March, 2004

Searching for Peer Group Effects: A Test of the Contagion Hypothesis

by

Laura M. Argys
University of Colorado, Denver
Department of Economics, CB 181
P.O. Box 173364
Denver, CO 80217-3364
Phone: (303) 556-3949 Fax: (303) 556-3547
E-mail: Laura.Argys@cudenver.edu

Daniel I. Rees
University of Colorado, Denver
Department of Economics, CB 181
P.O. Box 173364
Denver, CO 80217-3364
Phone: (303) 556-3348 Fax: (303) 556-3547
E-mail: DREES@carbon.cudenver.edu

Searching for Peer Group Effects: A Test of the Contagion Hypothesis

Abstract

Using state-level variation in kindergarten start dates coupled with information on birth and interview dates to generate an exogenous measure of the relative age of a student's peer group, we find that, controlling for age, females with older peers are more likely to use substances than females with younger peers. In contrast, there is little evidence that having older peers is related to the risky behavior of male adolescents. Because there is no reason to suspect that birth and kindergarten start dates should be correlated with the choice of school, the socioeconomic status of a child's peers, or neighborhood unobservables, we view our results with regard to females as providing support for the idea that peer behavior can be contagious.

1. Introduction

According to what have been labeled "epidemic models" of behavior, children are directly influenced by the aspirations and actions of their peers. Individuals growing up in a neighborhood or attending a school in which a particular behavior is prevalent are themselves more likely to engage in that behavior. Although the notion of contagious behavior has an intuitive appeal, attempts to document the precise role played by peers in the determination of child outcomes have not been entirely successful.

A researcher intent on testing the contagion hypothesis is faced with at least two major obstacles. The first has to do with the issue of selection. It has long been recognized by economists that families and children have some control over their environment. If unobserved parental inputs to a child's education or upbringing (such as effort or the level of supervision) are correlated with the choice of neighborhood and school, then the results from studies that treat peer behavior as exogenous are suspect.

In addition to addressing the selection issue, the researcher must distinguish the direct influence of peer behavior from what Manski (1993) termed "contextual effects" (or, in other words, the effects of predetermined peer characteristics), and from the effects associated with a shared environment. For instance, a correlation between the probability that a child smokes and the percentage of his or her peers who smoke may be evidence of a contagion effect, but it is also possible that this correlation is driven by peer socioeconomic status or by shared neighborhood attitudes towards smoking. As noted by Manski (1993) and Gaviria and Raphael (2001) distinguishing between these effects may have important policy implications.

In this research we analyze data drawn from the 1997 National Longitudinal Survey of Youth to test the contagion hypothesis. Our empirical strategy is inspired by the work of Angrist and Krueger (1991), and designed to avoid the methodological pitfalls described above. Specifically, we utilize state-level variation in kindergarten start dates coupled with information on birth and interview dates to generate an exogenous measure of the relative age of a student's peer group. We hypothesize that individuals who are young as compared to their peers will begin certain risky behaviors earlier than their counterparts who are old as compared to their peers.

Our results suggest that having older peers is associated with an increased likelihood that female adolescents use marijuana, drink alcohol, and smoke cigarettes. In contrast, we find little evidence that male adolescents with older peers are more likely to engage in these behaviors. Because there is no reason to suspect that the interaction of birth and kindergarten start dates should be correlated with the choice of school, the socioeconomic status of a child's peers, or neighborhood unobservables, we view our results with regard to females as providing support for the idea that peer behavior can be contagious.

2. Background

Researchers working in this area have typically assumed that peer characteristics and behavior can be treated as exogenous.[1] However, as noted above, if parents select their neighborhood (or the school to which they send their children) in part based on the characteristics of other children in the neighborhood (or other students in the school), then this assumption can lead to biased estimates of peer group effects. Evans et al. (1992) argued that the most likely direction of this bias is upwards. This is because parents who care enough to try to locate in a neighborhood where their children's peers will be of "high quality," are also more likely to possess unobservable attributes that contribute to the success of their offspring.

One possible solution to the problem of selection is to utilize an extensive set of family background controls in order to minimize the impact of parental unobservables. For instance, in addition to measures of parental income, occupation, and education, Zax and Rees (2002) included parental attitudes towards college as explanatory variables in their study. A similar approach was taken by Datcher (1982) who included measures of parental "feelings of efficacy

[1]Jencks and Mayer (1990) provide a thorough review of the earlier empirical work on peer effects.

and ambition" (p. 34) as explanatory variables. In general, additional controls at the family-level tend to be associated with smaller peer-effect estimates.[2]

An alternative solution is to explicitly model the choice of school or neighborhood. Evans et al. (1992) found that, if treated as exogenous, the percentage of economically disadvantaged students at a school was a good predictor of teen pregnancies and dropout behavior. However, this effect completely disappeared when the authors addressed the selection issue using a simultaneous equations framework. Although their choice of instruments has been questioned (Sacerdote 2001; Krauth 2002), the results presented by Evans et al. suggest that the upward bias due to selection can be severe.

In more recent work, Gaviria and Raphael (2001) attempted to address the selection issue by comparing peer effect estimates for long-term residents of a neighborhood with those of newcomers. Following Glaeser (1996), they reasoned that the selection problem should be more severe for newcomers than for long-term residents, who presumably made their location decisions taking into account past, as opposed to present, peer and school characteristics. They found that peer drug use had a larger effect for newcomers than for long-term residents, but found no differences with regard to peer drinking, smoking, dropping out or church attendance. They concluded that there was "no evidence of bias" for most of the outcomes analyzed (Gaviria and Raphael 2001, p. 266).

Even in the absence of selection it has, in practice, been difficult to distinguish the effects of current peer behavior from the predetermined characteristics of a child's peers or neighborhood. Gaviria and Raphael (2001) argued that because their measure of peer behavior was at the school as opposed to the neighborhood level, they could safely assume that peer

[2]In their review of the literature in this area, Jencks and Mayor (1990, p. 176) wrote,

> As a rule, the more aspects of family background we control, the smaller neighborhood and school effects look. Initially, for example, we thought that attending a low-SES High school substantially reduced twelfth graders' chances of attending college. Today, using more elaborate background measures we are reasonably certain that the effect is trivial.

background and neighborhood influences were of secondary importance. This approach, although innovative, is not entirely satisfactory. Many researchers have equated high school attendance areas with neighborhoods (Jencks and Mayer 1990, p. 112), and there is no way to be certain that Gaviria and Raphael's estimates are not a reflection of shared environmental factors (Eisenberg 2003, p. 9).

Another vein of research in this area utilizes roommate data from universities in an effort to pin down peer effects. Many universities randomly assign roommates in their freshman year. This natural experiment has allowed researchers to investigate the effect of one roommate's behavior and background on the other's without having to worry about selection or unobservables associated with having grown up in the same neighborhood.

Using data from Dartmouth College, Sacerdote (2001) found that a student's academic performance (as measured by their freshman year GPA) was positively related to his or her randomly assigned roommate's academic performance. He also found that a student's decision to join a fraternity was related to his or her roommate's decision to join a fraternity, although roommate background characteristics such as SAT scores and high school rank were weak predictors of these same outcomes. This pattern of results can be interpreted as evidence that contextual effects as defined by Manski (1993) are less important than contagion effects.[3]

Kremer and Levy (2001) examined roommate data from a "large, academically strong, state university." They found that males who were assigned roommates who drank in high school suffered a decrease in academic performance (as measured by GPA), although there was no effect of roommate high school grades or admission test scores. Zimmerman (2003) examined roommate data from Williams College. In contrast to Kremer and Levey's results, he found that roommate admission test scores did have an effect on academic performance. Specifically, he found that being assigned a roommate who performed poorly on the verbal section of the SAT

[3]In an earlier version of his 2001 article, Sacerdote wrote, "[t]he coefficient on roommate high school academic index is small and insignificant. The implication is that while there is a significant peer effect, it does not work through roommate's background. Instead the peer effect works through the roommates behavior and outcomes while at Dartmouth" (Sacerdote 1999, p. 19).

was associated with a decrease in cumulative GPA for students who scored in the middle 70% of the SAT distribution.[4]

The advantage of utilizing roommate data is that it does not rely on what can be interpreted as arbitrary exclusion restrictions, nor does it assume that the contextual effect is zero as did Gaviria and Raphael (2001). The disadvantage is that the results may not hold for the general population. Although the work by Sacerdote provides some evidence that contagion effects exist between university roommates, the profession still lacks conclusive evidence that contagion effects exist in other settings, for instance among high school students.[5] Much of the debate over vouchers and other educational reforms hinges on whether such effects exist, and a growing theoretical literature in economics is built upon the assumption of positive spillover effects within primary and secondary schools and within neighborhoods. For instance, de Bartolome (1990) showed that family location and community expenditure decisions will be inefficient when the standard Tiebout model is modified to include "large, favorable" peer effects on student achievement. Epple and Romano (1998) and Caucutt (2002) explored the effects of introducing educational vouchers on student welfare and enrollment patterns under the assumption of positive peer group effects.

