

Peer Effects on Adolescent Smoking: Are Popular Teens more Influential?

Juan David Robalino^{1,2} & Michael Macy³

¹Department of Economics, Cornell University

²IZA – Institute for the Study of Labor

³Department of Information Science, Cornell University

Abstract

Previous research on adolescent cigarette adoption has focused on peer influence and the perceived status gain from smoking but has ignored the status effects on peer influence. We analyze adolescent peer effects on cigarette consumption while considering the popularity of peers. The analysis is based on a four wave panel survey representative of American high school students. We measure peers' popularity by their eigenvector centrality in high school social networks. Using lagged peers' behavior, school fixed effects, and instrumental variables to control for homophily and contextual confounds, we find that the probability of smoking the following year increases with the mean popularity of smokers, while the popularity of non-smokers has the opposite effect. These effects persist seven and fourteen years later (wave 3 and 4 of the data). In addition, the probability of smoking increases with the smoking propensity of the 20% most popular teens and decreases with the smoking propensity of the bottom 80%. The results indicate the importance of knowing not only the smoking propensity within a school but also the location of smokers within the social hierarchy.

1 Introduction

Identifying the drivers of smoking adoption among youth remains a public health priority. Initiating smoking at a young age is correlated with smoking more cigarettes per day and with a lower probability of quitting later in life [11, 18]. More than 480,000 deaths are attributed to cigarette smoking every year in the US alone. At the current rate of adoption, 5.6 million Americans under 18 –about 1 of every 13– will die early from a smoking-related illness.¹

Previous research points to peer influence as an important cause of adolescent smoking [7, 12, 14, 17, 19, 23, 26]. Adolescents are especially vulnerable to social influence as they try to fit in with their peers [2, 13]. Numerous studies suggest that some smokers believe that smoking promotes social status [1, 10, 22, 30], and yet previous research has failed to test a key implication: that the propensity to smoke may increase with the popularity of smokers among peers. In other contexts, peer influence has been found to increase with their social status [25], but no previous studies have empirically tested the effects of the social status of adolescent smokers on the spread of smoking through peer networks. The need to empirically test this status-belief explanation is the starting point for our study.

Previous research on popularity and smoking

Several studies have analyzed the link between popularity of adolescents (usually measured by the number of incoming friendship nominations) and their smoking behavior. For example, Valente et al. [30] and Ennett et al. [10] found a positive relationship between popularity and smoking; Alexander et al. [1] found that popular students are more likely to smoke in schools with high smoking rates, and less likely in schools with low smoking rates. Michell and Amos [22] used qualitative data from Scottish schoolgirls and found that popular girls smoke to

¹Centers for Disease Control and Prevention.

maintain their image while some unpopular girls smoke with the hope of gaining social status, but no effect in the mid-popularity range.

These studies indicate that smoking may be used as a strategy to climb the social ladder, based on the belief that smoking will lead to an increase in social status. However, they do not test the peer effects of smokers' social status. To address this gap, we reverse the causal arrow: instead of testing whether smoking affects status, we test how the popularity of smokers affects their influence on the behavior of their peers. We posit that the relative popularity of smokers and non-smokers will condition whether smoking will be associated with social status in that population. To the best of our knowledge, this is the first study of the effect of smokers' popularity on peer influence for smoking.²

It is worth noting that there have been anti-smoking field interventions that relied implicitly on the hypothesis that we are testing. For example, to promote an anti-smoking campaign in high schools, a field experiment explicitly recruited individuals nominated by fellow students as "influential" [4], and their intervention proved successful in comparison to the smoking rates of the control high schools that did not recruit any students. In short, this research assumed status differences in influence but did not test that assumption. Our study is motivated by the need to empirically test an assumption that informs intervention strategies not only into smoking but into other public health, advertising, and electoral campaigns.

The reflection problem

Like most previous studies of peer influence on adolescent smoking, we rely on observational data. Seminal work by Manski [20] identifies the inherent difficulty in estimating peer effects with observational data, which he refers to as the reflection problem. According to Manski,

²In his book "The Tipping Point", Malcolm Gladwell [15] presents informal interviews from smokers describing their first impressions about cigarettes; most descriptions point towards a given person whom they admired in their youth and who smoked. In fact, this study is inspired by this book.

a correlation between the average behavior of the group and an individual's behavior can be attributed to three possible mechanisms:

- *endogenous effects*, where an individual's choice is influenced by the choices of the group, as occurs with peer effects on behavior;
- *exogenous (contextual) effects*, where individuals in a given group may behave similarly because the whole group has experienced an (unobserved) exogenous shock;
- *correlated effects*, where individuals in a group behave similarly because they have similar unobservable characteristics and self-select into the group.

The challenge with observational data is how to disentangle endogenous effects from contextual and correlated effects. Several identification approaches have been used. Sacerdote [28] studied peer effects on the Grade Point Average (GPA) of college students using exogenous group formation from random allocation to college dorms in order to control for peers' self-selection and homophily. He also used peers' lagged GPA (from high-school) to control for contextual confounds.

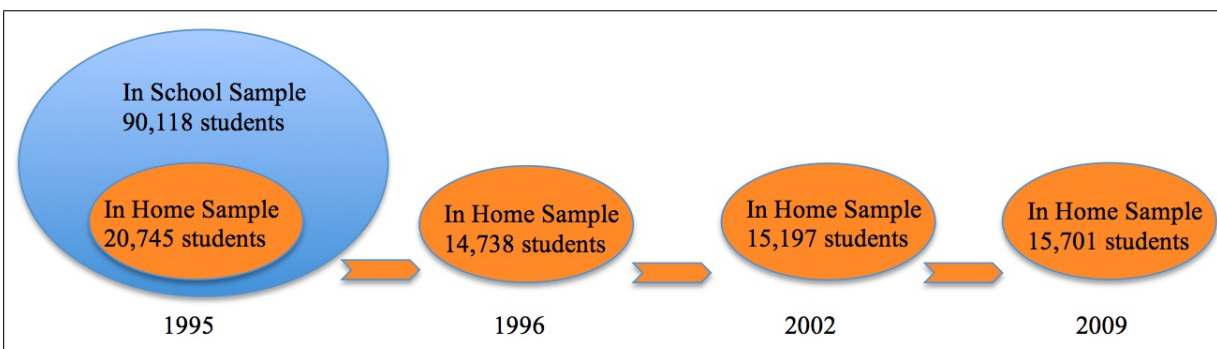
In the absence of random allocation of peers, several studies of peer effects on smoking have used broad pseudo-exogenous peer groups, controls for school level fixed effects, as well as instrumental variables for peers' behavior. Lundborg [19] used the smoking rate among classmates as peer referents; Fletcher [12] used the school grade; Clark and Lohéac [7] used lagged behavior at the grade level (i.e., the smoking rate during the previous year); Norton et al. [24], Gaviria and Raphael [14] and Powell et al. [26] used the whole school. To account for contextual cofounds, Gaviria and Raphael [14], and Powell et al. [26] controlled for cigarette prices and public policy variables, while Lundborg [19], Clark and Lohéac [7], and Fletcher [12] included school fixed effects. Norton et al. [24], Gaviria and Raphael [14], Powell et al. [26],

and Fletcher [12] also instrumented the behavior of peers using household or neighborhood characteristics. All these studies found strong evidence of peer influence on cigarettes use.³

