Using cross-classified multilevel models to disentangle school and neighborhood effects: An example focusing on smoking behaviors among adolescents in the United States

Erin C. Dunn [a,b,c,1], Tracy K. Richmond [d,2], Carly E. Milliren [e], S.V. Subramanian [f,*]

[a] Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge St, Simches Research Building, 6th Floor, Boston, MA 02114, USA
[b] Department of Psychiatry, Harvard Medical School, Boston, MA 02115, USA
[c] Stanley Center for Psychiatric Research, The Broad Institute, Cambridge, MA 02142, USA
[d] Division of Adolescent Medicine, 6th Floor, Boston Children's Hospital, 33 Longwood Avenue, Boston, MA 02115, USA
[e] Clinical Research Center, Department of Medicine, Boston Children's Hospital, 21 Autumn St. #304, Boston, MA 02115, USA
[f] Department of Social and Behavioral Sciences, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA

ARTICLE INFO

ABSTRACT

Background: Despite much interest in understanding the influence of contexts on health, most research has focused on one context at a time, ignoring the reality that individuals have simultaneous memberships in multiple settings.
Method: Using the example of smoking behavior among adolescents in the National Longitudinal Study of Adolescent Health, we applied cross-classified multilevel modeling (CCMM) to examine fixed and random effects for schools and neighborhoods. We compared the CCMM results with those obtained from a traditional multilevel model (MLM) focused on either the school and neighborhood separately.
Results: In the MLMs, 5.2% of the variation in smoking was due to differences between neighborhoods (when schools were ignored) and 6.3% of the variation in smoking was due to differences between schools (when neighborhoods were ignored). However in the CCMM examining neighborhood and school variation simultaneously, the neighborhood-level variation was reduced to 0.4%.
Conclusion: Results suggest that using MLM, instead of CCMM, could lead to overestimating the importance of certain contexts and could ultimately lead to targeting interventions or policies to the wrong settings.

1. Introduction

There is much interest among epidemiologists in understanding multilevel phenomena, or how features of the physical and psychosocial environment in which individuals live, learn, work, and play influence individual health, disease, and behavior (Pickett and Pearl, 2001; Mair et al., 2008). The growth and interest in multilevel analyses has been facilitated by conceptual developments in multilevel theory (Bronfenbrenner and Morris, 2006; Stokols, 1996; Krieger, 2001; Dunn et al., 2014) as well as statistical advancements in multilevel statistical modeling (Diez Roux, 1998; 2002; Raudenbush and Bryk, 2002; Subramanian et al., 2003). Although multilevel theory posits that multiple contexts (e.g., residential environments, schools, workplaces, hospitals, or other "areas") influence individual and population health simultaneously, most empirical applications have studied contexts in isolation and the majority of studies have focused on neighborhoods. An emphasis on single contexts, and more specifically neighborhoods, is problematic for at least two reasons. First, it ignores the reality that individuals simultaneously belong to multiple settings that could each independently affect their health. For example, focusing on the influence of neighborhood factors on adolescent health behaviors ignores the influence of schools, which may be a more influential context in teens' lives. Second, results from studies assessing a single context may be misleading, as the effect of one context can be over- or under-estimated depending on what context is ignored.

The objective of our study was to provide a methodological demonstration of cross-classified multilevel models (CCMM), also sometimes referred to as "multilevel cross-classified models, or

* Corresponding author.
E-mail addresses: edunn2@mgh.harvard.edu (E.C. Dunn),
tracy.k.richmond@childrens.harvard.edu (T.K. Richmond),
carly.milliren@childrens.harvard.edu (C.E. Milliren),
svsubram@hsph.harvard.edu (S.V. Subramanian).
[1] Tel.: +1 617 726 9387.
[2] Tel.: +1 617 355 4453.

'cross-classified random effect models", in disentangling the role of two critical influences on adolescent tobacco use: schools and neighborhoods. CCMM allows researchers to incorporate *non-hierarchical* nesting structures, where individuals are *simultaneously* nested within multiple non-hierarchical settings. Thus, rather than modeling the effect of either the school or neighborhood setting, as would be done in a traditional two-level multilevel model (MLM), application of a CCMM enables researchers to *simultaneously* examine the fixed and random effects corresponding to the school *and* neighborhood settings. Simultaneous examination of schools and neighborhoods, in particular, is important because both settings can influence health behaviors through multiple pathways, including policies, normative behaviors, access to resources, and the like (Kawachi and Berkman, 2003; Bonnell et al., 2013). In this paper, our intention was to provide a methodological demonstration of the CCMM method and show where and how results from a CCMM might deviate from a traditional multilevel model focused on a single context. A methodological application and applied example of CCMM is warranted, given the underrepresentation of CCMM in the epidemiology literature relative to MLM (see for example Leyland and Naess, 2008; Lloyd et al., 2010; Utter et al., 2011; Riva et al., 2009; Moore et al., 2013; Virtanen et al., 2010; Basile et al., 2012).

Our demonstration proceeded in several steps. We first modeled our outcome using a traditional multilevel modeling (MLM) approach that included only one context (neighborhood or school) per model. In other words, we modeled the school as a random effect, ignoring the neighborhood in one MLM and then modeled the neighborhood as a random effect, ignoring the school in a second MLM. We then modeled tobacco use using CCMM, which simultaneously accounts for the influence of schools and neighborhoods. We compared both the fixed (i.e., population average effects) and random effects (i.e., variance in the outcome) in models assuming a two level hierarchy (MLM) versus those allowing for multiple non-hierarchical memberships (CCMM). The fixed effects we examined were individual-, school-, and neighborhood-level demographic indicators, including socioeconomic status and race/ethnicity. These fixed effect estimates are informative for determining the extent to which both the predictor of interest is associated with the outcome and the degree to which the predictor of interest reduces between-level variation. Finally, in our CCMM, we compared the relative variance contribution of schools and neighborhoods. By comparing variance contributions (i.e., random effects) across models, we are able to evaluate the extent to which inclusion of the fixed effect variables helped to explain the observed between-school and between-neighborhood variation in smoking.

To increase the use of CCMM, and make the processes of analyzing cross-classified data more transparent, we provide readers with instructions on how to implement the CCMM in MlwIN (refer to Technical Appendix: Part 1 online) and through MlwIN as implemented via STATA (refer to Technical Appendix: Part 2 online).

2. Materials and methods

2.1. Data

Data for the study came from the National Longitudinal Study of Adolescent Health (Add Health), a nationally representative school-based longitudinal survey focusing on the health and behavior of adolescents in middle and high school (grades 7–12; ages 12–18) who were first interviewed in 1994–1995 (Wave 1) (Harris, 2013). To ensure selected schools were representative of US schools, researchers stratified schools by census region, urbanicity, size, type, and ethnic background of the student body (i.e., percent

White) prior to systematic random sampling. From a sampling frame of 26,666 schools, investigators selected a sample of 80 high schools and 52 middle schools for participation. A systematic random sample of high schools along with feeder schools (i.e., middle schools whose students matriculate at the selected high school) was selected. A total of 134 schools (79%) participated. These schools represented the spectrum of schools available to U.S. students, including private schools requiring tuition, parochial or religious based schools, and free public schools of which there are schools of choice (i.e., any student can attend), magnet schools (i.e., schools that typically are available to students via an exam) and neighborhood schools. An in-school survey was completed by 90,118 students. A random sample of these 90,118 students (as well as all students who were eligible to complete the in-school survey, but were absent on the day of administration) were invited to complete a more detailed in-home interview. 20,745 students completed the in-home interview (over 75% of those asked to participate did so). In addition, 17,670 caregivers provided information at Wave 1. AddHealth was an appropriate dataset to use in the current analysis, as the sample was large and comprised individuals who attended distinct school units and lived in distinct residential neighborhoods (as defined by Census tracts, or the relatively stable, geographic groupings of between 1200 and 8000 people, which vary by size, that are created by the United States Census Bureau https://www.census.gov/geo/reference/gtc/gtc_ct.html; defining neighborhoods based on Census tracts is commonplace among studies in the United States) (Dunn et al., 2014).

Our analyses are based on an analytic sample of 16,070 youth who attended 128 schools and lived in 2111 census tracts. This analytic sample was derived after eliminating youth ($n=660$) in the non-nationally-representative sample (i.e., who attended schools sampled for genetic analyses) or from schools that did not provide demographic data. We also removed youth who were missing data on the outcome measure ($n=139$ excluded) or predictors and covariates ($n=1425$). These 1425 participants did not differ with respect to smoking behaviors from those who were included in our analysis with respect to age sex, race, or parental education. However, participants missing information about public assistance tended to be smokers (34% smokers compared to 26% of non smokers were missing information on public assistance; $p=0.002$).

We restricted the analysis to youth who were white, black, or Hispanic, excluding 2451 students of other races, given that students in other racial/ethnic groups were not sufficiently represented in Add Health to obtain robust group estimates (Native Americans $n=104$, $<1\%$; Asians $n=1176$, 6%; Other $n=1171$, 6%). Although Add Health is a longitudinal study, we conducted a cross-sectional analysis because the majority of youth resided in the same neighborhood and school in Wave 2 as they did i n Wave 1, and because Wave 3 and Wave 4 were conducted when most youth had graduated from high school and so were no longer uniformly in two different contexts.

2.2. Measures

2.2.1. Outcomes: Smoking

In the Wave 1 In-Home survey, youth reported the number of days in the past month they smoked cigarettes. We coded responses to this question into both a continuous (ranging from 1 to 30, i.e. the number of days in past month) and binary outcome variable (yes/no had ever smoked in the past month), consistent with previous studies (Duncan and Rees, 2005; Alexander et al., 2001).

2.2.2. Predictors: Socioeconomic status (SES)

SES was obtained at the individual, school, and neighborhood level. At the individual-level, SES was determined based on parent

receipt of public assistance and parent education. We used data from either the youth or caregiver interview to capture receipt of public assistance (1=mother currently receiving public assistance, such as welfare; 0=not) and highest level of parent education (defined as the maximum level of education by the resident mother, resident father, or resident step-father/partner; 1=parent did not have at least a high school diploma or equivalent; 0=parent had completed high school/equivalent). At the school-level, we created a continuous measure of school-level SES using individual-level data. Use of individual-level data was required as information about school-level SES was not directly available. We calculated the proportion of students within each school whose mother had received public assistance or did not have a high school degree or equivalent. At the neighborhood level, we used data from the 1990 Census to create a neighborhood-level SES measure indicating the proportion of residents within each neighborhood who had received public assistance or did not have at least a high school degree or equivalent.

2.2.3. Covariates

Adjusted models included the following individual-level covariates: age (continuous), sex (male=0; female=1), and self-reported race/ethnicity (1=non-Hispanic white; 2=non-Hispanic black; 3=Hispanic). We also adjusted for the percentage of students in either the school or the neighborhood who were white.

2.3. Analysis

We fit three sets of models in the current analysis. The first two models used a traditional two-level multilevel model. As described in detail elsewhere (Diez Roux, 1998; 2002; Raudenbush and Bryk, 2002; Subramanian et al., 2003), the traditional multilevel model assumes a two-level multilevel data structure, where observations are *hierarchically* nested, such that members of the lower level (i.e., level one) are nested in *one and only* one entity at the higher level (i.e., level two). Thus, we began by fitting a two-level "school only" multilevel model (ignoring the neighborhood), where the outcome (denoted y) for a person (denoted i) nested in a given school (denoted j) was modeled as:

$$y_{ij} = \beta_0 + \beta_{xij} + u_{0j} + e_{0ij} \tag{1}$$

In Eq. (1), the fixed effect parameter β_0 refers to the overall mean outcome y across all schools and β_{xij} refers to a vector of individual-level covariates. The random effect parameter u_{oj} refers to the random effect for school (assumed to be normally distributed with a mean of 0 and variance σ^2_{uo}), and e_{oij} refers to the random effect for the individual.

Second, we ran a two-level "neighborhood only" multilevel model (ignoring school), where the outcome (denoted y) for a person (denoted i) nested in a given neighborhood (denoted k) was modeled as:

$$y_{ik} = \beta_0 + \beta_{xik} + u_{0k} + e_{0ik} \tag{2}$$

The fixed and random effect parameters in Eq. (2) have an identical interpretation as Eq. (1), except they now refer to neighborhoods (instead of schools). Though the two-level multi-level modeling strategy marks a significant advancement over traditional regression models in which context is ignored, it is not flexible enough to accommodate multiple non-nested contexts at once.

The third model we fit was the CCMM. In a CCMM, individuals (denoted i) simultaneously belong to two non-nested contexts, here school (denoted j) and neighborhood (denoted k). Thus our outcome (denoted y) for a person i nested in school j and neighborhood k is modeled as:

$$y_{i(jk)} = \beta_0 + u_{0j} + u_{0k} + e_{0i(jk)} \tag{3}$$

In Eq. (3), which presents a null or intercept-only CCMM (i.e., a model without covariates), the fixed effect parameter, β_0, refers to the overall mean outcome y across all schools and neighborhoods, u_{oj} refers to the random effect for schools, u_{ok} refers to the random effect for the neighborhood, and $e_{oi(jk)}$ refers to the random effect for the individual with the combination of j school and k neighborhood. Like a traditional multilevel model, the CCMM variance parameters are assumed to be independent of each other and normally distributed, with a mean of 0 and variance that is estimated (i.e., σ^2_{uo}).

This CCMM null model can be extended to include predictors or covariates (i.e., fixed effects) at each level of analysis. For example, we included two indicators of SES; the first was a measure of receipt of public assistance at the individual ($\beta_{1oi(jk)}$), school (β_{2oj}), and neighborhood-level (β_{3ok}):

$$y_{i(jk)} = \beta_0 + \beta_{1oi(jk)} + \beta_{2oj} + \beta_{3ok} + u_{0j} + u_{0k} + e_{0i(jk)} \tag{4}$$

The parameters in Eq. (4) are interpreted differently than Eq. (1). For example, β_{2oj} refers to the average effect of school-level parent receipt of public assistance controlling for individual and neighborhood parent receipt of public assistance, and β_{3ok} refers to the average effect of neighborhood-level parent receipt of public assistance controlling for individual and school parent receipt of public assistance. The interpretation of each random effect parameter is the same, except each estimate now controls for the parent receipt of public assistance predictors in the model.

Our analyses proceeded in three steps. First, to partition the variance in number of days smoked into within and between components and estimate an intraclass correlation coefficient (ICC; i.e., the proportion of variation in the outcome that was due to differences across schools and neighborhoods, rather than differences across students), we estimated a null model for each model type: school-only multilevel model (Eq. (1)), neighborhood-only multilevel model (Eq. (2)), and the CCMM (Eq. (3)) (Table 2, Model 1). These ICC estimates were obtained for the school-only multilevel model, neighborhood-only multilevel model, and CCMM. The ICCs in the school-only and neighborhood-only multilevel model were generated by dividing the between-level random effect by the total variance. In the CCMM, we calculated ICCs for the school and neighborhood level, which are referred to as the *intra-neighborhood* (i.e., correlation in outcome between two youth who live in the same neighborhood, but attend different schools; this was calculated by dividing the neighborhood-level random effect by the total variance, or the sum of the three variance components) and *intra-school correlation* coefficient (i.e., correlation in outcome between two youth who attend the same school, but live in different neighborhoods; this was calculated by dividing the school-level random effect by the total variance). We also calculated an *intracell correlation*, referring to the correlation in outcome of two students who live in the same neighborhood and attend the same school; this was calculated by summing the between-level variances, i.e., for the neighborhood and school, by the total variance. Subsequently, we estimated a model that contained individual-level predictors and covariates (Model 2). By including these individual-level variables, we were able to evaluate the extent to which the between-level variance estimates (i.e., random effect parameters) could be explained by the observed individual characteristics across schools and neighborhoods. We then fit a school-only model and CCMM containing individual-level variables as well as the *school-level* measures of SES and race/ethnicity (Model 3) and a neighborhood-only model and CCMM containing individual-level variables combined with the *neighborhood-level* measures of SES and race/ethnicity (Model 4). Finally, we fit a CCMM containing all individual-, school-, and neighborhood-level variables (Model 5).

Table 1
Descriptive Statistics on Individuals ($n = 16,070$) in the National Longitudinal Study of Adolescent Health (AddHealth).

Demographic characteristics	Total sample ($n = 16,070$)	Non-smokers 11,908 (74%)	Smokers 4,162 (26%)	p-Value[†]	Number of days smoked median (IQR)	p-Value[††]
Child sex						
Male	7,948 (49%)	5,847 (49%)	2,101 (50%)	0.13	15.0 (27.0)	0.74
Female	8,122 (51%)	6,061 (51%)	2,061 (50%)		15.0 (27.0)	
Mother received public assistance						
No	14,502 (90%)	10,753 (90%)	3,749 (90%)	0.68	15.0 (27.0)	0.005
Yes	1,568 (10%)	1,155 (10%)	413 (10%)		10.0 (28.0)	
Parent obtained high school degree						
No	2,120 (13%)	1,608 (14%)	512 (12%)	0.05	12.0 (28.0)	0.31
Yes	13,950 (87%)	10,300 (86%)	3,650 (88%)		15.0 (27.0)	
Child race						
White	9,387 (58%)	6,304 (53%)	3,083 (74%)	< .0001	20.0 (26.0)	< .0001
Black	3,817 (24%)	3,348 (28%)	469 (11%)		4.0 (12.0)	
Hispanic	2,866 (18%)	2,256 (19%)	610 (15%)		10.0 (27.0)	
Child age[†††]						
11–13	2,299 (14%)	1,991 (17%)	308 (7%)	< .0001	5.0 (18.0)	< .0001
14–16	8,352 (52%)	6,194 (52%)	2,158 (52%)		15.0 (27.0)	
17+	5,419 (34%)	3,723 (31%)	1,696 (41%)		20.0 (25.0)	

Cell entries are reported as n (%) comparing smokers to non-smokers. Number of days smoked among those who smoke within demographic groups is reported as median interquartile range (IQR) due to the bimodal distribution.

[†] Chi-square tests or t-test where appropriate comparing smokers to non-smokers within demographic groups.

[††] Wilcoxon rank-sum test or Kruskal–Wallis test where appropriate comparing median number of days reporting smoking across demographic groups.

[†††] Child age is presented here as categories, however, it was modeled as a continuous variable in our models.

All analyses were conducted in MLwiN version 2.29 via Stata version 13 (College Station, TX) with Bayesian estimation procedures as implemented via Markov Chain Monte Carlo (MCMC) methods using Metropolis–Hastings algorithm (Browne, 2004). We used a Bayesian estimation procedure with non-informative priors, because others have shown that estimates of random effects in binary models using maximum-likelihood procedures tend to be biased (Rodriguez and Goldman, 1995) and computationally intensive (Rabash and Browne, 2001) and that MCMC is preferable when two sets of random effects are estimated, when there is concern about the correlation between random effects, and when there are some instances of a small number of individuals per cluster. Model fit was evaluated using the Deviance Information Criterion (DIC), which is a test statistic produced by the MCMC procedure that refers to the model complexity and "badness of fit". Higher DIC values indicate a poorer fitting model. We calculated a linear regression for the continuous outcome and a logistic regression for the binary smoking variable; no other link functions (e.g., log) were available in MlwiN. Parameter estimates (betas or odds ratios, OR) and 95% credible intervals (CI, which are the confidence intervals generated using Bayesian procedures) are presented for fixed and random effect parameters. We examined residual plots at each level of analysis to evaluate model diagnostics on the variance parameter; this enabled us to test model assumptions, detect outliers and influence points on model fit. Analyses were conducted using unweighted data, as weighting techniques for cross-classified methods have not been established (Raudenbush and Bryk, 2002). A non-weighted analysis is also appropriate as our emphasis was on tests of association, rather than deriving nationally representative estimates, and we adjusted our analyses for sample characteristics and thus reduced the heterogeneity in the sample (Lee and Forthofer, 2006).

3. Results

The AddHealth data were suited to cross-classified analyses. An average of 125.5 (sd = 116.5) youth per school completed an In-Home survey (minimum = 18; maximum = 1012). In each neighborhood, an average of 7.6 (sd = 18.5) youth completed an In-Home survey (minimum = 1; maximum = 260). There were 970 (of 2111) census tracts that contained only one youth respondent. There was an average of 20.2 (sd = 22.0) census tracts per school (minimum = 1; maximum = 175), an average of 1.22 (sd = 0.42) schools per census tract (minimum = 1; maximum = 3), and 2584 different combinations of school and neighborhood contexts. Thus, there was no clear hierarchical nesting of schools within neighborhoods (or vice versa).

Table 1 presents descriptive statistics on individuals in the total sample as well as by smoking status. 26% of youth ($n = 4162$) smoked at least once in the past month, with smokers reporting smoking an average of 16.2 days per month (sd = 12.3). The sample was balanced in terms of sex (49% male). The majority of participants were white (58%) and were in mid-adolescence (mean age = 15.6; sd = 1.7). Parent SES was relatively high, as only 10% of mothers were currently receiving public assistance and only 13% had less than a high school education. Smokers were more likely than non-smokers to be white ($p < 0.001$), older in age; $p < 0.001$), and have a parent without at least a high school degree ($p = 0.05$). No differences were found when comparing smokers to non-smokers by sex or receipt of public assistance. Among smokers, whites reported smoking more days than either Blacks or Hispanics ($p < 0.001$) as did those whose parent *did not* receive public assistance ($p = 0.005$) and who were older ($p < 0.0001$).