3. Data and Research Strategy

The data for this analysis come from the first three rounds of the National Longitudinal Survey of Youth - 1997 cohort (NLSY97). The NLSY97 consists of a representative sample of the U.S. population between the ages of 12 and 16 on December 31, 1996, and a supplemental

---

[4]Neither Kremer and Levy(2001) nor Zimmerman (2003) attempted to distinguish between contextual and contagion effects.

[5]Hoxby (2000) documented the existence of peer effects on achievement among third through sixth graders in Texas. Hoxby identified exogenous variation in peer test scores based on year to year changes in gender and racial composition within elementary schools. She found evidence that peer test scores affect achievement, but did not focus on the "channels" through which these peer effects operate. These channels include students teaching one another, "but may even work through the way in which teachers or administrators react to students" (Hoxby 2000, p. 6). Evidence of similar peer effects on student achievement is provided by Hanushek et al. (2001).

over-sample of Hispanics and blacks belonging to this same cohort. It contains detailed information on family background, personal characteristics, and a variety of behaviors that can be considered risky or delinquent.

Because our focus is on the influence of grade-level peers, we restrict our analysis to respondents in grades 6 through 12 at the time of their interview. Most of our sample can be observed in each of the three waves under study (1997, 1998, and 1999). However, some individuals are observed less than three times due to natural attrition from the NLSY97, while others aged into or out of the sample.

Figures 1 through 4 show the proportion of NLSY97 respondents who engage in four risky behaviors by age. They demonstrate a strong positive link between these behaviors and age among adolescents living in the United States in the late 1990s. For instance, the probability of marijuana use increases, on average, by .034 for every year of adolescence. Or, to take another example, the probability that an adolescent is sexually active increases, on average, by .123 per year.[6] The positive relationship between age and risky behavior is evident whether or not one controls for personal characteristics, family background, or family structure. It exists for males and females, for blacks as well as whites and Hispanics, for children living in urban areas and children living in rural areas, and for adolescents in public and private schools. In fact, no matter how we cut the data, the basic age patterns shown in Figures 1 through 4 hold.[7]

The positive relationship between age and these four behaviors suggests that children will be exposed to different environments depending on the relative age of their classmates. For instance, if an experimenter placed a child chosen at random with older classmates he or she

[6]These probabilities are estimated coefficients from linear probability models of the following form:

$$R_i = \beta' Age_i + \varepsilon_i,$$

where R_i is equal to 1 if the individual engaged in the risky behavior in question and equal to 0 otherwise, and age is measured in years. Figures 1 through 4 show the proportion of teens who engage in these behaviors at each age.

[7] In keeping with these findings, the introduction of Gruber (2001) documents a positive relationship between grade and substance use among U.S. adolescents using data from the Youth Risk Behavior Survey.

would be exposed to more substance use than a similar child who was placed with younger classmates. If the contagion hypothesis is correct, then we would expect the first child to have a greater probability of using substances than the second.

Our research strategy mimics the experiment described above by taking advantage of state variation in statutorily determined kindergarten start dates. Between 1985 and 1989, the relevant time period for our sample, forty-three states and the District of Columbia used state-wide kindergarten start dates to determine when a child was eligible to begin kindergarten.[8] Children in these states could not enroll in public school unless they turned five by the dates specified in Table 1. Thus, abstracting for the moment from complicating factors such as grade retention and movement between states, one would expect a child born in California on December 1 to begin kindergarten the year he turned five, and be the youngest in his grade. Another child in California, born one day later, would start school one year later, when she was almost six years old, and be 364 days older than the youngest student in her grade. We hypothesize that children who were born shortly before their state's kindergarten start date, and who are therefore relatively young as compared to their classmates, will be more likely to engage in certain risky behaviors at every age than children born shortly after their state's kindergarten start date. In effect, we are arguing that the interaction of birth and kindergarten start dates sorts students into different peer groups.[9]

To explore whether this sorting is random, we created a variable equal to the difference between an individual's birth date and the kindergarten start date in his or her state of

---

[8]Seven states had start dates that were determined at the local level. Information on kindergarten start dates was obtained from the Education Commission of the States (1985), Wolf and Kessler (1987), and direct communications between the authors and the Idaho Department of Education, the Illinois State Board of Education, the Missouri Department of Elementary and Secondary Education, the Rhode Island Department of Elementary and Secondary Education, and the Utah State Office of Education.

[9]Eisenberg (2003) examined a similar natural experiment. Specifically, he noted that most 7th and 8th graders in the United States attend schools made up of 6th to 8th graders, but some attend schools with a wider range of grades (for instance, grades 7 through 12). Because "substance use rates increase significantly with grade level" (p.3), Eisenberg proposed using this variation in grade span as an instrument for peer behavior, but found no evidence of contagion effects. His approach utilized a broader definition of peers than we use here.

residence.[10] As shown in Figure 5, the distribution of the variable is almost uniform and is, therefore, consistent with random assignment.

Table 2 shows the results of a simple regression in which the number of days by which the child's birth date exceeds the kindergarten start date is regressed on measures of race and ethnicity, family structure, parental education, urban status, and school characteristics. If the sorting of children into peer groups is truly random, then we would expect these measures to have little explanatory power. For the most part this is the case, suggesting that we have identified what can be thought of as an exogenously determined variable. However, there is evidence of a negative relationship between being in high school and the birth/start date difference, suggesting that some children born close to the kindergarten start date began high school later than would be expected based on their age and the kindergarten start date in their state of residence.[11] The issue of grade retention is discussed below.

To control for the possibility that the interaction of birth and kindergarten start dates is somehow connected to socioeconomic status, we control for family background and personal characteristics in the regressions below. We also control for whether a student attended high school, whether a student attended a public (as opposed to private) school, and urban status. Estimates based on specifications without these controls produced very similar results and are available upon request.

One potential complicating factor in our analysis comes from the fact that most states have start dates in the fall.[12] Therefore, the difference between a child's birth and kindergarten start dates is potentially correlated with a child's own age as well as the age of his or her classmates. To avoid the obvious problems this correlation would entail, we control for age in

[10]Students who lived in states in which kindergarten start dates were determined at the local level were not included in this or subsequent analyses.

[11]When these regressions are run separately by gender, the coefficient of "High School" is statistically significant only for males.

[12]As shown in Table 1 there is a six month gap between the earliest kindergarten start date (July 1) and the latest (January 1).

months, which is identified both through variation in kindergarten start dates and through variation in interview dates.[13] Controlling for age in days does not change the basic pattern of results presented below.

Another complicating factor is that many children are not in the grade one would predict based on their current state of residence and age. This could come about for a variety of reasons. For instance, students who begin their schooling in one state but then move to another are not expected to repeat a grade if, by chance, they would have started kindergarten a year later in their adopted state; students who attend private schools or switch from private schools to public schools are not subject to the kindergarten start dates; parents can choose to delay their children's entry into kindergarten; and students may be held back or (less frequently) advanced a grade based on academic performance and other factors.

In the analysis below we include controls for whether a child is in their expected grade based on the kindergarten start date in the child's current state of residence and the child's age. If retention/advancement decisions were uncorrelated with a child's underlying propensity to engage in risky behavior and uncorrelated with the difference between the child's birth and kindergarten start dates, then one would expect a student placed above their predicted grade to have greater exposure to age-dependent behaviors such as substance use, and therefore a higher probability of engaging these same behaviors, controlling for absolute age. Likewise, one would expect an individual in a grade below their predicted grade to be less likely to engage in these behaviors. Moreover, estimates of the contagion effect based on children in the "correct" grade would be unbiased. If, however, grade placement is related to both the child's relative age and his or her underlying propensity to engage in risky behaviors, then a potentially important selection issue emerges.

There is strong evidence that retention decisions are, at least in part, based on relative age. For instance, Eide and Showalter (2001) show that children born 90 days before the

[13]During round 1, interviews spanned more than 12 months, but in the latter two rounds, all interviews took place during the school year, between September and May.

kindergarten cutoff date are much more likely to be held back as compared to their older peers in the same grade.[14] In addition, there is evidence that children from poorer backgrounds, racial minorities, and children who exhibit "problem" behaviors are more likely to be held back (Alexander et al. 2003, Chapter 5). If, as seems plausible, these same students are particularly prone to engaging in risky behaviors when they reach adolescence, then our research strategy may produce estimates of the contagion effect that are biased downwards. This issue is potentially more serious for males than for females: over one third of the males in our sample are in a grade lower than would be expected based on their age and the kindergarten start date in their state of residence, compared to just over one fourth of the females.