In line with previous studies, we use the lagged behavior of peers at the school grade level to control for homophily and include school fixed effects to control for contextual confounds. Nevertheless, the observed effects from multivariate models remain susceptible to unobserved heterogeneity. We therefore also use instrumental variables to test the robustness of the causal inferences.

Data and methods

Figure 1: Data structure.



The present study is based on data from the National Longitudinal Survey of Adolescent

³A more controversial approach to identify peer effects is that of Christakis and Fowler [5]: using data from the The Framingham Heart Study consisting of a large (sparse) social network, they relied on the direction of social tie nominations and the panel aspect of the data (the behavior of nominated peers at time t to predict the behavior of the focal individual at time $t + 1$) to capture peers' influence on the diffusion of obesity and subsequently on the diffusion of smoking [5, 6]. Although they found peer effects in both domains, their approach has been criticized for not accounting sufficiently for contextual effects and homophily [8].

Other studies identify smoking peer effects using Monte Carlo estimations. Two main types of models exist: the first are structural models from game theory where the influence of peers is estimated as the equilibrium outcome of strategic interactions [17, 23]; the second type consists of agent-based models capturing the co-evolution of friendship formation and peers' influence [21, 29].

Health (AddHealth), a four-wave panel (1995, 1996, 2002 and 2009) administered to a representative sample of American High Schools. This data set contains detailed information about substance use as well as rich social network data with which to measure the popularity of students. The multi-wave longitudinal data also includes identifiers for grade, year, school, and detailed measures of household and parents’ characteristics. The In-School survey conducted in 1995 covered 90,118 adolescents in 144 schools. A representative sub-sample of 20,745 students also completed the 1995 In-Home wave 1 survey which included more detailed information such as family income and parents’ smoking habits. Finally, about 15,000 students were surveyed again in 1996 (In-Home survey wave 2), as well as 2002 (In-Home wave 3) and 2009 (In-Home wave 4; see Fig 1). After excluding individuals with missing information and schools with insufficient social network data, we have a sample of about 66,000 peers and a core sample of about 7,500 individuals.

Popularity measure

AddHealth identifies peers by asking each respondent to name their five closest male friends within the school and their five closest female friends in order of importance (i.e., best friend, second best friend, etc). These nominations form the social network of the school and can then be used to measure the popularity of each respondent by their *eigenvector centrality*.⁴ Intuitively, the popularity of individual i will be proportional to the popularity of peers who nominated i as a friend, hence being nominated by popular teens makes you more popular. Formally, the centrality v_i of individual i in school s is

$$v_i = \frac{1}{\lambda} \sum_j v_j a_{ji}$$

⁴In the working paper [27] we obtain similar results using in-degree, out-degree, and Katz–Bonacich centrality; see Jackson [16] Chapter 2 for a discussion about network centrality measures.

where $a_{ij} = 1/n_{ij}$ if individual i nominated individual j as his/her n^{th} best friend and $a_{ij} = 0$ otherwise; or equivalently using vector notation we recover the eigenvector expression $\lambda v = A'_s v$, where A_s is the so called adjacency matrix of the social network in school s and λ is the largest eigenvalue –in order to ensure that all the elements of v are positive as guaranteed by the Perron–Frobenius theorem.⁵ This weighted-eigenvector centrality (i.e., weighted by the inverse of the friendship nomination order) is preferred as it also takes account of the difference between being the best friend of many people and being the fifth best friend of many people.⁶

We compare the centrality measure across school/grades in two ways: by standardizing the measure, and by classifying respondents into centrality quintiles for each grade level (from the 20% most popular to the 20% least popular).

Table 1: Correlates of popularity (weighted standardized eigenvector centrality).

Smoke - Tried 1995	0.033 (0.028)	Black	-0.367*** (0.046)	Candid (0/1)	0.003 (0.030)
Smoke - Everyday 1995	-0.022 (0.050)	Hispanic	-0.123*** (0.045)	Attractive personality (0/1)	-0.018 (0.038)
Cigs. available at home	-0.044 (0.030)	Asian	0.014 (0.068)	Well groomed (0/1)	0.060* (0.032)
Male	-0.002 (0.031)	Other	-0.13 (0.162)	Physically attractive (0/1)	0.182*** (0.037)
Age	0.298* (0.164)	Weekly earnings	-0.027 (0.018)	Physically mature (0/1)	0.118*** (0.032)
Age sq.	-0.009* (0.005)	Overweight (0/1)	-0.151*** (0.028)	Const.	-2.719* (1.390)
HH income	0.922*** (0.346)	Sports 1-2 times/week	0.055 (0.033)		
Foreign	-0.096** (0.039)	Sports 3-4 times/week	0.094** (0.040)	Degrees of freedom	148
New student	-0.152*** (0.033)	Sports 5+ times/week	0.198*** (0.039)	R ²	0.154
				N	7169

Table 1 describes ‘popular’ teens by regressing centrality on respondent attributes. The results show that popular teens tend to be older, white, from higher income households, well

⁵Note that we use the transpose of A_s so that the centrality is proportional to the centrality of the peers who nominate individual i , hence, ignoring individuals claiming to be friends with the popular teens.