The schools and neighborhoods in this sample were diverse. Schools varied significantly with respect to their SES. Across the 128 schools, the average percentage of students whose mothers had received public assistance was 10.4% (sd = 9.4; minimum = 0; maximum = 45.4%), the average percentage whose parents did not have at least a high school education was 11.5% (sd = 10.4; minimum = 0; maximum = 55.8%) and the average percentage reporting White race was 47.5% (sd = 25.5; minimum = 0; maximum = 85.9%). Similarly, across the 2111 neighborhoods, the average percentage of students whose mothers had received public assistance was 10.8% (sd = 10.0; minimum = 0; maximum = 67.5%), the average percentage of residents without at least a high school education was 29.2% (sd = 16.1; minimum = 0; maximum = 78.7%) and the average percentage reporting white race was 66.6%
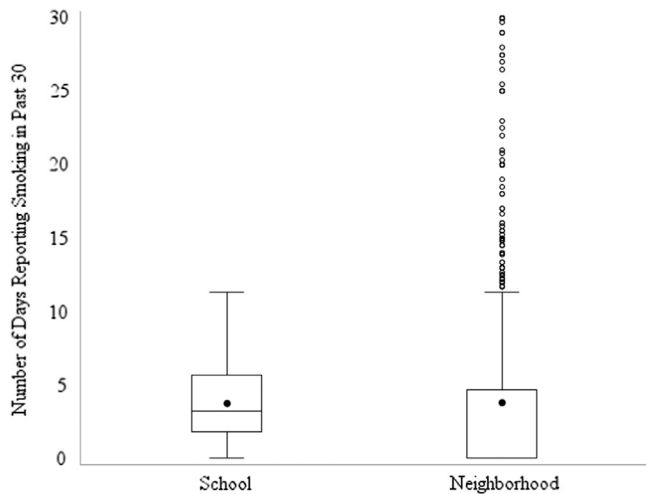
Fig. 1. Distribution of the number of days reported smoking in the Past 30 Days Across Schools ($N=128$) and neighborhoods ($N=2111$ neighborhoods). Box plots are presented depicting the number of days youth reported smoking in the past 30 across all schools and neighborhoods. The solid colored dot represents the mean. Horizontal lines in the boxes represent the first quartile, the median, and the third quartile of the distribution of smoking values. Vertical lines depict the 1.5 times the interquartile range of smoking for each school or neighborhood in the sample. Dots outside depict outlying values. The large number of outliers for neighborhoods was due largely to the neighborhoods including only a small number of respondents.
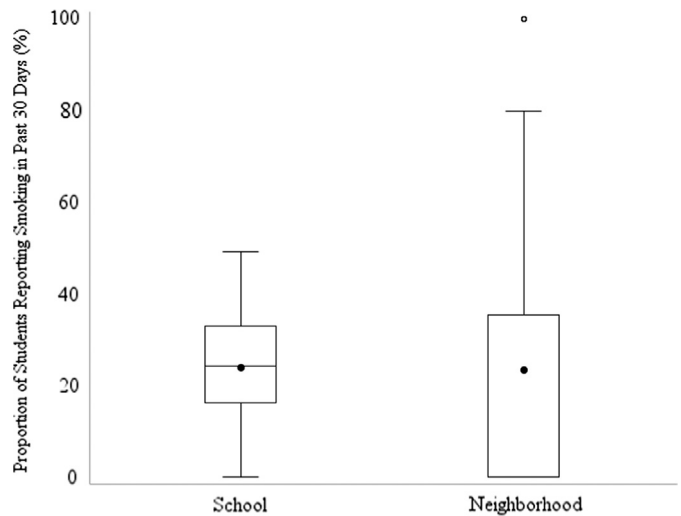
Fig. 2. Distribution of the proportion of youth reporting smoking in the Past 30 Days Across Schools ($N=128$) and Neighborhoods ($N=2111$ neighborhoods). Box plots are presented depicting the proportion of youth reporting smoking in the past 30 across all schools and neighborhoods. The solid colored dot represents the mean. Horizontal lines in the boxes represent the first quartile, the median, and the third quartile of the distribution of smoking values. Vertical lines depict the 1.5 times the interquartile range of smoking for each school or neighborhood in the sample. Dots outside depict outlying values.

(sd=33.3; minimum=0%; maximum=100%). The average number of days smoked (school mean=3.73; sd=2.50; neighborhood mean=3.76; sd=7.15) and percentage of respondents who were considered smokers (school mean=23.9%; sd=11.0; neighborhood mean=23.4%; sd=33.6) was similar across neighborhoods and schools. However, as shown in Figs. 1 and 2, there was considerable variability *within* and *between* these contexts with respect to youth smoking. For example, in some neighborhoods almost 80% of youth were smokers, whereas in others 0% were smokers (Fig. 2). Moreover, as shown in Fig. 3, schools differed with respect to both the average number of days youth smoked and the variability (minimum and maximum) in numbers of days smoked. Thus, smoking behaviors were not uniform, but rather varied as a result of neighborhood or school context.

The data were also suited to cross-classified analyses based on the observed differences in contexts youth experienced. That is, youth did not always attend schools that reflected the demographic makeup of their neighborhoods. Though there were moderate to high correlations between school and neighborhood demographic characteristics ($r=0.556$; $p<0.0001$ for school and neighborhood percent receiving public assistance; $r=0.587$; $p<0.001$ for school and neighborhood parental education; $r=0.705$; $p<0.001$ for school and neighborhood percent of population identified as white), the correlations varied across these indicators and provided evidence of some discordance between the neighborhood and school. Indeed, after dichotomizing the public assistance measure to differentiate *poor* schools and neighborhoods from *non-poor* schools and neighborhoods (using a cut-point of 10% of more of the population receiving public assistance to denote a high poverty setting, similar to previous studies; Leventhal and Brooks-Gunn, 2003), we found close to one-quarter of all youth lived in incongruent settings. That is, 13.7% ($n=2.200$) of youth lived in a low poverty neighborhood, but attended a high poverty school and 10% ($n=1.645$) lived in a high poverty neighborhood, but attended a low poverty school. A McNemar's chi-square test of the agreement between the neighborhood and school poverty measures was significant (chi-square 80.1; $p<0.0001$), indicating that there was not perfect agreement between school and neighborhood poverty.

3.1. Multilevel models with smoking as continuous outcome

Tables 2 and 3 present the results of a series of models for the school-only multilevel model, neighborhood-only multilevel model, and the CCMM predicting the number of days smoked. In our null model (Table 2, Model 1), the random effects for the school-only ($\sigma_{u0j}^2 = 5.44$) and neighborhood-only model ($\sigma_{u0k}^2 = 4.58$) were similar. However, in the CCMM, the between-level variance in smoking was driven largely by the school $\sigma_{u0j}^2 = 5.36$ and not by the neighborhood ($\sigma_{u0k}^2 = 0.46$). Comparable ICC values were obtained for the school-only (6.1%) and neighborhood-only multilevel model (5.2%), indicating that the majority of the variability was attributable to individual characteristics. In the CCMM, the intra-neighborhood correlation was 0.52% and the intra-school correlation was 6.1%. Finally, the intracell correlation coefficient, or the correlation in smoking behavior between two youth who live in the same neighborhood and attend the same school, was 6.5%.

When individual-level covariates were added to these three models (Table 2, Model 2), the between-level variance declined by more than half relative to Model 1 (school-only $\sigma_{u0j}^2 = 2.12$; neighborhood only $\sigma_{u0k}^2 = 1.59$; and CCMM ($\sigma_{u0j}^2 = 0.24$; $\sigma_{u0k}^2 = 2.08$). This decline suggests that the between-level variation in smoking was due largely to the observed individual characteristics across schools and neighborhoods.

Model 3 (see Table 3) introduced school-level SES indicators and covariates into the school-only multilevel model and CCMM. When these school-level variables were added, the school-level variance declined slightly compared to Model 2 for both the school-only ($\sigma_{u0j}^2 = 1.83$ from 2.12) and CCMM ($\sigma_{u0j}^2 = 1.78$ from 2.08). The neighborhood-level variance in the CCMM of Model 3 decreased slightly relative to the CCMM in Model 2 ($\sigma_{u0k}^2 = 0.17$ from 0.24). There was a modest association detected between the percentage of students in the school whose mother had received public assistance and number of cigarettes smoked in the school-only ($\beta=0.07$) or CCMM ($\beta=0.07$). There was no association detected between the percentage of students whose parents did not have a high school degree in the school-only multilevel model ($\beta=-0.04$) or CCMM ($\beta=-0.04$). There was also a very small association for the percentage of students who were white in the school-only multilevel ($\beta=0.02$) or CCMM ($\beta=0.02$).
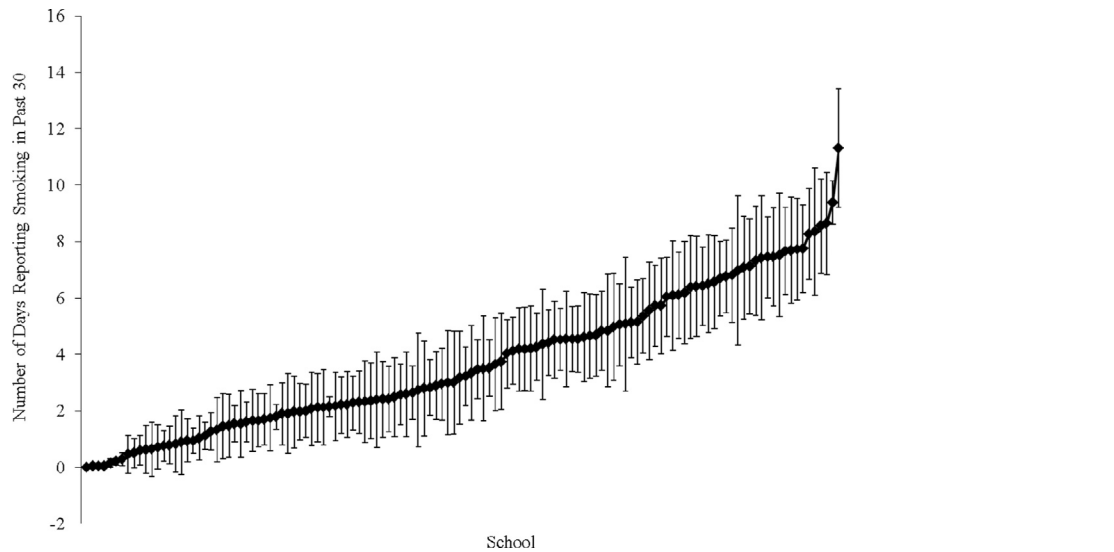
Fig. 3. Distribution of the Number of Days Youth Reported Smoking in the Past 30 Within and Between Schools ($N=128$ schools). Dots represent the mean number of days youth reported smoking within each school. 95% bounds around the mean based on the standard deviation (SD) of number of days smoked are also presented. Values are sorted from left to right by lowest school mean. 95% confidence intervals may extend into the negative range when the mean smoking days within a school is not significantly different from zero.

Table 2
Nested Models Describing Association between predictors and number of days smoked in the Past 30 Days in the National Longitudinal Study of Adolescent Health ($N=16,070$).

Fixed effect estimates	Model 1			Model 2		
	School only	Neighborhood only	Cross-classified	School only	Neighborhood only	Cross-classified
Intercept (SE)	3.8 (0.23)	3.94 (0.11)	3.88 (0.22)	−6.94 (0.84)	−7.42 (0.73)	−6.91 (0.84)
Individual-level						
Age				0.80 (0.70, 0.90)	0.85 (0.76, 0.93)	0.80 (0.70, 0.90)
Female				0.06 (−0.22, 0.34)	0.02 (−0.26, 0.30)	0.06 (−0.22, 0.34)
Public assistance				0.76 (0.27, 1.25)	0.84 (0.34, 1.34)	0.74 (0.26, 1.24)
High school degree				−0.21 (−0.66, 0.25)	−0.15 (−0.62, 0.29)	−0.21 (−0.67, 0.23)
Race						
White				Ref	Ref	Ref
Black				−4.28 (−4.73, −3.86)	−4.57 (−4.96, −4.16)	−4.29 (−4.71, −3.86)
Hispanic				−1.98 (−2.49, −1.48)	−2.86 (−3.32, −2.38)	−1.99 (−2.50, −1.48)
School-level						
Public assistance						
High school degree						
Percent White						
Neighborhood-level						
Public assistance						
High school degree						
Percent White						
Random effect estimates						
U3 neighborhood (SE)		4.58 (3.66, 5.65)*	0.46 (0.13, 0.88)*		1.59 (1.03, 2.26)*	0.24 (0.06, 0.61)*
U2 school (SE)	5.44 (4.04, 7.19)*		5.36 (3.95, 7.07)*	2.12 (1.48, 2.97)*		2.08 (1.42, 2.92)*
U1 individual (SE)	83.1 (81.3, 85.0)*	84.0 (82.1, 85.9)*	82.7 (80.9, 84.6)*	80.7 (79.0, 82.5)*	81.1 (79.3, 83.0)*	80.5 (78.8, 82.3)*
Fit Statistics						
DIC	–	–	116,736	–	–	116,259

Model 1 presents the results for a null model (i.e., no covariates) for each model type: school-only multilevel model, neighborhood-only multilevel model, and the cross-classified multilevel model. Model 2 presents the same models as Model 1, except Model 2 includes individual-level predictors and covariates. For the fixed effect estimates, cell entries are parameter (beta) estimates and credible intervals. The intercept is presented as parameter estimate and standard error (SE). Random effects are presented as estimate and credible intervals. DIC refers to Deviance Information Criterion, a measure of model fit and complexity and is only reported for the cross-classified models.

* Significant random effects are indicated by $p < 0.05$.

Model 4 (see Table 3) introduced neighborhood-level SES indicators and covariates into the neighborhood-only multilevel model and CCMM. Adding these variables led to a slight decrease in neighborhood-level variance for the neighborhood-only multilevel model ($\sigma^2_{u0k} = 1.52$ from 1.59) and a larger decrease in the CCMM ($\sigma^2_{u0k} = 0.16$ from 0.24). The school-level variance in the CCMM of Model 4 was nearly identical ($\sigma^2_{u0j} = 2.06$ compared to 2.08). Neighborhood-level values of public assistance were not associated with number of cigarettes smoked for either the neighborhood-only ($\beta=0.03$) or CCMM ($\beta=0.01$), nor were neighborhood values for residents without a high school degree for either the neighborhood-only ($\beta=-0.01$) or CCMM ($\beta=0.01$), or percentage white for either the neighborhood-only ($\beta=0.01$) or CCMM ($\beta=0.01$).

Table 3
Nested Models Describing Association between predictors and number of days smoked in the Past 30 Days in the National Longitudinal Study of Adolescent Health ($N=16,070$).

Fixed effect estimates	Model 3		Model 4		Model 5
	School only	Cross-classified	Neighborhood only	Cross-classified	Cross-classified
Intercept (SE)	-8.26 (0.96)	-8.24 (0.97)	-8.62 (0.86)	-7.67 (1.02)	-8.43 (1.04)
Individual-level					
Age	0.81 (0.71, 0.91)	0.81 (0.71, 0.91)	0.85 (0.76, 0.93)	0.80 (0.70, 0.90)	0.81 (0.72, 0.91)
Female	0.07 (-0.20, 0.34)	0.06 (-0.22, 0.34)	0.02 (-0.25, 0.3)	0.06 (-0.22, 0.34)	0.06 (-0.21, 0.34)
Public assistance	0.70 (0.20, 1.18)	0.69 (0.19, 1.20)	0.81 (0.31, 1.31)	0.72 (0.22, 1.23)	0.68 (0.19, 1.18)
High school degree	-0.24 (-0.69, 0.21)	-0.24 (-0.69, 0.22)	-0.19 (-0.64, 0.29)	-0.19 (-0.64, 0.28)	-0.21 (-0.68, 0.24)
Race					
White	Ref	Ref	Ref	Ref	Ref
Black	-4.10 (-4.55, -3.64)	-4.11 (-4.56, -3.63)	-4.19 (-4.70, -3.7)	-4.20 (-4.72, -3.68)	-4.11 (-4.63, -3.61)
Hispanic	-1.81 (-2.33, -1.29)	-1.80 (-2.32, -1.27)	-2.71 (-3.18, -2.24)	-2.00 (-2.51, -1.48)	-1.83 (-2.36, -1.30)
School-level					
Public assistance	0.07 (0.02, 0.12)	0.07 (0.02, 0.12)			0.07 (0.02, 0.12)
High school degree	-0.04 (-0.08, 0.01)	-0.04 (-0.08, 0.004)			-0.04 (-0.09, 0.0001)
Percent White	0.02 (0.003, 0.03)	0.02 (0.003, 0.03)			0.02 (0.002, 0.03)
Neighborhood-level					
Public assistance			0.03 (-0.01, 0.06)	0.01 (-0.03, 0.05)	-0.005 (-0.04, 0.03)
High school degree			-0.01 (-0.02, 0.01)	0.01 (-0.01, 0.03)	0.01 (-0.01, 0.03)
Percent White			0.01 (0.004, 0.02)	0.01 (-0.005, 0.02)	0.0002 (-0.01, 0.01)
Random effect estimates					
U3 neighborhood (SE)		0.17 (0.04, 0.42)*	1.52 (0.96, 2.18)*	0.16 (0.03, 0.37)*	0.25 (0.05, 0.52)*
U2 school (SE)	1.83 (1.22, 2.64)*	1.78 (1.15, 2.59)*		2.06 (1.39, 2.91)*	1.71 (1.10, 2.48)*
U1 individual (SE)	80.7 (79.0, 82.5)*	80.6 (78.8, 82.3)*	81.1 (79.3, 83.0)*	80.6 (78.8, 82.3)*	80.5 (78.8, 82.3)*
Fit statistics					
DIC	–	116,258	–	116,261	116,265

Model 3 presents the results of the school-only multilevel model and CCMM containing individual-level variables as well as the *school-level* measure of SES and race/ethnicity. Model 4 presents the results of the neighborhood-only multilevel model and CCMM containing individual-level variables combined with the *neighborhood-level* measure of SES and race/ethnicity. Model 5 presents the results of a CCMM containing all individual-, school-, and neighborhood-level variables. For the fixed effect estimates, cell entries are parameter (beta) estimates and credible intervals. The intercept is presented as parameter estimate and standard error (SE). Random effects are presented as estimate and credible intervals. DIC refers to Deviance Information Criterion, a measure of model fit and complexity and is only reported for the cross-classified models.

* Significant random effects are indicated by $p < 0.05$.

Results of Model 5 (see Table 3), which introduced both school- and neighborhood-level SES indicators and covariates into the CCMM, showed that the neighborhood-level variance increased compared to Model 4 ($\sigma_{u0k}^2 = 0.25$ from 0.16). The school-level variance declined slightly compared to Model 4 ($\sigma_{u0j}^2 = 1.71$ from 2.06). Modest associations between school-level public assistance ($\beta=0.07$) and percentage white ($\beta=0.02$) persisted, as did the lack of association between any of the neighborhood-level indicators.

Post-hoc evaluation of these models suggested the results were robust. With the exception of the neighborhood random effect estimates, all other standard errors were small relative to the parameter estimates, suggesting the variances were estimated with a good level of precision. However, the standard errors for the variance components were inflated in the multilevel models in particular. The fixed effect estimates were similar between the CCMM and multilevel models. Residual plots suggested that the model assumptions were not violated in our final model.

3.2. Multilevel models with smoking as binary outcome

Web Table 1 and Web Table 2 present the results of models with smoking as a binary outcome. The same overall conclusions can be drawn from these models as the models focusing on the outcome for number of days reported smoking. Fig. 4 presents the variance estimates at the school- and neighborhood-level for the school-only multilevel model, neighborhood-only multilevel model, and CCMM predicting odds of smoking. As seen in this figure, the school- (Model 3) and neighborhood-only multilevel models (Model 4) both identify a modest degree of between-level variation. In other words, the school-level multilevel model suggests there is a significant degree of between-school variation

in the odds of smoking, while the neighborhood-level multilevel model suggests there is a significant degree of between-neighborhood variation in the odds of smoking. However, the results of the CCMM (Model 5) suggest that the between-level variation in smoking is driven primarily by school-level variation and not neighborhood-level variation.

3.3. Sensitivity analysis

We ran a sensitivity analysis to address possible concerns regarding the distribution of students in schools versus neighborhoods. Using Model 5 for both outcomes, we evaluated whether similar results would be obtained if we included an indicator variable adjusting for the number of youth in each neighborhood (0=included more than one respondent; 1=included only one respondent). This allowed us to determine whether the neighborhoods with only one respondent ($n=970$) had a different effect on smoking relative to the neighborhoods that contained a greater number of respondents. As shown in Web Table 3, there was no relationship between the indicator variable and either the continuous ($\beta=0.44$; CI$=-0.17$, 1.05) or binary smoking outcome (OR$=1.06$; CI$=0.89$, 1.24). Fixed and random effect parameters in the models including the indicator variable were also nearly identical to the final CCMM. These findings suggest our results were not influenced by one-respondent neighborhoods.