4. The Empirical Model

Imagine that R_i^* represents an individual's propensity to engage in a particular risky activity, and that this propensity is related to a set of control variables, X_i, and the individual's relative age by the following equation:

$$(1) \quad R_i^* = \beta' X_i + \pi_1 Younger_i + \pi_2 Behind_i + \pi_3 Ahead_i + \varepsilon_i,$$

where Xi includes measures of age (in months), race and ethnicity, parental education, family structure, and a dichotomous variable equal to 1 if the student attended high school, and equal to 0 otherwise; "Younger" is a dichotomous variable equal to 1 if the respondent was born 182 days or less before the kindergarten start date in his or her state of residence, and equal to 0 otherwise; "Behind" is a dichotomous variable equal to 1 the respondent is in a lower grade than would be predicted based on his or her age and kindergarten start date, and equal 0 otherwise; "Ahead" is a dichotomous variable equal to 1 the respondent is in a higher grade than that predicted, and equal

[14]Our data show that children born soon after the cutoff date are more likely to be in a higher grade than would be expected based on their age and state kindergarten start date.

to 0 otherwise; and ε_i, is a random error term.[15] Although an individual's propensity to engage in risky behavior is latent, when $R_i^* > 0$ an indicator variable, R_i, is observed to equal 1, so that Prob ($R_i = 1$) = Prob ($\beta'X_i + \pi_1 Younger_i + \pi_2 Behind_i + \pi_3 Ahead_i + \varepsilon_i > 0$). If the error term is normally distributed, then the result is a standard single-equation probit model.

Our primary focus throughout the paper is on the estimate of π_1. A positive estimate might be viewed as evidence of a contagion effect. It would indicate that being, on average, 0.5 years younger than the other students in your grade, is associated with the adoption of the risky behavior under study.[16] A non-positive estimate could indicate that the behaviors under study are not subject to peer influences, or that grade retention and advancement decisions are creating a downward bias.

As noted in the previous section, we would expect the estimate of π_2 to be negative because being held back should result in less exposure to risky behaviors such as substance use and sexual activity. However, if problem children are more likely to be held back (and more likely to engage in these behaviors) then the estimate of π_2 may be positive. Likewise, the estimate of π_3 is expected to be positive to the extent that skipping a grade results in greater exposure to risky behaviors, but if the best-behaved students are skipped ahead then the opposite result may occur.

The basic model described above can be modified by interacting a student's relative age with the variables "Behind" and "Ahead." This modification allows us to interpret the coefficient of π_1 as the effect of having older peers only for those students in the "correct" grade:

---

[15]Table 1 of the appendix presents descriptive statistics by gender and the dichotomous variable "Younger." It is clear from this table that students born 0 to 182 days before their state's kindergarten start date are much more likely to be in a lower grade than would be expected based on their age and kindergarten start dates. These same students are much less likely to be in a higher grade than would be expected based on their age and kindergarten start dates.

[16]Students in the "Younger" category were, on average, born .24 years before the kindergarten start date in their state of residence. Students in the "Older" category were born, on average, .760 years before the kindergarten start date in their state of residence. The difference between these figures is actually .52 years. See Table 1 of the appendix.

$$(2) \quad R_i^* = \beta'X_i + \pi_1 Younger_i + \pi_2 Behind_i + \pi_3 Ahead_i + \pi_4 Younger_i * Behind_i + \pi_5 Younger_i * Ahead_i + \varepsilon_i.$$

The coefficients of the interaction terms, π_4 and π_5, can be interpreted as capturing the effect of being born less than 183 days before the kindergarten start date for students who were either held back or skipped ahead.

5. The Analysis

Tables 3 and 4 report estimated marginal probabilities and robust standard errors for the empirical models outlined above in which four outcomes are considered: the probability of having used marijuana, the probability of having consumed alcohol, the probability of having smoked cigarettes, and the probability of having had sex.[17] The first three of these outcomes are based on questions pertaining to the month prior to the interview, whereas the measure of sexual activity is based on reported behavior in the year prior to the interview.[18] Results for males and females are presented separately to allow for gender differences in the relationship between relative age and risky behavior.

In general, the results for males are inconclusive. The estimates of π_1 are consistently small and statistically insignificant; they provide no evidence that males who were relatively young for their grade modified their behavior based on the behavior of their older peers. In contrast, we find evidence that adolescent females may be influenced by the behavior of their

[17]Because respondents can appear in the sample as many as three separate years (1997, 1998, 1999), standard errors were corrected for clustering at the individual level.

[18] Specifically, respondents were asked:

1) "During the past 30 days, on how many days did you smoke a cigarette?"
2) "During the past 30 days, on how many days did you have one or more drinks of an alcoholic beverage?"
3) "On how many days have you used marijuana in the past 30 days?"
4) "About how many times have you had sexual intercourse in the past 12 months?"

Table 2 of the appendix reports means for the outcome variables used in the analysis. In general, the responses to these questions are in keeping with what we know from other national surveys of adolescent behavior done in the mid-90s. See, for instance, Gruber (2001).

older peers.

Focusing first on the model without interactions, females who were born 0 to 182 days before the kindergarten start date in their state of residence were more likely to drink and more likely to smoke than their counterparts born 183 to 364 days before the kindergarten start date. Specifically, being younger than one's "average" peer is associated with a .025 increase in the probability of alcohol use and a .024 increase in the probability of tobacco use. It is also associated with a .011 increase in the probability of marijuana use and a .019 increase in the probability of sexual activity, but these estimates are not significant at conventional levels (p-values = .14 and .21, respectively).

The model with interactions produces similar estimates. Being younger than one's "average" peer is associated with a .019 increase in the probability of marijuana use, a .026 increase in the probability of drinking, and a .031 increase in the probability of smoking. Again, the marginal effect of "Younger" in the sexual activity equation is positive but not precisely estimated (p-value = .19). It suggests that females born 0 to 182 days before the kindergarten start date were 2.4 percentage points more likely to be sexually active.

Upon first inspection these estimates may seem modest in terms of magnitude. However, they represent sizable increases as compared to, for instance, the effect of aging. Figure 1 showed that a one-year increase in age was associated with a .034 increase in the probability of using marijuana. Thus, the estimates contained in Table 3 suggest that having peers who are on average .5 years older is associated with the same change in the probability of marijuana use as aging .32 to .56 years. To take another example, having peers who are on average .5 years older is associated with the same change in the probability of smoking cigarettes as aging .44 to .56 years.

Moreover, if one is interested in the effect of skipping (or being held back) an entire grade, then these estimates should be doubled. They suggest that being placed among peers who are, on average, a full year older would cause an increase in the probability of marijuana use of .022 to .038, an increase in the probability of drinking of .050 to .052, an increase in the probability of smoking of .048 to .062, and an increase in the probability of being sexually active of .038 to .048.

Of course, the estimates of π_2 and π_3 provide an alternative method within this empirical framework for judging the impact of skipping a grade on risky behavior. The results show that males in a lower grade than expected were more likely to engage in a number of risky behaviors than their counterparts in the "correct" grade, a pattern of results consistent with the selection hypothesis discussed earlier. For females, the estimates of these parameters are more in keeping with the existence of peer effects.

Focusing on the model without interactions, females in a grade higher than would be expected based on their age and kindergarten start dates were, on average, .060 more likely to have smoked cigarettes. This figure is approximately twice the size of the marginal effect for "Younger" in the same equation. Thus, both estimates suggest similar effects of having peers who were, on average, one year older.

Likewise, the results suggest that females in a grade lower than would be predicted were, on average, .040 less likely to have consumed alcohol. This figure is also approximately twice the size of the corresponding marginal effect.[19] Finally, the results suggest that being in a grade higher than would be predicted based on a respondent's age and the kindergarten start date is associated with a .086 increase in the probability of having had sex. This figure is larger than would be expected based on the marginal effect of "Younger," but still suggests that females who are placed with younger peers modify their behavior to conform to the group norm.

6. Delinquent Activities

Up to this point in the analysis we have concentrated on adolescent substance use, including tobacco, and sexual activity. Each of these behaviors is positively related to age, as demonstrated in Figures 1 through 4, which allowed us to test the hypothesis that students adopt

[19]In neither case can we reject the hypothesis that the estimated coefficient is twice that of "Younger."

the behavior of their older peers.

In this section we turn our focus to an alternative set of outcomes, namely vandalism, theft, and fighting/assault.[20] Figures 6 through 8 show the relationship between these delinquent acts and age in the NLSY97. In each case, we fail to find the strong positive relationship found for substance use and sexual behavior.

For instance, the probability that the typical adolescent destroys property actually decreases by one percentage point for every year of adolescence. The probability of committing theft increases with age, but this increase is very small: only one-tenth of a percentage point for every year of adolescence. Likewise, the probably of being involved in a serious fight or committing assault increases by only 3 tenths of a percentage point per year.

This pattern suggests that, with regard to delinquent acts, children will *not* be exposed to substantially different environments depending on the relative age of their classmates, and therefore we cannot document a contagion effect using these outcomes. If we find evidence that the interaction of a child's birth and kindergarten start dates is related to the probability of engaging in delinquent acts, then this would suggest that the results of the previous section were generated by an alternative process.

Table 5 presents estimates of (1) and (2) in which R^* is replaced by D^*, the propensity to commit a delinquent act. In all other respects the estimating equations are the same as were used in the previous section. In no case do we find that having older peers is associated with the outcomes under study. The marginal effects associated with the variable "Younger" are statistically insignificant and, for the most part, much smaller in magnitude than what we found

---

[20]NLSY97 respondents were asked:

1) "How many times have you purposely damaged or destroyed property that did not belong to you in the last 12 months?"
2) "How many times have you stolen something from a store, person or house, or something that did not belong to you worth 50 dollars or more including stealing a car in the last 12 months?"
3) "How many times have you attacked someone or have had a situation end up in a serious fight or assault of some kind in the last 12 months?"