⁶About 15% of students were not properly matched to the nomination roster and thus their incoming friendship nominations were lost. A similar issue is that friendship nominations are truncated to a maximum of ten nominations. Costenbader and Valente [9] find that, when working with a subsample of nodes, the eigenvector centrality is the most robust among the eight common centrality measures they consider: across several data sets, eigenvector centralities computed with 50% of the sample have an average correlations higher than 0.95 with eigenvector centralities computed with the full sample.

groomed, play sports and they tend to be physically mature and attractive. In contrast, unpopular teens tend to be black or hispanic, overweight, foreign, and new to the school.⁷

Smoking measures

The items on smoking behavior differ slightly between the In-Home and In-School survey instruments. For the In-Home survey, we consider as regular smokers those who smoked *everyday* during the month prior to the interview. We also classify casual smokers as those who *tried* at least one cigarette in the 30 days prior to the interview. More precisely, for each wave we also measure the number of cigarettes consumed per month.⁸ In wave 4 (for which the interview took place fourteen years after wave 1) we also code respondents having *ever* smoked everyday for at least one month;⁹ the age at which individuals tried their first cigarette; and the

age at which individuals started smoking everyday. The top panel of Table 2 summarizes these measures. In 1996, 32% of students tried cigarettes and 11% smoked everyday; by 2009, 21% were regular smokers.¹⁰

Table 2: Summary statistics.

	Mean	SD	N
Smoking			
Tried by 1995	0.25	0.43	7620
Tried 1996	0.32	0.47	7620
Everyday 1996	0.11	0.31	7620
Everyday 2002	0.17	0.37	7620
Everyday 2009	0.21	0.41	6293
Everyday by 2009	0.43	0.49	6290
# cig./month 1996	157.01	238.05	1860
# cig./month 2002	302.8	260.08	1679
# cig./month 2009	274.74	281.78	2179
Age first cigarette	15.84	3.49	3955
Age smoked everyday	17.38	3.38	2661
Peers in 1995			
<i>Smoking rate among:</i>			
20% most popular	0.16	0.37	13394
80% least popular	0.16	0.37	52792
<i>Mean popularity of:</i>			
Smokers	-0.04	0.87	10657
Non-smokers	0.01	1.02	55529

Peer smokers are those who smoke at least “once or twice a week”. Popularity is defined by standardized weighted-eigenvector centrality.

⁷Note that smoking is not related to popularity; neither is genre.

⁸Number of cigarettes is computed by multiplying the answers to the questions “During the past 30 days, on how many days did you smoke cigarettes?” and “During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?”

⁹These measure of smoking should account for individuals who may have smoked only between interviews.

¹⁰Our estimation strategy assumes that the correlations between smoking in 1995 and in subsequent years are not so high as to introduce multicollinearity. In Table S1 in the supporting information we present these correlations and the first column shows that there is considerable variation between smoking in 1995 and subsequent years (ranging from a correlation with smoking in 2009 of 0.3, to a correlation with “trying cigarettes in 1996” of 0.53).

For peers in the In-School survey, we define as smokers those who reported smoking “once or twice a week” to “daily”. We intentionally use a high threshold of consumption for peers because we suspect that influence arises from regular users who may influence an (initially) soft consumption to the newly initiated individuals, which may eventually result in regular smoking later on.¹¹ The bottom panel of Table 2 summarizes the smoking patterns of peers classified by their standardized weighted-eigenvector centrality. On average, there is no difference in the smoking propensity of the 20% most popular and the 80% least popular (both have a smoking propensity of 0.16). The standardized popularity of smokers and non-smokers is near zero, indicating that both of these groups have average popularity (see the working paper [27] for detailed summary statistics).

Model specification

We test the effects of popularity on peer influence using three related models. With each model, we regress smoking in each of the last three panel waves on peers’ smoking in the first panel wave (the 1995 In-School survey) while controlling for school fixed effects.

Equation (1) models the probability of smoking as a probit regression of the smoking behavior $Y_{i,s,g}^t$ of individual i at time t on the mean popularity of smokers, $\bar{P}_{s,g,y=1}^{t_0}$, and the mean popularity of non-smokers, $\bar{P}_{s,g,y=0}^{t_0}$, controlling for the smoking propensity, $\bar{Y}_{s,g}^{t_0}$, at the grade level g in school s during the first wave, t_0 . Formally we estimate the following equation:

$$Probit(Y_{i,s,g}^t = 1 | X_i, s, g) = c + z_s + \beta X_i + \alpha_1 \bar{P}_{s,g,y=0}^{t_0} + \alpha_2 \bar{P}_{s,g,y=1}^{t_0} + \alpha_3 \bar{Y}_{s,g}^{t_0} \quad (1)$$

where X_i is a vector of demographics and household characteristics and z_s is a school fixed effect. Equation (1) formalizes the hypothesized peer effects as $\alpha_2 > 0 > \alpha_1$, that is, the greater the popularity of smokers, the more likely an individual is to smoke in the future; and

¹¹All results are similar when we measure peer’s smoking as having tried cigarettes; results available upon request.

the greater the popularity of non-smokers, the less likely an individual is to smoke in the future.

Equation (2) models the smoking propensity of the 20% most popular teens, $\bar{Y}_{s,g,Pop}^{t_0}$, and the smoking propensity of the 80% least popular teens, $\bar{Y}_{s,g,noPop}^{t_0}$, in school s and grade g :

$$Probit(Y_{i,s,g}^t = 1 | X_i, s, g) = c + z_s + \beta X_i + \alpha_1 \bar{Y}_{s,g,Pop}^{t_0} + \alpha_2 \bar{Y}_{s,g,noPop}^{t_0} \quad (2)$$

Equation (2) formalizes the hypothesized peer effects as $\alpha_1 > \alpha_2$, representing a stronger influence from the popular teens.

Equation (3) models the smoking propensity $\bar{Y}_{s,g,q}^{t_0}$ of the teens in each popularity quintile q in school s and grade g :

$$Probit(Y_{i,s,g}^t = 1 | X_i, s, g) = c + z_s + X_i \beta + \sum_{q=1}^{q=5} \alpha_q \bar{Y}_{s,g,q}^{t_0} \quad (3)$$

We hypothesize that $\alpha_5 > \alpha_{q \neq 5}$.

All three models are vulnerable to unobservable confounds, which we address using temporal lags and school fixed effects. Most teens go to their local public school and many parents have little flexibility to switch neighborhoods based only on school choices. Even when they consider the schools in their neighborhoods, they will presumably consider other aspects of the school, including academic performance, class sizes, and facilities, before considering the smoking propensity in the school. We control for new students and for parents who claimed to have chosen their neighborhood in part for the school quality, and we use lagged peers' behavior to model the direction of influence. We also control for detailed tobacco policy in 1995 –excise tax, funds per capita for tobacco control, and marketing restrictions on public transports and on billboards near schools. Most importantly, school fixed effects control for contextual factors such as the local price of cigarettes, the school's rules about smoking, and the local sentiment towards smoking. Thus, our variation is among cohorts within schools, which minimizes the effects of self-selection into schools and contextual confounds.

