American Journal of Epidemiology

Vol. 168, No. 10
DOI: 10.1093/aje/kwn241
Advance Access publication October 9, 2008

Original Contribution

Estimating Co-Occurring Behavioral Trajectories Within a Neighborhood Context: A Case Study of Multivariate Transition Models for Clustered Data

Magdalena Cerdá, Brisa N. Sánchez, Sandro Galea, Melissa Tracy, and Stephen L. Buka

Initially submitted July 27, 2007; accepted for publication July 14, 2008.

Comorbidity is well-documented in psychiatric and risk behavior epidemiology. The authors present a novel application of clustered multivariate transition models to study comorbidity within a clustered context. The authors used data from the Project on Human Development in Chicago Neighborhoods (1995–2002) to assess trajectories in substance use, problems with police, and antisocial behavior among 1,517 participants in 80 neighborhoods followed from ages 12–15 years through ages 18–21 years. The authors used pairwise odds ratios to quantify behavior comorbidity at the individual and neighborhood levels. Risk behaviors co-occurred within individuals at specific points in time: antisocial behavior and substance use were 3.37 times more likely to co-occur within an individual at wave 1, as compared with the co-occurrence of any 2 behaviors from different individuals, while substance use and police problems were 2.94 times more likely to co-occur than substance use and antisocial behavior at wave 2. The authors also evaluated sequential comorbidity. Antisocial behavior was sequentially comorbid with substance use and police problems: 31% of youths who had reported antisocial behavior at baseline reported police problems or drug use at wave 2. These models can prove instrumental in answering the persistent questions about possible sequential relations among problem behaviors.

adolescent; comorbidity; logistic models; regression analysis; residence characteristics; risk-taking; social environment; transition probability

Abbreviations: CI, confidence interval; OR, odds ratio; POR, pairwise odds ratio.

Comorbidity is well-documented in psychiatric and risk behavior epidemiology (1–7). Psychiatric disorders (5, 8–13) and risk behaviors (14–16) appear in clusters (concurrent comorbidity) and are associated with each other over time (sequential comorbidity). In the National Comorbidity Survey Replication, Kessler et al. (17) found that 27.7% of respondents had had 2 or more psychiatric disorders in their lifetime. Risk behaviors cluster as well—young people with offending records, for example, also have problems in school and with social relationships and more often use illegal substances (18–21). Persons with comorbid psychiatric disorders have more severe psychiatric symptoms and a lower level of social competence than those with a single disorder (18, 19, 22, 23). Therefore, comorbidity is a central consideration for research on the epidemiology of psychiatric disorders and associated risk behaviors.

Prior studies on psychiatric comorbidity have applied a range of methods, from traditional regression models for estimating associations between different disorders (20, 21, 24) to multinomial logistic models that compare combinations of pairs of comorbid disorders (25) to latent growth models that jointly estimate trajectories of behavior clusters (26, 27). Few such approaches examine comorbidity among clusters, rather than pairs, of conditions, and none of the existing methods are able to account for the influence of factors beyond the individual on the generation of comorbid patterns (13, 28, 29). Concurrent and sequential comorbidity may depend not only on individual, family, and peer characteristics but also on the context in which individuals reside (28). For example, adolescents who live in more disadvantaged neighborhoods with harsher policing strategies may be more likely to progress from initial conduct problems to

Correspondence to Dr. Magdalena Cerdá, Center for Urban Epidemiologic Studies, New York Academy of Medicine, 1216 Fifth Avenue, New York, NY 10029 (e-mail: mcerda@nyam.org).

difficulties with the police than adolescents who live in less disadvantaged contexts (30, 31).

In this paper, we present clustered multivariate transition models as a method of studying concurrent and sequential comorbidity of multiple risk behaviors within a clustered context. With multivariate transition models, we can model longitudinal relations between comorbid risk behaviors and thus both assess the temporal order between risk behaviors and capture potentially reciprocal associations. These models also use alternating logistic regression to model the correlation structure between observations within individuals. Through clustering parameters, we can quantify the magnitude of concurrent clustering, or comorbidity, between risk behaviors at different time points. Joint modeling of behaviors also produces potentially more efficient estimates than modeling of the different responses separately (32). Accounting for group-level clustering, in this case within neighborhoods, allows us to incorporate group-level characteristics as potential predictors of the prevalence of behaviors as well as of their sequential comorbidity. The models are population-averaged, since estimating equations are used to fit the models instead of random-effects models.

These models build upon prior work with longitudinal hierarchical models (33), cross-sectional models for multivariate outcomes (34), bivariate transition models (32), and univariate transition models (35, 36). We apply this method to the investigation of 3 related youth problem behaviors: antisocial behavior, substance use, and reports of problems with the police.

MATERIALS AND METHODS

Data were obtained from the Project on Human Development in Chicago Neighborhoods, a multilevel, prospective study on child development. Children within 6 months of birth and children and youths aged 3, 6, 9, 12, 15, and 18 years living in Chicago, Illinois, neighborhoods were selected to participate in the study on the basis of a sampling design documented elsewhere (37). Three waves of measurement were conducted: baseline in-person interviews (wave 1) took place in 1995–1997, a follow-up interview took place in 1998–1999 (wave 2), and a final interview took place in 2000–2002 (wave 3). All protocols were approved by the institutional review board of the University of Michigan School of Public Health, and participants gave informed consent.

The analyses used data from the 12- and 15-year-old cohorts to examine behavioral transitions at 3 critical turning points: entry into high school, graduation from high school, and young adulthood. The baseline sample included 1,517 adolescents from 80 neighborhoods; wave 2 included 87% of the baseline sample, and wave 3 included 80%. Multiple imputation of missing observations for respondent variables was performed through the sequential regression imputation method, using IVEWARE software (38, 39). All available data on study variables were used to impute missing observations for all 3 waves. Imputation presented an advantage over the use of a complete case analysis, since it assumed that data were missing at random rather than missing completely at random and allowed us to use more data, thus increasing statistical power.