Mean responses are shown in Table 2 of the appendix.

in the previous section.[21]

In addition, we find only limited evidence that being in the "wrong" grade is associated with the probability that an adolescent commits a delinquent act. The estimates of π_2 through π_5 are generally not statistically significant for females.[22] For males, being in a lower grade than would be predicted based on age and the kindergarten start date in a child's state of residence is associated with large increases in the likelihood of fighting/assault and the probability of engaging in theft. Specifically, being in a lower grade is associated with a .028 increase in the probability of having been in a serious fight in the past year or committing assault, and a .028 increase in the probability of having stolen something. These effects suggest that males selected by parents and educators to be held back are especially prone to engaging in delinquent acts.

In general, the results of this exercise support a peer effect interpretation of the results discussed in the previous section. They suggest that being placed with older peers is associated with female substance use because older peers are more likely to use substances; there is no evidence of a similar effect with regard to delinquent acts because the incidence of these acts is not positively related to age.

It is possible, however, to imagine other factors aside from peer behavior that might be correlated with the variable "Younger." For instance, teachers may treat the youngest students in a particular grade differently than their older peers. Alternatively, there may exist an effect of school grade on the likelihood of engaging in risky behavior not captured by the high school indicator. For instance, in some schools upper classmen are allowed to leave school grounds during the school day, whereas younger students must stay on campus. According to these scenarios, however, one would expect the variable "Younger" to be related to *both* substance use and delinquent behavior, which it is not.

[21]The marginal effect associated with "Younger" in the male vandalism equation is .021, which is comparable in magnitude to the effects we found with regard to substance use and sexual activity for females. This effect, however, is not statistically significant at the .10 level.

[22] Females in a grade below their expected grade are more likely to report having engaged in vandalism than their counterparts in the "correct" grade.

7. Sensitivity Analysis.

Although the results are not reported in Tables 3 through 5, we estimated our model using a number of different sub-samples in an attempt to test the sensitivity of our results. For instance, we divided our sample based on whether the individual attended pubic versus private school, and found some evidence that having older peers is associated with an increase in the probability of substance use for females in public schools, but not for females in private schools where state-specified kindergarten start dates are non-binding. We also found that the effect of older peers on substance use and sexual activity was more precisely estimated for girls in grades 10 through 12 than for girls in grades 6 through 9. When we estimated the parameters of our model only for individuals who were in the grade one would predict based on their age and the relevant kindergarten start date, the estimates were consistent with those reported in Table 3a. We found no evidence of contagion effects among males in any of these sub-samples.

We also experimented with adding state fixed effects and controls for season of birth to the model (Table 6). Adding fixed effects controls for any state-level unobservables correlated with kindergarten cutoff dates and risky behaviors. In this specification, the effect of "Younger" is identified only through variation in birth dates, which are assumed exogenous. The results from this exercise were in keeping with the findings discussed above. That is, females with older peers were more likely to use marijuana, drink, and smoke cigarettes. Again, the relationship between "Younger" and female sexual activity was positive, but not precisely estimated. Controlling for season of birth produced noticeably larger estimates of the relationship between female substance use and relative age than those reported in Table 3a.

Finally, we experimented with alternative measures of the relative age of an adolescent's peers. For instance, we redefined the variable "Younger" to equal 1 if the respondent was born 0 to 91 days before the kindergarten start date in his or her state of residence, and to equal 0 otherwise. We also replaced "Younger" with a continuous measure equal to the difference between an adolescent's birthday and the relevant kindergarten start date. For females, these experiments produced estimates of the effect of relative age that were in keeping with those reported in Tables 3 and 4, although they were less precise. For males, the estimates were consistently smaller than those we found for females and never statistically significant at conventional levels.[23]

8. Conclusion

Researchers have devoted a great deal of effort to documenting the existence of peer effects among adolescents, but have struggled with two potentially important empirical issues. The first of these issues arises because parents have some control over their offspring's environment through school and neighborhood choice. It has been argued that parents who care the most about selecting high-quality schools and neighborhoods will tend to be those who devote more time and energy to raising their children. Because it is difficult to measure parental inputs such as time and energy, researchers may unwittingly attribute their impact to peer influences.

Even controlling for selection of this sort, it is still difficult to distinguish the effects of peer behavior in the current period from pre-determined peer characteristics and difficult-to-observe shared neighborhood influences. For instance, a correlation between the probability that a child smokes and the percentage of his or her peers who smoke may indicate that children adopt the behavior of their peers, but it is also possible that this correlation is driven by peer socioeconomic status or by neighborhood attitudes towards smoking.

The most successful efforts at dealing with these twin issues have relied on data from university roommate assignments. A number of studies have taken advantage of the fact that

[23] These results are available from the authors upon request.

freshman roommate assignment at a number of large universities is random and therefore uncorrelated with parental decisions and neighborhood characteristics. The problem with this particular vein of research is that the results may not be transferable to other settings. A number of recently published articles assume the existence of peer effects within high schools, however empirical evidence for this assumption is sparse.

Here we adopt a new approach to documenting the existence of peer effects. This approach relies on the fact that adolescent substance use and sexual activity are positively related to age, and therefore we would expect that a child randomly placed among older peers to experience greater exposure to these behaviors than a similar child placed with younger peers. In an attempt to mimic an experiment of this sort, we examine the effect of being born 0 to 182 days before the kindergarten start date in a respondent's state of residence. Children born during this period have peers who are, on average, .5 years older than themselves. Children born 183 to 364 days before the kindergarten start date are on average .5 years older than their peers.

Controlling for age in months, we find little evidence that males who were relatively young for their grade were influenced by the behaviors of their older peers. However, females who were relatively young for their grade were significantly more likely to use substances. According to our preferred specification, female adolescents with older peers were 1.9 percentage points more likely to use marijuana, 2.6 percentage points more likely to use alcohol, and 3.1 percentage points more likely to use tobacco. We also find that females with older peers were 2.4 points more likely to have been sexually active in the past year, but this estimate is not statistically significant at conventional levels (p-value = .19).

Although adolescent substance use and sexual activity are positively related to age, other delinquent behaviors such as theft, vandalism and fighting/assault are no more prevalent among 18-year-old females than among 14-year-old females. We would therefore expect that being young for one's grade to be unrelated to these activities. In fact, our results indicate that the probability of engaging in delinquent acts such as theft, vandalism, and fighting/assault is unrelated to the relative age of a respondent's peers.

Because there is no reason to suspect that the difference between birth and kindergarten start dates is linked to socioeconomic status, school quality, or neighborhood unobservables, we view our results as consistent with the existence of a contagion effect. In other words, female adolescents seem to do as their peers do, at least with regard to substance use. Moreover, there is reason to believe that our estimates might be larger in the absence of grade retention.

Approximately 30 percent of our sample was in a grade lower than would be expected based on their age and the kindergarten start date in their state of residence. A number of studies have documented that problem students--students who have trouble socializing, or who have academic difficulties--are much more likely to be held back. Other studies have shown that students near to the cutoff are also more likely to be held back. If, as seems likely, problem students are more likely to use substances and become sexually active at an early age, then our preferred estimates can be thought of as understating the true effect of having older peers because they are based only on students in the "correct" grade. This problem is potentially more serious for males as compared to females: over one third of the males in our sample are in a grade lower than expected based on their age and the kindergarten start dates in their state of residence, as compared to just over one fourth of the females. Although we find no evidence of contagion effects among male adolescents, it is difficult to say whether this is the result of more males being held back than females, or whether males are in some sense better insulated from peer influences than females.[24]

[24] There have been a number of experimental studies reported in the psychology literature that investigate whether susceptibility to peer influences is related to gender. In general, the results from these studies suggest that females

In terms of policy, this work sheds little light on what would be an optimal kindergarten start date for the typical state. However, it does suggest a clear rationale for delaying a child's entry into school, especially females born near the kindergarten start date. To date, most empirical studies of grade retention have found negative effects on outcomes such as high school completion, self-esteem, and academic performance (Eide and Showalter 2001). Our results suggest that, at least for females, attending school with younger peers may lower the probability of engaging in a number of risky behaviors. Whether this benefit outweighs the potential cost in terms of academic performance is an open question.

---

are more susceptible to peer influences than their male counterparts. See, for instance, Steiner (1960), Iscoe et al. (1963), Santee et al. (1982), and Ellis et al. (1991).

References

Alexander, Karl L., Doris R. Entwisle, and Susan L. Dauber. 2003. On the Success of Failure: A Reassessment of the Effects of Retention in the Primary School Grades, Second Edition. Cambridge: Cambridge University Press

Angrist, Joshua D. and Alan B. Krueger. 1991. "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" The Quarterly Journal of Economics, Vol. 105, No. 4 (November), pp. 979-1014.

