		(0.0083)	(0.0113)	(0.0166)	(0.0226)	(0.0289)	(0.0380)
W/ feedback							
Low consumption		0.2023	0.0229	0.0024	0.0000	-0.0001	0.0001
		(0.0524)	(0.0112)	(0.0043)	(0.0022)	(0.0013)	(0.0006)
High consumption		0.3927	0.1056	0.0318	0.0099	0.0033	0.0011
		(0.1047)	(0.0306)	(0.0105)	(0.0039)	(0.0015)	(0.0006)
W/out feedback							
Low consumption		0.2023	0.0192	0.0013	-0.0002	-0.0002	0.0000
		(0.0524)	(0.0100)	(0.0039)	(0.0019)	(0.0011)	(0.0005)
High consumption		0.3927	0.0971	0.0278	0.0083	0.0027	0.0009
		(0.1047)	(0.0274)	(0.0088)	(0.0031)	(0.0012)	(0.0004)

WEB APPENDIX TABLE 10 (cont.)

	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	0.3174	0.3388	0.3640	0.3875	0.4010	0.4068	
	(0.0089)	(0.0082)	(0.0074)	(0.0074)	(0.0080)	(0.0083)	
High consumption	0.0713	0.1057	0.1336	0.1706	0.1977	0.2135	
	(0.0053)	(0.0056)	(0.0068)	(0.0077)	(0.0074)	(0.0078)	
W/ feedback							
Low consumption	0.0000	0.2573	0.0260	0.0020	0.0002	0.0000	
	(0.0000)	(0.0324)	(0.0069)	(0.0011)	(0.0002)	(0.0001)	
High consumption	14.1059	0.7627	0.1040	0.0152	0.0022	0.0004	
	(1.0431)	(0.1089)	(0.0249)	(0.0053)	(0.0011)	(0.0003)	
	Younger Siblings						
Baseline							
Low consumption		0.2984	0.3147	0.3468	0.3704	0.3823	0.3906
		(0.0078)	(0.0087)	(0.0083)	(0.0097)	(0.0110)	(0.0140)
High consumption		0.0568	0.0900	0.1223	0.1529	0.1746	0.1916
		(0.0045)	(0.0087)	(0.0139)	(0.0191)	(0.0240)	(0.0304)
W/ feedback							
Low consumption		0.1414	0.0164	0.0018	0.0002	0.0000	0.0000
		(0.0877)	(0.0160)	(0.0027)	(0.0006)	(0.0002)	(0.0001)
High consumption		0.3543	0.0427	0.0061	0.0008	0.0001	0.0000
		(0.2348)	(0.0457)	(0.0109)	(0.0025)	(0.0006)	(0.0002)
W/out feedback							
Low consumption		0.1414	0.0221	0.0032	0.0005	0.0001	0.0000
		(0.0877)	(0.0147)	(0.0024)	(0.0005)	(0.0001)	(0.0001)
High consumption		0.3543	0.0645	0.0136	0.0029	0.0006	0.0001
		(0.2348)	(0.0416)	(0.0091)	(0.0020)	(0.0005)	(0.0001)

WEB APPENDIX TABLE 10 (cont.)

