

ADOLESCENT CIGARETTE USE: EFFECTS OF RACE/ETHNICITY AND SCHOOL RACIAL/ETHNIC COMPOSITION

Robert A. Johnson
Abt Associates Inc.

Abstract

Research has shown there exists substantial variability in substance use among U.S. schools but has been unable to separate variability due to differences among schools from variability due to differences among the populations that are served by schools. To disentangle the effects of explanatory variables operating at the levels of individuals, families, and schools, we applied multilevel models (mixed models or hierarchical models) to self-report data on daily cigarette use collected in the eighth and tenth grade panels of the National Educational Longitudinal Study (NELS). The response variable is cigarette initiation (first use), coded "1" if the adolescent initiated daily cigarette use (at least one cigarette per day) during the interval between the baseline NELS interview and the reinterview conducted two years later; and coded "0" if the adolescent did not smoke on a daily basis at either wave. The final model used 18 explanatory variables- including 7 individual, 4 family, and 7 school variables- and incorporated "cross-level interaction effects"- interactions between school and family/individual variables- and variance components gauging differences among schools in the effects of individual and family variables.

This paper focuses on a subset of the results, those pertaining to the effects of racial/ethnic minority status and the interactions between minority status and the racial/ethnic composition of students in the school. The findings support research showing that Asian/Pacific Islander, black, and Hispanic adolescents are substantially less likely than white non-Hispanic adolescents to report cigarette use. In each of the NELS eighth and tenth grade panels, API and black adolescents are less than half as likely- and Hispanics about three-fourths as likely- to initiate daily cigarette use as other adolescents. The findings go beyond previous research by showing that the deterrent effects of minority status on cigarette use are much larger among minority adolescents who attend predominantly minority schools. Especially in the NELS tenth grade panel, where the effect is to reduce cigarette risk by one-half, minority schools appear to reinforce the effect of an individual's minority status, so minority students are at even lower risk of cigarette use if they attend schools with a high percentage of minority students. The results also suggest that minority schools help to mitigate the adverse effects on cigarette risk of two social conditions- fewer than two parents living at home and school dropout- that are prevalent within the largest racial/ethnic minority populations of the U.S.

1. Introduction. Research on adolescent cigarette use has been concerned mainly with the effects of individual and family variables. Many studies document associations between cigarette use and family attachment, school involvement, peer smoking, and other individual and family variables (e.g., Akers and Lee 1996; Ennett and Bauman, 1993). Recent research suggests that cigarette use also varies by type of school (Ennett et al. 1997; Skager and Fisher 1989).

A drawback is the failure to link school with individual and family explanatory variables. To reduce omitted-variables bias, factors operating at the different levels need to be included in the same model. School variables may also condition the effects of variables operating at the individual and family levels. For example, one might expect racial/ethnic differences in cigarette use to be intensified in schools with a high percentage of minority students. If minority adolescents are less likely than others to use cigarettes (Fendrich and Vaughan, 1994⁽¹⁾; Johnson and Larison, 1998), then predominantly minority schools may reinforce this effect by offering a normative climate in opposition to cigarette use, a "cross-level interaction effect" (Bryk and Raudenbush, 1992).

The multilevel modeling approach applied in this paper (Bryk and Raudenbush 1992; Goldstein 1995; Kreft and De Leeuw 1999) has a number of advantages, including variance estimates that take into account the data hierarchy, such as clustering of sample students within schools. For our purposes- and those of public policy research generally- the most important advantage may be that multilevel models can yield consistent estimates of cross-level interaction effects (see the next section). We applied multilevel models to the National Educational Longitudinal Study (NELS) to explore how individual, family, and school characteristics affect adolescent cigarette use. This paper reports our findings about the effects of race/ethnicity and school racial/ethnic composition.⁽²⁾

Adolescent cigarette use is only one of many social outcomes that are relevant to policies affecting the racial/ethnic composition of schools. Attempts to manipulate school composition as a matter of law and policy have a contentious history in the U.S. (Bok, 1996). Supreme Court rulings in the 1960s favored racial integration of the public schools, and the percentage of black students enrolled in predominantly (more than half) minority schools declined from about 77%

in 1968 to 64% in 1973. However, school segregation has remained roughly constant since the early 1970s (Orfeld, 1993). Based on NELS, the percentages of API, black, Hispanic, and other eighth graders who were in predominantly minority schools in 1988 were 34%, 64%, 65%, and 40%, respectively. The corresponding percentages of tenth graders in 1990 were 41%, 66%, 68%, and 37%.

The following sections discuss the advantages of multilevel models for policy research, the data and methods of our application, and the results.