Benabou, Roland. 1993. "Workings of a City: Location, Education, and Production." The Quarterly Journal of Economics, August, pp. 619-652.

Borjas, George J. 1995. "Ethnicity, Neighborhoods, and Human Capital Externalities." American Economic Review, Vol. 85, No. 3, pp. 365-390.

Datcher, Linda. 1982. "Effects of Community and Family Background on Achievement." The Review of Economics and Statistics, Vol. 64, No. 1 (February), pp. 32-41.

Durlauf, Steven N. 1994. "Spillovers, Stratification, and Inequality." European Economic Review, 38, pp. 836-845.

Education Commission of the States. 1985. Clearinghouse Notes, "State Characteristics: Kindergartens, November 1985." Denver CO: Education Commission of the States.

Eide, Eric R., and Mark H. Showalter. 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." Economics of Education Review, Vol. 20, No. 6 (December), pp. 563-76.

Eisenberg, Daniel. 2003. "Peer Effects for Adolescent Substance Use: Do They Really Exist?" Working Paper, UC-Berkeley School of Public Health.

Ellis, Engela G., Elizabeth M. Nel, and Justus Van Rooyen. 1991. "Conformity Behavior of Afrikaans- and English-Speaking Adolescents in South Africa." The Journal of Social Psychology, Vol. 131, No. 6 (December), pp. 875-879.

Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Group Effects: A Study of Teenage Behavior." Journal of Political Economy, Vol. 100, No. 5, pp. 966-991.

Gaviria, Alejandro and Steven Raphael. 2001. "School-Based Peer Effects and Juvenile Behavior." The Review of Economics and Statistics, Vol. 82, No. 2 (May), pp. 257-268.

Glaeser, E.L. 1996. "Discussion to 'Spatial Effects upon Employment Outcomes: The Case of New Jersey Teenagers' by O'Reagan, K., and J. Quigley." New England Economic Review: Federal Reserve Bank of Boston (May/June), pp. 58-54.

Gruber, Jonathan (ed.) 2001. Risky Behavior Among Youths: An Economic Analysis. Chicago: University of Chicago Press.

Haveman, Robert and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." Journal of Economic Literature, Vol 33, No. 4 (December), pp. 1829-1878.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2002. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement. NBER Working Paper #8741.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2001. "Does Peer Ability Affect Student Achievement?" NBER Working Paper #8502.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning From Gender and Race Variation." NBER Working Paper # 7867.

Iscoe, Ira, Martha Williams, and Jerry Harvey. 1963. "Modification of Children's Judgments by a Simulated Group Technique: A Normative Developmental Study." Child Development, Vol. 34, No. 4 (December), pp. 963-978.

Jencks, Christopher and Susan E. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood," in Laurence E. Lynn and Michael G. McGeary (eds.), Inner-City Poverty in the United States. Washington, DC: National Academy Press, pp. 111-186.

Krauth, Brian. 2002. "Simulation-Based Estimation of Peer Effects." Working paper, Simon Fraser University.

Kremer, Michael, and Dan M. Levy. 2001. "Peer Effects from Alcohol Use Among College Students." Working paper, Harvard University.

Manski, Charles F. 1993. "The Identification of Endogenous Social Effects: The Reflection Problem." Review of Economic Studies, Vol. 60, No. 3, pp. 531-542.

Sacerdote, Bruce. 1999. " Peer Effects with Random Assignment: Results for Dartmouth Roommates." Working paper, Dartmouth College.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." The Quarterly Journal of Economics, Vol. 116, No. 2, pp. 681-704.

Santee, Richard T., and Christina Maslach. 1982. "To Agree or Not to Agree: Personal Dissent Amid Social Pressure to Conform." The Journal of Personality and Social Psychology, Vol. 42, No. 4, pp. 690-700.

Steiner, Ivan D. 1960. "Sex Differences in the Resolution of A-B-X Conflicts." Journal of Personality, Vol 28, No. 1 (March), pp. 188-128.

Wolf, J.M. and A.L. Kessler. 1987. "Entrance to Kindergarten: What is the Best Age?" Arlington Virginia: Educational Research Service, ERS Monograph.

Zax, Jeffrey S. and Daniel I. Rees. 2002. "IQ, Academic Performance, Environment, and Earnings." Review of Economics and Statistics, Vol. 84, No. 4, pp. 600-616.

Zimmerman, David, J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." Review of Economics and Statistics, Vol. 85, No. 1, pp. 9-23.

Table 1. Kindergarten Start Dates by State, 1985-1989

States	*Date by Which Child Must Be 5 to Enter Kindergarten*
North Dakota, Washington	August 31
Arizona, Florida, Georgia, Kansas, Massachusetts, Minnesota, Mississippi, New Mexico, Oklahoma, South Dakota, Texas, Utah, Wisconsin, West Virginia	September 1
Missouri (1985 –86 school year) (1986 –87 school year) (1987-88 through 1989-90 school years)	September 1 August 1 July 1
Montana	September 10
Iowa, Wyoming	September 15
Nevada, Ohio, Tennessee, Virginia	September 30
Alabama, Arkansas, Kentucky	October 1
Idaho, Maine, Nebraska	October 15
North Carolina	October 16
South Carolina	November 1
Illinois (1985-86 through 1986-87 school years) (1987-88 school year) (1988-89 school year)	November 1 October 1 September 1
Alaska	November 2
Oregon (1985-86 school year) (1986-87 through 1989-90 school years)	November 15 September 1
California, Michigan, New York	December 1
Hawaii, Maryland, Rhode Island, Washington DC	December 31
Connecticut, Delaware	January 1

Sources: the Education Commission of the States (1985), Wolf and Kessler (1987), and direct communications between the authors and the Idaho Department of Education, the Illinois State Board of Education, the Missouri Department of Elementary and Secondary Education, the Rhode Island Department of Elementary and Secondary Education, and the Utah State Office of Education.

Note: Kindergarten start dates were determined at the school district level in the seven remaining states. Students from these states were excluded from the analysis.

Table 2. Personal and Family Background Variables as Predictors of the Difference Between Birth and Kindergarten Start Dates.

Male	1.42
	(2.64)
Black	.183
	(3.33)
Other Race	5.78
	(4.51)
Hispanic	-3.63
	(4.21)
Lived with Both Parents	3.64
	(2.65)
Parental Education	-.789
	(.492)
Urban	-3.54
	(3.08)
High School	-5.35**
	(1.96)
Private School	-2.35
	(5.21)
Observations	17384
R-Squared	.002

***p < 0.01; **p <0.05; * p<0.10

Notes: "Lived with Both Parents" is a dichotomous variable indicating that the respondent lived with both biological parents; "Parental Education" is defined as the number of years of education of the respondent's best educated parent; "High School" is a dichotomous variable indicating that the respondent attended high school; and "Private School" is a dichotomous variable indicating that the respondent attended Catholic or private school. Although not shown, controls for missing race, ethnicity, and parental education are also included in the regression. The sample includes both males and females.

Table 3a. Probit Results: Relative Age and Substance Use

	Males					
	Marijuana		**Alcohol**		**Tobacco**	
	(1)	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
Younger	- .001	-.001	-.005	-.002	.011	.017
	(.009)	(.011)	(.013)	(.016)	(.013)	(.016)
Behind	.018*	.019	-.018	-.008	.034**	.048**
	(.010)	(.014)	(.014)	(.021)	(.014)	(.021)
Ahead	-.002	-.006	.004	-.001	.039	.025
	(.023)	(.024)	(.033)	(.035)	(.033)	(.039)
Behind*Younger		-.001		-.016		-.023
	---	(.018)	---	(.027)	---	(.026)
Ahead*Younger		.002		.078		-.075
	---	(.069)	---	(.093)	---	(.075)
Age:						
Age (in months)	.014***	.014***	.017***	.018***	.010**	.010**
	(.004)	(.004)	(.005)	(.005)	(.005)	(.005)
Age Squared/100	-.0028***	-.0028***	-.0026*	-.0027*	-.0010	-.0010
	(.0009)	(.0009)	(.0014)	(.0014)	(.0013)	(.0013)
Racial/Ethnic Categories:						
Black	-.042***	-.042***	-.175***	-.175***	-.138***	-.140***
	(.009)	(.009)	(.013)	(.013)	(.013)	(.013)
Other Race	.001	.001	-.051**	-.051**	-.025	-.026
	(.014)	(.014)	(.019)	(.019)	(.020)	(.020)
Hispanic	-.013	-.013	-.001	-.001	-.072***	-.072***
	(.013)	(.013)	(.019)	(.019)	(.017)	(.017)
Background Measures:						
Parental Education	-.00004	-.00004	.0001	.0001	-.008***	-.008***
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Lived with Both Parents	-.074***	-.074***	-.062***	-.062***	-.110***	-.110***
	(.009)	(.009)	(.013)	(.013)	(.012)	(.013)
Urban	.022**	.022**	.018	.018	-.012	-.012
	(.009)	(.009)	(.015)	(.015)	(.014)	(.014)
School Characteristics:						
High School	.019*	.019*	.027	.027	-.010	-.010
	(.010)	(.010)	(.015)	(.015)	(.015)	(.015)
Private School	-.061	-.061	-.048	-.048	-.087***	-.086***
	(.012)	(.012)	(.024)	(.024)	(.021)	(.021)
Observations	8989		8759		8683	

***p < 0.01; **p <0.05; *p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. Although not shown, controls for year and missing race, ethnicity, and parental education are also included in the regressions.