4. Discussion

This paper demonstrated the value of cross-classified multi-level models (CCMM) to ascertain the quantitative importance of
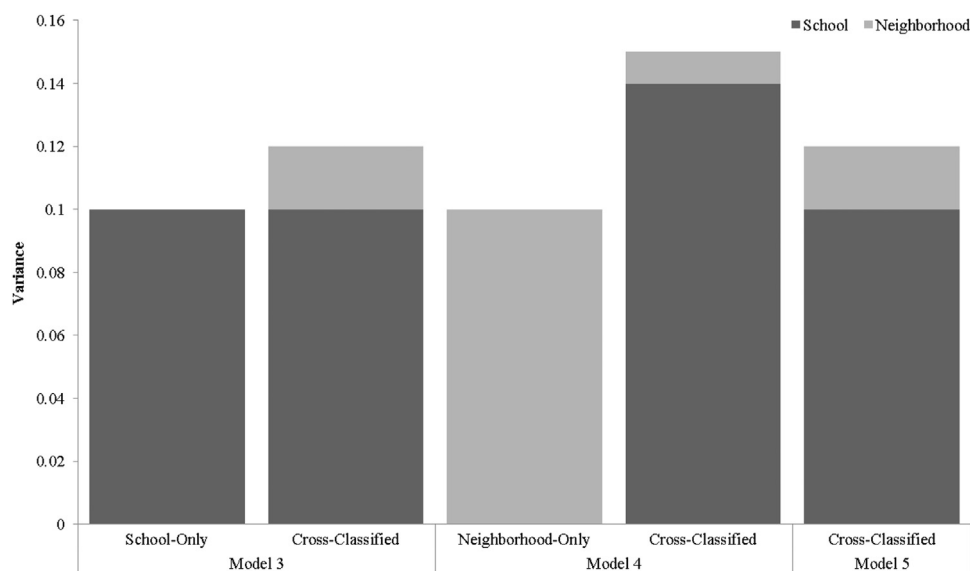
Fig. 4. Variance estimates at the school- and neighborhood-level for the school-only multilevel model, neighborhood-only multilevel model, and Cross-Classified Multilevel Model Predicting Odds of Smoking. Model 3 includes individual-level covariates as well as school-level SES indicators and covariates. Model 4 includes neighborhood-level covariates as well as neighborhood-level SES indicators and covariates. Model 5 includes individual-level, school-level, and neighborhood-level SES indicators and covariates. All variance estimates shown were statistically significant with the exception of the neighborhood effect in cross-classified Models 3 and 4.

more than one context simultaneously. One very salient finding emerged from the current study. We found that two-level multilevel models that do not account for other non-nested contexts in which individuals are embedded can produce misleading results. In the current case, which focused on adolescent smoking, we found that the role of neighborhoods was overestimated. Specifically, the two-level multilevel model that examined neighborhoods without considering schools suggested that 5.2% of the variation in smoking was due to differences between neighborhoods. However, the neighborhood variation was reduced to 0.52% after running a CCMM that considered the role of schools. Thus, had we estimated only a neighborhood-level multilevel model, and not estimated a CCMM that accounted for the random effect of both schools and neighborhoods, we would have misattributed more of the variance in smoking to neighborhoods and missed the contribution of schools. These results are consistent with previous studies, which have found that failing to account for cross-classified data structures will produce biased variance estimates, such that the variance associated with the omitted level (e.g., school) will be attributed to the included level (e.g., neighborhood) (Meyers and Beretvas, 2006). Model misspecification will be most problematic when there is substantial between-level variation in the higher level unit that is omitted (Goldstein, 1994) and when the two higher-level units are independent (Meyers and Beretvas, 2006).

Results from our methodological demonstration of CCMM should be evaluated in light of several limitations. First, the analyses are based on a nationally-representative sample of adolescents drawn from a study that selected youth using school-based sampling. As a result, the number of individuals per neighborhood was smaller than the number of individuals per school, with some neighborhoods only having one respondent. The fact that we did observe meaningful cross-classification of neighborhood and school units, combined with the finding that there was significant neighborhood-level variation in smoking to be explained in the two-level multilevel model, suggests that data sparseness was likely not an issue here. While it is possible our results reflect the greater number of individuals per school rather than neighborhood, we think this is unlikely, as preliminary analyses we conducted in Add Health using CCMM to examine other health outcomes did find meaningful effects for neighborhoods

(results available from the author). However, future research should conduct simulation work and also use the CCMM framework in neighborhood-based studies to evaluate the degree to which sampling methods contribute to these findings.

Second, AddHealth Wave 1 data were also collected more than a decade ago. Although these data may be considered old, AddHealth remains the only nationally-representative sample of adolescents in the US and one of very few in which school and neighborhood level data are available. AddHealth was therefore an appropriate dataset to demonstrate the value of the CCMM framework. Third, we did not complete a detailed investigation into the predictors of smoking, as our example was intended to be a methodological and practical illustration of the CCMM. Inclusion of other potentially relevant factors, when modeled as fixed effects, such as smoking-related policies, measures of social cohesion, will be important for shedding light on the determinants of smoking behavior as well as the between-school and between-neighborhood differences in smoking. Fourth, there are limitations with using Census tracts to define neighborhoods. Although Census Tracts are an imperfect measure to define neighborhoods, they are most commonly used in multilevel research (Dunn et al., 2014). Future studies should expand upon traditional boundaries of neighborhood to focus on other aspects of the residential environment (e.g., activity spaces; neighborhoods as defined by residents rather than administrative datasets). Finally, our school-level measures of socioeconomic status are limited in that they were aggregated directly from individual-level data. This may be problematic, particularly as our measure of public assistance was based on whether mothers' (and not fathers') reported receiving assistance. Though a common approach in multilevel research (Dunn et al., 2014), future studies should use administrative or school-level data when available in order to avoid the concerns about the indicators being a reflection of the students in the sample.

What implications can be drawn from our findings? First, our results emphasize the need to extend current analytic approaches from the basic multilevel model to examine cross-classification. Although there have been some studies using CCMM (see for example Leyland and Naess, 2008; Lloyd et al., 2010; Utter et al., 2011; Riva et al., 2009; Moore et al., 2013; Virtanen et al., 2010; Basile et al., 2012), the CCMM method is not being used nearly as much as it should be, most likely due to a lack of applied examples.

Greater use of CCMM in epidemiology is highly plausible, given the fact that many epidemiological studies collect data in a way that support investigation of cross classification.

Second, our results suggest that given the observed differences in results between the multilevel models and CCMM, scrutiny may be needed to evaluate the strength of prior epidemiological studies on the social determinants of health focusing solely on one context. In the AddHealth sample alone, the majority of studies have examined the role of neighborhoods on adolescent smoking, with much fewer focusing on schools. No studies that we are aware of have used CCMM applied to smoking in AddHealth. How should we interpret this extant research in light of our findings? Although much more research is needed, results of the current study reveal that it is possible that the fixed and random effects identified in previous studies are misleading. We think this is a possibility as even our null CCMM, which did not include any predictors at all, found that the variance attributed to smoking for neighborhoods and schools differed from the conclusions that would have been reached using a traditional MLM, where only one random effect was modeled.

The reasons why CCMM has not been widely are unclear. However, what is known is that few studies have tried to explicitly disentangle the role of one context from another, including neighborhoods and schools (Leventhal and Brooks-Gunn, 2000). Considerably more work in this area is needed, particularly to help guide the investment of limited public health resources. Without having clear data showing that one context is more important than another (or that both are equally important), the field runs the risk of implementing misguided policies and interventions to contexts that may not be capable of having large effects on reducing population health risk and promoting population health outcomes.

In summary, this paper provides a framework that we hope will spark more critical evaluations of multilevel analyses that do not account for other potentially relevant contexts. The CCMM methods described here can be extended to many other health outcomes, and at other stages in the lifecourse, to evaluate the role of social contexts on health and ultimately lead to the development of interventions to improve population-level health.

Acknowledgments

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.healthplace.2014.12.001.

References

Alexander, C., et al., 2001. Peers, schools, and adolescent cigarette smoking. J. Adolesc. Health 29, 22–30.

Basile, C., et al., 2012. Associations of supermarket characteristics with weight status and body fat: a multilevel analysis of individuals within supermarkets (RECORD Study). PLoS One 7 (4), 1–10.

Bonnell, C., et al., 2013. The effects of the school environment on student health: a systematic review of multi-level studies. Health Place 21, 180–191.

Bronfenbrenner, U., Morris, P.A., 2006. The bioecological model of human development. In: Lerner, R.M. (Ed.), Handbook of Child Psychology. John Wiley and Sons, Hoboken, NJ, pp. 793–828.

Browne, N.J., 2004. MCMC Estimation in MLwiN Version 2.0. Centre for Multilevel Modeling, Institute of Education, London, UK.

Diez Roux, A.V., 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Am. J. Public Health 88 (2), 216–222.

Diez Roux, A.V., 2002. A glossary for multilevel analysis. J. Epidemiol. Community Health 56, 588–594.

Duncan, B., Rees, D.I., 2005. Effect of smoking on depressive symptomatology: a reexamination of data from the National Longitudinal Study of Adolescent Health. Am. J. Epidemiol. 162 (5), 461–470.

Dunn, E.C., et al., 2014. Translating multilevel theory into multilevel research: challenge and opportunities for understanding the social determinants of psychiatric disorders. Soc. Psychiatry Psychiatr. Epidemiol. 49, 859–872.

Goldstein, H., 1994. Multilevel cross-classified models. Soc. Methods Res. 22, 364–375.

Kawachi, I., Berkman, L.F. (Eds.), 2003. Neighborhoods and Health. Oxford University Press, New York, NY.

Harris, K.M., 2013. The Add Health Study: Design and Accomplishments. Carolina Population Center, University of North Carolina at Chapel Hill.

Krieger, N., 2001. A glossary for social epidemiology. J. Epidemiol. Community Health 55, 693–700.

Lee, E.S., Forthofer, R.N., 2006. Analyzing Complex Survey Data. Sage, Thousand Oaks.

Leventhal, T., Brooks-Gunn, J., 2000. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol. Bull. 126 (2), 309–337.

Leventhal, T., Brooks-Gunn, J., 2003. Moving to opportunity: an experimental study of neighborhood effects on mental health. Am. J. Public Health 93 (9), 1576–1582.

Leyland, A.H., Naess, O., 2008. The effect of area of residence over the life course on subsequent mortality. J. R. Stat. Soc. Ser. A (Stat. Soc.) 172 (3), 555–578.

Lloyd, J.E.V., Li, L., Hertzman, C., 2010. Early experiences matter: lasting effect of concentrated disadvantage on children's language and cognitive outcoems. Health Place 16, 371–380.

Mair, C.F., Diez Roux, A.V., Galea, S., 2008. Are neighborhood characterisics associated with depressive symptoms? A critical review. J. Epidemiol. Community Health 62 (11), 940–946.

Meyers, J.L., Beretvas, S.N., 2006. The impact of inappropriate modeling of cross-classified data structures. Multivariate Behav. Res. 41 (4), 473–497.

Moore, K., et al., 2013. Home and work neighbourhood environments in relation to body mass index: the Multi-Ethnic Study of Atherosclerosis (MESA). J. Epidemiol. Community Health 67 (10), 846–853.

Pickett, K.E., Pearl, M., 2001. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. J. Epidemiol. Community Health 55, 111–122.

Rabash, J., Browne, W., 2001. Non-hierarchical multilevel modelsIn: Leyland, A.H., Goldstein, H. (Eds.), Modeling Non-hierarchical Structures. Wiley and Sons, New York, NY, pp. 93–105.

Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methodsecond ed. Sage Publications, Thousand Oaks, CA.

Riva, M., et al., 2009. Disentangling the relative influence of built and socio-economic environments on walking: the contribution of areas homogenous along exposures of interest. Soc. Sci. Med. 69, 1296–1305.

Rodriguez, G., Goldman, N., 1995. An assessment of estimation procedures for multilevel models with binary responses. J. R. Stat. Soc. Ser. A (Stat. Soc.) 158, 73–90.

Stokols, D., 1996. Translating social ecological theory into guidelines for community health promotion. Am. J. Health Promotion 10, 282–298.

Subramanian, S.V., Jones, K., Duncan, C., 2003. Multilevel methods for public health research. In: Kawachi, I., Berkman, L.F. (Eds.), Neighborhoods and Health. Oxford University Press, New York, NY, pp. 65–111.

Utter, J., et al., 2011. Social and physical contexts of schools and neighborhoods: associations with physical activity among young people in New Zealand. Am. J. Public Health 101 (9), 1690–1695.

Virtanen, M., et al., 2010. School neighborhood disadvantage as a predictor of long-term sick leave among teachers: prospective Cohort Study. Am. J. Epidemiol. 171 (7), 785–792.

Web Table 1. Nested Models Describing Association Between Predictors and Binary Smoking Outcome (0=not smoked in past 30 days; 1=smoked in past 30 days) in the National Longitudinal Study of Adolescent Health (N=16,070)

	Model 1			Model 2		
Fixed Effect Estimates	School Only	Neighborhood Only	Cross-Classified	School Only	Neighborhood Only	Cross-Classified
Intercept (SE)	-1.22 (0.06)	-1.20 (0.03)	-1.23 (0.06)	0.03 (0.01)	0.02 (0.005)	0.03 (0.01)
Individual-level						
Age				1.19 (1.16, 1.22)	1.21 (1.18, 1.24)	1.19 (1.17, 1.23)
Female				1.01 (0.94, 1.09)	1.00 (0.92, 1.07)	1.01 (0.94, 1.09)
Public Assistance				1.40 (1.22, 1.58)	1.43 (1.25, 1.63)	1.41 (1.23, 1.60)
High School degree				1.00 (0.90, 1.12)	1.01 (0.90, 1.15)	1.01 (0.89, 1.13)
Race						
White				Ref	Ref	Ref
Black				0.30 (0.26, 0.34)	0.28 (0.24, 0.31)	0.30 (0.26, 0.34)
Hispanic				0.68 (0.59, 0.78)	0.54 (0.48, 0.60)	0.68 (0.60, 0.78)
School-level						
Public Assistance						
High School Degree						
Percent White						
Neighborhood-level						
Public Assistance						
High School Degree						
Percent White						
Random Effect Estimates						
U3 neighborhood (SE)		0.31 (0.24, 0.38)*	0.05 (0.03, 0.09)*		0.10 (0.06, 0.15)*	0.01 (0.00, 0.02)
U2 school (SE)	0.36 (0.26, 0.48)*		0.36 (0.25, 0.49)*	0.14 (0.09, 0.20)*		0.13 (0.08, 0.20)*
Fit Statistics						
DIC			17522			17129

Model 1 presents the results for a null model (i.e., no covariates) for each model type: school-only multilevel model, neighborhood-only multilevel model, and the cross-classified multilevel model. Model 2 presents the same models as Model 1, except Model 2 includes individual-level predictors and covariates. For the fixed effect estimates, cell entries are odds ratio (OR) estimates and credible intervals. The intercept is presented as parameter estimate and standard error (SE). Random effects are presented as estimate and credible intervals. DIC refers to Deviance Information Criterion, a measure of model fit and complexity, and is only reported for the CCMM. Significant random effects are indicated by * (p<0.05).

Web Table 2. Nested Models Describing Association Between Covariates and Binary Smoking Outcome (0=not smoked in past 30 days; 1=smoked in past 30 days) in the National Longitudinal Study of Adolescent Health (N=16,070)

Fixed Effect Estimates	Model 3		Model 4		Model 5
	School Only	Cross-Classified	Neighborhood Only	Cross-Classified	Cross-Classified
Intercept (SE)	0.02 (0.003)	0.02 (0.003)	0.02 (0.004)	0.03 (0.01)	0.02 (0.004)
Individual-level					
Age	1.19 (1.16, 1.21)	1.19 (1.17, 1.21)	1.22 (1.20, 1.24)	1.18 (1.16, 1.20)	1.20 (1.17, 1.23)
Female	1.01 (0.94, 1.09)	1.01 (0.93, 1.09)	1.00 (0.93, 1.07)	1.01 (0.94, 1.09)	1.01 (0.94, 1.09)
Public Assistance	1.37 (1.20, 1.58)	1.37 (1.20, 1.56)	1.43 (1.25, 1.63)	1.39 (1.21, 1.60)	1.36 (1.18, 1.56)
High School degree	0.98 (0.87, 1.09)	0.99 (0.86, 1.12)	0.99 (0.87, 1.14)	1.00 (0.88, 1.16)	1.00 (0.88, 1.13)
Race					
White	Ref	Ref	Ref	Ref	Ref
Black	0.31 (0.27, 0.36)	0.32 (0.28, 0.36)	0.3 (0.25, 0.34)	0.3 (0.26, 0.36)	0.31 (0.26, 0.35)
Hispanic	0.72 (0.62, 0.81)	0.73 (0.64, 0.83)	0.55 (0.48, 0.63)	0.68 (0.59, 0.78)	0.71 (0.61, 0.81)
School-level					
Public Assistance	1.02 (1.01, 1.03)	1.02 (1.01, 1.03)			1.02 (1.01, 1.04)
High School Degree	0.99 (0.98, 1.00)	0.99 (0.98, 1.00)			0.99 (0.98, 1.00)
Percent White	1.01 (1.00, 1.01)	1.01 (1.00, 1.01)			1.01 (1.00, 1.01)
Neighborhood-level					
Public Assistance			1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	0.99 (0.98, 1.00)
High School Degree			1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Percent White			1.00 (0.99, 1.01)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Random Effect Estimates					
U3 neighborhood (SE)		0.02 (0.00, 0.04)	0.10 (0.06, 0.14)*	0.01 (0.00, 0.01)	0.02 (0.01, 0.04)*
U2 school (SE)	0.10 (0.06, 0.16)*	0.10 (0.06, 0.16)*		0.14 (0.09, 0.20)*	0.10 (0.06, 0.15)*
Fit Statistics					
DIC		17131		17129	17133

Model 3 presents the results of the school-only multilevel model and CCMM containing individual-level variables as well as the *school-level* measure of SES and race/ethnicity. Model 4 presents the results of the neighborhood-only multilevel model and CCMM containing individual-level variables combined with the *neighborhood-level* measure of SES and race/ethnicity. Model 5 presents the results of a CCMM containing all individual-, school-, and neighborhood-level variables. For the fixed effect estimates, cell entries

are odds ratio (OR) estimates and credible intervals. The intercept is presented as parameter estimate and standard error (SE). Random effects are presented as estimate and credible intervals. DIC refers to Deviance Information Criterion, a measure of model fit and complexity, and is only reported for the CCMM. Significant random effects are indicated by * (p<0.05).

Web Table 3. Nested Cross-Classified Multilevel Models (CCMM) Describing Association Between Predictors and Number of Days Smoked in the Past 30 Days and Binary Smoking Outcome in the National Longitudinal Study of Adolescent Health (N=16,070) adjusting for neighborhoods with one respondent (n=970)

	Number of Days Smoked in Past 30 Days	Binary Smoking
Fixed Effect Estimates		
Intercept (SE)	-8.40 (1.05)	0.02 (0.01)
Individual-level		
Age	0.81 (0.71, 0.90)	1.18 (1.15, 1.21)
Female	0.06 (-0.21, 0.35)	1.01 (0.94, 1.08)
Public Assistance	0.68 (0.18, 1.17)	1.36 (1.19, 1.56)
High School degree (parent)	-0.22 (-0.69, 0.25)	0.98 (0.86, 1.12)
Race		
White	Ref	
Black	-4.13 (-4.63, -3.59)	0.30 (0.26, 0.35)
Hispanic	-1.84 (-2.36, -1.32)	0.70 (0.61, 0.81)
Neighborhood has one respondent	0.44 (-0.17, 1.05)	1.06 (0.89, 1.24)
School-level		
Public Assistance	0.07 (0.02, 0.12)	1.02 (1.01, 1.03)
High School Degree	-0.04 (-0.09, 0.003)	0.99 (0.98, 1.00)
Percent White	0.02 (0.001, 0.03)	1.01 (1.00, 1.01)
Neighborhood-level		
Public Assistance	-0.005 (-0.04, 0.03)	0.99 (0.98, 1.01)
High School Degree	0.01 (-0.01, 0.03)	1.005 (1.00, 1.01)
Percent White	0.0002 (-0.01, 0.01)	0.998 (0.995, 1.00)
Random Effect Estimates		
U3 neighborhood (SE)	0.21 (0.05, 0.50)*	0.01 (0.003, 0.02)
U2 school (SE)	1.79 (1.17, 2.60)*	0.10 (0.06, 0.16)*
U1 individual (SE)	80.6 (78.8, 82.4)*	-

For the fixed effect estimates, cell entries are parameter (beta) estimates and credible intervals for continuous days smoked outcome and parameter (OR) estimates and credible intervals for binary smoking outcome. The intercept is presented as parameter estimate and standard error (SE). Random effects are presented as estimate and credible intervals. Significant random effects are indicated by * (p<0.05).

Technical Appendix Part 1:
Running Cross-Classified Multilevel Models in MLWiN

**If you used this appendix for your analysis, please cite us:

Dunn, E.C., Richmond, T.K., Milliren, C.E., & Subramanian, S.V. Using Cross-Classified Multilevel Models to Disentangle School and Neighborhood Effects: An Example Focusing on Smoking Behaviors among Adolescents in the United States. *Health and Place*

Introduction

This first technical appendix is intended to show MLwiN users how to fit cross-classified multilevel models in MLwiN. A second technical appendix shows users how MLwiN can be executed through STATA.

Here, we provide detailed instruction on how to fit a cross-classified model with a continuous outcome followed by a brief overview of how to fit a cross-classified model with a binary outcome. Although the same general set of steps are taken for either model, we think it is easier to understand the procedures to fit a linear model and thus recommend readers start analyzing cross-classified models with continuous outcomes.

Our illustration uses data previously constructed, cleaned, and analyzed in the paper by Dunn and colleagues noted above. We used MLwiN version 2.26 for all analyses (Rasbash et al. 2012, Center for Multilevel Modelling).

In our analysis, all data manipulation (e.g., creating derived variables, recoding variables, etc.) was performed prior to importing the data into MLwiN. For ease of implementation, we recommend all data manipulation (e.g., creating derived variables, recoding variables, etc) occur in other programs (e.g., SAS, STATA) outside of MLwiN. Readers interested in specific data manipulation capabilities should refer to the MLwiN manual (Rasbash et al. 2012).

We would also like to note that MLwiN requires a constant variable consisting of a vector of 1's for all observations. The constant variable is necessary for modeling the intercept and thus is required to fit a model with random intercepts. We recommend you create this constant variable before importing the data into the program.