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	0.0724	0.0732	0.0748	0.0743	0.0703	0.0673	
	(0.0045)	(0.0044)	(0.0043)	(0.0038)	(0.0033)	(0.0036)	
High consumption	0.0761	0.0917	0.0987	0.0987	0.0921	0.0872	
	(0.0057)	(0.0066)	(0.0063)	(0.0061)	(0.0062)	(0.0064)	
W/ feedback							
Low consumption	0.0000	0.5650	0.0612	0.0086	0.0024	0.0003	
	(0.0000)	(0.0761)	(0.0162)	(0.0040)	(0.0016)	(0.0007)	
High consumption	13.2128	0.8917	0.1584	0.0341	0.0079	0.0019	
	(1.0024)	(0.1226)	(0.0351)	(0.0108)	(0.0034)	(0.0011)	
	Younger Siblings						
Baseline							
Low consumption		0.0735	0.0785	0.0810	0.0830	0.0829	0.0805
		(0.0041)	(0.0048)	(0.0060)	(0.0076)	(0.0092)	(0.0118)
High consumption		0.0685	0.0933	0.1096	0.1189	0.1238	0.1247
		(0.0050)	(0.0096)	(0.0169)	(0.0246)	(0.0327)	(0.0432)
W/ feedback							
Low consumption		0.1520	0.0210	0.0046	0.0011	0.0004	0.0001
		(0.1505)	(0.0223)	(0.0063)	(0.0022)	(0.0011)	(0.0006)
High consumption		0.2358	0.0623	0.0183	0.0060	0.0020	0.0007
		(0.2386)	(0.0636)	(0.0198)	(0.0068)	(0.0025)	(0.0010)
W/out feedback							
Low consumption		0.1520	0.0196	0.0042	0.0011	0.0004	0.0001
		(0.1505)	(0.0209)	(0.0057)	(0.0018)	(0.0009)	(0.0005)
High consumption		0.2358	0.0606	0.0180	0.0058	0.0020	0.0007
		(0.2386)	(0.0615)	(0.0184)	(0.0062)	(0.0022)	(0.0008)

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. Parametric boostrap standard errors based on 150 replications in parentheses.

WEB APPENDIX TABLE 11

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 (Based on Dynamic Ordered Probit Model, Error B)

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	0.0988	0.0869	0.0871	0.0883	0.0888	0.0894	
	(0.0063)	(0.0052)	(0.0055)	(0.0055)	(0.0057)	(0.0059)	
High consumption	0.2104	0.2424	0.2648	0.2803	0.2893	0.2941	
	(0.0093)	(0.0088)	(0.0084)	(0.0081)	(0.0083)	(0.0092)	
W/ feedback							
Low consumption	0.0000	0.2360	0.0156	0.0002	-0.0006	-0.0003	
	(0.0000)	(0.0335)	(0.0105)	(0.0035)	(0.0016)	(0.0009)	
High consumption	4.7618	0.5140	0.1053	0.0243	0.0061	0.0016	
	(0.2129)	(0.0663)	(0.0231)	(0.0076)	(0.0025)	(0.0008)	
	Younger Siblings						
Baseline							
Low consumption		0.1004	0.0876	0.0877	0.0903	0.0908	0.0912
		(0.0067)	(0.0057)	(0.0059)	(0.0062)	(0.0068)	(0.0075)
High consumption		0.1695	0.2145	0.2443	0.2695	0.2810	0.2884
		(0.0074)	(0.0112)	(0.0172)	(0.0236)	(0.0314)	(0.0419)
W/ feedback							
Low consumption		0.1485	0.0189	0.0020	-0.0003	0.0001	0.0000
		(0.0927)	(0.0121)	(0.0035)	(0.0021)	(0.0012)	(0.0006)
High consumption		0.2923	0.0759	0.0221	0.0068	0.0021	0.0007
		(0.1847)	(0.0428)	(0.0123)	(0.0039)	(0.0014)	(0.0005)
W/out feedback							
Low consumption		0.1485	0.0151	0.0011	-0.0005	0.0000	0.0000
		(0.0927)	(0.0117)	(0.0029)	(0.0016)	(0.0010)	(0.0006)
High consumption		0.2923	0.0670	0.0181	0.0053	0.0016	0.0005
		(0.1847)	(0.0425)	(0.0119)	(0.0036)	(0.0012)	(0.0004)

WEB APPENDIX TABLE 11 (cont.)