2. Multilevel models in policy research. The standard single-level regression model has long been the model of choice in policy-oriented research. Let y_{ij} be a continuous response measured for the i -th student in the j -th school; x_{ij} an individual-level explanatory variable measured for the same student; and z_j a school-level variable measured for the j -th school. The model can be written

$$y_{ij} = a + b x_{ij} + c z_j + d (x_{ij} z_j) + e_{ij}, \quad (1)$$

where e_{ij} is student-level error with zero mean; and a , b , c , and d are regression coefficients. The error e_{ij} represents student-level variables that are not included in the model and that affect y_{ij} .

It is instructive to write the single-level model as a "pseudo two-level model" by defining a school-specific intercept a_j equal to $(a + c z_j)$:

$$\text{Level 1 (students): } y_{ij} = a_j + b x_{ij} + d (x_{ij} z_j) + e_{ij}$$

$$\text{Level 2 (schools): } a_j = a + c z_j \quad (1')$$

But (1) is not a true multilevel model because there is no random error at Level 2. The single-level model allows unmeasured variables at the student level, but not at the school level.

The multilevel approach introduces the idea of separate regressions in each school or context:

$$\text{Level 1 (students): } y_{ij} = a_j + b_j x_{ij} + e_{ij}$$

$$\text{Level 2 (schools): } a_j = a + c z_j + u_{1j}$$

$$b_j = b + d z_j + u_{2j} \quad (2)$$

The key property of the level-1 equation is that the regression intercept and slope of y_{ij} on x_{ij} each have a subscript "j," which implies that these parameters can vary across schools. The level-1 regressions are linked by a level-2 model, where the regression coefficients of the level-1 model are themselves regressed on the school explanatory variable z_j . As in the single-level model, additional assumptions are needed to estimate the model, the main ones being that the level-1 error (e_{ij}) and level-2 errors (u_{1j} and u_{2j}) are uncorrelated with each other and with the explanatory variables.

For comparison with the single-level model, we can write the two-level model as a single equation by substituting the right-hand-sides of the level-2 equations for a_j and b_j in the level-1 equation:

$$y_{ij} = a + b x_{ij} + c z_j + d (x_{ij} z_j) + (e_{ij} + u_{1j} + x_{ij} u_{2j}). \quad (2)$$

Comparing (1) with (2) shows the only difference is in the assumed error structure.

An important parameter for policy is d , the cross-level interaction. Individuals and families in the U.S. are afforded many legal protections, so schools are the principal lever of drug prevention policy. Cross-level interactions may be critical paths by which school policies can impact individual behavior. Yet, if d is estimated using (1) when the true

error is that of (2), the estimate is inconsistent, because the error in (2) is correlated with $(x_{ij}; z_j)$. Thus, if there exist unmeasured school variables- as there almost surely are- good estimates of cross-level interaction effects might not be possible using a single-level model.

3. Data and methods

a. Sample and data collection design. The longitudinal design of NELS (NCES, 1992) allows us to gauge changes in cigarette use between measurement waves and to control for whether or not respondents used cigarettes at the prior wave. Most research on cigarette use in the U.S. has used cross-sectional rather than longitudinal data, perhaps because the National Household Survey on Drug Abuse (NHSDA) and Monitoring the Future (MTF)-two major surveys designed to measure substance use-are cross-sectional in design. Yet retrospective reporting can bias responses about past drug use obtained from cross-sectional surveys (Fendrich and Vaughn, 1994; Johnson et al. 1998).

We split the NELS longitudinal file into two panels to investigate cigarette use separately among eighth graders in 1988 and tenth graders in 1990. Eighth grade panel members were first interviewed as eighth graders in Fall 1988 (Wave 1) and reinterviewed two years later (Wave 2). Tenth grade panel members were first interviewed as tenth graders in Fall 1990 (Wave 1) and reinterviewed two years later (Wave 2).⁽³⁾ Both panels followed up school drop-outs. About 6.8% of eighth grade panel members and 10.4% of tenth grade panel members dropped out before Wave 2. The adolescent interviews used traditional personal interviewing techniques. Personal interviews of parents and school administrators were also conducted at Wave 1 of each panel.