Getting Started in MLwiN

Before opening a dataset, the worksheet size should be adjusted to accommodate the size of data (i.e., number of observations) by clicking on *Worksheet* under the **Options** menu.If the number of cells listed in the worksheet is less than the actual number of observations in the data, some observations may be lost in the imported dataset.
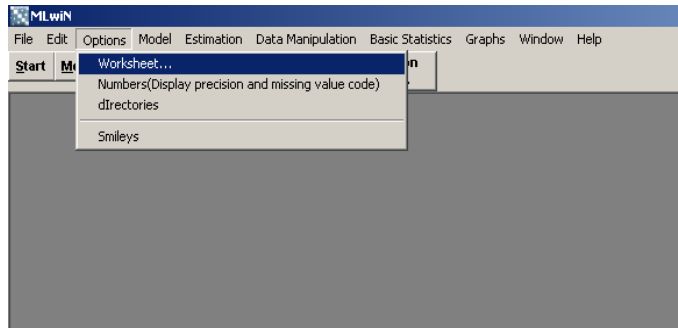
On the Worksheet tab, adjust the number of cells in the worksheet (i.e., number of observations), maximum number of levels (e.g. 2 for individual and school), number of columns (i.e., variables), number of explanatory variables allowed in a single model, and number of groups (i.e., number of categories within the categorical variables) to fit the dataset. The number of cells must be at least equal to the number of observations in the dataset. Adding more cells, levels, or other features to the worksheet, that exceed the actual number in the dataset, will not affect model specification or results in any way. Thus, it is reasonable to always choose numbers larger than what you think you will need. After all adjustments are made, click the "**Done**" button for the changes to take effect.
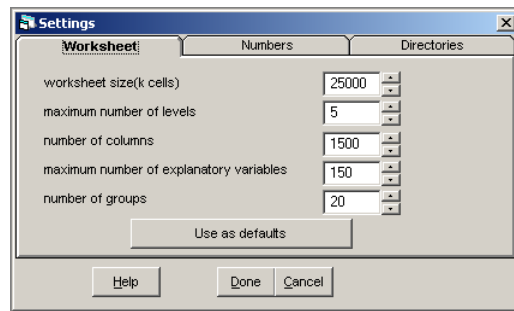
In order to open your dataset, go to the **File** menu and select *Open worksheet*.
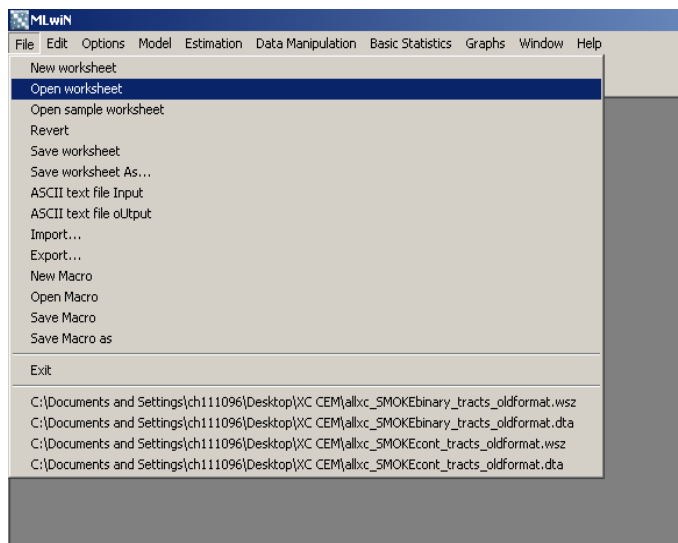
You can open previously saved MLwiN worksheets (.wsz or .ws) as well as datasets from other programs (Stata, SAS, SPSS) by selecting the file type from the drop-down menu
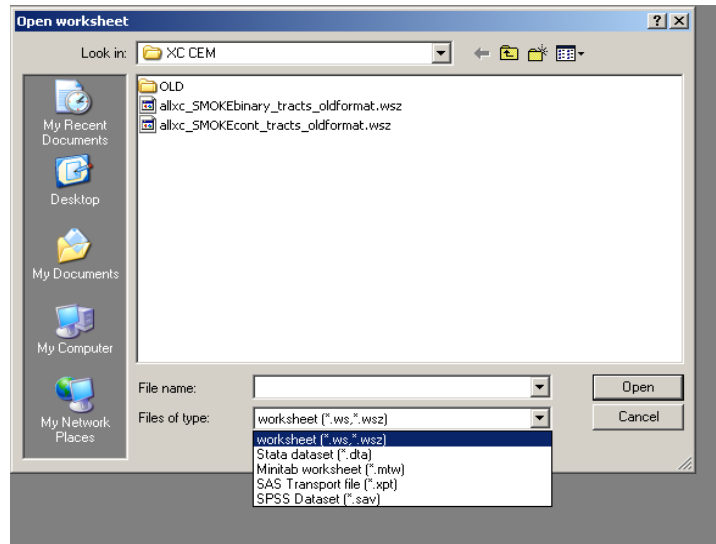
Once a dataset has been opened, the variable names and attributes are displayed in the Names window. A 'categorical' column indicates whether MLwiN recognizes the variable as categorical (indicated by True in the Categorical column) or continuous (indicated by False in the Categorical column) based on the values. This attribute may be changed using the "**Toggle Categorical**" button. Miscategorization of variables will result in continuous variables being treated as categorical and vice versa.

Name	Cn	n	missing	min	max	categorical	description
aid	1	16553	0	0	16552	True	RESPONDENT IDENTIFIER
scid	2	16553	0	0	131	True	SCHOOL ID-W1
age	3	16553	0	11	21	False	Age (years)
female	4	16553	0	0	1	True	Female Gender
publicassist	5	16553	0	0	1	True	Receiving Public Assistance
highschool	6	16553	0	0	1	True	Parental Education >=High Sc...
smoking	7	16553	0	0	30	False	Number of days smoked in pa...
tract_assista...	8	16553	0	0	0.6745562	False	Tract Proportion on Public As...
w1nhood1	9	16553	0	5	2602	False	NEIGHBORHOOD GROUPING L...
sch_publicasst	10	16553	0	0	45.39474	False	% of kids with public assistance
cons	11	16553	0	1	1	False	Constant for MLwiN
black	12	16553	0	0	1	False	
hispanic	13	16553	0	0	1	False	
c14	14	0	0	0	0	False	
c15	15	0	0	0	0	False	
c16	16	0	0	0	0	False	
c17	17	0	0	0	0	False	
c18	18	0	0	0	0	False	
c19	19	0	0	0	0	False	
c20	20	0	0	0	0	False	
c21	21	0	0	0	0	False	
c22	22	0	0	0	0	False	
c23	23	0	0	0	0	False	
c24	24	0	0	0	0	False	
c25	25	0	0	0	0	False	
c26	26	0	0	0	0	False	
c27	27	0	0	0	0	False	
c28	28	0	0	0	0	False	
c29	29	0	0	0	0	False	
c30	30	0	0	0	0	False	
c31	31	0	0	0	0	False	
c32	32	0	0	0	0	False	
c33	33	0	0	0	0	False	
c34	34	0	0	0	0	False	
c35	35	0	0	0	0	False	
c36	36	0	0	0	0	False	

To implement multilevel models, the data must be sorted in ascending order by the identifier variable within each level. Thus, the level 1 identifier (e.g., individual subject identification number) must be sorted within level 2 (e.g., school identification number), level 2 within level 3, etc.. The *Sort* function can be accessed from the **Data Manipulation** menu.
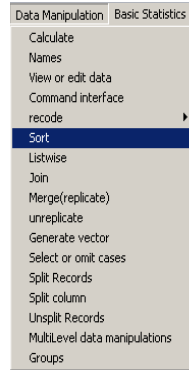
To sort the data, first select the number of levels to sort on (MlwIN refers to this as the number of 'keys' to sort on) and choose the level identifiers from the drop-down menu (under 'key code columns'). The highest level in the hierarchy (e.g. school) should be the first variable and lowest level should be last (e.g. individual). The level identifier variables to be sorted on should be selected next under the Input Columns. Variables can be chosen individually, or click the "**Select All**" button to sort all variables.
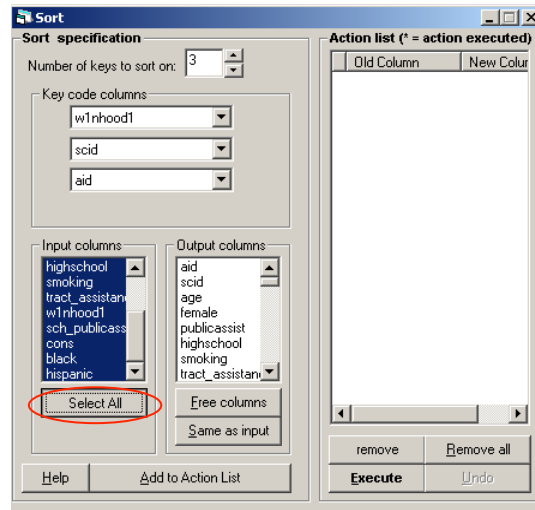
To save these sort changes, the data must be "output." To output all of the sorted variables without creating copies of the variables (recommended), click the "**Same as input**" button under Output columns. The sorted variables can also be output to empty columns by clicking "**Free columns**" instead.
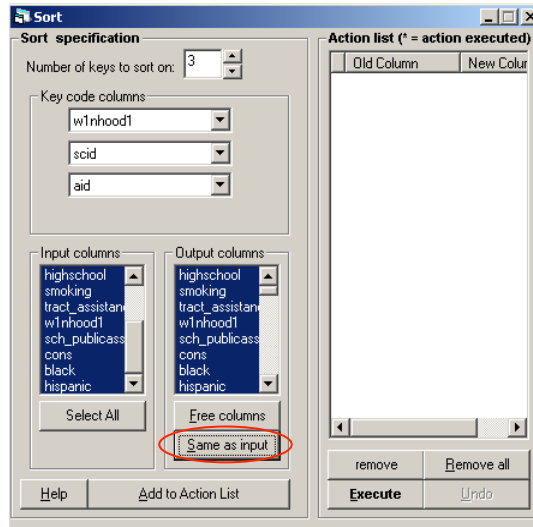
Clicking the "**Add to Action List**" button will create a queue for the variables to be sorted.
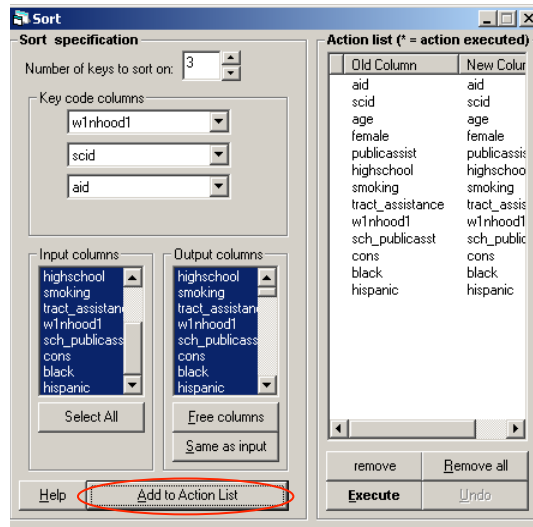
To perform the sort, click the "**Execute**" button. Once the sort is complete, an asterisk will appear next to all the variable names that have been sorted.
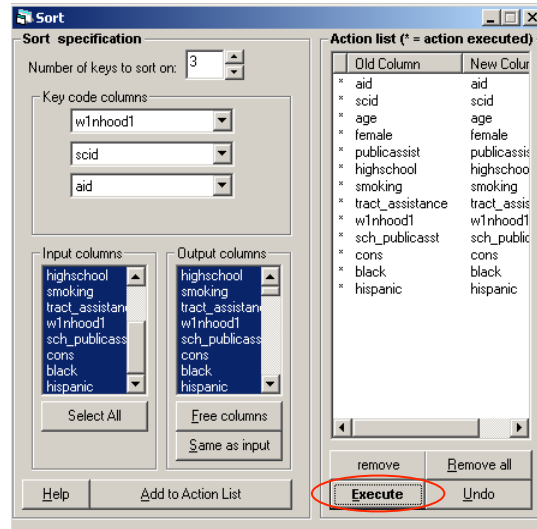
Specifying the Model: Example with a Continuous Outcome

Note: When fitting any cross-classified multilevel model in MLwiN, the model must first be specified assuming a nested hierarchical structure (i.e., individuals clustered in schools and schools clustered in neighborhoods). This initial model provides starting values. The model must then be refit with the cross-classified structure, using the parameters obtained from the first ("nested structure") model as starting values. MLwiN uses the Iterative Generalized Least Squares method to implement hierarchically nested models while cross-classified models must be implemented using Markov Chain Monte Carlo (MCMC) methods. MCMC is a simulation-based method estimating the parameters by re-sampling the data to produce more accurate estimates of the unknown parameters. IGLS and traditional least-squares regression methods give point estimates for unknown parameters calculated from the sample data; however there is no re-sampling performed. MCMC allows for more complex models to be fitted, including cross-classified models. For more information on MCMC, see Chapter 1 of the MCMC manual (Browne, 2012).

To specify the first model, choose *Equations* from the **Model** menu.
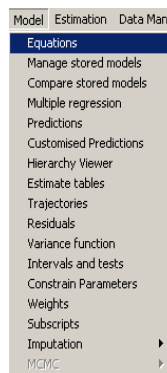
This will bring up the Equations window, where the first model can be specified. Here, you will specify the dependent and independent variables, distribution function of the dependent variable, and the number of levels to model. Red text in the Equations window indicates elements of the model that have not yet been specified. The nested structure model is run using Iterative Generalized Least Squares regression which is indicated by the IGLS under the toolbar. This regression estimation method is the default when MLwiN is opened.
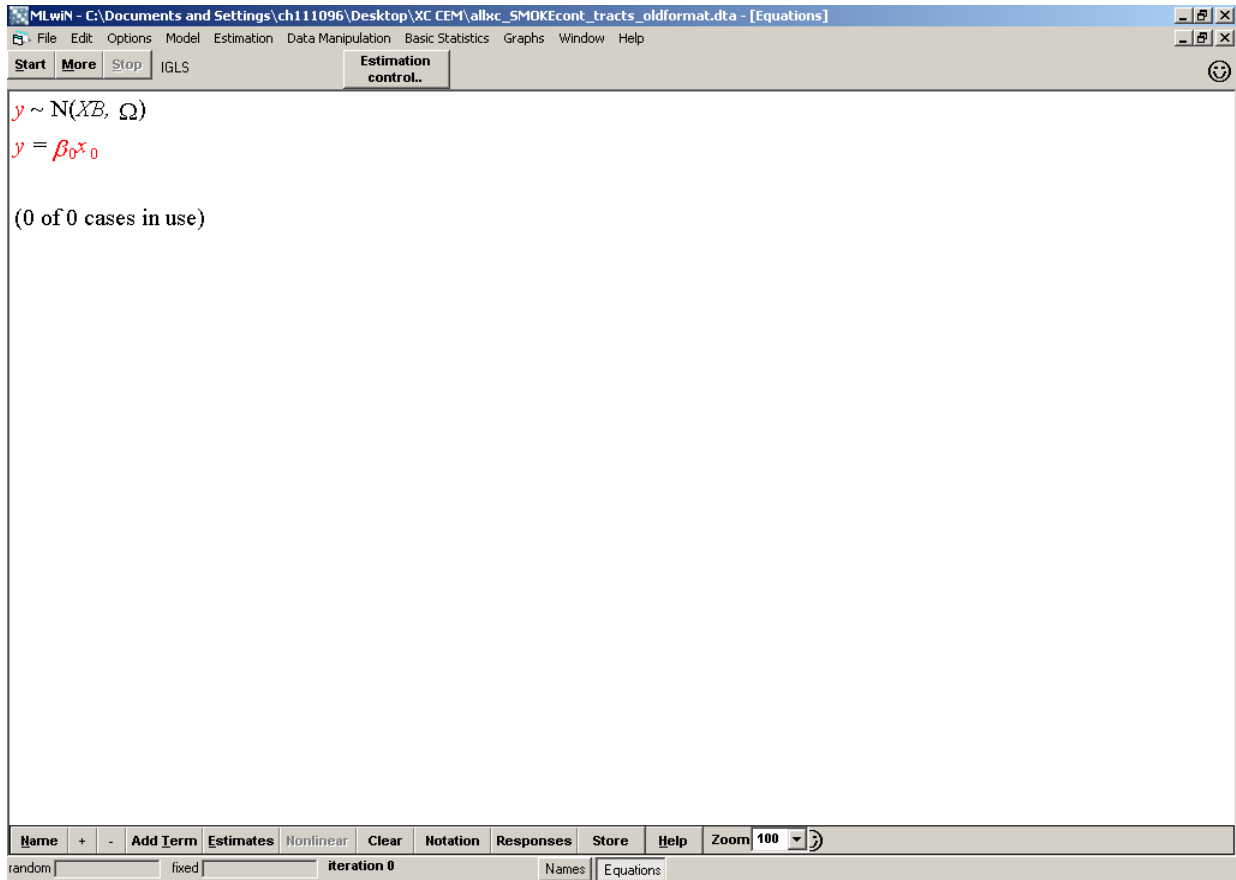
To specify the model, click on the red "y" on the left side of the equation to specify the dependent variable as well as the number of levels. Choose the outcome variable from the drop-down list. Then choose the appropriate number of levels to include in the model.
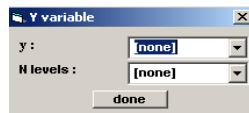
Once the number of levels has been specified, drop-down menus will appear for each level. The level identifiers are specified here. Level 3 is the highest level unit (in our example it is the neighborhood, coded as: w1nhood1), while level 1 is the lowest level (it is the individual, coded as: aid in our example).
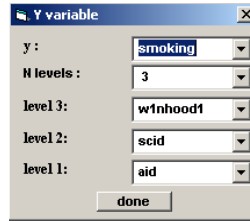
Here the variable *smoking* has been specified as the dependent variable, being modeled at 3 levels neighborhood: winhood1, school: scid, and individual: aid. Click done when the dependent variables and levels have been specified. Clicking the "**Notation**" button, we can change the way subscripts are displayed for the levels.
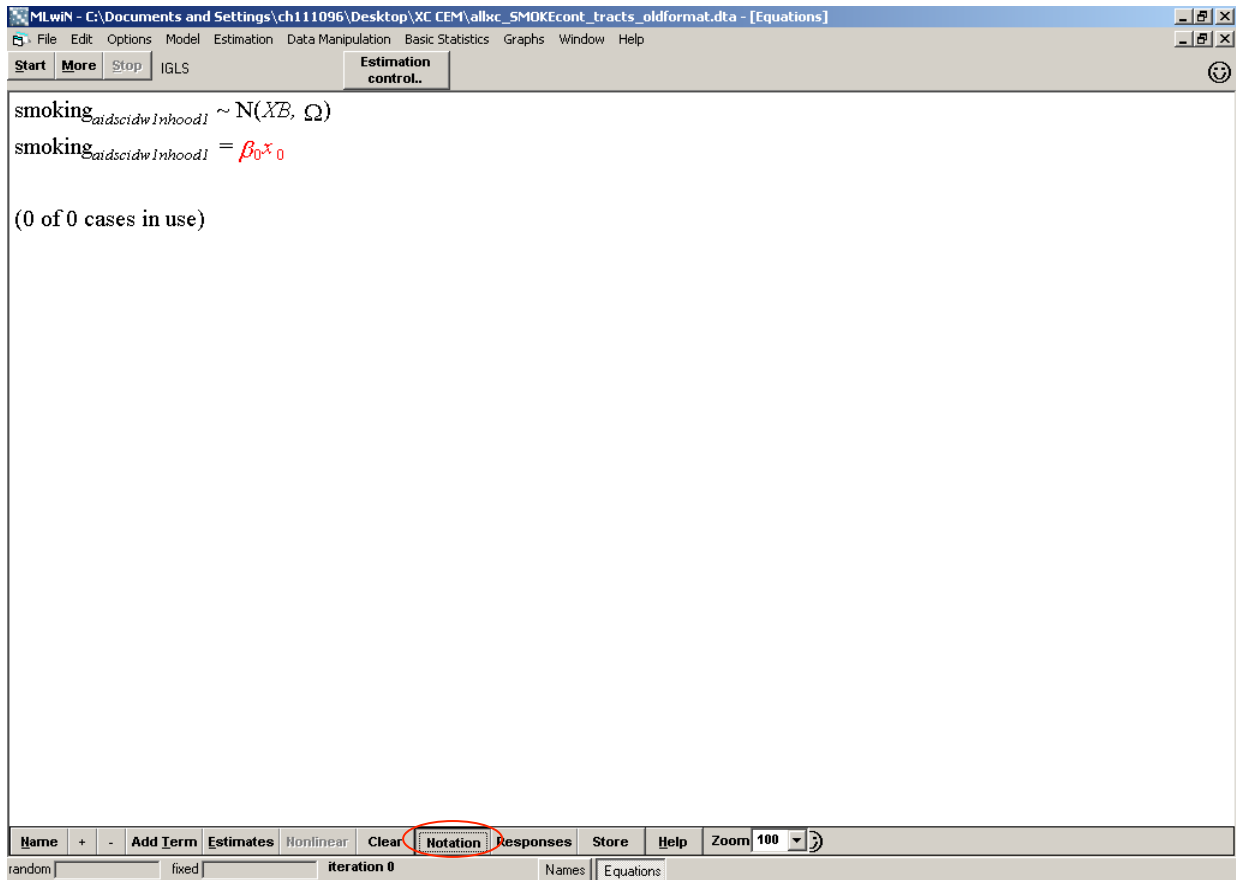
It is particularly useful when implementing the cross-classified model to display a single letter subscripts for all levels. This is because the levels are no longer nested which is not properly accounted for by the traditional hierarchical notation (i[th] student in j[th] school in k[th] neighborhood). There are more complex relationships between school and neighborhood in a cross-classified model (i.e. there can be both crossed and nested relationships between the two levels) which cannot easily be indicated in the traditional notation. For this reason, we use a classification notation with a single subscript to indicate that the levels are not entirely nested.   To display letter subscripts for all levels, uncheck the box for "subscripts as names." To display a single "i" for all three levels, unclick 'multiple subscripts,' to further simplify the display of the model.
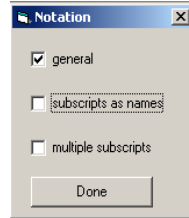
Now the single subscript is displayed instead of the level identifiers.