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	0.3206	0.3430	0.3699	0.3938	0.4095	0.4144	
	(0.0127)	(0.0111)	(0.0109)	(0.0102)	(0.0103)	(0.0102)	
High consumption	0.0718	0.1029	0.1302	0.1662	0.1951	0.2102	
	(0.0083)	(0.0093)	(0.0109)	(0.0128)	(0.0139)	(0.0153)	
W/ feedback							
Low consumption	0.0000	0.2612	0.0273	0.0024	0.0002	0.0000	
	(0.0000)	(0.0329)	(0.0063)	(0.0011)	(0.0003)	(0.0001)	
High consumption	14.1160	0.7688	0.1028	0.0148	0.0022	0.0004	
	(1.6318)	(0.0967)	(0.0220)	(0.0045)	(0.0009)	(0.0002)	
	Younger Siblings						
Baseline							
Low consumption		0.2796	0.3173	0.3461	0.3700	0.3823	0.3901
		(0.0148)	(0.0107)	(0.0119)	(0.0137)	(0.0159)	(0.0197)
High consumption		0.0668	0.0880	0.1253	0.1567	0.1811	0.2011
		(0.0076)	(0.0109)	(0.0185)	(0.0256)	(0.0327)	(0.0420)
W/ feedback							
Low consumption		0.1661	0.0192	0.0022	0.0002	0.0000	0.0000
		(0.1083)	(0.0166)	(0.0025)	(0.0005)	(0.0002)	(0.0001)
High consumption		0.4186	0.0535	0.0072	0.0011	0.0002	0.0000
		(0.2843)	(0.0484)	(0.0095)	(0.0021)	(0.0005)	(0.0001)
W/out feedback							
Low consumption		0.1661	0.0251	0.0035	0.0005	0.0001	0.0000
		(0.1083)	(0.0170)	(0.0026)	(0.0005)	(0.0001)	(0.0001)
High consumption		0.4186	0.0762	0.0146	0.0031	0.0006	0.0001
		(0.2843)	(0.0503)	(0.0099)	(0.0022)	(0.0005)	(0.0001)

WEB APPENDIX TABLE 11 (cont.)

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	0.0727	0.0735	0.0746	0.0742	0.0707	0.0670	
	(0.0054)	(0.0035)	(0.0032)	(0.0033)	(0.0036)	(0.0041)	
High consumption	0.0761	0.0897	0.0972	0.0982	0.0915	0.0859	
	(0.0064)	(0.0066)	(0.0063)	(0.0063)	(0.0065)	(0.0070)	
W/ feedback							
Low consumption	0.0000	0.5188	0.0532	0.0076	0.0018	0.0003	
	(0.0000)	(0.0848)	(0.0175)	(0.0041)	(0.0017)	(0.0007)	
High consumption	13.2317	0.8267	0.1387	0.0279	0.0062	0.0015	
	(1.1041)	(0.1390)	(0.0376)	(0.0108)	(0.0030)	(0.0009)	
	Younger Siblings						
Baseline							
Low consumption		0.0777	0.0781	0.0802	0.0815	0.0819	0.0800
		(0.0048)	(0.0043)	(0.0057)	(0.0073)	(0.0090)	(0.0117)
High consumption		0.0646	0.0913	0.1089	0.1176	0.1232	0.1242
		(0.0052)	(0.0085)	(0.0156)	(0.0231)	(0.0306)	(0.0404)
W/ feedback							
Low consumption		0.1936	0.0243	0.0043	0.0010	0.0003	0.0001
		(0.1831)	(0.0219)	(0.0048)	(0.0018)	(0.0008)	(0.0005)
High consumption		0.3081	0.0664	0.0172	0.0050	0.0015	0.0005
		(0.2932)	(0.0601)	(0.0161)	(0.0051)	(0.0017)	(0.0006)
W/out feedback							
Low consumption		0.1936	0.0225	0.0039	0.0008	0.0002	0.0001
		(0.1831)	(0.0222)	(0.0043)	(0.0015)	(0.0007)	(0.0003)
High consumption		0.3081	0.0637	0.0164	0.0048	0.0015	0.0005
		(0.2932)	(0.0610)	(0.0162)	(0.0050)	(0.0017)	(0.0006)

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. Parametric boostrap standard errors based on 150 replications in parentheses.