Problem behaviors

Three behaviors were measured at each wave: substance use, reports of problems with the police, and antisocial behavior. Participants were asked about substance use through the Substance Use Interview (40–42). Substance use was measured as any use of marijuana or another illicit drug reported during the year prior to assessment. Problems with the police were measured at each wave as a binary indicator: "Have you had any problems with the police in the past year?", taken from the Self-Report of Offending instrument (43). Finally, antisocial behavior was measured with the Child Behavior Checklist (44), the Youth Self-Report (45), and the Young Adult Self-Report (46). At baseline, primary caregivers rated participants' antisocial behavior using the Child Behavior Checklist. At wave 2, participants received a reduced version of the Youth Self-Report; and at wave 3, the older cohort received the reduced Young Adult Self-Report (47–49) (the younger cohort received the Youth Self-Report). Participants who scored at or above the top decile on the basis of published norm samples (44) were categorized as antisocial (50).

Individual and neighborhood covariates

Individual covariates considered included sex, cohort, race/ethnicity, socioeconomic status, and neighborhood concentrated disadvantage. Socioeconomic status was measured as the first principal component of parental education, parental occupation, and household income, standardized with a mean of 0 and a standard deviation of 1 (51). Socioeconomic status was measured at waves 1 and 2. Neighborhood concentrated disadvantage was derived from US Census data and was measured as the first principal component of the proportion of neighborhood residents who were below the poverty line, receiving public assistance, or unemployed (51). Data from the 1990 Census were used for wave 1, and data from the 2000 Census were used for waves 2 and 3.

Statistical models

The regression models used in this investigation model the effects of prior states of the behaviors reported above on the probability of reporting a behavior at the current wave, as well as the cross-sectional association among behaviors in any given wave. The models also account for clustering of individuals within neighborhoods.

General modeling framework. Let Y^r_{ijt} be a binary variable taking the value 1 if individual j in neighborhood i reported the behavior $r = 1, \ldots, R$ at wave $t = 1, \ldots, T$. For each of the R behaviors (in this example, $R = 3$), a marginal model for the probability of reporting behavior r at time $t = 2, \ldots, T$, given behaviors reported at wave $t - 1$, is given by

$$\operatorname{logit}\Pr(Y_{ijt}^{r}=1|\,Y_{ij(t-1)}^{1},\ \ldots,\ Y_{ij(t-1)}^{r},\ \ldots,\ Y_{ij(t-1)}^{R},\mathbf{x}_{ijt})$$
$$=\ \beta_{ort}+\mathbf{x}_{ijt}^{'}\boldsymbol{\beta}_{r}+\gamma_{r1t}Y_{ij(t-1)}^{1}$$
$$+\ \cdots\ +\ \gamma_{rrt}Y_{ij(t-1)}^{r}+\cdots$$
$$+\gamma_{rRt}Y_{ij(t-1)}^{R}, \tag{1.1}$$

where β_{ort} is the log odds of the probability of onset of behavior r at wave t for an individual with average covariates, assuming no other behaviors were present at the previous wave, and $\mathbf{x}'_{ijt}$ is a mean-centered covariate vector. Parameter $\gamma_{rr^{'}t}$ represents the effect on the log odds of the probability of reporting behavior r at wave t due to reporting behavior r' at the previous wave. The subscript t in $\gamma_{rr^{'}t}$ indicates that this effect can vary with wave. Equation 1.1 can be used to calculate the probability of onset of behavior r at wave t (i.e., $\Pr(Y_{ijt}^{r}=1|\,Y_{ij(t-1)}^{1},\ldots,Y_{ij(t-1)}^{r}=0,\ldots,Y_{ij(t-1)}^{R},\mathbf{x}_{ijt})$) or, similarly, its persistence and desistance. The equation represents a first-order transition model, because the probability of reporting the behavior at time t depends only on the responses at the previous time point. Larger-order models can be obtained by including the $k>1$ previous outcomes as part of the right-hand side of the equation. In addition, differential effects of previous behaviors for given levels of a covariate (individual- or cluster-level) may be modeled by including corresponding interaction terms.

We use pairwise odds ratios (PORs) to model the association between pairs of observations (Y_{ijt}^{r}, $Y_{ij^{'}t^{'}}^{r^{'}}$) in neighborhood i. The POR is defined as the odds that Y_{ijt}^{r} is 1 given that $Y_{ij^{'}t^{'}}^{r^{'}}$ is 1, divided by the odds that Y_{ijt}^{r} is 1 given that $Y_{ij^{'}t^{'}}^{r^{'}}$ is 0 (33). In a clustered setting, PORs do not enable us to partition the variance in the outcome explained by different levels of nesting (52), but they do provide information about the strength and structure of the association between observations (53). A model for how the POR depends on the characteristics shared by Y_{ijt}^{r}, $Y_{ij^{'}t^{'}}^{r^{'}}$ is given by

$$\begin{aligned}\operatorname{logPOR}(Y_{ijt}^{r},Y_{ij^{'}t^{'}}^{r^{'}})&=\alpha_0+\alpha_2 I(j=j^{'},t=t^{'}=2)\\&+\sum_{\substack{\text{behavior}\\\text{pairs}(m,l)}}\alpha_{2,ml}I\big(j=j^{'},t=t^{'}=2,r=m,r^{'}=l\big)\\&+\alpha_3 I\big(j=j^{'},t=t^{'}=3\big)\\&+\sum_{\substack{\text{behavior}\\\text{pairs}(m,l)}}\alpha_{3,ml}I\big(j=j^{'},t=t^{'}=3,r=m,r^{'}=l\big)\\&+\cdots+\alpha_T I\big(j=j^{'},t=t^{'}=T\big)\\&+\sum_{\substack{\text{behavior}\\\text{pairs}(m,l)}}\alpha_{T,ml}I\big(j=j^{'},t=t^{'}=T,r=m,r^{'}=l\big)\end{aligned} \tag{1.2}$$