The default distribution function for the dependent variable is a normal distribution, which is indicated by the N on the right side of the top equation. This indicates a normal distribution (N=normal) for the fixed estimate, $X\beta$ and a random part indicated by Ω.   The distribution of the dependent variable can be changed by double-clicking the N.
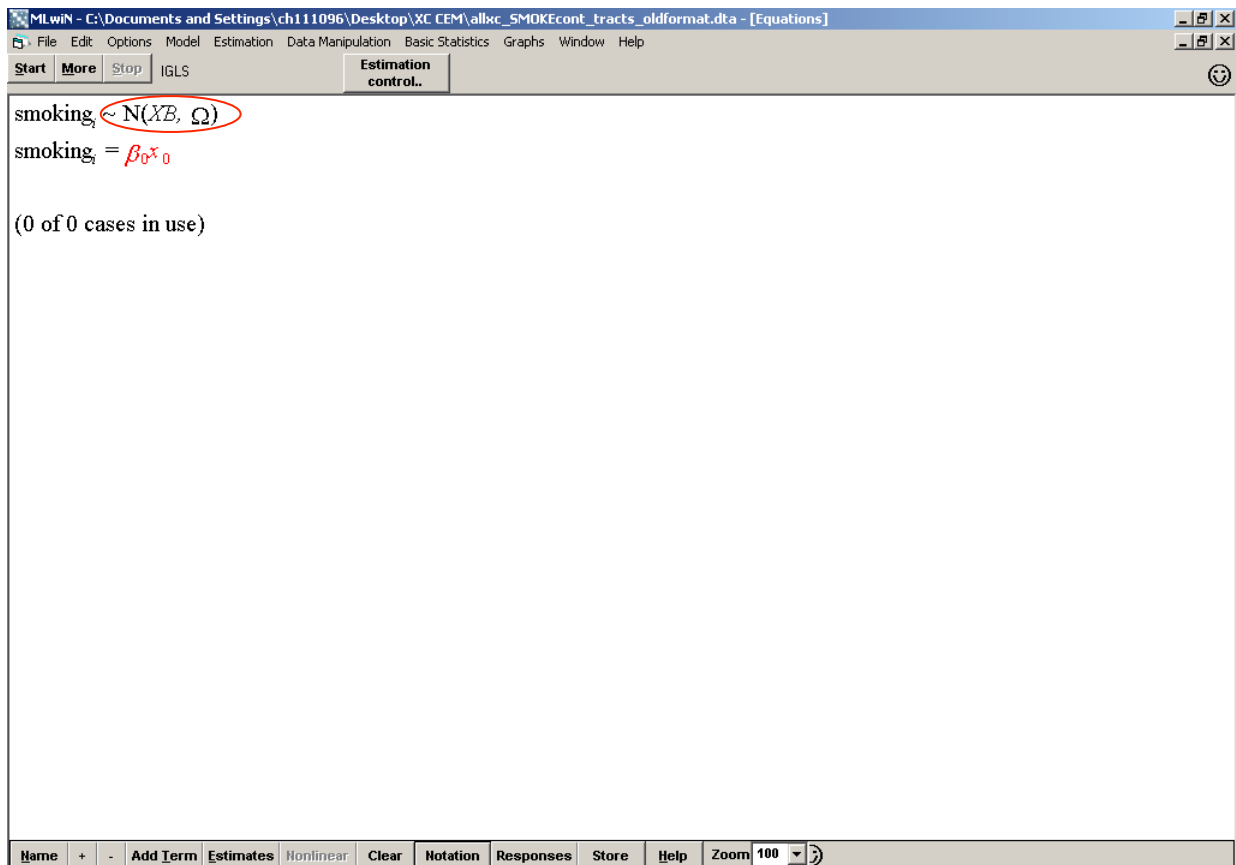
In our example, the outcome variable *smoking* is continuous and for illustrative purposes will be treated as though it is normally distributed. Therefore the model will be specified using the normal distribution. Binomial, Poisson, negative binomial, and multinomial distributions are other options depending on the distribution of the outcome. Once the distribution has been selected, click the "**Done**" button to continue specifiying the model.
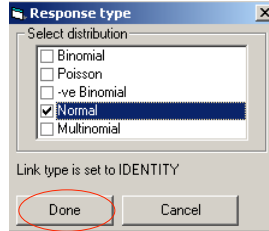
To add a random intercept to the model, double click on the red $\beta_0 x_0$.
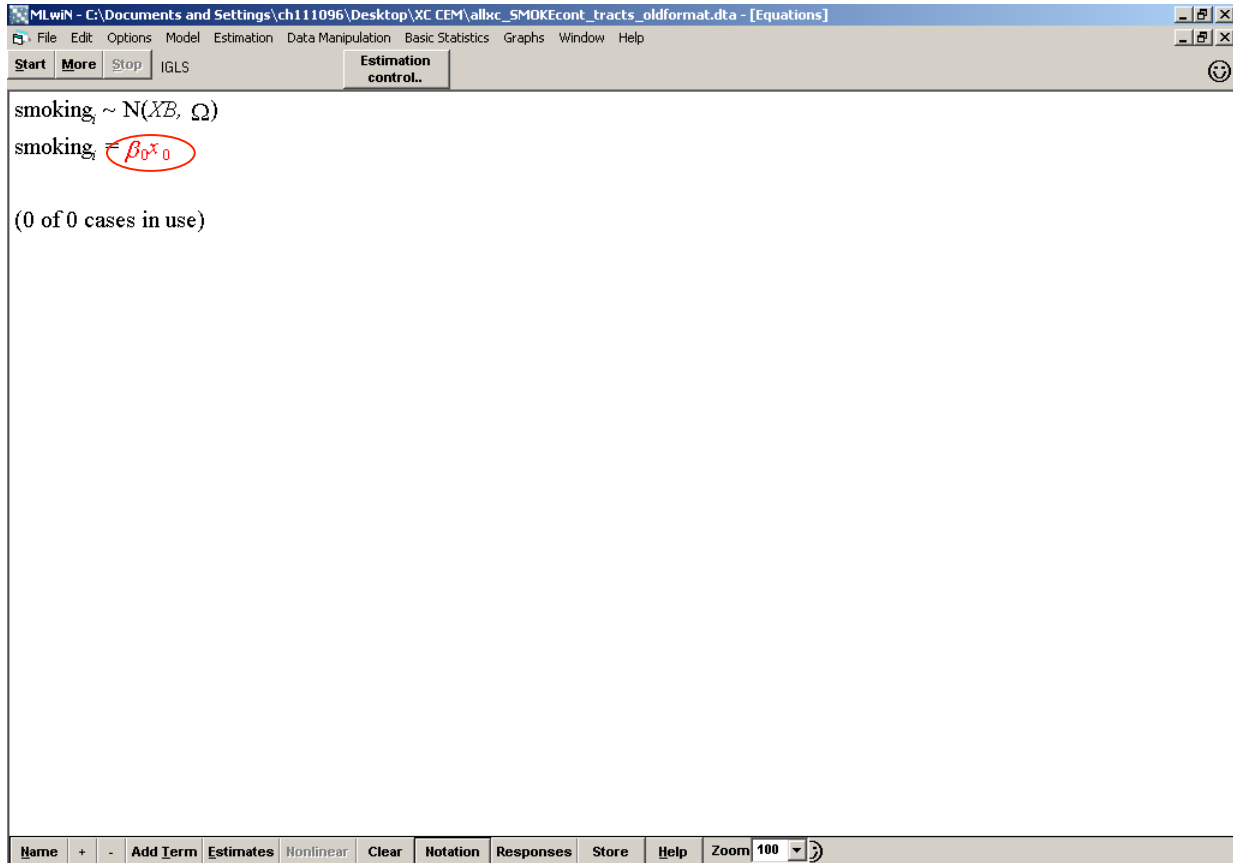
Choose the constant variable (i.e. the vector of 1's) from the drop-down menu (we called this variable *cons*). As noted previously, the constant variable is necessary for fitting the intercept(s). For a random-intercepts model, which allows for random intercepts at all levels, check the boxes for each level as well as the Fixed Parameter. Click done when you are finished.
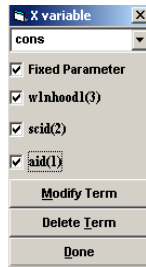
Now that the null random-intercepts model has been fully specified, the equations are black indicating that specification is complete. The number of cases in use is also populated. Compare the number of cases here to the sample size in your dataset to ensure there are no missing values and all cases are used. The model can be now be run using IGLS by clicking the "**Start**" button. To view the estimates, click the "**Estimates**" button twice. Doing so will allow you to view the full algebraic specification of the model.
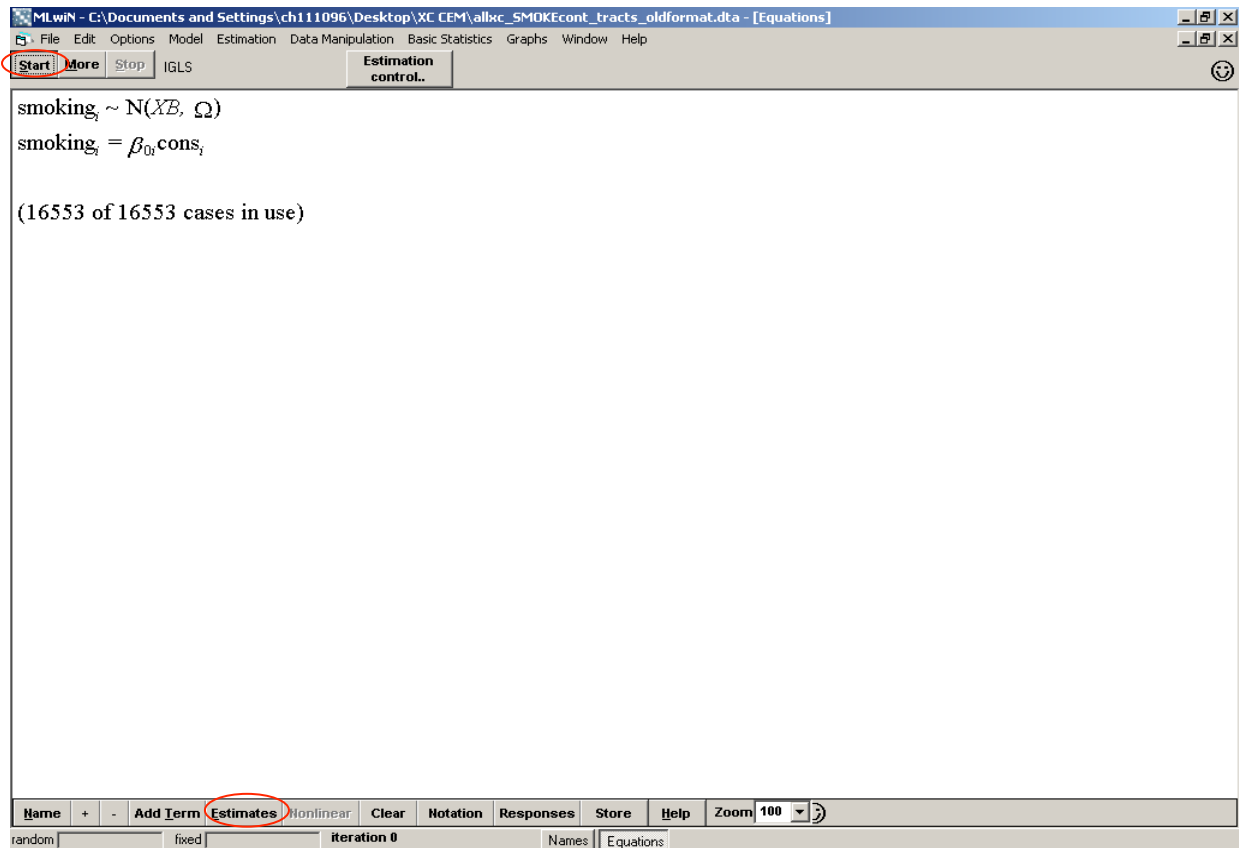
After clicking the "**Estimates**" button once, the specification of the fixed and random parts of the model can be viewed.

$\text{smoking}_i \sim N(XB, \ \Omega)$

$\text{smoking}_i = \beta_{0i}\text{cons}_i$

$\beta_{0i} = \beta_0 + u^{(3)}_{0,wlnhood1(i)} + u^{(2)}_{0,scid(i)} + e_{0i}$

$\left[u^{(3)}_{0,wlnhood1(i)} \right] \sim N(0, \ \Omega^{(3)}_u) \ : \ \Omega^{(3)}_u = \left[\Omega^{(3)}_{u0,0} \right]$

$\left[u^{(2)}_{0,scid(i)} \right] \sim N(0, \ \Omega^{(2)}_u) \ : \ \Omega^{(2)}_u = \left[\Omega^{(2)}_{u0,0} \right]$

$\left[e_{0i} \right] \sim N(0, \ \Omega_e) \ : \ \Omega_e = \left[\Omega_{e0,0} \right]$

$-2*loglikelihood(IGLS\ Deviance) = 120951.462(16553\ of\ 16553\ cases\ in\ use)$

Note the subscripts are uniformly *i* above, as we had indicated earlier.

Clicking the "**Estimates**" button again, we can see the converged estimates, specifically the parameter estimate, (standard error), the variance components for each level in brackets, along with the standard error in parentheses. **Note**: The model is currently assuming a hierarchical structure *not* cross-classified. The estimates should not be interpreted. It is necessary to obtain these estimates as starting values for the cross-classified structure.

To check the hierarchical structure, or the nesting structure of the levels, go to the *Hierarchy Viewer* on the **Model** menu.

Here we see the nested structure of 16,553 students (aid) in 2,647 schools (scid) in 2,142 neighborhoods (w1nhood). As defined in the non-cross classified hierarhical model, the nesting structure is treating different combinations of school and neighborhood as though they are unique schools (e.g, if a single school has individuals from 2 different neighborhoods, this hierarchical nested model would count schools as 2 separate schools). This is not what you want for a cross-classified model.

To treat the levels as cross-classified instead of nested, the model needs to be re-run using MCMC reestimation which can be selected from the **Estimation** menu.
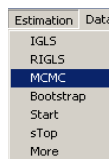
The resampling strategy (e.g. burn in length and number of iterations of the respamples MCMC is taking) can be changed in the Estimation control menu in the MCMC tab. After making any changes, click the "**Done**" button.
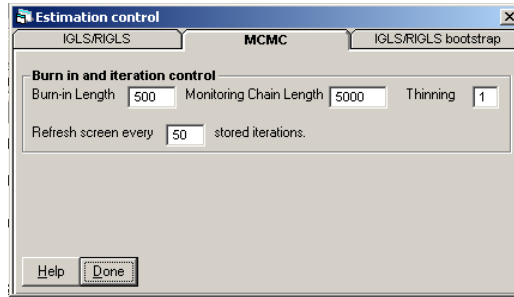
In order to treat neighborhood and school levels as cross-classified instead of hierarchically nested, we need to change the structure in Classifications Information which can be accessed by choosing *Classifications* from the **Model** menu.

Check the box "Treat levels as cross-classified" in the Classification Information window to run the model with a cross-classified structure.

Looking at the Hierarchy Viewer again, we see there are now 16,553 students nested in 132 schools from 2,142 neighborhoods. The number of schools is reduced by treating schools and neighborhoods as cross-classified levels.
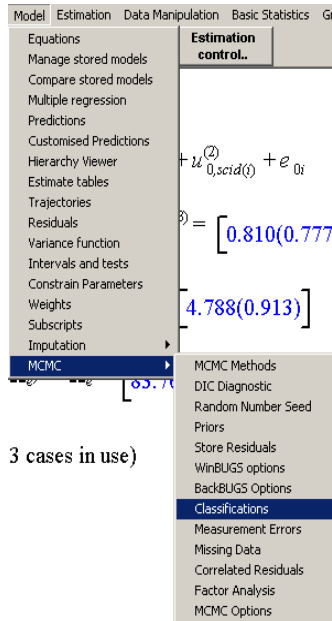
Now the model can be run with a cross-classified structure using MCMC re-estimation by clicking the "**Start**" button again. The MCMC re-estimation may take a few minutes depending on the sample size and extent of cross-classification in your data. The estimates are changed from the ones we computed using IGLS. In particular, you could notice a change in the variance components. With the model correctly specified and the cross-classification of the levels accounted for, these estimates are fully interpretable.   However, the Deviance statistic reported is not interpretable for these models and should be ignored. This is because the Deviance statistic does not utilize the information from the MCMC re-sampling. The Deviance Information Criterion or DIC is a better diagnostic of model fit in cross-classified models which is discussed on the next page.

The Deviance Information Criterion (DIC) can be used to assess model fit in cross-classified models; the Deviance, reported in the Equations window, should be ignored when examining the cross-classified model. To obtain this statistic, choose "DIC Diagnostic" from the *MCMC* drop-down menu in the **Model** menu.
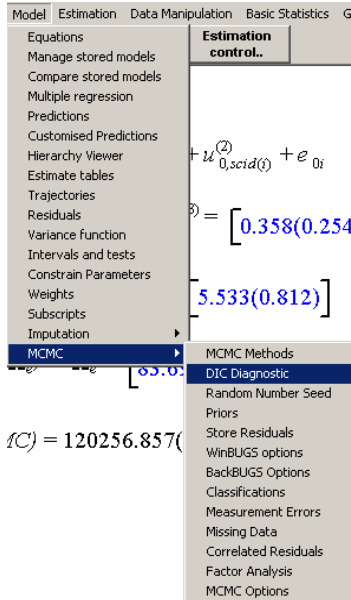
Here the DIC is 120419.9. We use this DIC to compare the fit across cross-classified models. A lower value of the DIC indicates better model fit. We will come back to this value later to compare the null random-intercepts cross-classified model to a random-intercepts model with more predictors.

To add more predictors to the model, we need to switch the estimation back to IGLS from MCMC by selecting *IGLS* from the **Estimation** menu.
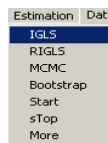
Click the "**Done**" button on the Estimation control window to use the default options in IGLS.
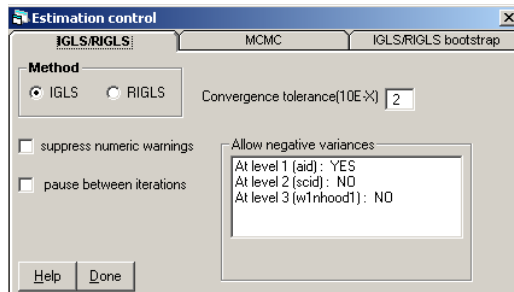
Now more predictors can be added to the model by clicking the "**Add Term**" button on the bottom toolbar.
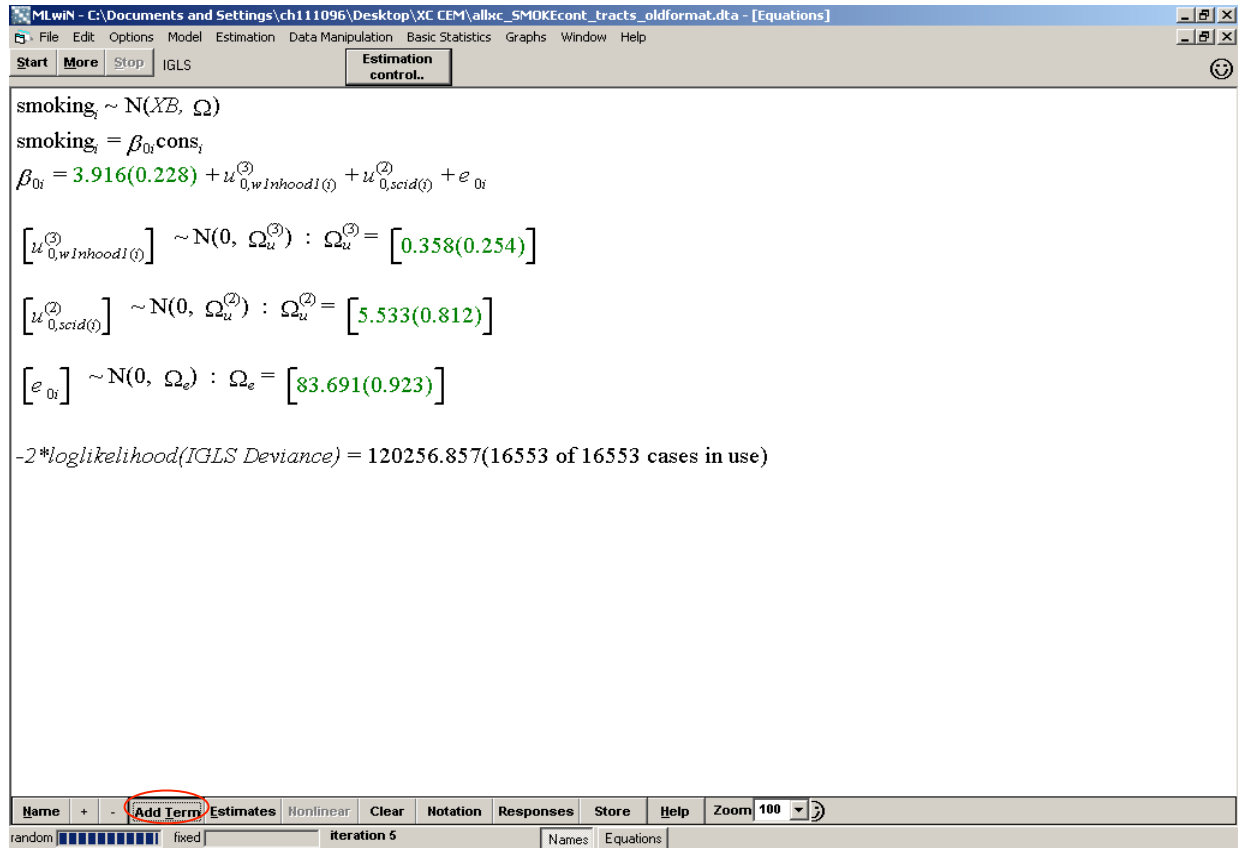
Choose the variable name from the 'variable' drop-down menu in the Specify term window to indicate the term you want to add to the model. Depending on whether the variable is continuous or categorical, options to center the variable or choose the reference category will appear. With a continuous variable, choose whether values should be uncentered or centered, and if so, around what value. Click the "Done" button when finished choosing options for the variable. Here uncentered *age* is added to the model.
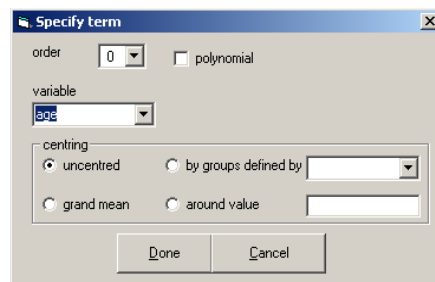
The reference category can be specified for categorical variables. Here the variable for sex called *female* is added to the model with males as the reference category.
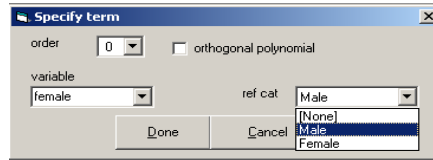
Once all variables have been added to the model, each variable can be modeled as a fixed effect or random effect, which would allow the slopes to vary at different levels. The default is to treat all predictors as fixed parameters. By double-clicking on the variable name, a random slope can be fitted for the variable(s). If two variables with the same formatting/labeling (e.g. "Yes"/"No") are added to a model, an error message stating "Category name clash, using extended names" will appear. This can be fixed by using a more descriptive name for one of the categories.