WEB APPENDIX TABLE 12

Parameter Estimates of the Joint Dynamic Ordered Probit Model with 5 Categories (Error A)

		Smoking cigarettes	Drinking alcohol	Smoking marijuana
State dependence parameters				
Older sibling	g_1^1	0.043 ***	0.047 ***	0.046 ***
		(0.011)	(0.006)	(0.010)
	g_2^1	0.026 *	0.018 **	0.041 **
		(0.014)	(0.009)	(0.016)
	g_3^1	0.017	0.010	0.011
		(0.011)	(0.010)	(0.015)
Younger sibling	g_1^2	0.076 ***	0.076 ***	0.085 ***
		(0.011)	(0.006)	(0.011)
	g_2^2	0.020	0.034 ***	0.027
		(0.015)	(0.010)	(0.018)
	g_3^2	0.023 **	-0.024 *	0.008
		(0.012)	(0.012)	(0.015)
Sibling's influence parameters				
Initial condition	$l_{1,0}^2$	0.014	0.011	0.039 *
		(0.025)	(0.011)	(0.023)
	$l_{2,0}^2$	0.044	0.015	0.035
		(0.039)	(0.022)	(0.040)
	$l_{3,0}^2$	-0.014	-0.005	-0.056
		(0.030)	(0.042)	(0.038)
Later periods	l_1^2	0.012	-0.003	0.012
		(0.012)	(0.006)	(0.012)
	l_2^2	-0.005	-0.013	-0.035 *
		(0.018)	(0.009)	(0.020)
	l_3^2	0.003	0.005	0.018
		(0.014)	(0.012)	(0.018)

WEB APPENDIX TABLE 12 (cont.)

	Smoking cigarettes	Drinking alcohol	Smoking marijuana
Standard deviation of error term specific to:			
Family (σ_{ε})	0.796 ***	0.527 ***	0.653 ***
	(0.055)	(0.026)	(0.051)
Older sibling (σ_{v^1})	1.265 ***	0.515 ***	0.831 ***
	(0.070)	(0.041)	(0.072)
Younger sibling (σ_{v^2})	0.874 ***	0.478 ***	0.652 ***
	(0.072)	(0.043)	(0.075)
Thresholds:			
q_1	-0.257	-0.387	0.710
	(0.462)	(0.345)	(0.473)
q_2	0.322	1.216 ***	1.508 ***
	(0.462)	(0.345)	(0.474)
q_3	0.476	1.805 ***	1.697 ***
	(0.461)	(0.345)	(0.475)
q_4	0.736	2.556 ***	2.064 ***
	(0.461)	(0.347)	(0.476)
Log likelihood value	-11961.94	-15445.85	-8536.23

Note: See Table 6. In this specification, all factor loadings are normalized to 1.

WEB APPENDIX TABLE 13

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 Based on the Joint Dynamic Ordered Probit Model with Five Consumption Categories (Error A)

Smoking cigarettes - Older siblings

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
Baseline							
1-7 days	0.0952	0.0863	0.0874	0.0886	0.0891	0.0897	
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	
8-14 days	0.0221	0.0203	0.0208	0.0213	0.0216	0.0217	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
15-22 days	0.0344	0.0321	0.0332	0.0341	0.0347	0.0351	
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	
23-30 days	0.1623	0.1904	0.2101	0.2235	0.2306	0.2361	
	(0.009)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	
W/ feedback							
1-7 days	0.0000	0.2402	0.0082	-0.0034	-0.0010	-0.0006	
	(0.000)	(0.032)	(0.010)	(0.004)	(0.002)	(0.001)	
8-14 days	0.0000	0.3350	0.0429	0.0085	0.0018	0.0003	
	(0.000)	(0.053)	(0.022)	(0.010)	(0.005)	(0.003)	
15-22 days	14.6708	0.3817	0.0656	0.0140	0.0038	0.0011	
	(1.425)	(0.048)	(0.019)	(0.009)	(0.004)	(0.002)	
23-30 days	3.0914	0.5508	0.1199	0.0290	0.0074	0.0020	
	(0.173)	(0.060)	(0.022)	(0.007)	(0.003)	(0.001)	