Both panels are based on a two-stage national probability sample of U.S. students: Stratified random sampling of schools was followed by random sampling of eligible students within schools. In our analysis, the eighth grade panel consists of 17,424 adolescents in 1,014 schools who responded to both interviews. The tenth grade panel consists of 16,542 adolescents in 1,464 schools who responded to both interviews. We used standard NELS weights (NCES, 1992) to adjust for unit nonresponse and unequal selection probabilities. We used techniques described in Pfeiffermann et al. (1997)- as implemented in the program MLWIN (www.ioe.ac.uk/mlwin) to incorporate the NELS weights in the multilevel model estimation.

b. Measurement of daily cigarette use. Daily cigarette use at each wave of each panel was measured based on responses to the question "How many cigarettes do you usually smoke in a day?" We collapsed the response categories at each wave to form a binary variable: 1 = One or more cigarettes per day; 0 = Not a daily smoker. Item nonresponse was small, ranging from 2.3% at Wave 1 of the eighth grade panel to 7.1% at Wave 2 of the tenth grade panel. To impute the missing data, we used techniques for multilevel models described by Schafer (1996, 1997).⁽⁴⁾

c. Measurement of race/ethnicity and school racial/ethnic composition. The adolescent's race/ethnicity was measured at Wave 1 using response categories similar to the 1990 Census. This report distinguishes four categories- Asian and Pacific Islander (API), black, Hispanic, and other.⁽⁵⁾ School racial/ethnic composition is based on the Wave-1 school administrator interviews. The administrators were asked about the percentage of students who were non-white or Hispanic. To reduce measurement error, we recoded the responses to form a binary variable, equal to 0 if the minority percentage was less than 50% and equal to 1 otherwise. Missing values were imputed based on the distributions of races and ethnicities reported by sample students in the school.

d. Measurement of other explanatory variables. There are seven additional explanatory variables at the individual and family levels: gender (based on Wave-1 adolescent interviews); family income (Wave 1 parental interviews); two biological parents at home (Wave-1 adolescents)⁽⁶⁾; school dropout (Wave-2 follow up); negative peer relations scale (Wave-1 adolescents- number of affirmative answers to five questions about how school peers viewed the respondent, e.g., as a poor student); school participation scale (Wave-1 adolescents- number out of nine activities, e.g., athletics, music); and parental support scale (Wave-1 adolescents- number of affirmative answers to ten questions about the respondent's parent(s), e.g., whether a parent helped with home work).⁽⁷⁾ There are six additional explanatory variables at the school level: region; type of place (central city v. other metro v. nonmetro); type of school (public v. Catholic v. other); school size; average beginning teacher salary; and student-teacher ratio. The first four are from the school sampling frame (NCES, 1992); the last two from administrator interviews.⁽⁸⁾ Descriptive statistics for all variables-

means, variances, and intercorrelations- are in Johnson and Hoffmann (1999). Prior to analysis, continuous variables were "centered" by subtracting their means (Bryk and Raudenbush, 1992).

e. Statistical models. This paper presents results based on two multilevel models- called Model 1 and Model 2. Model 1 is a 2-level "variance-components model" with a logit-linked binary response and one fixed covariate. We use Model 1 to underscore the importance of cigarette initiation as a response variable, so it is convenient to present both the model and the results in this section. Let y_{ij} denote a binary (0-1) response variable indicating daily cigarette use at Wave 2 and let p_{ij} denote the corresponding probability of using cigarettes daily at Wave 2. Let x_{ij} denote a binary (0-1) response variable indicating daily cigarette use at Wave 1. The model is written

$$\begin{aligned} \text{Level 1 (adolescents): } & y_{ij} = p_{ij} + e_{ij} \\ & \text{logit}(p_{ij}) = a_j + b x_{ij} \\ \text{Level 2 (schools): } & a_j = a + u_j, \end{aligned} \quad (3)$$

where $\text{logit}(p_{ij}) = \log(p_{ij}/(1 - p_{ij}))$; "log" denotes the natural logarithm; e_{ij} is a level-1 random error; and u_j is a level-2 random error. We assume that y_{ij} is distributed as an extra-Bernoulli variable with mean p_{ij} , so e_{ij} has mean 0 and variance $\text{SIGMA } e^2 = k p_{ij} (1 - p_{ij})$.⁽⁹⁾ We also assume that u_j is normal with mean 0 and variance u^2 and that the level-1 and level-2 errors are independent.⁽¹⁰⁾

Table 1 shows the Model 1 parameter estimates.⁽¹¹⁾ The slope parameter b gauges the dependence of daily cigarette use on daily cigarette use two years earlier. For the eighth grade panel, the estimated b of 2.77 corresponds to an odds-ratio of current relative to past smoking of about $\exp(2.77) = 16$. That is, an eighth grade panel member is about 16 times more likely to be a daily smoker if he (she) was a daily smoker two years ago than if he was not. For the tenth grade panel, the corresponding estimate equals about $\exp(2.99) = 20$. The increase in the odds-ratio may reflect that addiction becomes more severe the longer an individual uses cigarettes. If so, it makes sense to try to prevent adolescents from ever using cigarettes for the first time.⁽¹²⁾ Past cigarette use is such a strong predictor of current use that it might be misleading to include past smokers and nonsmokers in the same model. We examined separate models for cigarette initiation and cessation and found that most explanatory variables interact with past use. NELS data on initiation are more plentiful than data on cessation, so this paper focuses on initiation.