By clicking "**Ok,**" the variable will be added to the model. The full name of the variable will appear. For categorical variables, the full variable name of the variable will appear, followed by a colon and the non-referent category. In this example the variables *publicassist* and *highschool* were both coded "Yes" and "No" with "No" as the reference category.
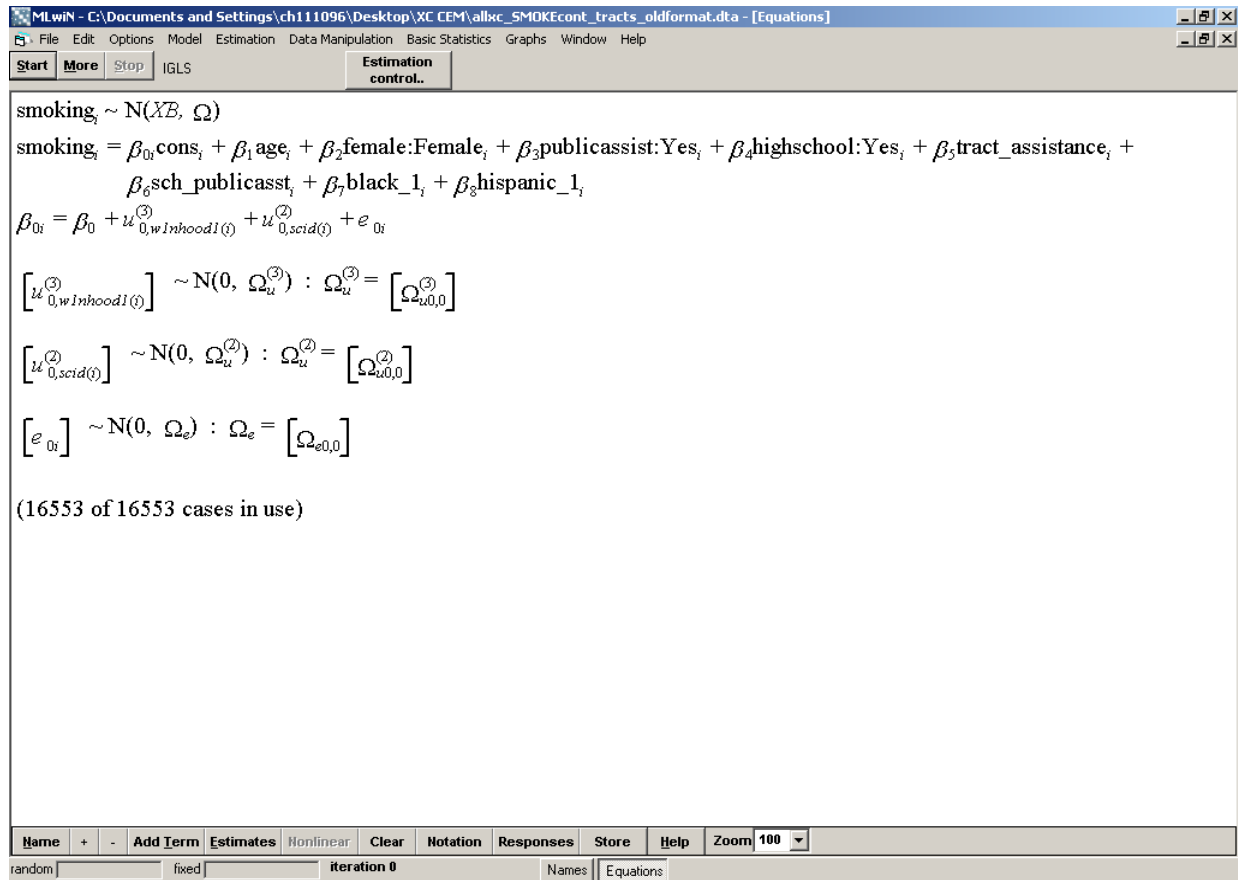
To fit a random slope for a specific level, check the box next to that level identifier. In these models, we are only modeling a random intercept, not random slopes.
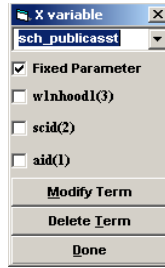
Once you are finished adding variables to the model, fit the model in IGLS to get the estimates assuming a nested structure.

$\text{smoking}_i \sim N(XB, \ \Omega)$

$\text{smoking}_i = \beta_{0i}\text{cons}_i + 0.841(0.044)\text{age}_i + 0.014(0.142)\text{female:Female}_i + 0.789(0.257)\text{publicassist:Yes}_i +$

$\qquad -0.066(0.233)\text{highschool:Yes}_i + -0.313(1.309)\text{tract\_assistance}_i + 0.021(0.012)\text{sch\_publicasst}_i +$

$\qquad -4.687(0.212)\text{black\_1}_i + -3.030(0.229)\text{hispanic\_1}_i$

$\beta_{0i} = -7.508(0.761) + u^{(3)}_{0,\text{wlnhood1}(i)} + u^{(2)}_{0,\text{scid}(i)} + e_{0i}$

$\left[u^{(3)}_{0,\text{wlnhood1}(i)} \right] \sim N(0, \ \Omega^{(3)}_u) \ : \ \Omega^{(3)}_u = \left[0.250(0.452) \right]$

$\left[u^{(2)}_{0,\text{scid}(i)} \right] \sim N(0, \ \Omega^{(2)}_u) \ : \ \Omega^{(2)}_u = \left[1.569(0.537) \right]$

$\left[e_{0i} \right] \sim N(0, \ \Omega_e) \ : \ \Omega_e = \left[81.595(0.922) \right]$

$-2*loglikelihood(\text{IGLS Deviance}) = 120124.857(16553 \text{ of } 16553 \text{ cases in use})$

Run the model again using MCMC re-estimation to fit the model with a cross-classified structure with the starting values obtained from IGLS which assume a nested structure. As noted previously, ignore the Deviance statistic and instead use the DIC, which can be accessed from the **Model** menu, under *MCMC*, and "DIC").
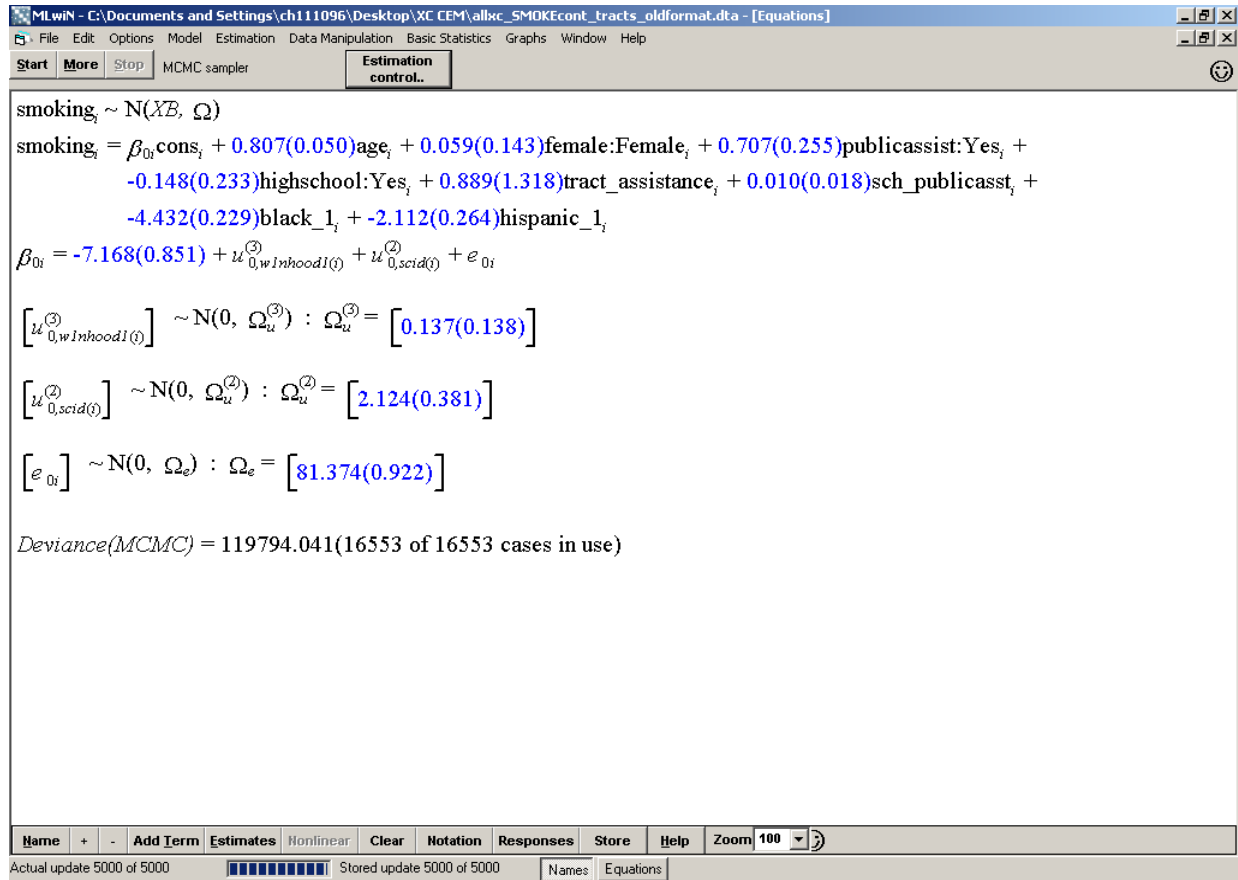
The DIC for the full model with all predictors is less than the null model. This indicates that the full model is a better fitting model.

```
->BDIC
Bayesian Deviance Information Criterion (DIC)
   Dbar    D(thetabar)      pD        DIC
119794.07   119670.46    123.60   119917.67
```

To check the assumptions of the linear model, we can look at the residuals which can be accessed from the **Model** menu under *Residuals*.
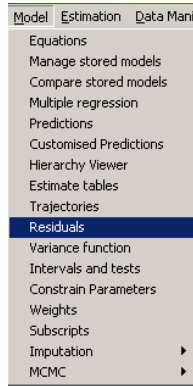
This will open the Residuals window. Residuals can be calculated and plotted for each level to look for issues with overall model fit as well as outliers and influential observations. The level on which to calculate residuals must be chosen first from the drop-down menu on the Settings tab.
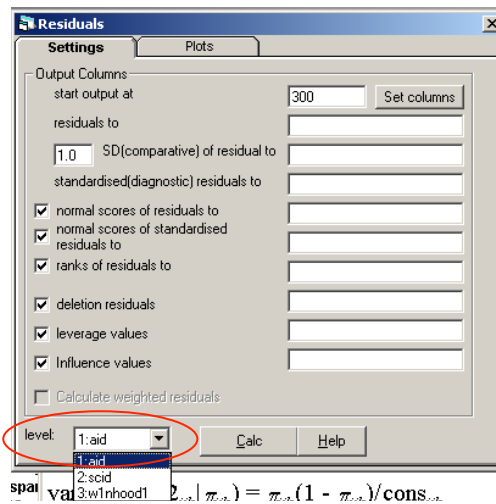
Here the residuals at the school-level are calculated by choosing the school identifier and clicking the "**Calc**" button.
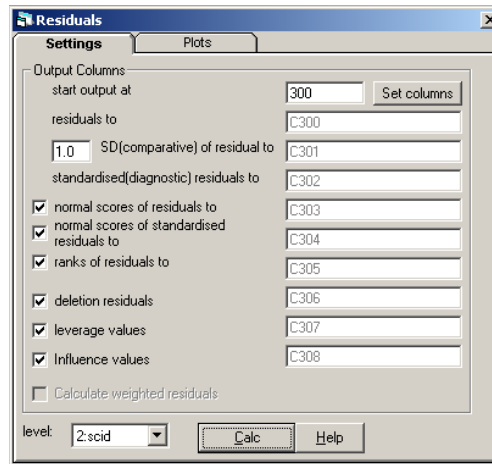
Once the residuals have been calculated at the desired level, different plots may be selected from the Plots tab. Here we will plot the school residuals by their rank on the x-axis. This will produce a plot with the smallest residuals on the left up to the largest residuals on the right. Once the desired plot is selected, click "**Apply**."
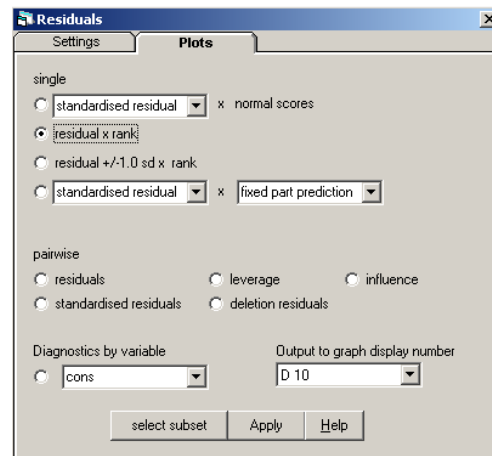
The following residual plot is produced for the school level Each triangle represents a different school. The residual plot indicates good model fit if there is an overall linear trend with residuals close to zero (indicating perfect correlation between model-predicted outcome and observed outcome for a given school). A slight s-shaped curve toward the minimum and maxium is common as there is usually sparse data. Observations with large residuals may be concerning. Here we may be concerned about the two outlying schools with large positive residuals indicating that the observed average smoking days for these schools were greater than what was predicted by the model. By double-clicking on one of these extreme observations, we can determine which school it represents and highlight it in the graph or remove it from the model altogether.
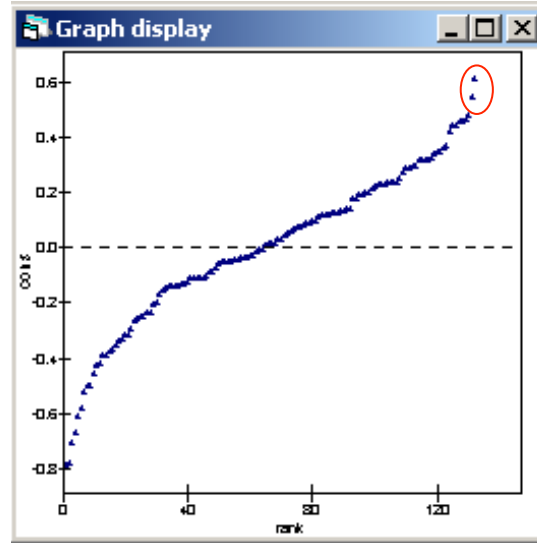
Here we see that the largest positive residual comes from school 31 (SCID=31) which is the 32nd school in the model (j=32 because SCID starts at 0). By making changes to the "In graphs" inset, we can leave this school in the residual plot, leave it out entirely, or highlight it using a different color if desired. Additionally, we could choose to leave this outlying school out of the model or model this school as a dummy variable to try to capture the excess variability contributed by this school. For more information on model diagnostics and residuals in MLwiN, see Chapter 15 of the manual for general information on diagnostics (Rasbash, et al. 2012), and Chapter 15 of the MCMC manual for information specific to cross-classified models (Browne, 2012).
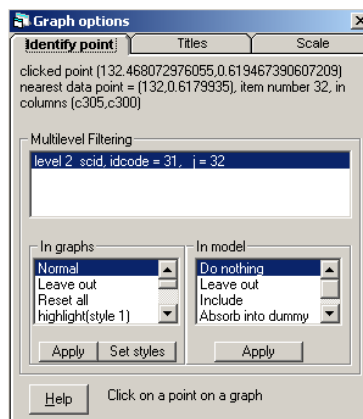
Specifying the Model: Example with a Binary Outcome

Fitting a model with a binary outcome is very similar to fitting a model with a continuous outcome in MLwiN. The major difference is that the distribution function will change. There are a few minor differences between fitting the binary model that are highlighted here. Specifying the dependent variable, running the model, and adding predictors are practically identical to fitting a model with a continuous outcome. As before, the data must be sorted by level using the *Sort* function (Click Remove All before starting the Sort) and any categorical variables should be classified as such using the "**Toggle Categorical**" button in the Names window. Additionally, a constant variable (vector of 1's for all observations) is also necessary when fitting a binary model and random intercepts. **Note**: This appendix provides instructions for modeling binary data. For modeling proportions at a higher level, see the MLwiN User Manual.