where $I(a)$ takes the value 1 if statement a is true and 0 otherwise. In equation 1.2, α_0 quantifies the association between any 2 behaviors from different individuals within a neighborhood, where $\exp(\alpha_0)$ represents the POR; α_t, $t=2,\ldots,T$, determines the additional association between a reference pair of behaviors within an individual at wave t; and $\alpha_{t,ml}$ indicates differential association between behaviors m and l, in comparison with an excluded reference pair, at wave t. Thus, $\exp(\alpha_t)$ and $\exp(\alpha_{t,ml})$ represent multiplicative changes in the POR (ΔPOR). Together, equations 1.1 and 1.2 represent a multivariate transition model for clustered data; equation 1.2 can be extended or modified to include other types of clustering.

Important special cases of the model are the *univariate* transition model (54), which models the transitions of only 1 behavior at a time, and the *multivariate* model for cluster-correlated longitudinal data. The second is a combination of the model presented by Preisser et al. (33) for longitudinal clustered data and the model presented by Das et al. (34) for cross-sectional data with multivariate outcomes. These models are useful in model-building (see Appendix 1).

In our analysis, we used the alternating logistic regression algorithm in SAS PROC GENMOD (55) to fit all models. The POR models (e.g., equation 1.2) were specified by creating a data set with the ZDATA option on the REPEATED statement (33). The data set consisted of 1 line of "data" for each pair of observations $Y_{ijt}^{r}, Y_{ij^{'}t^{'}}^{r^{'}}$ within a neighborhood. Each line consists of variables that describe the relation between Y_{ijt}^{r}, $Y_{ij^{'}t^{'}}^{r^{'}}$, that is, equation 1.2 (see Appendix 2).

We employed the PROC MIANALYZE procedure in SAS to combine the model estimates and their standard errors from the 5 imputed data sets with standard approaches (56).

Modeling strategy. We first constructed a 3-wave multivariate model for cluster-correlated longitudinal data (model 1) to estimate the level of clustering between behavior pairs at each wave, as well as the differential relation between sociodemographic characteristics and each of the 3 behaviors. The correlation structure accounted for wave-specific associations between behavior pairs within individuals and clustering of any 2 observations from different individuals within neighborhoods.

We then modeled the association of behaviors over time by estimating their concurrent and sequential comorbidity. We fitted a population-averaged multivariate transition model with clustering parameters for the association of observations within neighborhoods and for the association between behaviors within persons (model 2). The model also allowed the states of the other 2 behaviors to influence the future state of a third behavior.

RESULTS

Table 1 presents the characteristics of the study participants across the 3 study waves. On average, participants were aged 13.5 years at baseline (54.1% of the participants were aged 12 years, while 45.9% were aged 15 years) and aged 18.2 years (range, 15.3–22.3 years) at wave 3. In terms of race/ethnicity, they were 14.4% non-Hispanic white, 36.8% non-Hispanic black, 44.9% Hispanic, and 3.8% other. At baseline, participants had lived an average of 6.7 years at their current address. Between waves 1 and 3, 20% of the participants were lost to follow-up. Those participants

Table 1. Observed Characteristics of Participants Before Multiple Imputation, by Wave, When Measures Were Administered, Project on Human Development in Chicago Neighborhoods, 1995–2002[a]

	Wave 1			Wave 2			Wave 3		
	No.	Mean (SD)	%	No.	Mean (SD)	%	No.	Mean (SD)	%
Sample size	1,517			1,315			1,210		
Age, years	**1,517**	**13.5 (1.5)**		**1,315**	**15.6 (1.6)**		**1,210**	**18.1 (1.6)**	
Cohort									
12-year-olds	**821**		**54.1**	718		54.6	650		53.7
15-year-olds	**696**		**45.9**	597		45.4	560		46.3
Sex									
Male	**745**		**49.1**	651		49.5	589		48.7
Female	772		**50.9**	664		50.5	621		51.3
Race/ethnicity									
Non-Hispanic black	557		36.8	**430**		**28.4**	367		32.4
Non-Hispanic white	217		14.4	**189**		**12.5**	170		15.0
Hispanic	680		44.9	**519**		**34.3**	453		39.9
Other	58		3.8	**165**		**10.9**	153		13.5
Socioeconomic status[b]	**1,502**	−0.1 (1.4)		**1,307**	**−0.3 (1.4)**		1,197	−0.1 (1.4)	
No. of years of living at current address	**1,491**	**6.7 (7.3)**		1,295	6.9 (7.5)		1,191	6.9 (7.3)	

Abbreviation: SD, standard deviation.
[a] Estimates highlighted in bold indicate the wave at which each of the measures was used in the models.
[b] Composite measure from principal-components analysis (see text).

were not significantly different from those retained, except that they had higher levels of antisocial behavior at baseline.

Figure 1 presents the predicted probabilities for each of the 3 problem behaviors, by study wave and cohort. The prevalence of antisocial behavior peaked at wave 2, at 40% for participants in the 12-year-old cohort (who were aged 14 years on average) and at 38% for those in the 15-year-old cohort (who were approximately aged 17 years). In contrast, as predicted by previous studies, substance use and problems with the police gradually increased until they reached the highest measured levels at 37%–39% and 43%–44%, respectively, at wave 3. Few differences in these probabilities existed between the 12-year-old and 15-year-old cohorts. Given sample size restrictions and few differences between cohorts, we estimated models jointly for the 12- and 15-year-old cohorts.