Model 2 uses daily cigarette initiation between Waves 1 and 2 -rather than cigarette use at Wave 2-as the response variable. That is, y_{ij} equals 1 if the adolescent began daily cigarette use between Waves 1 and 2; and y_{ij} equals 0 if the adolescent was a daily nonsmoker at both waves. The analysis is restricted to daily nonsmokers at Wave 1, which reduces the sample size from 17,424 to 16,454 in the eighth grade panel and from 16,542 to 13,840 in the tenth grade panel. Model 2 also extends Model 1 by adding individual and family explanatory variables at level 1 and school explanatory variables at level 2. We assume P level-1 explanatory variables, denoted x_{pij} , $p = 1, \dots, P$; and Q level-2 explanatory variables, denoted w_{qj} , $q = 1, \dots, Q$. Model 2 is written:

$$\begin{aligned} \text{Level 1 (adolescents): } & y_{ij} = p_{ij} + e_{ij} \\ & \text{logit}(p_{ij}) = a_j + \text{SUM}_p b_{pj} x_{pij} \\ \text{Level 2 (schools): } & a_j = a + \text{SUM}_q c_{0q} w_{qj} + u_{0j} \\ & b_{1j} = b_1 + \text{SUM}_q c_{1q} w_{qj} + u_{1j} \\ & \dots \\ & b_{Pj} = b_P + \text{SUM}_q c_{Pq} w_{qj} + u_{Pj}, \end{aligned} \quad (4)$$

where the summations extend from $p = 1$ to $p = P$ at Level 1 and from $q = 1$ to $q = Q$ at Level 2.

The first level of (4) is similar to (3), except that p_{ij} - the probability of initiation- depends upon a school-specific intercept- a_j - and upon school-specific slopes- b_{1j} through b_{pj} . In the $(P + 1)$ level-2 equations, the level-1 regression intercept and slopes are themselves treated as response variables. Each is regressed on Q school-level explanatory variables. For example, in the equation for b_{1j} , b_1 is the average across schools of the slope of $\text{logit}(p_{ij})$ on x_{1ij} ; c_{11} is the effect on b_{1j} of a unit increase in w_{1j} ; and u_{1j} is the level-2 random error associated with b_{1j} .⁽¹³⁾

Substituting the right-hand-side of each level-2 equation of (4) into Level 1 expresses $\text{logit}(p_{ij})$ as a function of the x_{pij} 's, the w_{qj} 's, and their products- the $x_{pij}w_{qj}$'s:

$$\text{logit}(p_{ij}) = a_j + \sum_p b_{pj} x_{pij} + \sum_q c_{0q} w_{qj} + \sum_p \sum_q c_{0pq} (x_{pij} w_{qj}) + u_{0j} + \sum_p x_{pij} u_{pj} \quad (4)$$

The coefficients of the $x_{pij}w_{qj}$'s - the c_{pq} 's - gauge the cross-level interaction effects, showing how school variables amplify or dampen the effects of individual and family variables.

The next section presents a simplified version of Model 2 in which interactions and school-level variances and covariances that were not statistically significant in either NELLS panel have been omitted from the model. In analyzing each panel, we tested each fixed and random parameter in the full model using Wald tests and found that only nine cross-level interactions and five school-level variance components were significant in one or both panels.

4. Results. Table 2 presents parameter estimates for Model 2. Results are presented both on the logit scale and on the scale of odds-ratios. For example, on the logit scale, the effect of dropout status on cigarette initiation equals 1.13 in the eighth grade panel. This implies that, after controlling for other explanatory variables, dropouts in the eighth grade panel were about $\exp(1.13) = 3.1$ times as likely as nondropouts to initiate daily use between Wave 1 and Wave 2.

Table 2 supports previous research showing that racial/ethnic minority status, two-parent families, parental support, school participation, urban residence, residence in the West, and predominantly minority schools are associated with reduced cigarette risk. Negative peer associations, school dropout, Catholic schools, and higher student/teacher ratios are associated with increased risk. The effects of family income, school size, and beginning teacher salary are not significantly different from zero in either panel. Most significant associations appear fairly stable across panels, with gender being the most notable exception.⁽¹⁴⁾

Yet the estimated main effects can be misleading unless cross-level interactions and school-level variance components are taken into account. The cross-level interactions (fourth set of estimates in Table 2) show that four school variables- school minority percentage, residence in the West, student/teacher ratio, and urban residence- condition the effects of one or more individual and family explanatory variables in one or both panels. The most important effects of school variables appear not as main effects but as cross-level interactions. The school variance components (fifth set) show there is significant unexplained school-level variation in the overall level of cigarette initiation and in the effects on initiation of parental support, school participation, gender, and minority status.