Instrumental variables

We use instrumental variables to provide additional support for the causal inferences based on temporal lags and school fixed-effects. For model 1 we instrument the popularity of smokers and non-smokers. We need instruments that determine the popularity of peers but do not influence the future smoking choices of others. Based on the correlates of popularity found in Table 1, we instrument peer smokers' mean popularity using the percent of smokers who are new to the school, overweight, physically attractive, physically mature, well groomed, white, and foreign, as well as their mean household income and weekly earnings. We also instrument popularity among non-smokers using these same measures. These measures are only available for the In-Home sample; thus, we use the corresponding averages in the In-Home sample to instrument for the popularity of peers in the In-School sample (see Fig 1). Note that school fixed effects are also included in the first stage regression ensuring that instruments such as percentage of white students and mean household income are uncorrelated with the error term in the second stage regression. The percentage of physically attractive smokers should only affect the smoking choices of others through the resulting prestige and popularity of these smokers.

For model 2 we instrument the smoking propensity among popular and non-popular students. The instruments should affect the smoking choices of peers but not the smoking choices of other individuals. We use the percentage of parents who smoke, percentage of households with smokers, the percentage of parents who are home when school finishes, the percentage of black students, percentage of students with older siblings, the average household income and the average weekly earnings among the 20% most popular and, analogously, among the 80% least popular students. Peers' parents smoking should only affect smoking choices of peers but not of other students.

Results

Using peers' lagged behavior, school fixed effects, and instrumental variables to control for contextual confounds and school selection, results for model 1 show that the probability that the respondent will take up smoking increases with a lagged measure of the mean popularity of peer smokers as measured by network centrality. Conversely, higher mean popularity of non-smokers decreases this probability. These patterns persist seven and fourteen years after peers' behavior was measured (i.e., in waves 3 and 4 of the data). Similarly, by decomposing the smoking propensity of peers into the propensity of the 20% most popular teens and that of the 80% least popular, results for model 2 show that peer effects are mainly driven by the influence of the 20% most popular teenagers. Furthermore, in the long run we find a negative influence from the smoking propensity of the 80% least popular peers (in waves 3 and 4 of the data). Similar results apply to the number of cigarettes smoked per month as well as to the age of initiation (see Tables S10 and S11 in the supporting information for details). These patterns suggest that social influence on smoking among adolescents is conditioned by the social status of peers: individuals tend to follow the smoking behavior of popular peers and avoid the behavior of unpopular peers.

Popularity of smokers and non-smokers

Table 3 summarizes results for model 1. Consistent with published results from previous studies of peer influence, the third row of the table shows an effect of the aggregate smoking propensity. In 1996, switching from a school grade where no one smokes to one where 25% of students smoke increases the probability of trying cigarettes by 4.8 percentage points and that of smoking every day by 4 percentage points, comparable to the effect of having access to cigarettes at home, and similar in magnitude to results reported in previous studies [12]. Yet, this peer effect vanishes by 2002 and 2009.

Table 3: Probability of smoking and popularity of smokers/non-smokers – probit average marginal effects.

	Tried 1996	1996	2002	2009	by 2009
Mean popularity of smokers	0.045*** (0.006)	0.009 (0.009)	0.026*** (0.008)	0.026*** (0.007)	0.033*** (0.009)
Mean popularity of non-smokers	-0.059*** (0.015)	-0.013 (0.011)	-0.032*** (0.012)	-0.028* (0.016)	-0.059*** (0.018)
% smokers in grade	0.191* (0.099)	0.163** (0.071)	-0.053 (0.074)	-0.02 (0.100)	-0.116 (0.103)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2 in the supporting information.

Our main result is presented in the first two rows of Table 3: the mean popularity of smokers increases the probability of an individual smoking, and the mean popularity of non-smokers decreases this probability. In contrast to the aggregate peer effect, the effect of peers’ popularity persists in 2002 and in 2009, suggesting that in the long run the popularity of peer smokers is more important than how many peers smoked. The left panel suggests that an increase of a standard deviation in the mean popularity of smokers results in a increase of 4.5 percentage points in the probability of trying cigarettes in 1996, while the same increase in the mean popularity of non-smokers reduces this probability by 5.9 percentage points. The fourth panel suggests that an increase of a standard deviation in the popularity of smokers will result in an increase of 2.6 percentage points in the probability of smoking everyday in 2009, fourteen years after having interacted with those peers, and the same variation in the popularity of non-smokers will reduce this probability by 2.8 percentage points. Similarly, the fourth panel suggests that an increase of a standard deviation in the mean popularity of smokers results in an increase of 3.3 percentage points in the probability of ever smoking daily by 2009 and the same increase in the popularity of non-smokers results in a decrease of 5.9 percentage points in the probability of ever smoking.¹²

¹²Note that the probability of smoking *everyday* in 1996 is not significantly affected by the popularity of peers. It seems that influence from peers’ popularity only produced casual smoking at that time (i.e. during high school years). Yet, the smoking rate in the grade is significant in this regression. We note that in 1996 the average age

Robustness One possible explanation for our results is that popular peers may not be more influential due to their social status, but may simply reach more people and that their influence is a simple one, just like that of other peers, but towards more friends. To test this, we include the smoking propensity of all nominated friends for each individual. Table S4 (in the supporting information) summarizes the results. The smoking propensity among direct friends absorbs the effect of the smoking propensity in the grade and it is still significant in 2002 and in 2009 (although there is potential self-selection). In contrast, the effects of the mean popularity of smokers and non-smokers remain essentially unchanged. This is consistent with the hypothesis that popular peers have a stronger influence than non-popular peers.

The persistence of peer effects may be due to addiction, which can be expected to induce serial correlation in smoking measures for latter periods. To control for serial correlation of smoking in the latter waves, we consider separately subsamples of smokers and non-smokers in early waves and predict their probability of starting/quitting smoking in latter waves. The results reported in Table S5 suggest that addiction may eventually override some peer effects in the long term, but not all. The first three columns indicate that individuals who were not regular smokers in 1996, and even those who had never been regular smokers by 2002, continue to be influenced by the popularity of high school smokers/non-smokers to adopt smoking by 2002 or 2009. The last two columns indicate that, among those who smoked in high school (1996), those who were exposed to popular smokers, compared to those who were not exposed, were still more likely to be smoking as late as 2002, but not after. And among those who smoked in 2002, we find no residual effect of the 1995 popularity of smokers by 2009. A reasonable conclusion is that by 2009, those who began smoking seven or more years earlier may indeed

in the sample was 16 years old, thus, most smokers could not yet buy cigarettes legally. This may explain why the aggregate smoking propensity was more important for regular smokers in this period as it may have facilitated the supply of cigarettes among peers. If we only consider the subsample of students who had access to cigarettes at home as a proxy of students with weaker supply constraints, the effect from the smoking rate has no statistical significance and we get a significant effect from the popularity of smokers (see Table S3).

be hooked and their probability of quitting may be less susceptible to the popularity of smokers and non-smokers in high school.