Note: "Baseline" corresponds to probabilities of simulated behaviors using the dynamic ordered probit model with 5 consumption categories. "W/ feedback" corresponds to an exogenous shift of the older sibling's behavior from zero to one of the highest two consumption categories 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 one of the two highest consumption categories in the first period, setting the older sibling's behavior in the later periods to the baseline level. Both of these simulations are performed so that half of the sample of older brothers is in the 15-22 days category and the other half in the 23-30 days in the first period. 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. Parametric boostrap standard errors based on 150 replications in parentheses.

WEB APPENDIX TABLE 13 (cont.)

Smoking cigarettes - Younger siblings

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
1-7 days		0.0957	0.0874	0.0891	0.0919	0.0920	0.0928
		(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
8-14 days		0.0211	0.0199	0.0209	0.0218	0.0222	0.0225
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
15-22 days		0.0318	0.0312	0.0329	0.0349	0.0357	0.0364
		(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
23-30 days		0.1213	0.1655	0.1975	0.2219	0.2358	0.2448
		(0.007)	(0.010)	(0.015)	(0.021)	(0.028)	(0.036)
W/ feedback							
1-7 days		0.2000	0.0180	-0.0013	-0.0018	-0.0009	-0.0003
		(0.065)	(0.013)	(0.004)	(0.003)	(0.001)	(0.001)
8-14 days		0.2623	0.0495	0.0091	0.0022	0.0010	0.0001
		(0.092)	(0.031)	(0.012)	(0.006)	(0.003)	(0.002)
15-22 days		0.2962	0.0671	0.0178	0.0053	0.0020	0.0008
		(0.105)	(0.029)	(0.012)	(0.005)	(0.003)	(0.002)
23-30 days		0.4587	0.1236	0.0373	0.0120	0.0040	0.0014
		(0.171)	(0.047)	(0.016)	(0.006)	(0.002)	(0.001)
W/out feedback							
1-7 days		0.2000	0.0139	-0.0020	-0.0018	-0.0008	-0.0003
		(0.065)	(0.011)	(0.004)	(0.002)	(0.001)	(0.001)
8-14 days		0.2623	0.0449	0.0071	0.0021	0.0005	0.0001
		(0.092)	(0.027)	(0.010)	(0.005)	(0.003)	(0.002)
15-22 days		0.2962	0.0625	0.0155	0.0045	0.0017	0.0006
		(0.105)	(0.028)	(0.011)	(0.004)	(0.003)	(0.001)
23-30 days		0.4587	0.1149	0.0332	0.0103	0.0034	0.0012
		(0.171)	(0.042)	(0.013)	(0.004)	(0.002)	(0.001)

WEB APPENDIX TABLE 13 (cont.)

Drinking alcohol - Older siblings

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
Baseline							
1-7 days	0.3201	0.3412	0.3680	0.3917	0.4043	0.4115	
	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	
8-14 days	0.0431	0.0568	0.0690	0.0823	0.0918	0.0985	
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	
15-22 days	0.0216	0.0333	0.0433	0.0551	0.0639	0.0706	
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	
23-30 days	0.0070	0.0141	0.0201	0.0287	0.0352	0.0408	
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	
W/ feedback							
1-7 days	0.0000	0.2778	0.0275	0.0021	0.0002	0.0000	
	(0.000)	(0.032)	(0.007)	(0.001)	(0.000)	(0.000)	
8-14 days	0.0000	0.7598	0.0946	0.0133	0.0018	0.0003	
	(0.000)	(0.107)	(0.020)	(0.004)	(0.001)	(0.000)	
15-22 days	23.4131	0.9940	0.1265	0.0180	0.0027	0.0005	
	(2.141)	(0.148)	(0.027)	(0.006)	(0.001)	(0.000)	
23-30 days	72.6749	1.2578	0.1706	0.0247	0.0037	0.0006	
	(10.538)	(0.202)	(0.041)	(0.009)	(0.002)	(0.000)	

WEB APPENDIX TABLE 13 (cont.)