Tables 3 and 4 re-express Table 2's results on the effects of school racial/ethnic composition in the forms of (a) percentages initiating cigarette use and (b) odds-ratios comparing minority with other adolescents. Table 3 shows that, in predominantly minority schools, the deterrent effect of racial/ethnic minority status on cigarette initiation becomes stronger. This is especially true in the tenth grade panel, where the percentage of black adolescents initiating daily cigarette use increases from 1.9% among those attending minority schools to 4.7% among those attending non-minority schools. The advantage of odds ratios over percentages is that they do not depend on the values of other explanatory variables (see Table 3 footnote). In the tenth grade panel, the overall odds on initiation of a black adolescent equals 0.31 times the odds of an other (non-API, non-black, and non-Hispanic) adolescent, but this odds ratio declines to 0.19

among those attending minority schools.

Table 4 presents a more detailed portrait of school composition effects, taking into account differences by region, dropout status, family type, and gender. The statistics presented are the percentages of minority adolescents initiating cigarette use and the odds ratios on initiation of minority (API, black, or Hispanic) relative to other adolescents. Both statistics are based upon Model 2's estimates of the main effects of six variables- minority status, school minority composition, region, dropout status, family type, and gender- and the interaction effects involving one or both of minority status and school minority composition. Since the school composition variable does not distinguish among predominantly API, black, and Hispanic schools, the percentages and odds ratios in Table 4 are weighted averages across API, black, and Hispanic adolescents. The positive Minority-by-West interaction (see Table 2) does indicate higher cigarette risk among minority adolescents in the West, where Hispanics are most heavily concentrated.

Comparisons of the percentages in Table 4 suggest that minority schools reduce the adverse effects of two social conditions- fewer than two parents at home and school dropout. For example, among minority female eighth graders who reside in non-Western regions, the effect of having fewer than two parents at home is to increase the percentage initiating cigarette use by 2.2 percentage points- from 6.2% to 8.4%- if the adolescent attends a non-minority school, but the increase is only 0.4 percentage points- from 4.8% to 5.2%- if she attends a minority school. In the same subpopulation, the effect of dropping out is to increase the percentage by 23.8 percentage points- from 8.4% to 32.2%- if the minority female attends a non-minority school but the increase is only 8.1 percentage points- from 5.2% to 13.3%- if she attends a minority school. The odds ratios suggest male dropouts are an exception to the generalization that minority adolescents are less likely to initiate cigarette use. Yet, regardless of gender, the adverse effect of school dropout is greatly reduced among those who attended minority schools.

In conclusion, the results suggest that schools with supportive environments for minority students can override some adverse influences arising from the adolescents' homes and neighborhoods or from the larger society. Yet two limitations of our measures and models point to the need for further research. First, our measure of negative peer relations pertains strictly to the respondent's perceptions of how peers in the same school regard him or her. Second, NELS provided no measure of parental cigarette use. Incorporating measures of peer cigarette use, relations with peers who do not attend the same school, and parental cigarette use might improve the model.

5. References

- Akers, R and G Lee 1996 "A longitudinal test of social learning theory: adolescent smoking." *Journal of Drug Issues* 26:317-343.
- Bok, D, 1996. *The State of the Nation*. Cambridge: Harvard U.
- Bryk, A. and S. Raudenbush. (1992). *Hierarchical Linear Models*. Newbury Park, CA: Sage.
- Ennett, S. and K. Bauman. 1993. "Peer group structure and adolescent cigarette smoking: a social network analysis." *Journal of Health and Social Behavior* 34: 226-236.
- Ennett S, R Flewelling, R Lindrooth, E Norton 1997. "School characteristics associated with school rates of alcohol, cigarette, and marijuana use." *J. Health and Social Behavior* 38: 55-71.
- Fendrich, M. and C. Vaughn, 1994. "Diminished lifetime substance use over time: an inquiry into differential underreporting." *Public Opinion Quarterly* 58: 96-123.
- Goldstein, H. 1995. *Multilevel Statistical Models*. London: Edward Arnold..
- Johnson, R. and D. Gerstein. 1998. "Initiation of use of alcohol, cigarettes, marijuana, cocaine and other substances in US birth cohorts since 1919." *Am. J. Public Health* 88:27-33.
- Johnson, R., D. Gerstein, and K. Rasinski. 1998. "Adjusting survey estimates for response bias" *Public Opinion Quarterly* 62:354-377.

Johnson, R. and C. Larison (1998). *Prevalence of Substance Use among Racial and Ethnic Subgroups in the United States, 1991-93*. Rockville, MD: SAMHSA/OAS.