Table S6 (in the supporting information) reports results of a robustness test using instrumental variables to infer causality. The results confirm those reported in Table 3, although, as expected, we lose some statistical significance. The instruments are highly significant in the first stage regression (J statistics between 13.2 and 22.2 corresponding to p-values of zero for the instruments' joint significance in the first stage regressions); over-identification tests fail to reject the validity of our instruments and Wald test of exogeneity fail to reject the exogeneity of the popularity of smokers and non-smokers' suggesting that a single equation probit model may be more appropriate.¹³

Smoking propensity of popular peers

Table S7 (in the supporting information) reports results estimated from equation (2). The results suggest that peer influence from smokers is driven mainly by the 20% most popular peers.¹⁴ If all of the 20% most popular teens in the grade smoked, the probability of trying cigarettes the following year would increase by 17.8 percentage points (statistically significant at the .01 level), while if all of non-popular teens smoked, resulting in a much larger population of smokers, the increase in this probability would not be statistically different from zero. In fact, in 2002 and 2009, the bottom 80% seem to have a negative influence, that is, the more of these less-popular peers smoke, the less likely an individual is to pick up smoking later in life (although the effect is no longer statistically significant in 2009).

¹³We can reject exogeneity at the 10% level in one regression: having ever smoked regularly by 2009. Results still hold.

¹⁴We find a similar pattern when considering the 10% most popular and the 90% least popular peers, albeit statistical significance is somewhat weaker. Results available upon request.

Robustness Table S8 (in the supporting information) reports the results for the instrumental regression for Model 2. The results are qualitatively similar, but the peer influence of popular students is not statistically significant. While the model fails the over-identification test only in the 1996 measures, we cannot reject the null hypothesis that the smoking rates of smokers/non-smokers are exogenous, suggesting that a probit model may be more appropriate.

Smoking propensity by popularity quintiles

Table S9 (in the supporting information) reports results for the effects estimated in equation (3). The effects are similar to those in Table S7, although the more granular cut of peer groups makes the pattern less striking. The top popularity quintile is the main driver of influence, and we find *negative* influence from the smoking rates of the least popular quintiles.

Heterogeneity

We tested for possible interaction effects between the influence from smokers' status and individual characteristics (gender, age, age relative to grade mean, physical attractiveness rated by interviewer, a dummy for new students, a dummy for foreigners and a dummy for parents claiming to have chosen the neighborhood partly for schools' quality), as well as individual's own network variables (the popularity of the respondent; a dummy signaling whether the respondent's best friend also nominated him/her as a best friend; and the overall percentage of reciprocal ties from the individual's nominations), and school social network characteristics (network density of the school). No statistically significant interaction effects were observed, suggesting that all students are vulnerable to the influence of popular smokers.¹⁵

¹⁵For continuous variables, we regressed separately continuous interactions as well as a dummy signaling values above and below the median. See [27] for details. All results available upon request.

Conclusion

There is a general consensus in the literature that adolescents are susceptible to peer influences, to engage in risky behaviors, including smoking. Previous studies have identified the desire for social status as an important motivation. The clear implication is that influence on smoking behavior from popular peers is stronger than influence from unpopular peers. We test this hypothesis using four waves of panel data from AddHealth. These data are unusual in providing a nearly complete social network of the school population, with which to measure popularity as the respondent's eigenvector centrality in the network. An important limitation of these data is "the reflection problem." We address challenges to causal inference posed by selection and contextual effects using school fixed effects, lagged measures of smoking adoption, and instrumental variables.

The results show that the probability of smoking increases with the lagged popularity of smokers, while the popularity of non-smokers has the opposite effect. These effects persist up to fourteen years after exposure to peers' behavior. Drilling down, we find that most of the aggregate peer effects come from the smoking propensity of the 20% most popular peers. Importantly, we also find evidence of negative social influence, in which respondents avoid the behavior of unpopular teens. The smoking propensity of the 80% least popular peers has a negative influence on individuals' smoking in the long run (seven and fourteen years later).

These findings underscore the importance of the popularity of peers in predicting their influence. Not only is it important to know the smoking propensity in schools but also where the smokers are located in the social hierarchy. For example, our results suggest that two schools where 20% of teens smoke may experience very different trajectories in the incidence of smoking: If those 20% who smoke are the most popular teens, the probability of smoking for students the next year can be expected to increase by 18 percentage points, while if they are among the

least popular, the probability would not increase by a statistically significant amount.¹⁶

These results point to several possible extensions. Although status has been shown to condition peer effects on the diffusion of a number of social contagions, more research is needed to know if popularity conditions peer influence on a range of other risky behaviors, including binge drinking, drug abuse, and stigmatized sexual behaviors. Equally important, we need to know if these status effects extend as well to positive adolescent behaviors such as academic achievement and sports.

¹⁶Collecting the entire social network of schools to compute popularity with a network centrality may be costly; yet, Banerjee et al. [3] find that asking a subsample of individuals who are the influential people can lead to accurate identification of central people.

References

- [1] C. Alexander, M. Piazza, D. Mekos, and T. Valente. Peers, Schools, and Adolescent Cigarette Smoking. *Journal of Adolescent Health*, 29(1):22–30, 2001.
- [2] S. Aral and D. Walker. Identifying Influential and Susceptible Members of Social Networks. *Science*, 337(6092):337–341, July 2012.
- [3] A. V. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson. Gossip: Identifying Central Individuals in a Social Network. *arXiv.org*, June 2014.
- [4] R. Campbell, F. Starkey, J. Holliday, S. Audrey, M. Bloor, N. Parry-Langdon, R. Hughes, and L. Moore. An Informal School-Based Peer-Led Intervention for Smoking Prevention in Adolescence (ASSIST): a Cluster Randomised Trial. *The Lancet*, 371(9624):1595–1602, May 2008.
- [5] N. A. Christakis and J. H. Fowler. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, 357(4):370–379, 2007.
- [6] N. A. Christakis and J. H. Fowler. The Collective Dynamics of Smoking in a Large Social Network. *New England Journal of Medicine*, 358(21):2249–2258, 2008.
- [7] A. E. Clark and Y. Lohéac. “It Wasn’t Me, It Was Them!” Social Influence in Risky Behavior by Adolescents. *Journal of Health Economics*, 26(4):763–784, July 2007.
- [8] E. Cohen-Cole and J. M. Fletcher. Detecting Implausible Social Network Effects in Acne, Height, and Headaches: Longitudinal Analysis. *BMJ*, 337(dec04 2):a2533–a2533, Dec. 2008.
- [9] E. Costenbader and T. W. Valente. The Stability of Centrality Measures When Networks Are Sampled. *Social Networks*, 25(4):283–307, 2003.