Drinking alcohol - Younger siblings

	$t^2_{min}-1$	t^2_{min}	$t^2_{min}+1$	$t^2_{min}+2$	$t^2_{min}+3$	$t^2_{min}+4$	$t^2_{min}+5$
1-7 days		0.3016	0.3150	0.3464	0.3703	0.3822	0.3910
		(0.011)	(0.009)	(0.010)	(0.012)	(0.014)	(0.017)
8-14 days		0.0362	0.0494	0.0631	0.0748	0.0837	0.0906
		(0.002)	(0.003)	(0.005)	(0.007)	(0.009)	(0.011)
15-22 days		0.0168	0.0285	0.0397	0.0493	0.0575	0.0644
		(0.002)	(0.003)	(0.005)	(0.007)	(0.009)	(0.012)
23-30 days		0.0050	0.0119	0.0185	0.0249	0.0301	0.0355
		(0.001)	(0.002)	(0.004)	(0.006)	(0.008)	(0.011)
W/ feedback							
1-7 days		0.1144	0.0150	0.0017	0.0002	0.0000	0.0000
		(0.157)	(0.026)	(0.004)	(0.001)	(0.000)	(0.000)
8-14 days		0.3117	0.0334	0.0053	0.0008	0.0002	0.0000
		(0.372)	(0.064)	(0.014)	(0.003)	(0.001)	(0.000)
15-22 days		0.4321	0.0402	0.0049	0.0005	0.0000	0.0000
		(0.509)	(0.082)	(0.017)	(0.004)	(0.001)	(0.000)
23-30 days		0.6191	0.0366	0.0032	0.0004	0.0000	0.0001
		(0.728)	(0.093)	(0.018)	(0.004)	(0.001)	(0.000)
W/out feedback							
1-7 days		0.1144	0.0202	0.0030	0.0005	0.0001	0.0000
		(0.157)	(0.025)	(0.004)	(0.001)	(0.000)	(0.000)
8-14 days		0.3117	0.0479	0.0102	0.0020	0.0005	0.0001
		(0.372)	(0.063)	(0.013)	(0.003)	(0.001)	(0.000)
15-22 days		0.4321	0.0600	0.0116	0.0024	0.0004	0.0001
		(0.509)	(0.080)	(0.016)	(0.003)	(0.001)	(0.000)
23-30 days		0.6191	0.0657	0.0119	0.0026	0.0006	0.0001
		(0.728)	(0.090)	(0.016)	(0.004)	(0.001)	(0.000)

WEB APPENDIX TABLE 13 (cont.)

Smoking marijuana- Older siblings

	$t_{\min}^2-1$	$t_{\min}^2$	$t_{\min}^2+1$	$t_{\min}^2+2$	$t_{\min}^2+3$	$t_{\min}^2+4$	$t_{\min}^2+5$
Baseline							
1-7 days	0.0881	0.0883	0.0907	0.0897	0.0845	0.0805	
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	
8-14 days	0.0129	0.0137	0.0143	0.0143	0.0134	0.0127	
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	
15-22 days	0.0189	0.0209	0.0222	0.0224	0.0209	0.0197	
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
23-30 days	0.0283	0.0404	0.0461	0.0470	0.0434	0.0408	
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	
W/ feedback							
1-7 days	0.0000	0.7693	0.0857	0.0139	0.0033	0.0008	
	(0.000)	(0.095)	(0.023)	(0.007)	(0.002)	(0.001)	
8-14 days	0.0000	1.0609	0.1718	0.0394	0.0110	0.0035	
	(0.000)	(0.153)	(0.046)	(0.018)	(0.009)	(0.005)	
15-22 days	26.7656	1.1969	0.2199	0.0548	0.0143	0.0043	
	(2.857)	(0.176)	(0.053)	(0.020)	(0.008)	(0.004)	
23-30 days	17.8476	1.4203	0.3163	0.0829	0.0246	0.0072	
	(1.865)	(0.193)	(0.066)	(0.024)	(0.010)	(0.004)	