Johnson, R. and J. Hoffmann, 1999. "School, family, and individual risk factors in adolescent cigarette use: a multilevel analysis." Under review by *Journal of Health and Social Behavior*.

Kreft, I. and J. De Leeuw, 1999. *Introducing Multilevel Models*. London: Sage

National Center for Educational Statistics. 1992. *NELS 1988 User's Manual: First Followup Student Component*, NCES 1992-030. Washington, DC: U.S. Department of Education.

Orfeld, G., 1993. *The Growth of Segregation in American Schools: Changing Patterns of Separation and Poverty since 1968*. Cambridge: Harvard U.

Pfeffermann, D., C. Skinner, D. Holmes, H. Goldstein, and J. Rasbash, 1997. "Weighting for unequal selection probabilities in multilevel models." *J. Royal Stat. Society, B*, 23-40.

Schafer, J. 1996. *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.

Schafer, J. 1997. "Imputation of missing covariates under a multivariate linear mixed model." Department of Statistics. University Park, PA: Pennsylvania State University.

Skager, R. and D. Fisher. 1989. "Substance use among high school students in relation to school characteristics." *Addictive Behaviors* 14: 129-138.

Table 1. Model 1 estimates. Daily cigarette use at Wave 2.

a = average school intercept.

b = slope of Wave 1 cigarette use.

s_u^2 = school-level variance.

Standard errors in parentheses. NELS 1988.

<i>Eighth grade panel</i>			<i>Tenth grade panel</i>		
a	b	s_u^2	a	b	s_u^2
-1.77 (.03)	2.77 (.08)	0.09 (.02)	-1.95 (.03)	2.99 (.06)	0.03 (.01)

Table 2. Model 2 estimates. Initiation of daily cigarette use between Waves 1 and 2. NELS 1988.

Odds ratios (OR) gauging change between panels.

<i>Parameter</i>	<i>Eighth grade</i>	<i>Tenth grade</i>	<i>Change between panels</i>
<i>1. Individual and family variables</i>			
Intercept	-1.35_(.06)	-1.85_(.07)	Base %: _ 21% to 14%*
Race/ethnicity: API	-1.00_(.15)	-0.67_(.15)	OR: _0.36 to 0.51
Black	-1.32_(.13)	-0.98_(.14)	OR: _0.27 to 0.38
Hispanic	-0.53_(.12)	-0.10_(.13)	OR: _0.59 to 0.90
Male	-0.23_(.06)	0.16_(.06)	OR: _0.79 to 1.17*
Negative peers (max v. average)	0.92_(.08)	0.42_(.11)	OR: _2.51 to 1.52*
Participation (add 2 activities)	-0.07_(.04)	-0.21_(.05)	OR: _0.93 to 0.81*
Dropout before Wave 2	1.13_(.12)	0.98_(.12)	OR: _3.10 to 2.66
2 biological parents at home	-0.31_(.05)	-0.24_(.06)	OR: _0.73 to 0.79

Parental support (max v. average)	-0.14_(.04)	-0.10_(.04)	OR:_.087 to 0.90
Family income- lowest quartile	0.07_(.06)	-0.12_(.07)	OR:_.107 to 0.89*
_____ - highest quartile	-0.03_(.06)	0.08_(.07)	OR:_.097 to 1.08
<i>2. Interactions between individual and family variables</i>			
Parental support by Participation	-0.11_(.04)	-0.07_(.05)	OR:_.090 to 0.93
Male by Minority by Dropout	1.08_(.24)	0.80_(.22)	OR:_.294 to 2.22
<i>3. School variables</i>			
West region	-0.44_(.09)	-0.17_(.10)	OR:_.064 to 0.84
Urban place	-0.08_(.07)	-0.18_(.09)	OR:_.092 to 0.84
Catholic school	0.29_(.13)	0.24_(.15)	OR:_.134 to 1.27
School % minority > 50%	-0.30_(.17)	-0.23_(.17)	OR:_.074 to 0.79
Small size: < 600 students	-0.11_(.07)	-0.04_(.11)	OR:_.090 to 0.96
Beginning teach salary: \$1000s	-0.01_(.01)	0.01_(.01)	OR:_.099 to 1.01
Stud-teach ratio (increase by 10)	0.08_(.04)	0.04_(.06)	OR:_.092 to 0.96
<i>4. Cross level interactions</i>			
Minority X School % minority	-0.21_(.17)	-0.69_(.17)	OR:_.081 to 0.50*
Male X School % minority	0.26_(.14)	0.32_(.15)	OR:_.130 to 1.38
2 biol. pars X School % minority	0.23_(.13)	0.34_(.15)	OR:_.126 to 1.40
Dropout X School % minority	-0.62_(.23)	-0.29_(.21)	OR:_.054 to 0.75
Minority X West region	0.48_(.16)	0.14_(.17)	OR:_.162 to 1.15
Parental support X Stud/teach	-0.04_(.05)	-0.13_(.06)	OR:_.096 to 0.88
Participation X Stud/teach	0.08_(.04)	0.07_(.07)	OR:_.108 to 1.07
Parental support X Urban place	0.17_(.08)	0.11_(.08)	OR:_.119 to 1.12
Dropout X Urban place	0.52_(.22)	0.24_(.21)	OR:_.168 to 1.27
<i>5. Random parameters- Variability among schools</i>			
Var(Intercept)_____	0.23_(.04)	0.08_(.02)	Difference = -0.15 (.04)*
Var(Parental support)	0.13_(.03)	0.12_(.02)	Difference = -0.01 (.04)
Var(School participation)	0.15_(.03)	0.13_(.03)	Difference = -0.02 (.04)
Var(Male)	0.34_(.07)	0.19_(.05)	Difference = -0.15 (.09)
Var(Minority)	0.45_(.11)	0.13_(.06)	Difference = -0.32 (.12)*
Corr(Intercept, Male)	-0.43_(.15)	-0.38_(.24)	Difference = 0.05 (.28)