Table 2 presents results from a 3-wave population-averaged multivariate clustered model, as described in Appendix 1. Adjusting for basic sociodemographic characteristics, the odds of reporting antisocial behavior did not increase significantly over time. The odds of substance use (odds ratio (OR) = 2.91, 95% confidence interval (CI): 1.39, 6.12) and reporting problems with the police (OR = 2.24, 95% CI: 1.07, 4.67) at wave 3 did increase in comparison with baseline.

There were different associations between individual sociodemographic characteristics and the 3 problem behaviors. Girls were 1.72 (95% CI: 1.47, 2.00) times more likely to report problems with the police than boys. Compared with the 15-year-old cohort, the 12-year-old cohort had a significantly lower likelihood of reporting substance use (OR = 0.42, 95% CI: 0.30, 0.60) and marginally significantly lower odds of reporting problems with the police (OR = 0.58, 95% CI: 0.34, 1.01) but not antisocial behaviors. Finally, being Hispanic was associated with lower odds of both substance use (OR = 0.50, 95% CI: 0.25, 0.99) and police problems (OR = 0.46, 95% CI: 0.31, 0.69) in comparison with being non-Hispanic white.

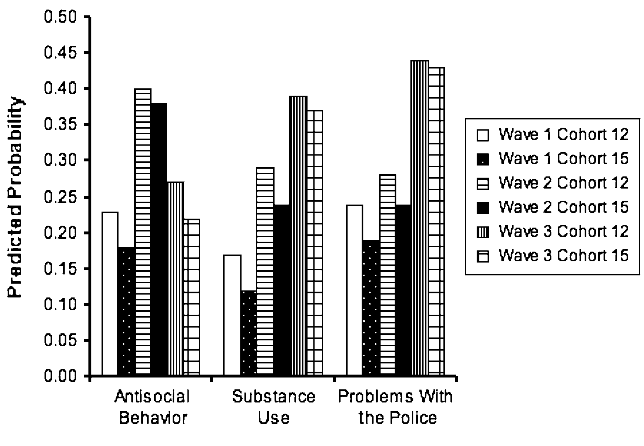

Figure 1. Predicted behavioral probabilities by study wave and cohort (based on a crude 3-wave multivariate model), Project on Human Development in Chicago Neighborhoods, 1995–2002.

Table 2. Behavioral Odds Ratios Based on a 3-Wave Multivariate Population-Averaged Model (Model 1) and Model Clustering Parameters, Project on Human Development in Chicago Neighborhoods, 1995–2002

	Antisocial Behavior		Substance Use		Problems With the Police	
Behavioral Odds Ratios	OR	95% CI	OR	95% CI	OR	95% CI
Wave (referent: wave 1)						
Wave 2	1.97	0.76, 5.16	1.66	0.95, 2.90	1.08	0.69, 1.70
Wave 3	1.11	0.38, 3.29	2.91	1.39, 6.12	2.24	1.07, 4.67
12-year-old cohort (referent: 15-year-old cohort)	0.79	0.24, 2.60	0.42	0.30, 0.60	0.58	0.34, 1.01
Female sex (referent: male)	0.87	0.75, 1.02	1.17	0.73, 1.90	1.72	1.47, 2.00
Socioeconomic status[a]	0.93	0.76, 1.14	1.03	0.87, 1.22	0.98	0.90, 1.07
Race/ethnicity (referent: non-Hispanic white)						
Non-Hispanic black	0.86	0.52, 1.42	0.59	0.29, 1.21	0.76	0.49, 1.18
Hispanic	0.65	0.36, 1.16	0.50	0.25, 0.99	0.46	0.31, 0.69
Other	0.80	0.52, 1.23	0.61	0.31, 1.21	0.86	0.58, 1.28
Neighborhood disadvantage[a]	1.09	0.78, 1.53	1.13	0.78, 1.65	1.12	0.77, 1.61

Model Clustering Parameters[b]	α (SE)	POR[c] or ΔPOR	95% CI
Neighborhood correlation	0.84 (0.33)	2.32	1.21, 4.42
Individual-level correlations			
Wave 1			
Antisocial behavior and substance use	1.21 (0.15)	3.35	2.50, 4.50
Differential correlation between antisocial behavior and police problems	−0.36 (0.17)	0.70	0.50, 0.97
Differential correlation between substance use and police problems	0.50 (0.28)	1.65	0.95, 2.85
Wave 2			
Antisocial behavior and substance use	0.08 (0.25)	1.08	0.66, 1.77
Differential correlation between antisocial behavior and police problems	0.20 (0.34)	1.22	0.63, 2.38
Differential correlation between substance use and police problems	1.08 (0.37)	2.94	1.43, 6.08
Wave 3			
Antisocial behavior and substance use	0.74 (0.37)	2.10	1.01, 4.33
Differential correlation between antisocial behavior and police problems	−0.09 (0.69)	0.91	0.24, 3.53
Differential correlation between substance use and police problems	−0.03 (0.51)	0.97	0.36, 2.64

Abbreviations: CI, confidence interval; OR, odds ratio: POR, pairwise odds ratio; SE, standard error.
[a] First principal component standardized to a mean of 0 and a standard deviation of 1 (see text).
[b] The POR model is given by equation 1.2 plus the terms described in Appendix 1.
[c] The POR and the differential POR (ΔPOR) are the exponentiated **α** parameters.