The model with a binary outcome can again be specified from the Equations window (if there is an existing equation you can click on terms; a menu will then appear, allowing you to 'delete term'. The variable *smoking2* (0=non-smoker, 1=smoker) is specified as the dependent variable for the binary model. The number of levels is chosen and the level identifiers are selected from the drop-down menus.
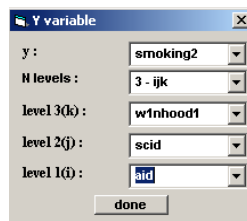
The dependent variable distribution can be changed from the Response Type window by double-clicking the *N* in the Equations window. For the binary model, the distribution function is binomial with a logit link function.
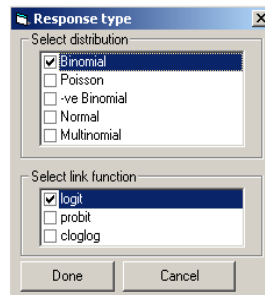
Now instead of the *N* for normal distribution, we see the dependent variable is modeled using a binomial distribution. This indicates that the binary smoking variable follows a binomial distribution with parameters n and π at each of the three levels. The red *n* indicates the denominator, which is a 1 for all observations in the case of binary data (all individuals can take on only one of two values). If the data had been binomial (proportions) the denominator would instead be the total number on which the proportion is based (e.g. modeling proportion of smokers at the school level the denominator would be the total number of students in each school). Double-clicking the red *n* makes the the Specify Denominator window appear.
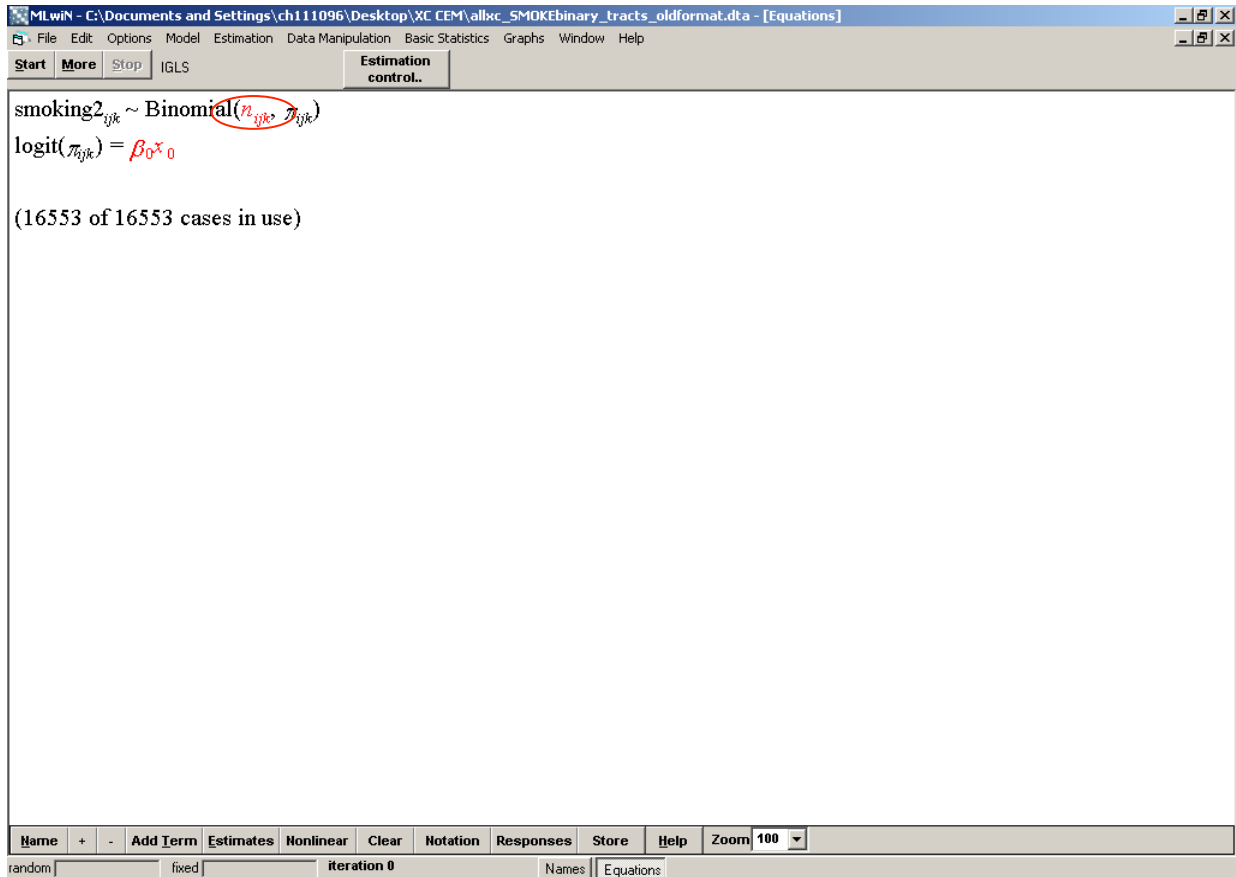
Since we already have a variable which is a 1 for each observation (*cons*), we can specify this variable as the denominator as well.
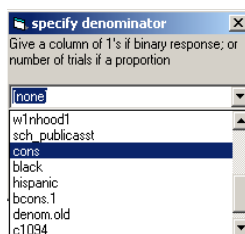
Clicking the "**Nonlinear**" button on the bottom toolbar opens the Nonlinear Estimation window where the default options for assumptions (e.g., binomial, extra binomial), linearization (e.g., first order, second order), and estimation (e.g., MQL, PQL) can be changed.
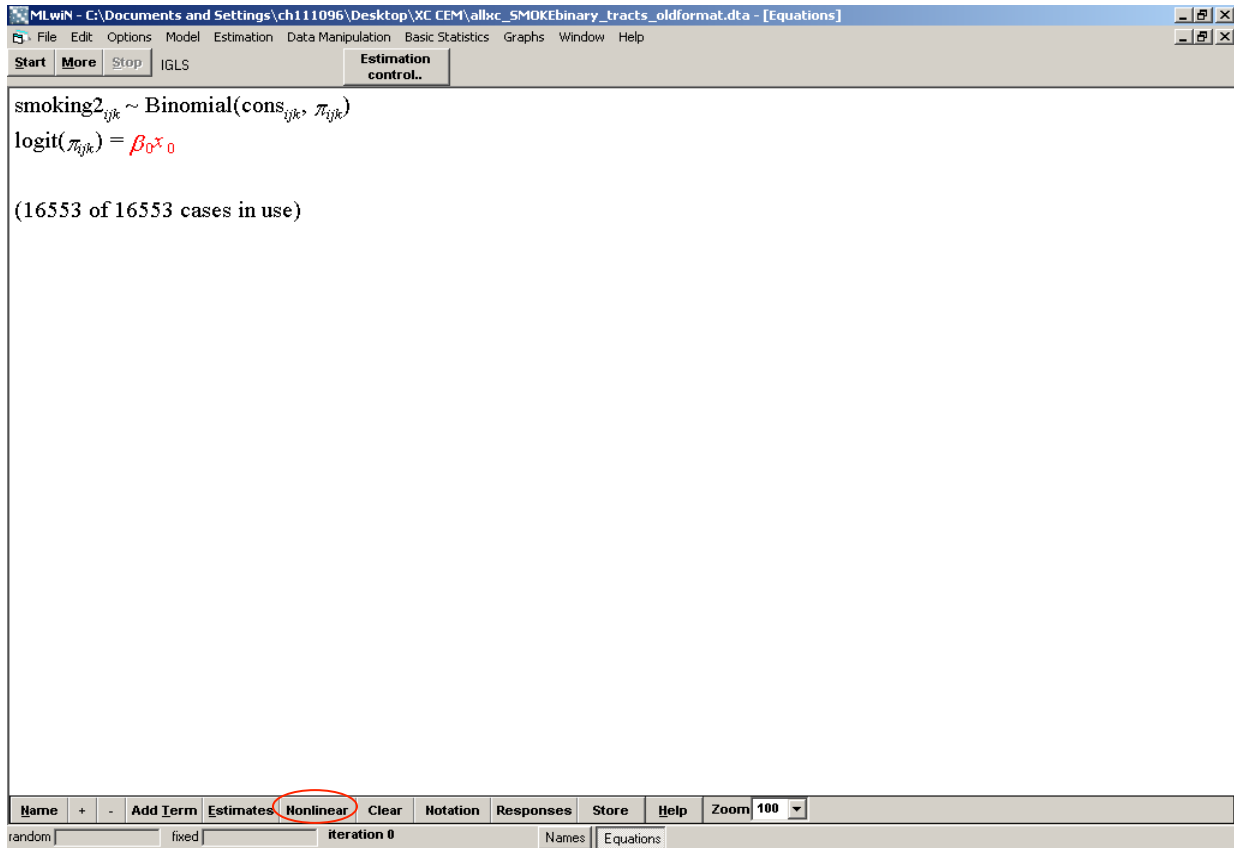
The default distributional assumption is binomial with a 1st order linearization and Marginal Quasi-Likelihood (MQL) estimation. The default settings are sufficient for this model. Once any necessary changes are made, click the "**Done**" button. In this model, we are using the default nonlinear estimation methods. For more information on the other methods, see the MLwiN MCMC manual, Chapter 10 (Browne, 2012).
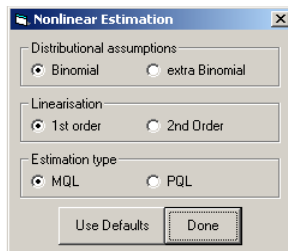
To add a random intercept to the model, double click on the red β_0x_0.
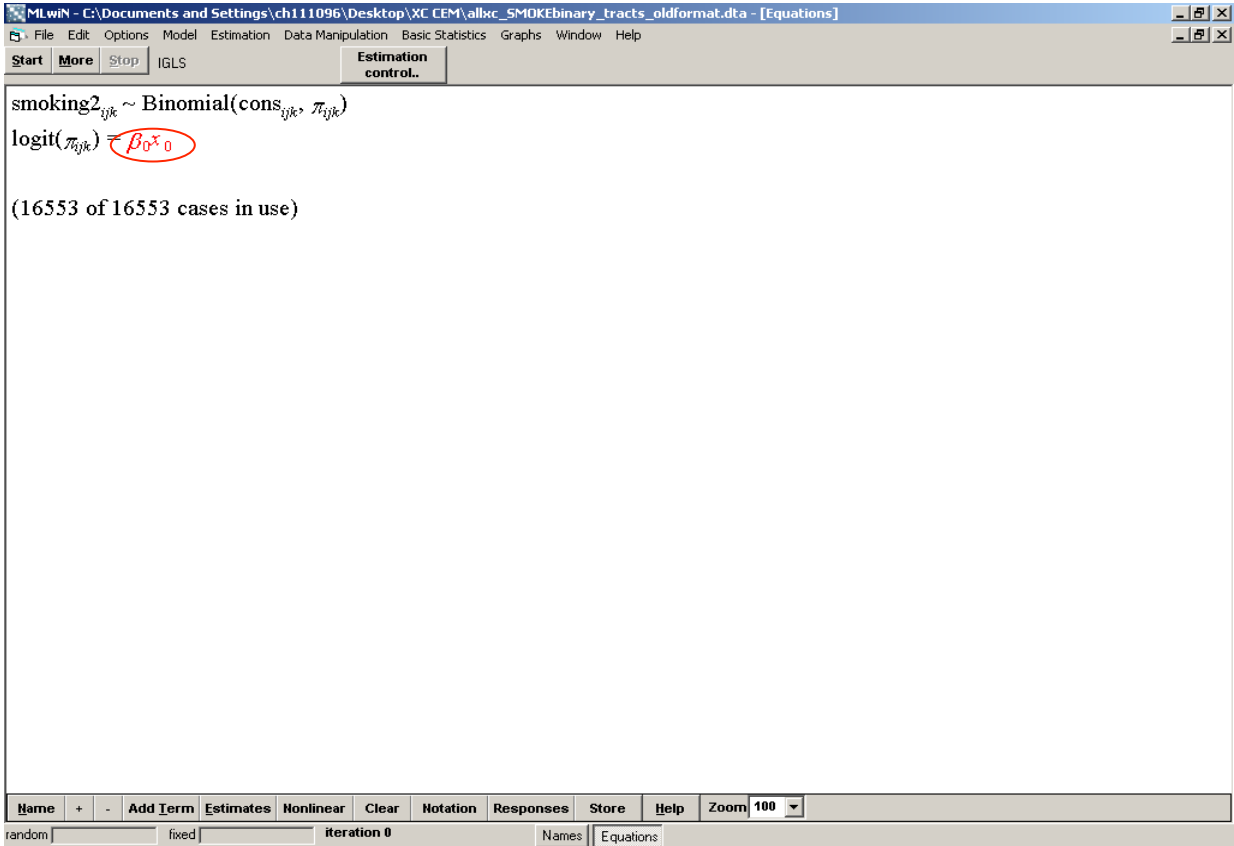
The constant variable is chosen from the drop-down menu as both a fixed effect and a random effect at all three levels. This allows the model to be fit with a random intercept at each level.

Click the '**Start**" button to fit the model in IGLS. **Note**: Just like with the models estimating a continuous outcome, fitting the model in IGLS assumes a nested structure. These parameter estimates are not interpretable. We will need to refit the model using MCMC re-estimation to fit a cross-classified model.
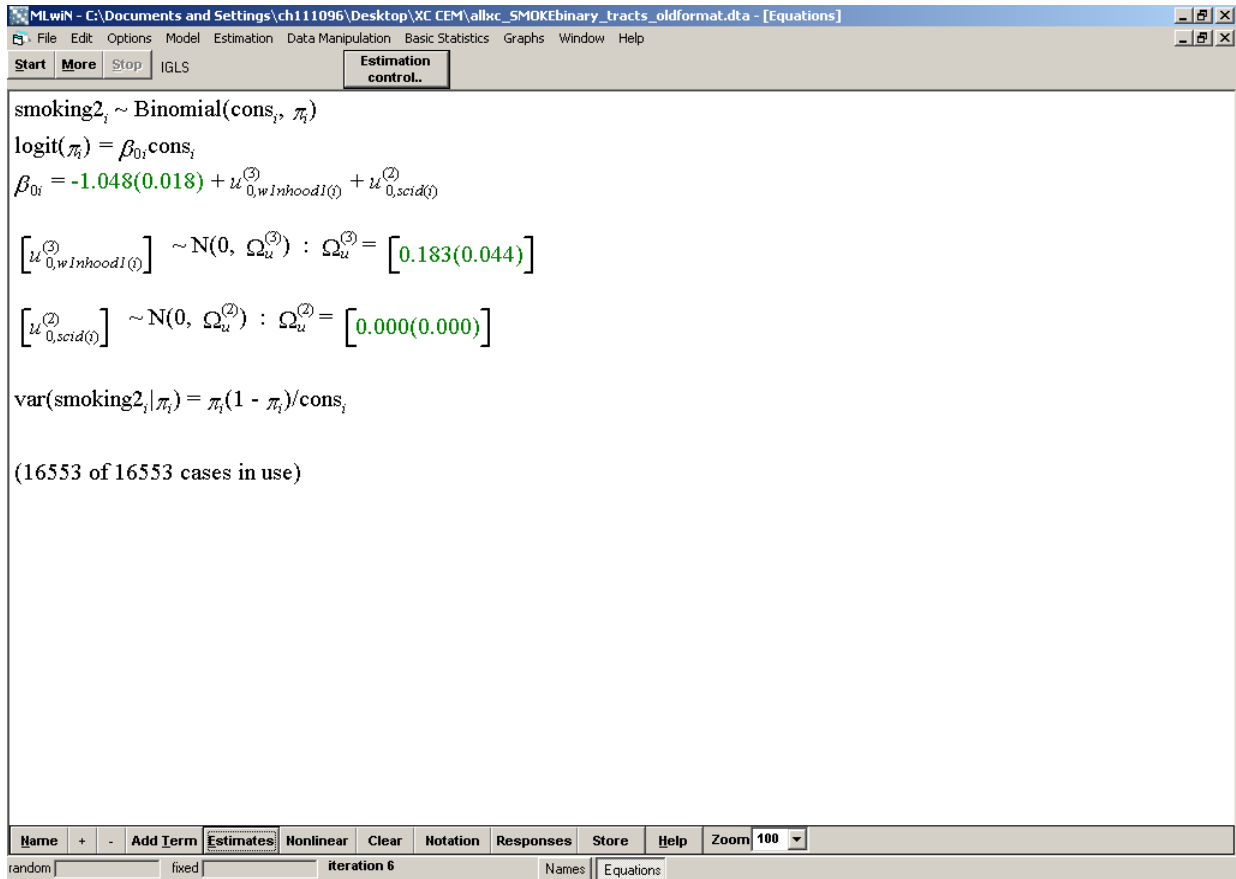
To switch to MCMC estimation, choose *MCMC* from the **Estimation** menu. Make any changes to the burn in length and iteration control in the Estimation Control window and click "**Done**."
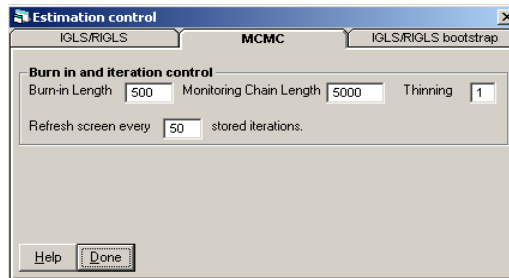
Change the model structure to treat the levels as cross-classified by accessing the Classifications Information by choosing *Classifications* from the **Model** menu. Check the box "Treat levels as cross-classified" in the Classification Information window to run the model with a cross-classified structure.
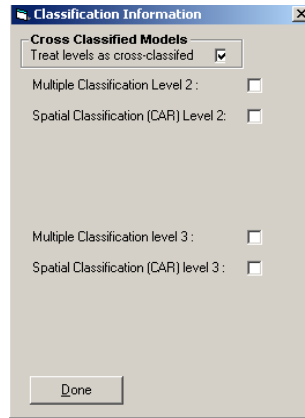
The model can be re-run with the cross-classified structure in MCMC by clicking the "**Start**" button. In a model with a binary outcome, individual-level variance is no longer computed. The estimates given are the parameter estimate (standard error) and variance estimate (standard error) in brackets. The estimates are now interpretable because the levels have been treated as cross-classified. Odds ratios can be computed by exponentiating the parameter estimates. The Deviance statistic is an inadequate measure of model fit for cross-classified models and should be ignored.
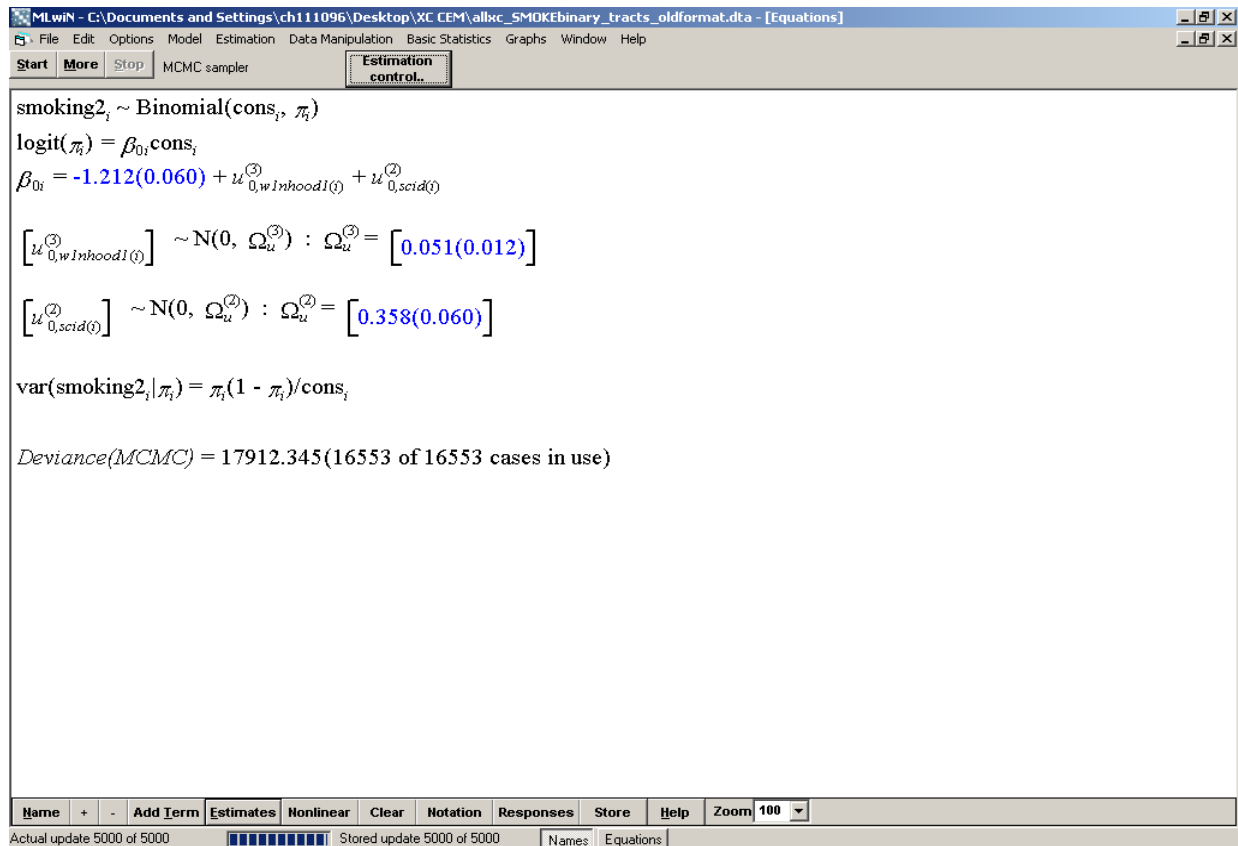
We can request the DIC again from the *MCMC* drop-down menu in the **Model** menu. The DIC for the null random-intercepts binary model is 18111.5.

```
->BDIC
Bayesian Deviance Information Criterion (DIC)
   Dbar    D(thetabar)     pD        DIC
17912.35   17711.23    201.12 18113.46
```

To add more predictors to the model, we again return to IGLS by selecting *IGLS* from the **Estimation** menu. This will make the Estimation Control window appear. You can then make changes to the estimation. To keep defaults, click the "**Done**" button.
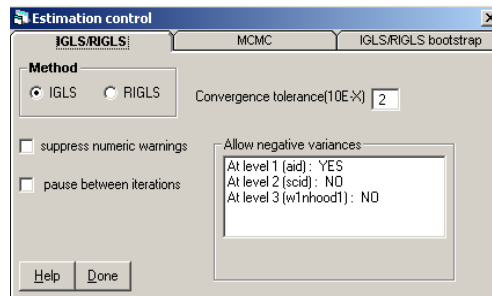
Click the "**Add Term**" button in the bottom toolbar to specify variables to add to the model.