- [10] S. T. Ennett, K. E. Bauman, A. Hussong, R. Faris, V. A. Foshee, L. Cai, and R. H. DuRant. The Peer Context of Adolescent Substance Use: Findings from Social Network Analysis. *Journal of Research on Adolescence*, 16(2):159–186, 2006.
- [11] S. A. Everett, C. W. Warren, D. Sharp, L. Kann, C. G. Husten, and L. S. Crossett. Initiation of Cigarette Smoking and Subsequent Smoking Behavior among U.S. High School Students. *Preventive Medicine*, 29(5):327–333, Nov. 1999.
- [12] J. M. Fletcher. Social Interactions and Smoking: Evidence Using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects. *Health Economics*, 19(4):466–484, Apr. 2010.
- [13] M. Gardner and L. Steinberg. Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: an Experimental Study. *Developmental Psychology*, 41(4):625–635, July 2005.
- [14] A. Gaviria and S. Raphael. School-Based Peer Effects and Juvenile Behavior. *Review of Economics and Statistics*, 83(2):257–268, 2001.
- [15] M. Gladwell. *The Tipping Point: How Little Things Can Make a Big Difference*. Little, Brown, 2006.
- [16] M. O. Jackson. *Social and Economic Networks*. Princeton, NJ : Princeton University Press, 2008.
- [17] B. V. Krauth. Simulation-Based Estimation of Peer Effects. *Journal of Econometrics*, 133(1):243–271, July 2006.
- [18] H. A. Lando, D. T. Thai, D. M. Murray, L. A. Robinson, R. W. Jeffery, N. E. Sherwood,

- and D. J. Henrikus. Age of Initiation, Smoking Patterns, and Risk in a Population of Working Adults. *Preventive Medicine*, 29(6 Pt 1):590–598, Dec. 1999.
- [19] P. Lundborg. Having the Wrong Friends? Peer Effects in Adolescent Substance Use. *Journal of Health Economics*, 25(2):214–233, Mar. 2006.
- [20] C. F. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531, 1993.
- [21] L. Mercken, T. A. B. Snijders, C. Steglich, E. Vartiainen, and H. de Vries. Dynamics of Adolescent Friendship Networks and Smoking Behavior. *Social Networks*, 32(1):72–81, Jan. 2010.
- [22] L. Michell and A. Amos. Girls, Pecking Order and Smoking. *Social Science & Medicine*, 44(12):1861–1869, 1997.
- [23] R. Nakajima. Measuring Peer Effects on Youth Smoking Behaviour. *Review of Economic Studies*, 74(3):897–935, 2007.
- [24] E. C. Norton, R. C. Lindrooth, and S. T. Ennett. Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use. *Health Economics*, 7(5):439–453, 1998.
- [25] J. A. Oldmeadow, M. J. Platow, M. Foddy, and D. Anderson. Self-categorization, status, and social influence. *Social Psychology Quarterly*, pages 138–152, 2003.
- [26] L. M. Powell, J. A. Tauras, and H. Ross. The Importance of Peer Effects, Cigarette Prices and Tobacco Control Policies for Youth Smoking Behavior. *Journal of Health Economics*, 24(5):950–968, Sept. 2005.

- [27] J. D. Robalino. Smoking peer effects among adolescents: Are popular teens more influential? *IZA DP*, (9714), 2016.
- [28] B. Sacerdote. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, 116(2):681–704, 2001.
- [29] C. Steglich and T. Snijders. Dynamic Networks and Behavior: Separating Selection from Influence. *Sociological Methodology*, 40(1):329–393, 2010.
- [30] T. W. Valente, J. B. Unger, and C. A. Johnson. Do Popular Students Smoke? The Association between Popularity and Smoking among Middle School Students. *Journal of Adolescent Health*, 37(4):323–329, Oct. 2005.

Supporting information

Table S1: Correlation across measures of smoking.

	Tried by 1995	Tried 1996	1996	2002	2009
Tried by 1995	1				
Tried 1996	0.535	1			
1996	0.492	0.523	1		
2002	0.349	0.391	0.427	1	
2009	0.305	0.338	0.367	0.572	1
by 2009	0.43	0.488	0.399	0.56	0.59

Table S2: Probability of smoking – probit average marginal effects.