Clustering at the neighborhood level was significant. The POR between any 2 behaviors of any 2 different individuals was 2.32 (95% CI: 1.21, 4.42). There was significant clustering of reports at the individual level at wave 1. Relative to the neighborhood POR, the multiplicative increase in the POR between antisocial behavior and substance use within the individual was 3.35 (i.e., differential POR (ΔPOR = 3.35, 95% CI: 2.50, 4.50)). Antisocial behavior had lower concurrent comorbidity with police problems than with substance use (ΔPOR = 0.70, 95% CI: 0.50, 0.97), whereas the comorbidity between substance use and police problems was marginally higher than that between antisocial behavior and substance use (ΔPOR = 1.65, 95% CI: 0.95, 2.85). At wave 2, differential additional clustering was found between the reporting of substance use and police problems (ΔPOR = 2.94, 95% CI: 1.43, 6.08), so that participants were more likely to jointly report substance use and police problems, instead of antisocial behavior and substance use.

Model 2 (Table 3) allowed us to estimate the timing of behavior initiation and persistence. At wave 2, 30% of participants initiated antisocial behaviors, 18% initiated substance use, and 19% started reporting police problems. At the same time, 65% of those who had been antisocial at wave 1 persisted at wave 2, 60% of wave 1 substance users persisted, and 42% persisted in having police problems from the previous wave.

Model 2 also presented instances of sequential comorbidity, when the presence of behaviors at a previous time point contributed to the reporting of comorbid behaviors at a subsequent point. Substance use and police problems showed the same transition patterns at wave 2: 31% of participants who reported antisocial behavior at wave 1 also reported substance use or police problems at wave 2, while 42% of those who had police problems at wave 1 also had police problems or used substances at wave 2. At wave 3, 37% of participants with antisocial behavior at wave 2 also reported substance use.

Concurrent clustering of behaviors decreased substantially once we introduced between-behavioral transitions into the model. Relative to the neighborhood POR, the multiplicative increase in the POR between antisocial behavior and substance use within the individual was 1.03 (95% CI: 0.99, 1.07). At wave 3, differential additional clustering existed between reports of antisocial behavior and problems with the police at wave 3 (ΔPOR = 2.29, 95% CI: 1.15, 4.55). Concurrent clustering was explained by past behaviors.

DISCUSSION

This study confirms the previously reported association between antisocial behavior, substance use, and problems with the police (57, 58), and it provides an example of multivariate transition models as a method for modeling concurrent and sequential comorbidity among problem behavior clusters. This is one of the first studies to present multivariate transition models within a clustered context and to apply this statistical technique to actual (nonsimulated) data.

These models make several contributions to the field of developmental epidemiology. First, this method builds on the contributions made using latent growth curve analysis to study connections between different behavior trajectories (59), since it allows us to estimate the magnitude of clustering between comorbid pairs at different points in time. For example, by constructing a population-averaged multivariate transition model, we found that participants with antisocial behavior were more likely to concurrently report substance use and less likely to report police problems at wave 1, and substance use was more likely to co-occur with police problems than with antisocial behavior at wave 2. Such information can prove instrumental in understanding how comorbidity evolves over life stages.

Second, multivariate transition models allow us to compare the timing of onset and the persistence of multiple co-occurring behaviors and to evaluate sequential comorbidity. For example, in this study, we found that reporting police problems and substance use at wave 2 depended on reporting antisocial behavior at the previous wave. Antisocial problems, in contrast, depended only on having had such problems in the prior wave. These results suggest that comorbidity may occur from antisocial behavior to substance use and police problems, but not in the reverse direction. Such models can prove instrumental in addressing the persistent debate about the possible sequential relations among problem behaviors (60–62).

Third, these models offer a way to investigate multiple levels of influence (e.g., individual, family, school, and neighborhood) on trajectories for different behaviors. This feature addresses the need in developmental epidemiology to adopt an ecologic approach and to study the developmental contributions of factors from different levels of the environment (63).

There are several key statistical features of the models applied to this work. First, the generalized estimating equations framework was used to estimate model parameters and to accommodate the nested structure of the data. With this estimation approach, parameter estimates have a population interpretation. The nested structure of the data could instead be modeled with random effects for neighborhood and person and random effects shared by behaviors within-person. Available estimation packages could be used to estimate model parameters in such cases—for example, SAS PROC GLIMMIX (64) or Mplus software (65).

Second, within the generalized estimating equations framework, we used the POR as the measure with which to model and test for clustering of behaviors within neighborhoods and within persons. There are advantages and disadvantages to using the POR as the measure of association in contrast to other measures such as the intraclass correlation coefficient, the kappa statistic (52, 66), or the median odds ratio (67), but it was not our objective in the current manuscript to compare these methods.

Third, alternating logistic regression relies on empirical sandwich variance estimation. This makes inferences about the beta and gamma parameters, or the transition probabilities, valid even if the POR model is misspecified.

There were limitations to this work. Transition models, by using past outcomes as predictors, may induce spurious associations between covariates and outcomes. While we recognize this limitation, we are specifically interested in the association between past and current behaviors, as this is a key component of sequential comorbidity. Further, an association between antisocial behavior and attrition suggests the possibility of informative missingness that may bias our results. The use of multiple imputation has extended the applicability of the methods to attrition that is missing at random (68). In addition, larger samples would allow us to include potential modifiers of the transitions within and between behaviors in the model. Had we used continuous or categorical measures of risk behaviors, we would have been able to describe the full spectrum of each risk behavior instead of dichotomous states. Binary transition models thus constitute a first step; the application of clustered multivariate continuous or categorical transition models in

Table 3. Estimated Probabilities of Behavioral Transition (Dependent on Previous Behavioral State) Based on a Fully Adjusted Multivariate Model (Model 2), Project on Human Development in Chicago Neighborhoods, 1997–2002[a,b]