Table 3b. Probit Results: Relative Age and Substance Use

	Females					
	Marijuana		**Alcohol**		**Tobacco**	
	(1)	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
Younger	.011	.019**	.025**	.026*	.024*	.031**
	(.007)	(.009)	(.013)	(.015)	(.012)	(.015)
Behind	-.003	.011	-.040***	-.041**	.009	(.015)
	(.009)	(.015)	(.015)	(.022)	.025	(.023)
Ahead	.023	.041	.033	.043	.060**	.068*
	(.020)	(.026)	(.031)	(.036)	(.032)	(.039)
Behind*Younger		-.021		.002		-.025
	---	(.015)	---	(.029)	---	(.027)
Ahead*Younger		-.051*		-.042		-.024
	---	(.018)	---	(.062)	---	(.058)
Age:						
Age (in months)	.017***	.018***	.033***	.033***	.015***	.015***
	(.003)	(.003)	(.006)	(.006)	(.005)	(.005)
Age Squared/100	-.0042***	-.0042***	-.0072***	-.0072***	-.0026**	-.0027**
	(.0009)	(.0009)	(.0014)	(.0014)	(.0013)	(.0013)
Racial/Ethnic Categories:						
Black	-.066***	-.066***	-.156***	-.157***	-.187***	-.187***
	(.007)	(.007)	(.013)	(.013)	(.012)	(.012)
Other Race	-.015	-.015	-.057***	-.057***	-.032	-.033
	(.011)	(.011)	(.019)	(.019)	(.019)	(.019)
Hispanic	-.028**	-.028**	-.018	-.018	.107***	(.016)
	(.010)	(.010)	(.019)	(.019)	-.107***	(.016)
Background Measures:						
Parental Education	-.001	-.001	.006**	.006**	-.004*	-.004*
	(.001)	(.001)	(.002)	(.002)	(.002)	(.002)
Lived with Both Parents	-.058***	-.058***	-.076***	-.076***	-.100***	-.100***
	(.008)	(.008)	(.013)	(.013)	(.021)	(.021)
Urban	.034***	.034***	.022	.022	-.025*	-.025*
	(.008)	(.008)	(.014)	(.014)	(.015)	(.015)
School Characteristics:						
High School	.017*	.017*	.014	.014	-.016	-.016
	(.009)	(.009)	(.016)	(.016)	(-.015)	(.015)
Private School	-.017	-.018	-.006	-.007	.008	.008
	(.013)	(.012)	(.024)	(.025)	(.025)	(.025)
Observations	8395		8202		8087	

***p < 0.01; **p <0.05; * p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. Although not shown, controls for year and missing race, ethnicity, and parental education are also included in the regressions.

Table 4. Probit Results: Relative Age and Sexual Activity

	Males (1)	*Males* (2)	*Females* (1)	*Females* (2)
Younger	.013	.023	.019	.024
	(.017)	(.020)	(.016)	(.018)
Behind	.057***	.079***	-.012	-.001
	(.019)	(.029)	(.018)	(.029)
Ahead	.014	-.012	.086**	.083**
	(.042)	(.045)	(.039)	(.043)
Behind*Younger		-.038		-.019
	---	(.034)	---	(.034)
Ahead*Younger		.162		.019
	---	(.111)	---	(.092)
Age:				
Age (in months)	.018	.018	.023*	.023*
	(.012)	(.012)	(.012)	(.012)
Age Squared/100	-.0012	-.0014	-.0030	-.0030
	(.0029)	(.0029)	(.0030)	(.0031)
Racial/Ethnic Categories:				
Black	.224***	.233***	.049	.049
	(.021)	(.020)	(.019)	(.019)
Other Race	.008	.007	-.012	-.013
	(.028)	(.028)	(.025)	(.025)
Hispanic	.047*	.047*	-.061**	-.061**
	(.026)	(.026)	(.023)	(.023)
Background Measures:				
Parental Education	-.013	-.013	-.007**	-.007**
	(.003)	(.003)	(.003)	(.003)
Lived with Both	-.159***	-.159***	-.152***	-.152***
Parents	(.016)	(.016)	(.015)	(.015)
Urban	-.007	-.007	-.010	-.010
	(.018)	(.018)	(.018)	(.018)
School Characteristics:				
High School	.028	.028	.012	.012
	(.020)	(.020)	(.021)	(.020)
Private School	-.064*	-.063*	-.116***	-.117***
	(.032)	(.032)	(.025)	(.025)
Observations	6502		6189	

***p < 0.01; **p <0.05; *p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. Although not shown, controls for year, missing race, ethnicity, and parental education are also included in the regressions.

Table 5a. Probit Results: Relative Age and Delinquency

	Vandalism (1)	Vandalism (2)	Fighting (1)	Fighting (2)	Theft (1)	Theft (2)
Males						
Younger	.013	.021	-.004	.012	-.001	.010
	(.012)	(.014)	(.010)	(.013)	(.009)	(.011)
Behind	-.006	.006	.028**	.043**	.028***	.042***
	(.013)	(.019)	(.012)	(.018)	(.010)	(.015)
Ahead	-.024	-.017	-.011	-.012	.032	.051*
	(.026)	(.030)	(.024)	(.028)	(.026)	(.031)
Behind*Younger		-.019		-.023		-.023
	---	(.023)	---	(.020)	---	(.015)
Ahead*Younger		-.027		.017		-.052
	---	(.058)	---	(.062)	---	(.026)
Age:						
Age (in months)	.013***	.013***	.002	.002	.011***	.011***
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
Age Squared/100	-.0035***	-.0035***	-.0006	-.0006	-.0027***	-.0027***
	(.0011)	(.0011)	(.0010)	(.0010)	(.0009)	(.0009)
Racial/Ethnic Categories:						
Black	-.072***	-.072***	.022*	.021*	-.027***	-.027***
	(.013)	(.013)	(.013)	(.013)	(.009)	(.009)
Other Race	-.003	-.004	-.017	-.018	-.013	-.013
	(.018)	(.018)	(.015)	(.015)	(.013)	(.013)
Hispanic	-.020	-.020	-.008	-.009	-.003	-.003
	(.016)	(.016)	(.015)	(.015)	(.013)	(.013)
Background Measures:						
Parental Education	.004**	.004**	-.006	-.006	-.002	-.002
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Lived with Both Parents	-.052***	-.052***	-.070***	-.070***	-.055***	-.055***
	(.012)	(.012)	(.010)	(.010)	(.009)	(.009)
Urban	.032**	.031**	.030***	.030***	.023**	.023**
	(.012)	(.012)	(.011)	(.011)	(.009)	(.009)
School Characteristics:						
High School	-.006	-.006	.019	.019	.007	.007
	(.013)	(.013)	(.012)	(.011)	(.010)	(.010)
Private School	-.030	-.030	-.043**	-.043**	-.025	-.025
	(.021)	(.021)	(.018)	(.018)	(.015)	(.015)
Observations	8029		8530		6123	

***p < 0.01; **p <0.05; *p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. Although not shown, controls for year and missing race, ethnicity, and parental education are also included in the regressions.

Table 5b. Probit Results: Relative Age and Delinquency

	Females					
	Vandalism		**Fighting**		**Theft**	
	(1)	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
Younger	.007	.006	.001	-.003	.007	.006
	(.008)	(.009)	(.008)	(.009)	(.005)	(.006)
Behind	.008	.003	.020**	.014	.003	-.002
	(.010)	(.015)	(.010)	(.014)	(.014)	(.008)
Ahead	.002	.006	-.017	-.023	.016	.018
	(.018)	(.024)	(.016)	(.018)	(.014)	(.017)
Behind*Younger		.008		.010		.008
	---	(.019)	---	(.018)	---	(.013)
Ahead*Younger		-.020		.037		-.005
	---	(.038)	---	(.058)	---	(.021)
Age:						
Age (in months)	.006*	.006*	.061**	.060***	.009***	.009***
	(.003)	(.003)	(.003)	(.003)	(.002)	(.002)
Age Squared/100	-.0019**	-.0019**	-.0017**	-.0017**	-.0023***	-.0023***
	(.0009)	(.0009)	(.0008)	(.0008)	(.0007)	(.0007)
Racial/Ethnic Categories:						
Black	-.040***	-.038***	.009	.009	-.015**	-.014**
	(.008)	(.008)	(.009)	(.010)	(.005)	(.013)
Other Race	-.016	-.016	.009	.009	.020**	-.020**
	(.012)	(.012)	(.014)	(.014)	(.011)	(.011)
Hispanic	.019	.019	-.012	-.012	-.018**	-.018**
	(.011)	(.011)	(.012)	(.012)	(.006)	(.006)
Background Measures:						
Parental Education	-.005***	-.004***	-.002*	-.002*	.0003	.0002
	(.001)	(.001)	(.001)	(.001)	(.0009)	(.0009)
Lived with Both Parents	-.044***	-.044***	-.060***	-.060***	-.026***	-.026***
	(.008)	(.008)	(.009)	(.008)	(.005)	(.005)
Urban	.018**	.018**	.017*	.017*	.018***	.018***
	(.008)	(.008)	(.008)	(.008)	(.005)	(.005)
School Characteristics:						
High School	-.005	-.005	.009	.009	-.004	-.004
	(.010)	(.010)	(.009)	(.010)	(.007)	(.007)
Private School	-.007	-.009	-.015	-.015	-.010	-.010***
	(.015)	(.015)	(.013)	(.013)	(.007)	(.007)
Observations	7972		8303		6209	

***p < 0.01; **p <0.05; * p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. Although not shown, controls for year and missing race, ethnicity, and parental education are also included in the regressions.