WEB APPENDIX TABLE 13 (cont.)

Smoking marijuana - Younger siblings

	$t^2_{\min}-1$	$t^2_{\min}$	$t^2_{\min}+1$	$t^2_{\min}+2$	$t^2_{\min}+3$	$t^2_{\min}+4$	$t^2_{\min}+5$
1-7 days		0.0874	0.0958	0.1010	0.1036	0.1051	0.1040
		(0.005)	(0.006)	(0.007)	(0.009)	(0.011)	(0.013)
8-14 days		0.0126	0.0148	0.0163	0.0172	0.0178	0.0179
		(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)
15-22 days		0.0180	0.0226	0.0254	0.0272	0.0284	0.0288
		(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.007)
23-30 days		0.0259	0.0415	0.0527	0.0599	0.0650	0.0688
		(0.003)	(0.005)	(0.010)	(0.015)	(0.021)	(0.029)
W/ feedback							
1-7 days		0.1776	0.0222	0.0033	0.0005	0.0001	0.0000
		(0.153)	(0.028)	(0.007)	(0.002)	(0.001)	(0.001)
8-14 days		0.2434	0.0441	0.0097	0.0023	0.0004	-0.0002
		(0.211)	(0.055)	(0.018)	(0.008)	(0.004)	(0.002)
15-22 days		0.2844	0.0539	0.0108	0.0029	0.0009	0.0004
		(0.235)	(0.056)	(0.019)	(0.009)	(0.003)	(0.002)
23-30 days		0.3895	0.0714	0.0167	0.0039	0.0010	0.0002
		(0.317)	(0.084)	(0.027)	(0.009)	(0.003)	(0.001)
W/out feedback							
1-7 days		0.1776	0.0258	0.0049	0.0011	0.0003	0.0001
		(0.153)	(0.023)	(0.005)	(0.002)	(0.001)	(0.000)
8-14 days		0.2434	0.0480	0.0119	0.0028	0.0010	0.0005
		(0.211)	(0.047)	(0.012)	(0.005)	(0.002)	(0.001)
15-22 days		0.2844	0.0599	0.0141	0.0046	0.0015	0.0004
		(0.235)	(0.052)	(0.015)	(0.006)	(0.002)	(0.001)
23-30 days		0.3895	0.0819	0.0229	0.0067	0.0021	0.0006
		(0.317)	(0.075)	(0.022)	(0.007)	(0.002)	(0.001)

WEB APPENDIX TABLE 14

Estimates of Dynamic Probit Model Allowing the Sibling Effect to Depend on the Gender Mix (Error B)