* Significant based on two-sample two-tail t-test assuming independent samples, alpha = .05

Table 3. Model 2 estimates.
Percentages of API, black, Hispanic, and other adolescents initiating daily cigarette use between Waves 1 and 2.
Odds ratios. By school racial/ethnic composition.* NELS 1988.

<i>School racial/ethnic composition</i>	<i>Race/ethnicity</i>			
	<i>API %__OR</i>	<i>Black %__OR</i>	<i>Hispanic %__OR</i>	<i>Other %__OR</i>
<i>Eighth grade panel</i>				

Total schools	7.6%__0.37	5.4%__0.26	11.1%__0.56	18.0%__1.0
Minority % > 50	5.0%__0.30	3.7%__0.22	7.8%__0.48	14.9%__1.0
Minority % 50 or less	8.0%__0.37	6.0%__0.27	12.2%__0.59	19.0%__1.0
<i>Tenth grade panel</i>				
Total schools	5.0%__0.43	3.7%__0.31	8.4%__0.76	10.8%__1.0
Minority % > 50	2.6%__0.26	1.9%__0.19	4.5%__0.46	9.3%__1.0
Minority % 50 or less	6.3%__0.51	4.7%__0.38	10.4%__0.90	11.4%__1.0

*Since continuous covariates have been centered at their means, percentages pertain to an average non-dropout female with fewer than two parents who attended a non-Catholic school in an urban area outside the West.

Table 4. Model 2 Estimates. Percentages of minority adolescents initiating daily cigarette use and odds ratios of minority vs. other.

By school minority composition, gender, region, dropout, and family type. *

Region	Dropout status	Family type	School % minority 50%		School % minority > 50%	
			Female %__OR	Male %__OR	Female %__OR	Male %__OR
<i>Eighth grade panel</i>						
NE, NC, S	Non-drop	< 2 parents	8.4%__0.38	6.7%__0.38	5.2%_0.31	5.3%_0.31
		2 parents	6.2%__0.38	5.0%__0.38	4.8%_0.31	5.0%_0.31
	Dropout	< 2 parents	32.2%_0.4	52.6%_1.1	13.3%_0.3	31.8%_0.91
		2 parents	25.9%_0.4	44.9%_1.1	12.4%_0.3	30.0%_0.91
West	Non-drop	< 2 parents	8.7%__0.61	7.0%__0.61	5.4%_0.50	5.6%_0.50
		2 parents	6.5%__0.61	5.2%__0.61	5.0%_0.50	5.2%_0.50
	Dropout	< 2 parents	33.2%_0.6	53.6%_1.8	13.8%_0.5	32.7%_1.5
		2 parents	26.7%_0.6	45.9%_1.8	12.8%_0.5	30.9%_1.5
<i>Tenth grade panel</i>						
NE, NC, S	Non-drop	< 2 parents	6.5%__0.52	7.5%__0.52	2.7%_0.26	4.3%_0.26
		2 parents	5.1%__0.52	6.0%__0.52	2.9%_0.26	4.7%_0.26
	Dropout	< 2 parents	18.9%_0.5	37.7%_1.2	6.5%_0.3	20.1%_0.6
		2 parents	15.4%_0.5	32.2%_1.2	7.1%_0.3	21.6%_0.6
West	Non-drop	< 2 parents	6.2%__0.60	7.2%__0.60	2.6%_0.30	4.1%_0.30
		2 parents	4.9%__0.60	5.8%__0.60	2.8%_0.30	4.5%_0.30
	Dropout	< 2 parents	18.3%_0.6	36.8%_1.3	6.3%_0.30	19.5%_0.7
		2 parents	15.0%_0.6	31.3%_1.3	6.9%_0.30	21.0%_0.7

*Percentages pertain to an average adolescent who attended a non-Catholic school in an urban area.