	Tried 1996	1996	2002	2009	by 2009
Mean popularity of smokers	0.045*** (0.006)	0.009 (0.009)	0.026*** (0.008)	0.026*** (0.007)	0.033*** (0.009)
Mean popularity of non-smokers	-0.059*** (0.015)	-0.013 (0.011)	-0.032*** (0.012)	-0.028* (0.016)	-0.059*** (0.018)
% smokers in grade	0.191* (0.099)	0.163** (0.071)	-0.053 (0.074)	-0.02 (0.100)	-0.116 (0.103)
Had tried smoking in 1995	0.383*** (0.006)	0.184*** (0.005)	0.152*** (0.009)	0.208*** (0.011)	0.390*** (0.012)
Male	-0.002 (0.009)	0.000 (0.007)	0.034*** (0.007)	0.049*** (0.010)	0.071*** (0.012)
Age (+ Age squ.)	0.176*** (0.062)	0.101** (0.045)	0.078* (0.045)	0.086 (0.062)	0.179*** (0.069)
White	<i>Omitted</i>				
Black	-0.108*** (0.015)	-0.081*** (0.010)	-0.099*** (0.013)	-0.035* (0.019)	-0.066*** (0.022)
Hispanic	0.001 (0.018)	-0.018 (0.014)	-0.066*** (0.018)	-0.069*** (0.018)	-0.02 (0.027)
Asian	-0.022 (0.027)	-0.030** (0.014)	-0.025 (0.022)	-0.035 (0.034)	0.019 (0.032)
Other	0.021 (0.038)	0.054 (0.041)	-0.082*** (0.030)	0.008 (0.057)	0.011 (0.054)
Foreign	-0.023 (0.025)	-0.019 (0.024)	-0.005 (0.021)	-0.028 (0.026)	-0.041* (0.022)
Out of school in 1996	0.062*** (0.023)	0.063*** (0.013)	0.066*** (0.018)	0.057** (0.024)	0.076*** (0.028)
New student in 1995	0.004 (0.011)	0.004 (0.007)	-0.009 (0.011)	0.020* (0.012)	0.024 (0.016)
Weekly earnings (\$100)	0.01 (0.007)	0.011*** (0.004)	0.002 (0.004)	0.006 (0.006)	0.015* (0.008)
Household Income (\$1000 000)	-0.057 (0.119)	-0.231* (0.134)	-0.286* (0.148)	-0.237 (0.203)	-0.316** (0.145)
Moved partly for school quality	-0.01 (0.010)	-0.013* (0.007)	-0.014 (0.008)	-0.017 (0.011)	-0.012 (0.012)
Mother smokes	0.007 (0.012)	0.028*** (0.008)	0.014 (0.011)	0.030*** (0.011)	0.033** (0.015)
Father smokes	0.023** (0.011)	0.016** (0.006)	0.035*** (0.009)	0.014 (0.010)	0.009 (0.012)
Cigarettes at home	0.046*** (0.012)	0.043*** (0.007)	0.033*** (0.009)	0.055*** (0.011)	0.069*** (0.015)
Excise tax per cigarette pack	-0.012*** (0.003)	-0.012 (0.008)	0.005 (0.004)	0.000 (0.003)	-0.002 (0.005)
Tobacco policy funds per capita	0.666*** (0.240)	1.730* (1.012)	1.566** (0.700)	1.028*** (0.391)	2.482*** (0.668)
Marketing prohibited near schools	0.01 (0.126)	-0.027 (0.118)	0.225 (0.157)	0.098 (0.147)	0.094 (0.167)
Marketing restricted on public transport	-0.996*** (0.186)	-1.466 (1.077)	-0.566 (0.670)	-0.151 (0.401)	-2.552*** (0.668)
Pseudo R sq.	0.258	0.363	0.153	0.154	0.193
N	7611	7151	7580	6243	6292

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995.

Table S3: Probability of smoking every day in 1996 for students with and without access to cigarettes at home – probit average marginal effects.

	Access to cigarettes at home	
	Yes	No
Mean popularity of smokers	0.035** (0.015)	0.002 (0.008)
Mean popularity of non-smokers	-0.015 (0.019)	-0.023** (0.011)
% smokers	0.102 (0.126)	0.210*** (0.075)
N	2092	4625

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S4: Probability of smoking controlling for smoking of direct friends – probit average marginal effects.

	Tried 1996	1996	2002	2009	by 2009
Mean popularity of smokers	0.052*** (0.007)	0.007 (0.009)	0.028*** (0.008)	0.024*** (0.007)	0.037*** (0.009)
Mean popularity of non-smokers	-0.065*** (0.016)	-0.012 (0.009)	-0.031*** (0.014)	-0.026 (0.019)	-0.053*** (0.019)
% Smokers	0.117 (0.104)	0.087 (0.069)	-0.156** (0.077)	-0.103 (0.102)	-0.237** (0.107)
% Direct friends smoking	0.117*** (0.024)	0.096*** (0.011)	0.113*** (0.020)	0.128*** (0.026)	0.172*** (0.036)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S5: Subsamples of smokers/non-smokers to test for serial correlation in smoking measures – probit average marginal effects.

Subsample – regular smoker in 1996:	No			Yes	
	2002	2009	by 2009	2002	2009
Mean popularity of smokers	0.015** (0.007)	0.019** (0.009)	0.031*** (0.011)	0.288*** (0.085)	0.049 (0.106)
Mean popularity of non-smokers	-0.011 (0.011)	-0.028* (0.016)	-0.056*** (0.019)	-0.274*** (0.076)	0.014 (0.074)
N	6699	5583	5628	759	599

Subsample – regular smoker by 2002:	No		Yes
	2009	by 2009	2009
Mean popularity of smokers	0.017 (0.018)	0.001 (0.016)	0.031* (0.017)
Mean popularity of non-smokers	-0.040** (0.018)	-0.049** (0.021)	-0.006 (0.029)
N	2550	3215	2033

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S6: Instrumental variable probit regression model 1: popularity of smokers and non-smokers.

	Tried 1996	1996	2002	2009	by 2009
<i>Probit Coef.</i>					
Mean popularity of smokers	0.145 (0.198)	0.314 (0.298)	0.578*** (0.224)	0.123 (0.204)	0.343* (0.202)
Mean popularity of non-smokers	-0.570** (0.245)	-0.345 (0.314)	-0.349 (0.263)	-0.282 (0.291)	-0.640** (0.263)
<i>Marginal effect</i>					
Mean popularity of smokers	0.038	0.04	0.123	0.03	0.105
Mean popularity of non-smokers	-0.148	-0.044	-0.075	-0.069	-0.197
p-value Wald exogeneity test	0.282	0.286	0.11	0.731	0.082
J statistic	13.213	22.279	19.362	15.227	19.399
p-value J statistic (over-identification)	0.657	0.134	0.25	0.508	0.249
F statistic (Mean pop. smokers)	40.937	40.937	41.19	34.869	34.43
p-value F statistic (Mean pop. smokers)	0	0	0	0	0
F statistic (Mean pop. non-smokers)	30.453	30.453	30.332	22.887	23.158
p-value F statistic (Mean pop. non-smokers)	0	0	0	0	0
N	7073	6654	6996	5788	5830

Instrumental variables: percentage of white, foreigners, new students, overweight students, physically attractive, physically mature, well groomed, mean household income, and weekly earnings. IVs are computed both among smokers and among non-smokers. These variables are only available in the InHome survey, hence, we compute them from the InHome sample to instrument the mean popularity in the InSchool sample (see Fig 1). Regressions include school fixed effects. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S7: Probability of smoking and smoking rates among popular and non-popular students – probit average marginal effects.