		Smoking	Drinking	Marijuana
State dependence	Older Sibling (γ^1)	0.832 ***	0.636 ***	0.633 ***
		(0.067)	(0.056)	(0.066)
	Younger sibling (γ^2)	0.912 ***	0.613 ***	0.714 ***
		(0.073)	(0.059)	(0.068)
Sibling's influence				
Initial condition	Brothers $(\lambda^2_{mm,0})$	-0.163	0.255	0.193
		(0.236)	(0.189)	(0.217)
	Sisters $(\lambda^2_{ff,0})$	0.200	0.229	0.469 *
		(0.256)	(0.196)	(0.244)
	Mixed Pair $(\lambda^2_{mf,0})$	0.004	0.286 *	0.185
		(0.208)	(0.154)	(0.174)
Later periods	Brothers (λ^2_{mm})	0.028	0.017	-0.177
		(0.121)	(0.105)	(0.127)
	Sisters (λ^2_{ff})	0.331 **	0.113	0.145
		(0.138)	(0.109)	(0.151)
	Mixed Pair (λ^2_{mf})	0.091	0.019	-0.012
		(0.097)	(0.080)	(0.093)
Family-specific error term				
Standard deviation (σ_ε)		0.825 ***	0.761 ***	0.526 ***
		(0.126)	(0.109)	(0.102)
Older sibling, later periods (α^1)		1.103 ***	0.859 ***	1.260 ***
		(0.189)	(0.145)	(0.244)
Younger sib, initial period (α^2_0)		1.052 ***	0.814 ***	1.293 ***
		(0.296)	(0.188)	(0.354)
Younger sib, later periods (α^2)		0.741 ***	0.749 ***	1.219 ***
		(0.172)	(0.156)	(0.355)
Individual-specific error term				
Standard deviation (σ_v)		0.754 ***	0.454 ***	0.555 ***
		(0.071)	(0.067)	(0.077)
Older sib, later periods (δ^1)		1.464 ***	1.207 ***	1.397 ***
		(0.177)	(0.269)	(0.259)
Younger sib, later periods (δ^2)		1.343 ***	1.720 ***	1.353 ***
		(0.158)	(0.277)	(0.255)
Log likelihood value		-7482.66	-8365.49	-6925.83

Note: See Table 4b. In this specification, $\alpha^1_0, \delta^1_0, \delta^2_0$ are normalized to 1.

WEB APPENDIX TABLE 15

Estimates of Dynamic Probit Model Allowing the Sibling's Influence to Depend on the Age Gap (Error B)

		Smoking	Drinking	Marijuana
State dependence	Older Sibling (γ^1)	0.833 ***	0.636 ***	0.632 ***
		(0.067)	(0.056)	(0.066)
	Younger sibling (γ^2)	0.910 ***	0.612 ***	0.712 ***
		(0.073)	(0.059)	(0.068)
Sibling's influence				
Initial condition	Main effect (λ_0^2)	-0.001	0.243 *	0.207
		(0.185)	(0.136)	(0.158)
	Age gap > 2 yrs $(\lambda_{2+,0}^2)$	0.047	0.127	0.114
		(0.210)	(0.184)	(0.216)
Later periods	Main effect (λ^2)	0.106	0.034	-0.076
		(0.082)	(0.065)	(0.079)
	Age gap > 2 yrs (λ_{2+}^2)	0.063	0.018	0.178
		(0.115)	(0.096)	(0.142)
Family-specific error term				
Standard deviation (σ_ε)		0.821 ***	0.759 ***	0.524 ***
		(0.124)	(0.108)	(0.103)
Older sibling, later periods (α^1)		1.113 ***	0.865 ***	1.266 ***
		(0.189)	(0.145)	(0.248)
Younger sib, initial period (α_0^2)		1.034 ***	0.805 ***	1.304 ***
		(0.284)	(0.186)	(0.360)
Younger sib, later periods (α^2)		0.754 ***	0.754 ***	1.219 ***
		(0.174)	(0.156)	(0.352)
Individual-specific error term				
Standard deviation (σ_v)		0.755 ***	0.455 ***	0.555 ***
		(0.071)	(0.066)	(0.077)
Older sib, later periods (δ^1)		1.458 ***	1.197 ***	1.398 ***
		(0.178)	(0.269)	(0.262)
Younger sib, later periods (δ^2)		1.343 ***	1.722 ***	1.358 ***
		(0.158)	(0.274)	(0.254)
Log likelihood value		-7486.32	-8366.92	-6929.61

Note: See Table 4b. In this specification, $\alpha_0^1, \delta_0^1, \delta_0^2$ are normalized to 1.