	Antisocial Behavior	Substance Use	Problems With the Police
Transition Probability			
	Probability of Behavior at Wave 2		
Behavior reported at wave 1			
Antisocial behavior			
No	0.3	0.22	0.22
Yes	0.65*	0.31*	0.31*
Substance use			
No	0.35	0.18	0.21
Yes	0.53	0.6	0.47
Police problems			
No	0.34	0.19	0.19
Yes	0.48	0.42*	0.42*
	Probability of Behavior at Wave 3		
Behavior reported at wave 2			
Antisocial behavior			
No	0.22	0.35	0.4
Yes	0.26	0.37*	0.43
Substance use			
No	0.23	0.3	0.39
Yes	0.26	0.53	0.48
Police problems			
No	0.23	0.33	0.37
Yes	0.26	0.43	0.54
Clustering Parameters[c]			
	α (SE)	POR[d] or ΔPOR	95% CI
Neighborhood	0.19 (0.68)	1.21	0.32, 4.59
Individual-level correlations			
Wave 2			
Antisocial behavior and substance use	0.03 (0.02)	1.03	0.99, 1.07
Differential correlation between antisocial behavior and police problems	0.43 (0.34)	1.54	0.79, 2.99
Differential correlation between substance use and police problems	−0.19 (0.42)	0.83	0.36, 1.88
Wave 3			
Antisocial behavior and substance use	0.55 (0.49)	1.73	0.66, 4.53
Differential correlation between antisocial behavior and police problems	0.83 (0.35)	2.29	1.15, 4.55
Differential correlation between substance use and police problems	−0.09 (0.75)	0.91	0.21, 3.97

Abbreviations: CI, confidence interval; POR, pairwise odds ratio; SE, standard error.

* $P < 0.05$.

[a] Obtained from model-predicted probabilities by stratifying the sample by wave, current risk behavior, and behavioral state in the previous wave and obtaining the mean predicted probability for that stratum.

[b] The model included adjustment for cohort (12-year-olds vs. 15-year-olds), sex, socioeconomic status (composite), and race/ethnicity.

[c] The POR model is given by equation 1.2.

[d] The POR and the differential POR (ΔPOR) are the exponentiated α parameters.

epidemiology remains to be explored. The models also included adjustment for a limited set of confounders, because of the large number of clustering parameters and covariates already in the model. This study serves as an application of a promising method, rather than a definitive investigation into the causal links between antisocial behavior, substance use, and police problems. Further, the models did not allow for change of neighborhood over time; however, since participants moved to the same types of neighborhoods throughout the study, we doubt that failing to account for residential movement affected our conclusions.

Notwithstanding these limitations, this study demonstrates the value of adopting a comprehensive approach to understanding multiple behaviors, where we can jointly study the transitions people experience in comorbid behaviors. By modeling transitions not only within 1 behavior but also across behaviors, we can track the pathways people follow into potentially increasingly severe profiles (24). These models can be applied to desired disorders and health-related behaviors, such as depression, exercise, and diet. Furthermore, the clustered structure of the models allows us to understand the roles that contexts such as neighborhoods or workplaces play in modifying the onset, prevalence, and persistence of co-occurring behaviors (63).

ACKNOWLEDGMENTS

Author affiliations: Center for Urban Epidemiologic Studies, New York Academy of Medicine, New York, New York (Magdalena Cerdá, Sandro Galea); Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan (Sandro Galea, Melissa Tracy); Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan (Brisa N. Sánchez); Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York (Sandro Galea); and Department of Community Health, Brown University, Providence, Rhode Island (Stephen L. Buka).

Funding for this work was provided by the Robert Wood Johnson Foundation Health and Society Scholars Program (M. C.) and by National Institutes of Health grants MH 078152, DA 017642, and DA 022720 (S. G., M. T.).

The authors acknowledge the essential roles of Drs. Albert Reiss, Jr., and Robert Sampson in the inception of this project, as well the contributions made by Drs. Arnold Sameroff and Jack Kabfleisch.

The data for this study were provided the Inter-University Consortium for Social and Political Research at the University of Michigan.

Conflict of interest: none declared.

REFERENCES

1. Bourdon KH, Rae DS, Locke BZ, et al. Estimating the prevalence of mental disorders in U.S. adults from the Epidemiologic Catchment Area Survey. *Public Health Rep*. 1992; 107(6):663–668.
2. Burke JD, Loeber R, Lahey BB, et al. Developmental transitions among affective and behavioral disorders in adolescent boys. *J Child Psychol Psychiatry*. 2005;46(11): 1200–1210.
3. de Graaf R, Bijl RV, ten Have M, et al. Rapid onset of comorbidity of common mental disorders: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatr Scand*. 2004;109(1):55–63.
4. Husler G, Blakeney R, Werlen E. Adolescent risk: the co-occurrence of illness, suicidality, and substance use. *J Youth Adolesc*. 2005;34(6):547–557.
5. Kessler RC, Chiu WT, Demler O, et al. Prevalence, severity and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey. *Arch Gen Psychiatry*. 2005; 62(6):617–627.
6. Kessler RC, Crum RM, Warner LA, et al. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry*. 1997;54(4):313–321.
7. Kessler RC, Nelson CB, McGonagle KA, et al. The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. *Am J Orthopsychiatry*. 1996;66(1):17–31.
8. Angold A, Costello EJ, Erkanli A. Comorbidity. *J Child Psychol Psychiatry*. 1999;40(1):57–87.
9. Angst J. Comorbidity of mood disorders: a longitudinal prospective study. *Br J Psychiatry Suppl*. 1996;(30):31–37.
10. Costello EJ, Egger HL, Angold A. The developmental epidemiology of anxiety disorders: phenomenology, prevalence, and comorbidity. *Child Adolesc Psychiatr Clin N Am*. 2005; 14(4):631–648, vii.
11. de Graaf R, Bijl RV, Smit F, et al. Risk factors for 12-month comorbidity of mood, anxiety, and substance use disorders: findings from the Netherlands Mental Health Survey and Incidence Study. *Am J Psychiatry*. 2002;159(4): 620–629.
12. Schuckit MA. Comorbidity between substance use disorders and psychiatric conditions. *Addiction*. 2006;101(suppl 1): 76–88.
13. Windle M, Davies PT. Depression and heavy alcohol use among adolescents: concurrent and prospective relations. *Dev Psychopathol*. 1999;11(4):823–844.
14. Crockett LJ, Moilanen KL, Raffaelli M, et al. Psychological profiles and adolescent adjustment: a person-centered approach. *Dev Psychopathol*. 2006;18(1):195–214.
15. Siebenbruner J, Englund MM, Egeland B, et al. Developmental antecedents of late adolescence substance use patterns. *Dev Psychopathol*. 2006;18(2):551–571.
16. Sullivan TN, Farrell AD, Kliewer W. Peer victimization in early adolescence: association between physical and relational victimization and drug use, aggression, and delinquent behaviors among urban middle school students. *Dev Psychopathol*. 2006;18(1):119–137.
17. Kessler RC, Berglund P, Demler O, et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):593–602.
18. Renouf AG, Kovacs M, Mukerji P. Relationship of depressive, conduct, and comorbid disorders and social functioning in childhood. *J Am Acad Child Adolesc Psychiatry*. 1997;36(7): 998–1004.
19. Schoevers RA, Deeg DJ, van Tilburg W, et al. Depression and generalized anxiety disorder—co-occurrence and longitudinal patterns in elderly patients. *Am J Geriatr Psychiatry*. 2005; 13(1):31–39.