Choose the variable to add from the drop-down menu and if continuous, whether to center and around which value.
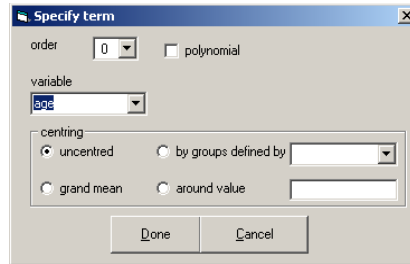
If the predictor being added to the model is categorical, choose the reference category.
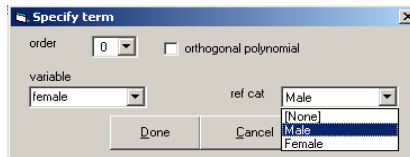
Fit the predictors as either fixed effects or random effects by double-clicking on each variable name and choosing the levels at which to model random slopes. The default is to treat all variables as fixed effects. Once any desired changes have been made, run the model by clicking "**Start**."

smoking2$_i$ ~ Binomial(cons$_i$, π_i)

logit(π_i) = β_{0i}cons$_i$ + 0.193(0.011)age$_i$ + 0.005(0.037)female:Female$_i$ + 0.349(0.067)publicassist:Yes$_i$ + 0.017(0.062)highschool:Yes$_i$ + -1.348(0.059)black_1$_i$ + -0.761(0.057)hispanic_1$_i$ + 0.005(0.003)sch_publicasst$_i$ + -0.196(0.330)tract_assistance$_i$

β_{0i} = -3.776(0.196) + $u^{(3)}_{0,wlnhood1(i)}$ + $u^{(2)}_{0,scid(i)}$

$\left[u^{(3)}_{0,wlnhood1(i)} \right]$ ~ N(0, $\Omega^{(3)}_u$) : $\Omega^{(3)}_u$ = $\left[0.091(0.040) \right]$

$\left[u^{(2)}_{0,scid(i)} \right]$ ~ N(0, $\Omega^{(2)}_u$) : $\Omega^{(2)}_u$ = $\left[0.000(0.000) \right]$

var(smoking2$_i$|π_i) = π_i(1 - π_i)/cons$_i$

(16553 of 16553 cases in use)

To run the full model treating the levels as cross-classified, choose *MCMC* from the **Estimation** menu. Make any changes to the burn in length and iteration control in the Estimation Control window and click "**Done.**" Then click the "**Start**" button on the Equations window to re-estimate the model.
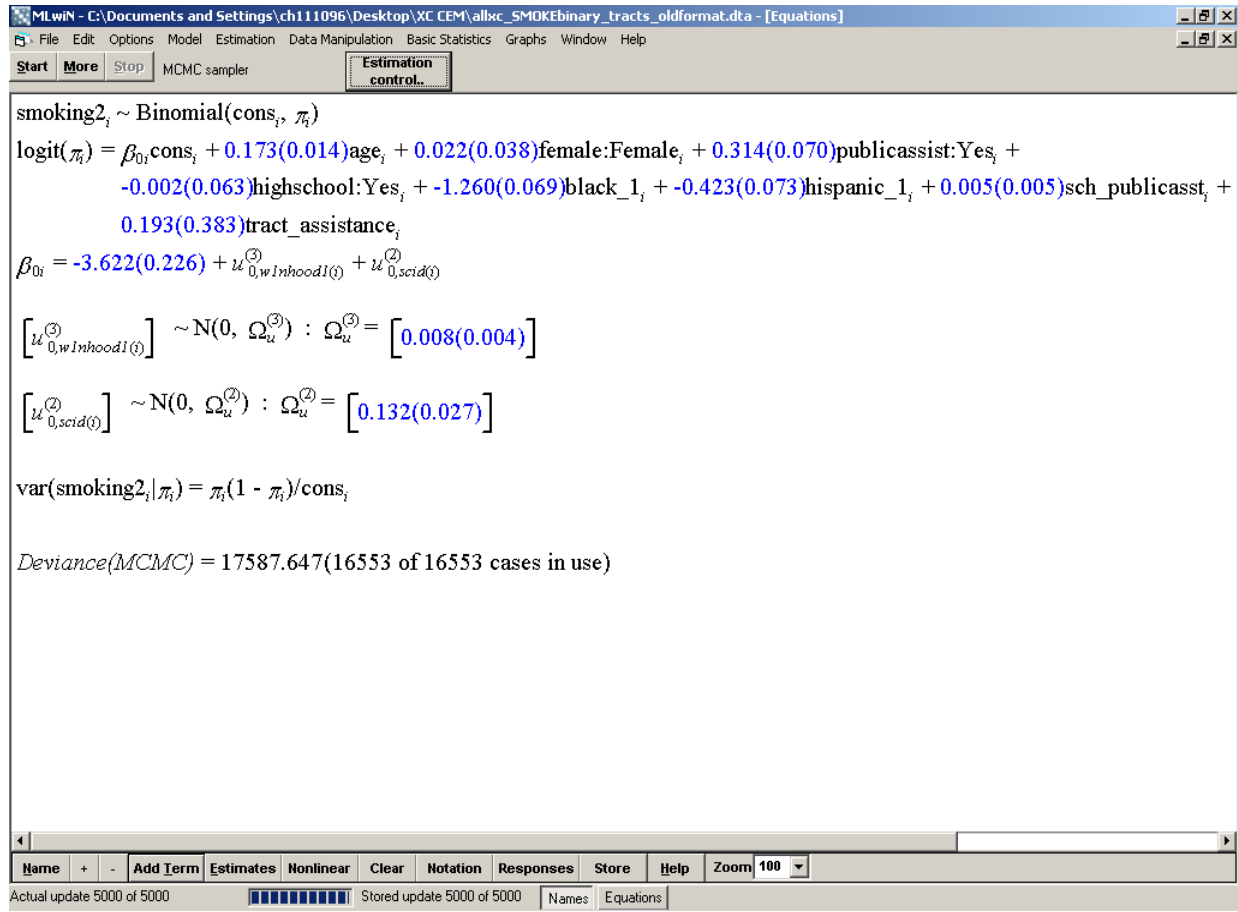
The DIC for the full binary model is lower than the null model indicating that the full model is a better fit.

```
->BDIC
Bayesian Deviance Information Criterion (DIC)
   Dbar     D(thetabar)      pD       DIC
 17587.65   17474.16      113.48 17701.13
```

References

Software

Browne, W.J. (2009) MCMC Estimation in MLwiN v2.1. Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., Charlton, C., Browne, W.J., Healy, M. and Cameron, B. (2012) *MLwiN Version 2.26*. Centre for Multilevel Modelling, University of Bristol.

Manuals

Browne, W.J. (2012) *MCMC Estimation in MLwiN, v2.26*. Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., Steele, F., Browne, W.J. and Goldstein, H. (2012) *A User's Guide to MLwiN, v2.26*. Centre for Multilevel Modelling, University of Bristol.

Subramanian, S. V. and Jones, K. (2010) *Multilevel Statistical Models: Concepts and Applications*. Boston, MA: Harvard School of Public Health/Bristol, UK: Centre for Multileve Modelling, University of Bristol.

Technical Appendix Part 2:
Running Cross-Classified Multilevel Models in MLwiN through STATA using the runmlwin package

**If you used this appendix for your analysis, please cite us:

Dunn, E.C., Richmond, T.K., Milliren, C.E., & Subramanian, S.V. Using Cross-Classified Multilevel Models to Disentangle School and Neighborhood Effects: An Example Focusing on Smoking Behaviors among Adolescents in the United States. *Health and Place*.

Introduction

This second technical appendix is intended to show users how to fit cross-classified multilevel models in MLwiN via STATA. With both STATA and MLwiN installed, MLwiN can be executed through STATA, which is convenient for data management as well as looking at and interpreting output. The following tutorial uses MLwiN version 2.29 and STATA version 13.1 (College Station, TX). Additional details are also available here:

> Leckie, G, C Charlton. Runmlwin: A program to run the MLwiN multilevel modeling software from within Stata. 2012. Journal of Statistical Software 52(11): 1-40.

The runmlwin package must first be installed from the Statistical Software Components (SSC) archive, which is a repository of user contributed STATA commands. The following accesses the SSC and installs the runmlwin command.

```
ssc install runmlwin
```

The filepath where MLwiN is located must be specified so that STATA can find the program to execute any runmlwin commands. The following command specifies the filepath where MLwiN is located on the user's computer (users should substitute their own filepath).

```
global MLwiN_path "C:\Program Files\MLwiN v2.29\i386\mlwin.exe"
```

Linear Models
The following command fits a two-level hierarchical (i.e., multilevel) linear null model with random intercepts for school and individual predicting the number of days smoked in the past 30. The data must first be sorted by the level identifiers just as it would be sorted within MLwiN. Additionally, a constant variable (here called *cons*) with a value of 1 for every observation must first be constructed in order to fit the intercept. The level 2 school identifier is *scid* and level 1 individual identifier is *aid*.

```
sort scid aid
runmlwin smoking cons, level2(scid:cons) level1(aid:cons) nopause
```

The following output generated by the above command appears in the STATA results window after MLwiN opens, runs, and closes.

```
MLwiN 2.29 multilevel model                    Number of obs    =    16070
Normal response model
Estimation algorithm: IGLS
```

| | No. of | Observations per Group | | |
Level Variable	Groups	Minimum	Average	Maximum
scid	128	18	125.5	1012

```
Run time (seconds)   =         5.94
Number of iterations =            3
Log likelihood       =  -58446.645
Deviance             =   116893.29
```

smoking	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
cons	3.825463	.220559	17.34	0.000	3.393175	4.257751

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
Level 2: scid				
var(cons)	5.330074	.7726001	3.815806	6.844343
Level 1: aid				
var(cons)	83.08062	.9305611	81.25675	84.90449

The output gives information about the number of level 2 groups, model run time, iterations, and fit statistics as well as the parameter estimate, test statistic and 95% confidence interval for the fixed effects (in this case just the intercept). Additionally, the random effect parameter estimates for school and individual variance in the intercept appear in the last output table.

To fit a cross-classified 3-level null model predicting number of smoking days with random intercepts for neighborhood, school, and individual the model must first be fit using the iterative generalized least squares (IGLS) algorithm with a hierarchical structure and re-run in the Bayesian Markov Chain Monte Carlo (MCMC) framework in order to properly account for the cross-classified data structure. The parameter estimates from the naïve IGLS model are used as starting values or priors for the MCMC cross-classified model. Again, the data should first be sorted by the level identifiers. The level 3 neighborhood identifier is *w1nhood1*, level 2 school identifier is *scid*, and level 1 individual identifier is *aid*.

The first model is quietly run (**quietly runmlwin**) and fits the data using a hierarchical structure. The second model is fit using the cross-classified structure in a Bayesian framework using MCMC. The option **mcmc(cc)** indicates that the data are cross-classified (cc) and the model should be fit using MCMC while the option **initsprevious** indicates that the parameter estimates from the first model (assuming hierarchical structure) should be used as starting values for the second model. The output from the first model will not appear in the output window if the **quietly** option is used.

```
sort w1nhood1 scid aid
quietly runmlwin smoking cons, level3(w1nhood1:cons) level2(scid:cons)
level1(aid:cons) nopause

runmlwin smoking cons, level3(w1nhood1:cons) level2(scid:cons)
level1(aid:cons) mcmc(cc) initsprevious nopause
```

After submitting the above command, the cross-classified results on the following page appear in the STATA results window.

The MCMC cross-classified model output gives information on the number of observations nested within each level, how long the model was run and for how many iterations (burnin, chain, thinning, etc.) and model fit statistics (deviance, DIC) in addition to the parameter estimates for the fixed and random effects at each level along with 95% credible intervals (Bayesian confidence interval).

```
MLwiN 2.29 multilevel model                        Number of obs      =      16070
Normal response model
Estimation algorithm: MCMC
```

Level Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
w1nhood1	2111	1	7.6	260
scid	128	18	125.5	1012

```
Burnin                       =          500
Chain                        =         5000
Thinning                     =            1
Run time (seconds)           =         20.6
Deviance (dbar)              =    116566.02
Deviance (thetabar)          =    116396.16
Effective no. of pars (pd)   =       169.87
Bayesian DIC                 =    116735.89
```

smoking	Mean	Std. Dev.	ESS	P	[95% Cred. Interval]	
cons	3.881897	.2225641	293	0.000	3.463005	4.319552

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred. Int]	
Level 3: w1nhood1					
var(cons)	.4572102	.1893233	14	.1262233	.8794164
Level 2: scid					
var(cons)	5.362256	.8018588	2527	3.952652	7.065619
Level 1: aid					
var(cons)	82.72698	.9336718	3899	80.93797	84.58237

Additional fixed effects can be added to the model by including them after the outcome variable specification. The following example adds additional predictors at the individual, school, and neighborhood levels first fitting a naïve hierarchical model in IGLS and then using the parameter estimates as starting values for the cross-classified MCMC model.

```
sort w1nhood1 scid aid
quietly runmlwin smoking age female publicassist highschool black
hispanic sch_publicasst sch_hsed sch_propwhite tract_assistance
tractprop_hs tractprop_white cons, level3(w1nhood1:cons)
level2(scid:cons) level1(aid:cons) nopause

runmlwin smoking age female publicassist highschool black hispanic
sch_publicasst sch_hsed sch_propwhite tract_assistance tractprop_hs
```

```
tractprop_white cons, level3(w1nhood1:cons) level2(scid:cons)
level1(aid:cons) mcmc(cc) initsprevious nopause
```

Output for the cross-classified model with additional fixed effects can be found on the following page.

```
MLwiN 2.29 multilevel model                    Number of obs      =      16070
Normal response model
Estimation algorithm: MCMC
```

Level Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
w1nhood1	2111	1	7.6	260
scid	128	18	125.5	1012

```
Burnin                          =          500
Chain                           =         5000
Thinning                        =            1
Run time (seconds)              =         34.5
Deviance (dbar)                 =    116131.02
Deviance (thetabar)             =    115998.01
Effective no. of pars (pd)      =       133.01
Bayesian DIC                    =    116264.02
```

smoking	Mean	Std. Dev.	ESS	P	[95% Cred. Interval]	
age	.8099638	.0505673	2956	0.000	.7131925	.9083905
female	.0631583	.1415275	4474	0.333	-.2096058	.3383548
publicassist	.6808348	.2573457	5083	0.003	.1830963	1.18176
highschool	-.2111972	.2340137	5037	0.180	-.6790206	.2505704
black	-4.112608	.2591757	4401	0.000	-4.630536	-3.612959
hispanic	-1.835435	.2730815	4027	0.000	-2.364895	-1.316999
sch_public~t	.0680946	.0251765	861	0.004	.0184816	.1170105
sch_hsed	-.0431506	.0222876	888	0.026	-.0859315	.0008327
sch_propwh~e	.0184719	.008447	885	0.014	.001701	.0349013
tract_assi~e	-.0049191	.0192253	2964	0.406	-.0425568	.0325781
tractprop_hs	.0107189	.0101577	2632	0.144	-.0089962	.0307047
tractprop_~e	.0002768	.0055375	2564	0.485	-.0105045	.010997
cons	-8.445673	1.036984	1900	0.000	-10.43999	-6.442746

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred. Int]	
Level 3: w1nhood1					
var(cons)	.2507093	.1310143	14	.0556763	.5244083
Level 2: scid					
var(cons)	1.719584	.3491465	1228	1.123541	2.485247
Level 1: aid					
var(cons)	80.53305	.9046208	4257	78.81626	82.29888

Logistic Models

Fitting cross-classified logistic models with a binary outcome is very similar to the linear models outlined above. The distribution and link function must be specified as binomial and logit, respectively instead of the default normal distribution used in the linear models (NOTE: MLwiN has other distribution and link functions not outlined in this tutorial). An additional constant variable for the denominator must be constructed. In the following example the denominator variable is called *denom* and has a value of 1 for all observations.

The following command fits a hierarchical null model with school as level 2 and individual as level 1 predicting smoking status (*smoking2* coded 0=non-smoker, 1=smoker). A random intercept is fitted for school. The **discrete** option specifies that the outcome variable is discrete and not continuous and that the distribution function is binomial (**dist(binomial)**) with a logit link function (**link(logit)**). The **denom** option specifies the variable that should be used to calculate the denominators for the higher level units which is called *denom*.

```
sort scid aid
runmlwin smoking2 cons, level2(scid:cons) level1(aid:)
discrete(dist(binomial) link(logit) denom(denom)) nopause
```

Output from the above command is shown below. The output gives information on the number of lower level units nested in the higher level units, run time and number of iterations as well as parameter estimates and 95% confidence intervals for the fixed and random effects. With logistic models, there is no variance parameter output for the lowest level (in this case individual) because individual variance in a binary outcome is a function of the proportion of individuals who have the outcome.

```
MLwiN 2.29 multilevel model                    Number of obs    =     16070
Binomial logit response model
Estimation algorithm: IGLS, MQL1
```

Level Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
scid	128	18	125.5	1012

```
Run time (seconds)   =      4.95
Number of iterations =        5
```

smoking2	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
cons	-1.130177	.0517373	-21.84	0.000	-1.23158	-1.028774

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
Level 2: scid				
var(cons)	.2845968	.0425854	.2011309	.3680626

Next a 3-level cross-classified null model with random intercepts for school and neighborhood is fit. Just as with the linear models, the logistic models must first be run in IGLS assuming a hierarchical structure and then refit in MCMC to account for the cross-classified structure using the parameter estimates from the first model as the starting values for the Bayesian MCMC. Similar to the linear models, the naïve hierarchical model is quietly run to suppress the output followed by the cross-classified model.

```
sort w1nhood1 scid aid
```

```
quietly runmlwin smoking2 cons, level3(w1nhood1:cons)
level2(scid:cons) level1(aid:) discrete(dist(binomial) link(logit)
denom(denom)) nopause

runmlwin smoking2 cons, level3(w1nhood1:cons) level2(scid:cons)
level1(aid:) discrete(dist(binomial) link(logit) denom(denom))
mcmc(cc) initsprevious nopause
```

The output for the cross-classified null logistic model can be found below.

```
MLwiN 2.29 multilevel model                    Number of obs      =      16070
Binomial logit response model
Estimation algorithm: MCMC


                    No. of        Observations per Group
   Level Variable   Groups     Minimum     Average     Maximum

         w1nhood1     2111           1         7.6         260
             scid      128          18       125.5        1012


Burnin                          =          500
Chain                           =         5000
Thinning                        =            1
Run time (seconds)              =         63.6
Deviance (dbar)                 =     17331.04
Deviance (thetabar)             =     17140.24
Effective no. of pars (pd)      =       190.80
Bayesian DIC                    =     17521.84

      smoking2  |     Mean     Std. Dev.      ESS       P       [95% Cred. Interval]

          cons  |  -1.226344   .0564287        92    0.000     -1.337117    -1.120508


   Random-effects Parameters  |     Mean     Std. Dev.    ESS       [95% Cred. Int]

Level 3: w1nhood1
                   var(cons)  |   .0482382   .0168734      11     .0256405    .0872135

Level 2: scid
                   var(cons)  |   .3561811   .0605399     1012    .2525054    .4894393
```

More predictors can be added to the model by including additional variables after the outcome variable.
Here fixed effects are added at the individual, school, and neighborhood levels. The last line in the cross-
classified model requests that a table of odds ratios be output in addition to the table of logit parameter
estimates.

```
sort w1nhood1 scid aid
quietly runmlwin smoking2 age female publicassist highschool black
hispanic sch_publicasst sch_hsed sch_propwhite tract_assistance
tractprop_hs tractprop_white cons, level3(w1nhood1:cons)
level2(scid:cons) level1(aid:) discrete(dist(binomial) link(logit)
denom(denom)) nopause
```

```
runmlwin smoking2 age female publicassist highschool black hispanic
sch_publicasst sch_hsed sch_propwhite tract_assistance tractprop_hs
tractprop_white cons, level3(w1nhood1:cons) level2(scid:cons)
level1(aid:) discrete(dist(binomial) link(logit) denom(denom))
mcmc(cc) initsprevious nopause
runmlwin, noheader noretable or
```

The output for the model with additional fixed effects can be found on the following page. Using the odds ratio option will give two tables of parameter estimates, one of the logits and one with odds ratios.

```
MLwiN 2.29 multilevel model                         Number of obs    =      16070
Binomial logit response model
Estimation algorithm: MCMC
```

Level Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
w1nhood1	2111	1	7.6	260
scid	128	18	125.5	1012

```
Burnin                         =          500
Chain                          =         5000
Thinning                       =            1
Run time (seconds)             =          153
Deviance (dbar)                =     17004.39
Deviance (thetabar)            =     16875.57
Effective no. of pars (pd)     =       128.82
Bayesian DIC                   =     17133.21
```

smoking2	Mean	Std. Dev.	ESS	P	[95% Cred. Interval]	
age	.1833469	.0123686	8	0.000	.1554669	.2065977
female	.0099364	.0381066	577	0.398	-.0617398	.0852968
publicassist	.3061968	.0719862	628	0.000	.1652309	.4457451
highschool	-.0023873	.0615281	85	0.479	-.1260161	.1207112
black	-1.184582	.0784246	175	0.000	-1.341672	-1.040694
hispanic	-.3479473	.0710086	313	0.000	-.4930292	-.2129315
sch_public~t	.0235929	.006303	48	0.000	.0121545	.0363701
sch_hsed	-.0129069	.0056384	45	0.012	-.0234648	-.0011173
sch_propwh~e	.0065937	.0021187	33	0.000	.0026996	.0109378
tract_assi~e	-.006447	.0058402	48	0.144	-.0167468	.0049644
tractprop_hs	.0049704	.0028578	46	0.039	-.0003789	.0107634
tractprop_~e	-.0012386	.0016749	19	0.250	-.0052289	.0017178
cons	-4.126022	.2606254	6	0.000	-4.566837	-3.59991

Random-effects Parameters	Mean	Std. Dev.	ESS	[95% Cred. Int]	
Level 3: w1nhood1					
var(cons)	.020705	.0072227	9	.0103887	.0371728
Level 2: scid					
var(cons)	.0988218	.0236526	296	.0589949	.1508862

```
. runmlwin, noheader noretable or
```

smoking2	Odds Ratio	Std. Dev.	ESS	P	[95% Cred. Interval]	
age	1.201323	.014823	8	0.000	1.168203	1.229488
female	1.01072	.0385453	578	0.398	.9401275	1.08904
publicassist	1.361772	.0980714	631	0.000	1.179665	1.561653
highschool	.9995055	.061538	85	0.479	.8816007	1.128299
black	.3068144	.0240339	177	0.000	.2614083	.3532096
hispanic	.7079151	.0501854	320	0.000	.6107734	.8082115
sch_public~t	1.023894	.006457	48	0.000	1.012229	1.03704
sch_hsed	.9871917	.0055684	45	0.012	.9768084	.9988834
sch_propwh~e	1.006618	.0021332	33	0.000	1.002703	1.010998
tract_assi~e	.9935907	.0058048	48	0.144	.9833926	1.004977
tractprop_hs	1.004987	.0028725	46	0.039	.9996211	1.010822
tractprop_~e	.9987636	.0016721	19	0.250	.9947848	1.001719
cons	.0167114	.0044696	7	0.000	.0103908	.0273262