1. Fendrich and Vaughan (1994) concluded- based on the National Longitudinal Study of Youth- that lower substance use among blacks may be due partly to differential underreporting.

2. A full discussion of results is in Johnson and Hoffmann (1999).

3. The samples of schools in the eighth and tenth grade panels are disjoint, because middle and high schools do not overlap, but the samples of students are overlapping. About 90 percent of responding students in each panel were also

responding students in the other panel. Given the positive correlation of cigarette use at different stages of adolescence, the sample overlap results in increased precision for comparing panels. The significance test results presented in this paper are conservative in that we treat the two panels as independent samples. The significance tests are also conservative-relative to single-level analyses-because the standard errors of test statistics correctly reflect the clustering of students within schools in the sample design.

4. For each panel, we first generated predicted values using a bivariate normal multilevel model with two response variables-daily cigarette use at waves 1 and 2-and twelve explanatory variables, including family structure, dropout status, parental support, school participation, negative peer associations, race/ethnicity, region, type of school, percent of minority students, student-teacher ratio, size of school, and teacher salary level. The effects of six individual/family variables were treated as random at the school level. The continuous imputed values were rounded to 0 or 1. We also generated three sets of imputations for each panel, and, using multiple imputation techniques (Schafer, 1997), found that the additional uncertainty contributed by the imputation amounted to less than 10% of each standard error reported in Table 1.

5. The "other" category is more than 98% white non-Hispanic in both panels, with the rest Native American. Based on NHSDA, Native Americans- like whites- are high in cigarette use (Johnson and Larison, 1998).

6. Living arrangements were measured at Wave 1 based on adolescents' responses to the question "Which of the following people live in the same household with you?" Preliminary analyses showed that detailed family types, such as mother only and mother-stepfather, did not differ significantly in their effects on adolescent cigarette use and the presence or absence of siblings was not a significant predictor, so our analysis distinguishes only those who lived with two biological parents (coded 1) from all other arrangements (coded 0).. Missing values were few-1.5% of eighth graders and 0.7% of tenth graders- and these were imputed as 0's.

7. Each of the three scales has high internal reliability, with Cronbach's alpha greater than 0.65 in each panel. Except for family income (10% missing in the eighth grade panel, 17% in the tenth grade panel), missing data rates of individual-level explanatory variables were less than 3% in each panel. We imputed missing values of *scale* items using the mode of respondents with nonmissing values. We imputed missing values of family income using predicted values from a linear mixed model (Schafer, 1997), treating the log of income as normally distributed and allowing the intercept to vary among schools. Details are in Johnson and Hoffmann (1999).

8. Region, type of place, type of school, and school size had no missing data in either panel, but missing data rates of salary and student-teacher ratio were appreciable in the tenth grade panel: 27% for salary and 7% for student-teacher ratio. The missing values on these two variables were imputed using mean imputation within imputation cells defined by region, type of place, and type of school.

9. In both panels, k was estimated to be very close to 1.0.

10. Inspection of residual plots suggested that the assumptions of normality and constant variance are reasonable for the level-2 errors in both models presented in this paper.

11. Multilevel parameter estimates reported in this paper are second-order penalized quasi-likelihood estimates ("PQL2"), as discussed in Goldstein (1995) and implemented in MLWIN. The estimates were corroborated using two alternative methods- bootstrap and Markov Chain Monte Carlo- also in MLWIN.

12. Another finding from Table 1 is that the school variance in daily cigarette use- after controlling for past use- is much larger among eighth graders than among tenth graders- 0.09 vs. 0.03. This suggests that opportunities for school interventions to prevent cigarette use are greater in middle schools than in high schools.

13. Model 2 also assumes 1) the level-1 random error e_{ij} is independent of each of the level-2 random errors- u_{0j} through u_{pj} ; 2) level-1 and level-2 random errors are independent of all measured explanatory variables; and 3) the vector of level-2 random errors u_{pj} , $p = 0, 1, \dots, P$ - is multinormal with zero means; variances $SIGMA_{up}^2$, $p = 0, 1, \dots, P$; and covariances $SIGMA_{upp}$, where p and p range from 0 to P and p does not equal p .

14. The effect of male gender shifts from negative in the eighth grade panel- -0.23- to positive in the tenth grade panel- 0.16. This finding is consistent with NHSDA data suggesting that, beginning in the 1980s, females tended to initiate cigarette use at earlier ages than males (Johnson and Gerstein, 1998).