Table 6a. Probit Results: Adding State Fixed Effects and Season of Birth to the Model

	Males							
	Marijuana		**Alcohol**		**Tobacco**		**Sexual Activity**	
Younger	.0001	-.020	-.0004	.003	.019	.005	.020	.005
	(.011)	(.013)	(.016)	(.019)	(.016)	(.019)	(.021)	(.024)
Behind	.022	.021	-.004	-.007	.043**	.051**	.075	.081
	(.014)	(.014)	(.021)	(.022)	(.021)	(.022)	(.028)	(.028)
Ahead	.003	-.007	.010	.012	.034	.021	-.020	-.014
	(.027)	(.023)	(.039)	(.036)	(.042)	(.038)	(.048)	(.046)
Behind*Younger	-.004	-.003	-.019	.017	-.020	.028	-.031	-.039
	(.017)	(.018)	(.027)	(.027)	(.026)	(.026)	(.034)	(.034)
Ahead*Younger	.025	.025	.074	.079	.088	.077	.135	.165
	(.070)	(.071)	(.097)	(.094)	(.077)	(.075)	(.112)	(.111)
State Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Season of Birth Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	8966	8989	8758	8759	8674	8683	6496	6502

***p < 0.01; **p <0.05; *p<0.10

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. See Table 3 for a full list of the controls. Season of birth is measured using a set of three dichotomous variables indicating if the respondent was born in the winter (December 21 through March 20), spring (March 21 through June 20), or summer (June 21 through September 21).

Table 6b. Probit Results: Adding State Fixed Effects and Season of Birth to the Model

	Females							
	Marijuana		**Alcohol**		**Tobacco**		**Sexual Activity**	
Younger	.018**	.035***	.026*	.038**	.032**	.050***	.018	.009
	(.009)	(.010)	(.015)	(.018)	(.015)	(.017)	(.018)	(.020)
Behind	.009	.010	-.040*	-.043*	.024	.027	-.003	-.0001
	(.014)	(.014)	(.022)	(.022)	(.023)	(.023)	(.028)	(.028)
Ahead	.031	.044*	.036	.047	.070**	.067*	.043	.081**
	(.025)	(.026)	(.037)	(.036)	(.038)	(.038)	(.043)	(.043)
Behind*Younger	-.020	-.023	.004	.003	-.022	-.032	-.011	-.020
	(.015)	(.015)	(.029)	(.030)	(.025)	(.027)	(.035)	(.035)
Ahead*Younger	.040	.051*	-.046	-.044	-.004	-.025	-.009	.022
	(.021)	(.018)	(.066)	(.062)	(.064)	(.058)	(.091)	(.092)
State Fixed Effects	Yes	No	Yes	No	Yes	No	Yes	No
Season of Birth Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	8310	8395	8186	8202	8077	8087	6185	6189

***$p < 0.01$; **$p < 0.05$; *$p < 0.10$

Notes: Marginal probabilities from a single-equation probit model are reported. Standard errors are corrected for clustering at the individual level and are in parentheses. See Table 3 for a full list of the controls. Season of birth is measured using a set of three dichotomous variables indicating if the respondent was born in the winter (December 21 through March 20), spring (March 21 through June 20), or summer (June 21 through September 21).

Appendix Table 1. Descriptive Statistics

	Males			*Females*		
	Full Sample (n=8988)	**Younger Than Peers** (n=4402)	**Older Than Peers** (n=4587)	**Full Sample** (n=8395)	**Younger Than Peers** (n=4127)	**Older Than Peers** (n=4268)
Difference (between kindergarten entrance and birth dates in years)	.507 (.295)	.244 (.144)	.760 (.142)	.502 (.297)	.237 (.145)	.758 (.142)
Behind	.346	.465	.233	.252	.317	.190
Ahead	.036	.015	.057	.048	.019	.075
Age (in months)	188.43 (18.53)	187.63 (18.35)	189.19 (18.67)	188.28 (18.25)	187.40 (17.89)	189.14 (18.56)
<u>Racial/Ethnic Categories:</u>						
Black	.267	.268	.266	.276	.273	.279
Other Race	.146	.146	.145	.147	.147	.146
Hispanic	.219	.211	.226	.224	.235	.213
<u>Background Measures:</u>						
Parental Education	13.03 (3.02)	13.17 (3.04)	12.89 (2.98)	12.98 (3.00)	12.96 (3.03)	12.99 (2.98)
Lived with Both Parents	.515	.501	.529	.484	.491	.478
Urban	.728	.740	.716	.741	.744	.739
<u>School Characteristics:</u>						
High School	.652	.657	.648	.683	.698	.668
Private School	.060	.060	.060	.062	.066	.059

Notes: standard deviations of continuous variables in parentheses.

Appendix Table 2. Means of Outcome Variables

	Males			*Females*		
	Full Sample	**Younger Than Peers**	**Older Than Peers**	**Full Sample**	**Younger Than Peers**	**Older Than Peers**
Marijuana (past month)	.137 (8989)	.137 (4402)	.136 (4587)	.099 (8395)	.103 (4127)	.095 (4268)
Alcohol (past month)	.314 (8759)	.322 (4472)	.305 (4287)	.286 (8202)	.291 (4064)	.282 (4138)
Tobacco (past month)	.254 (8683)	.259 (4255)	.249 (4428)	.229 (8087)	.235 (3985)	.224 (4102)
Sexually Active (past year)	.357 (6502)	.359 (3152)	.354 (3350)	.299 (6189)	.294 (3042)	.304 (3147)
Vandalism (past year)	.192 (8029)	.201 (3891)	.184 (4138)	.093 (7972)	.097 (3934)	.087 (4038)
Fighting/Assault (past year)	.160 (8530)	.167 (4147)	.154 (4383)	.091 (8303)	.093 (4081)	.089 (4222)
Theft (past year)	.089 (6123)	.092 (2981)	.086 (3142)	.037 (6209)	.041 (3012)	.033 (3197)

Note: Samples sizes in parentheses.