20. Costello EJ, Erkanli A, Federman E, et al. Development of psychiatric comorbidity with substance abuse in adolescents: effects of timing and sex. *J Clin Child Psychol*. 1999;28(3): 298–311.
21. Loeber R, Stouthamer-Loeber M, White HR. Developmental aspects of delinquency and internalizing problems and their association with persistent juvenile substance use between ages 7 and 18. *J Clin Child Psychol*. 1999;28(3):322–332.
22. Bakken K, Landheim AS, Vaglum P. Axis I and II disorders as long-term predictors of mental distress: a six-year prospective follow-up of substance-dependent patients. *BMC Psychiatry.* 2007;7:29–41.
23. de Graaf R, Bijl RV, ten Have M, et al. Pathways to comorbidity: the transition of pure mood, anxiety and substance use disorders into comorbid conditions in a longitudinal population-based study. *J Affect Disord*. 2004;82(3):461–467.
24. Ingoldsby E, Kohl G, McMahon R, et al. Conduct problems, depressive symptomatology and their co-occurring presentation in childhood as predictors of adjustment in early adolescence. *J Abnorm Child Psychol*. 2006;34(5):603–621.
25. Aseltine RH Jr, Gore S, Colten ME. The co-occurrence of depression and substance abuse in late adolescence. *Dev Psychopathol*. 1998;10(3):549–570.
26. Jaffe A, Shoptaw S, Stein JA, et al. Depression ratings, reported sexual risk behaviors, and methamphetamine use: latent growth curve models of positive change among gay and bisexual men in an outpatient treatment program. *Exp Clin Psychopharmacol*. 2007;15(3):301–307.
27. Measelle JR, Slice E, Hogansen JM. Developmental trajectories of co-occurring depressive, eating, antisocial and substance abuse problems in female adolescents. *J Abnorm Psychol*. 2006;115(3):524–538.
28. Chuang YC, Ennett ST, Bauman KE, et al. Neighborhood influences on adolescent cigarette and alcohol use: mediating effects through parent and peer behaviors. *J Health Soc Behav.* 2005;46(2):187–204.
29. Sellstrom E, Bremberg S. The significance of neighbourhood context to child and adolescent health and well-being: a systematic review of multilevel studies. *Scand J Public Health*. 2006;34(5):544–554.
30. Kane RJ. Compromised police legitimacy as a predictor of violent crime in structurally disadvantaged communities. *Criminololgy.* 2005;43(2):469–498.
31. Sellstrom E, Bremberg S. The significance of neighbourhood context to child and adolescent health and well-being: a systematic review of multilevel studies. *Scand J Public Health*. 2006;34(5):544–554.
32. Ten Have T, Morabia A. An assessment of non-randomized medical treatment of long-term schizophrenia relapse using bivariate binary-response transition models. *Biostatistics*. 2002;3(1):119–131.
33. Preisser J, Arcury T, Quandt S. Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data. *Am J Epidemiol*. 2003;158(5): 495–501.
34. Das A, Poole WK, Bada HS. A repeated measures approach for simultaneous modeling of multiple neurobehavioral outcomes in newborns exposed to cocaine in utero. *Am J Epidemiol*. 2004;159(9):891–899.
35. Diggle P, Heagerty P, Liang K-Y, et al. *Analysis of Longitudinal Data*. 2nd ed. New York, NY: Oxford University Press; 2002.
36. Zeger S, Qaquish B. Markov regression models for time series: a quasi-likelihood approach. *Biometrics*. 1988;44(4):1019–1031.
37. Earls F, Buka SL. *Project on Human Development in Chicago Neighborhoods*. Washington, DC: National Institute of Justice; 1997.
38. Raghunathan TE, Lepkowski JM, Van Hoewyk J, et al. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv Methodol*. 2001; 27(1):85–95.
39. Raghunathan T, Van Hoewyk J, Solenberger P. *IVEware: Imputation and Variance Estimation Software User Guide*. Ann Arbor, MI: University of Michigan; 2002.
40. Earls F, Brooks-Gunn J, Raudenbush S, et al. *Project on Human Development in Chicago Neighborhoods (PHDCN): Substance Use, Wave 1, 1994–1997* [computer file]. Cambridge, MA: Harvard Medical School; 2002. (http://www.icpsr.umich.edu/PHDCN/about.html).
41. Earls F, Brooks-Gunn J, Raudenbush S, et al. *Project on Human Development in Chicago Neighborhoods (PHDCN): Substance Use, Wave 2, 1997–2000* [computer file]. Cambridge, MA: Harvard Medical School; 2002. (http://www.icpsr.umich.edu/PHDCN/about.html).
42. Earls F, Brooks-Gunn J, Raudenbush S, et al. *Project on Human Development in Chicago Neighborhoods (PHDCN): Substance Use, Wave 3, 2000–2002* [computer file]. Cambridge, MA: Harvard Medical School; 2002. (http://www.icpsr.umich.edu/PHDCN/about.html).
43. Elliott DS, Huizinga D, Morse B. Self-reported violent offending: a descriptive analysis of juvenile violent offenders and their offending careers. *J Interpers Violence*. 1986;1(4): 472–514.
44. Achenbach TM. *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont; 1991.
45. Achenbach TM. *Manual for the Youth Self Report and 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont; 1991.
46. Achenbach TM. *Young Adult Self Report*. Burlington, VT: Department of Psychiatry, University of Vermont; 1997.
47. Huijbregts SC, Séguin JR, Zoccolillo M, et al. Associations of maternal prenatal smoking with early childhood physical aggression, hyperactivity-impulsivity, and their co-occurrence. *J Abnorm Child Psychol*. 2007;35(2):203–215.
48. Romano E, Tremblay RE, Boulerice B, et al. Multilevel correlates of childhood physical aggression and prosocial behavior. *J Abnorm Child Psychol*. 2005;33(5):565–578.
49. Wanner B, Vitaro F, Ladouceur R, et al. Joint trajectories of gambling, alcohol and marijuana use during adolescence: a person- and variable-centered developmental approach. *Addict Behav*. 2006;31(4):566–580.
50. Molnar BE, Cerdá M, Roberts AL, et al. Effects of neighborhood resources on aggressive and delinquent behaviors among urban youths. *Am J Public Health*. 2008;98(6):1086–1093.
51. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997;277(5328):918–924.
52. Goldstein H, Browne W, Rasbash J. *Partitioning Variation in Multilevel Models*. London, United Kingdom: Institute of Education, University of London; 2002.
53. Preisser J. [Re: "Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data."] The first author replies [letter]. *Am J Epidemiol*. 2004;160(5):506–507.
54. Diggle PJ, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data*. New York, NY: Oxford University Press; 1994.
55. SAS Institute Inc. *SAS/STAT User's Guide*. Cary, NC: SAS Institute Inc; 1999.