	Tried 1996	1996	2002	2009	by 2009
<i>Smoking propensity:</i>					
20% most popular	0.178*** (0.056)	0.055 (0.040)	0.170*** (0.037)	0.022 (0.063)	0.152** (0.067)
80% least popular	0.04 (0.092)	0.107** (0.052)	-0.221*** (0.066)	-0.045 (0.095)	-0.330*** (0.098)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S8: Instrumental variable probit regression model 2: smoking rate among popular and non-popular.

	Tried 1996	1996	2002	2009	by 2009
<i>Probit Coef.</i>					
smoking rate among 20% most pop.	0.879 (0.834)	0.532 (1.093)	0.664 (0.911)	1.258 (0.881)	0.753 (0.799)
smoking rate among 80% lest pop.	-0.284 (0.904)	0.021 (1.145)	-0.837 (0.952)	0.051 (1.014)	0.083 (0.923)
<i>Marginal effect</i>					
smoking rate among 20% most pop.	0.228	0.066	0.141	0.305	0.235
smoking rate among 80% lest pop.	-0.074	0.003	-0.177	0.012	0.026
p-value Wald exogeneity test	0.869	0.667	0.964	0.236	0.26
J statistic	26.762	23.349	16.108	12.045	17.173
p-value J statistic (over-identification)	0.021	0.055	0.307	0.603	0.247
F statistic (sm. rate 20% most pop.)	35.284	35.284	35.033	32.804	32.794
p-value F statistic (sm. rate 20% most pop.)	0	0	0	0	0
F statistic (sm. rate 80% least pop.)	105.28	105.28	104.445	89.464	89.433
p-value F statistic (sm. rate 80% least pop.)	0	0	0	0	0
N	7769	7232	7736	6387	6433

Instrumental variables: percentage of black students, students with older siblings, parents who smoke, households with smokers, parents home after school finishes, average household income and average weekly earnings. IVs are computed both among the 20% most popular and the 80% least popular students. These variables are only available in the InHome survey, hence, we compute them from the InHome sample to instrument the mean popularity in the InSchool sample (see Fig 1). Regressions include school fixed effects. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Table S9: Probability of smoking and smoking propensity by popularity quintiles of peers – probit average marginal effects.

	Tried 1996	1996	2002	2009	by 2009
<i>Smoking propensity:</i>					
Top pop. quintile	0.161*** (0.057)	0.051 (0.040)	0.179*** (0.039)	0.02 (0.064)	0.144** (0.067)
2nd pop. quintile	0.072 (0.058)	0.038 (0.038)	-0.092 (0.058)	-0.024 (0.051)	-0.077 (0.071)
3rd pop. quintile	0.078 (0.051)	0.064 (0.051)	-0.024 (0.057)	0.102* (0.057)	0.067 (0.070)
4th pop. quintile	-0.094* (0.050)	-0.006 (0.034)	-0.026 (0.050)	-0.062 (0.055)	-0.242*** (0.052)
5th pop. quintile	-0.018 (0.053)	0.008 (0.030)	-0.074* (0.045)	-0.033 (0.058)	-0.053 (0.069)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Number of cigarettes per month and age of initiation

In Table S10 we regress the number of cigarettes consumed per month, as well as the age at which individuals tried their first cigarette and the age at which they smoked everyday for the first time. To account for non-smokers, we estimate a Tobit censored equation analogous to equation (1). We find that the same patterns reported above regarding peers' popularity apply to the number of cigarettes and the age of initiation. Conditional on smoking, a standard deviation increase in the mean popularity of smokers increases consumption by 7 cigarettes per month in 1996, 50 cigarettes per month in 2002 and 8 per month in 2009; and the same variation in the popularity of non-smokers decreases consumption by 8 cigarettes per month in 1996, 80 cigarettes per month in 2002 and 20 in 2009.

Table S10: Number of cigarettes smoked per month and age of initiation – Tobit average marginal effects.

	Number of cigarettes/ month			Age	
	1996	2002	2009	1st cig.	smoked everyday
Mean popularity smokers	7.067*** (1.849)	50.330*** (2.457)	8.235** (3.361)	-0.822*** (0.227)	-0.641*** (0.158)
Mean popularity non-smokers	-8.142** (3.465)	-79.163*** (3.037)	-19.214*** (7.193)	1.061*** (0.387)	1.101*** (0.309)
% smokers	87.190*** (25.474)	89.301*** (15.055)	-64.441* (37.196)	1.472 (2.724)	2.631 (2.023)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least "once or twice a week" in 1995. Peer variables are at the grade level. Includes all covariates from Table S2.

Regarding the age of initiation, conditional on ever trying cigarettes, we find that a standard deviation above the mean popularity of smokers advances the age at which individuals try their first cigarette by 10 months; and the same variation in the popularity of non-smokers delays the age at which individuals try their first cigarette by over a year. Similarly, conditional on ever smoking everyday, we find that a standard deviation above the mean popularity of smokers advances the age at which individuals started smoking everyday by almost 8 months; and the same variation in the popularity of non-smokers delays the age at which individuals started

smoking everyday by over a year.

Table S11: Number of cigarettes smoked per month and age of initiation – model 2 Tobit average marginal effects conditional on smoking.

	Number of cigarettes/ month			Age	
	1996	2002	2009	1st cig.	smoked everyday
<i>Smoking propensity:</i>					
20% most popular	25.972** (11.446)	334.335*** (23.399)	-1.993 (26.049)	-2.829** (1.252)	-2.461** (1.242)
80% least popular	60.358*** (21.144)	-119.426*** (5.754)	-64.777* (35.779)	3.569 (2.478)	5.843*** (1.785)

Regressions include school fixed effects. Standard errors clustered at the school level are shown in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Peer variables are at the grade level. Includes all covariates from S2 Table .

Table S11 reports estimates for Tobit regressions analogous to model (2). If all of the 20% most popular students smoked, cigarettes consumption would increase by 26 cigarettes in 1996, and by 334 cigarettes in 2002. The same variation advances the age of the first cigarette by almost three years and the age when an individual first started smoking everyday by 2.5 years. If 25% of the 80% least popular teens (an equivalent 20% of the total students) cigarettes consumption would increase by 14 cigarettes in 1996, but it would decrease by 25 cigarettes in 2002. The same variation delays the age when first started smoking everyday by almost 1.5 years.¹⁷

These patterns are of particular importance in light of the existing evidence that smoking at a younger age has a strong impact in the number of cigarettes smoked and the probability of quitting smoking later in life (Everett et al., 1999; Lando et al., 1999). Popularity of peer smokers during the teenage years seems to be an important driver of these findings.

¹⁷We find similar patterns using peers’ smoking propensity by popularity quintiles; results available upon request.