Figure 1. Current Marijuana Use by Age

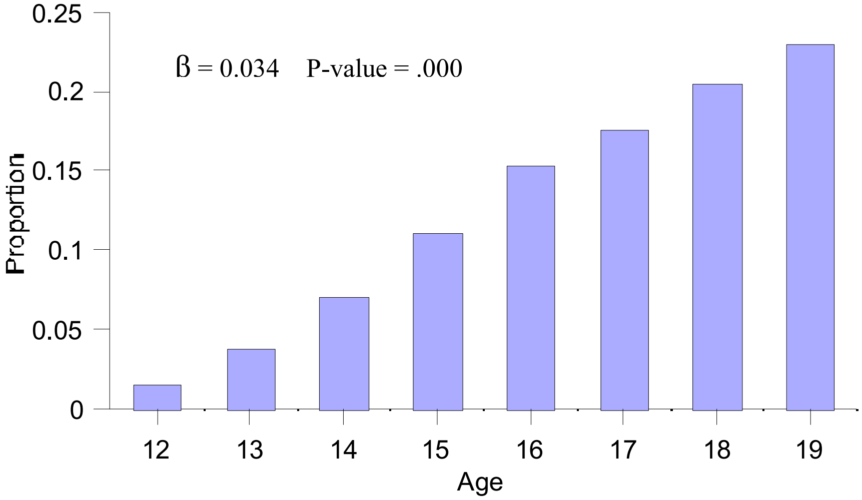

Figure 2. Current Alcohol Use by Age

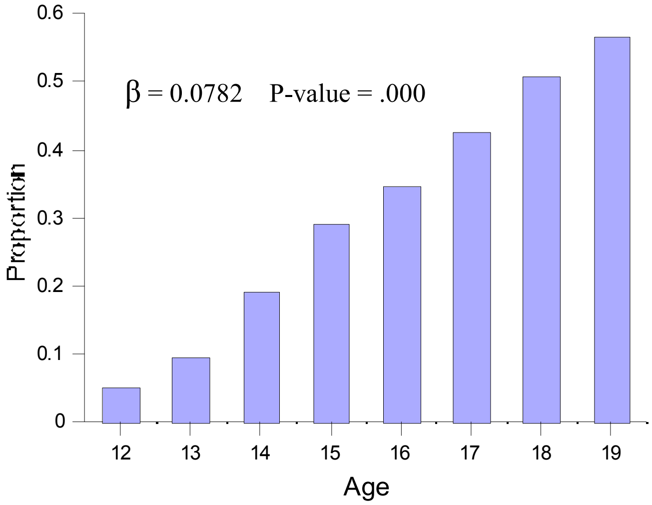

Figure 3. Current Smoking by Age

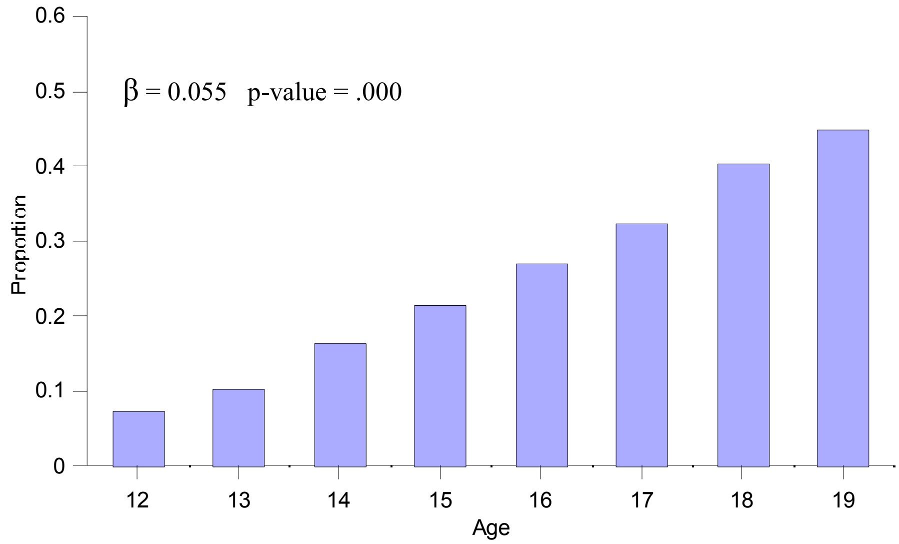

Figure 4. Sexual Activity by Age

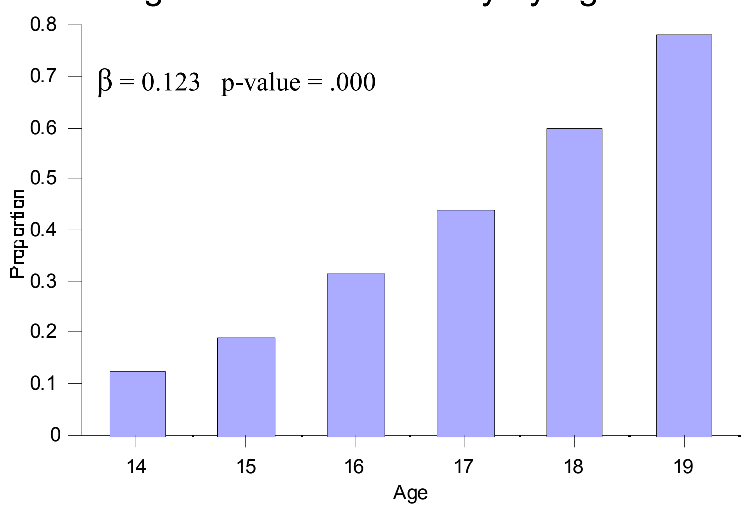

Figure 5. Distribution of months before kindergarten start date

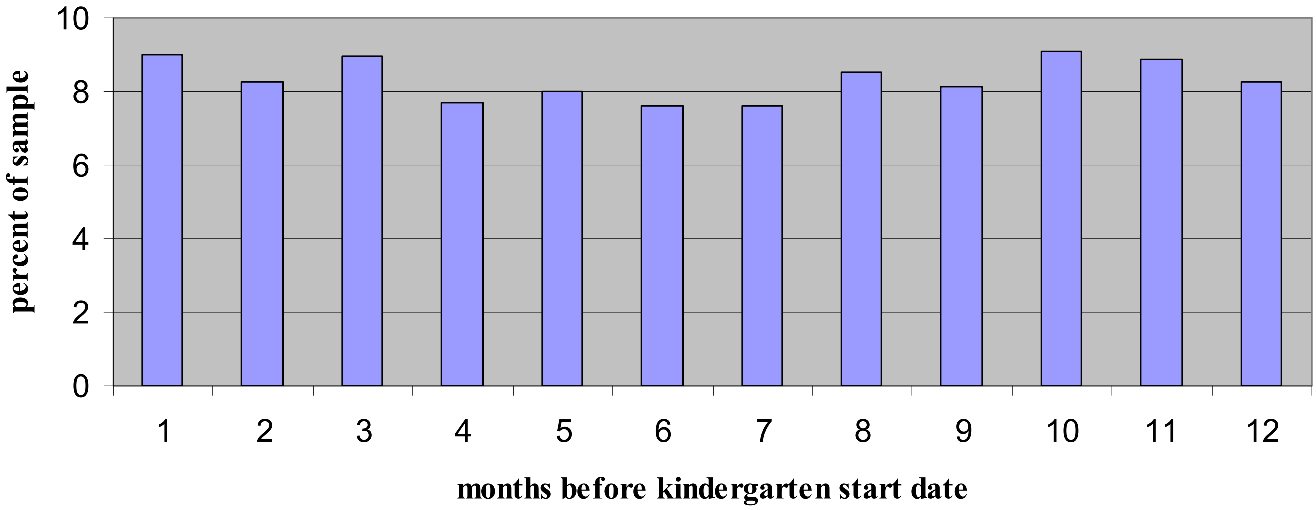

Figure 6. Vandalism by Age

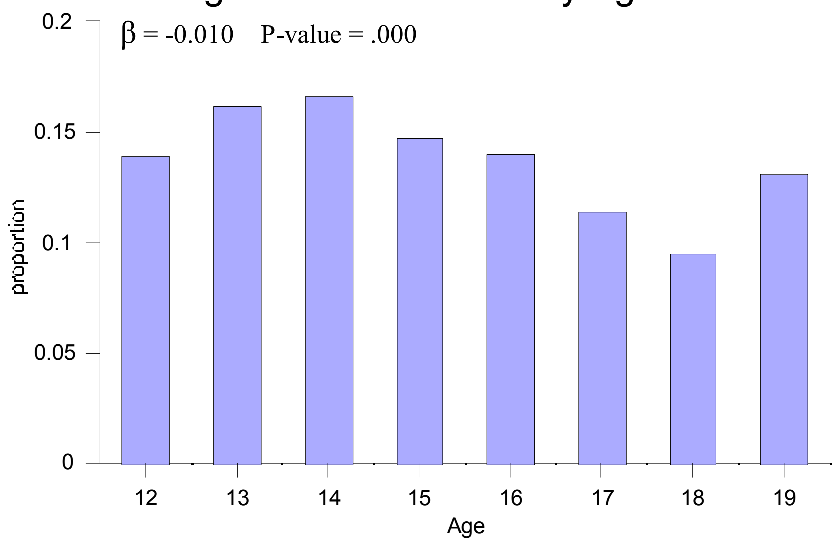

Figure 7. Theft by Age

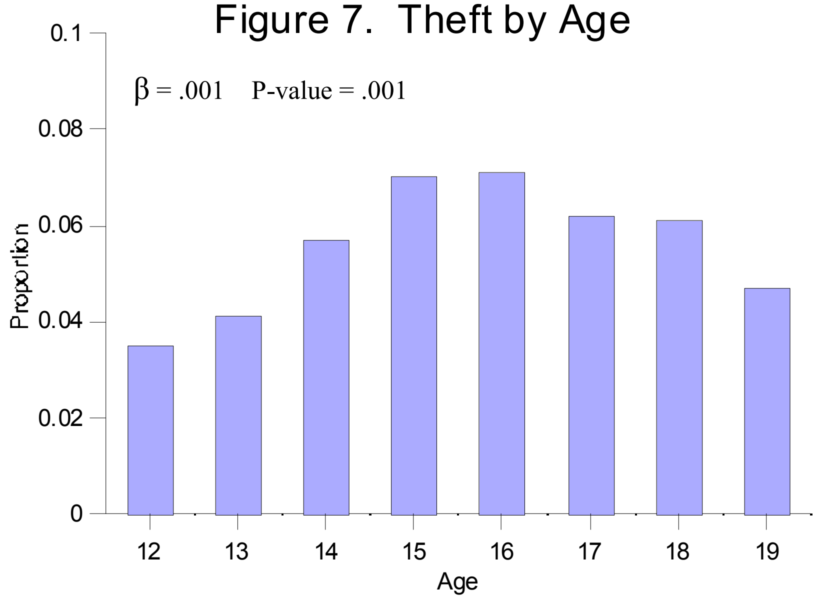

Figure 8. Assault by Age