56. SAS Institute Inc. The MIANALYZE procedure. In: *SAS OnlineDoc, Version 8*. Cary, NC: SAS Institute Inc; 2007. 201–233. (http://v8doc.sas.com/sashtml/).
57. Schulenberg J, Wadsworth KN, O'Malley PM, et al. Adolescent risk factors for binge drinking during the transition to young adulthood: variable and pattern-centered approaches to change. *Dev Psychol*. 1996;32(4):659–674.
58. Wiesner M, Kim HK, Capaldi DM. Developmental trajectories of offending: validation and prediction to young adult alcohol use, drug use, and depressive symptoms. *Dev Psychopathol*. 2005;17(1):251–270.
59. Needham BL. Gender differences in trajectories of depressive symptomatology and substance use during the transition from adolescence to young adulthood. *Soc Sci Med*. 2007;65(6):1166–1179.
60. Donovan J. Problem-behavior theory and the explanation of adolescent marijuana use. *J Drug Issues*. 1996;26(2):379–404.
61. Elliott D, Huizinga D, Menard S. *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems*. New York, NY: Springer-Verlag; 1989.
62. Loeber R, Farrington D, Southamer-Loeber M, et al. Multiple risk factors for multiproblem boys: co-occurrence of delinquency, substance use, attention deficit, conduct problems, physical aggression, covert behavior, depressed mood, and shy/withdrawn behavior. In: Jessor R, ed. *New Perspectives on Adolescent Risk Behavior*. Cambridge, United Kingdom: Cambridge University Press; 1998: 90–149.
63. Glantz MD, Leshner AI. Drug abuse and developmental psychopathology. *Dev Psychopathol*. 2000;12(4):795–814.
64. SAS Institute Inc. *Base SAS 9.13 Procedures Guide*. 2nd ed., vols. 1–4. Cary, NC: SAS Institute Inc; 2006.
65. Muthén LK, Muthén BO. *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén; 2004.
66. Merlo J. Multilevel analytical approaches in social epidemiology: measures of health variation compared with traditional measures of association. *J Epidemiol Community Health*. 2003;57(8):550–552.
67. Larsen K, Petersen J, Budtz-Jorgensen E, et al. Interpreting parameters in the logistic regression model with random effects. *Biometrics*. 2000;56(3):909–914.
68. Little R, Rubin D. *Statistical Analysis With Missing Data*. 2nd ed. New York, NY: John Wiley & Sons, Inc; 2002.

APPENDIX 1

An important special case of equations 1.1 and 1.2 is the univariate transition model (48), which models the transitions of only 1 behavior at a time. Such a model can be obtained from equations 1.1 and 1.2 by setting $R = 1$. Fitting univariate models for each behavior constitutes an important step in model-building, as it helps us understand how jointly estimating related behaviors can provide different behavior-specific transition profiles from the transition patterns traditionally found when models are fitted separately for each behavior.

Another special case is obtained by setting all parameters $\gamma_{rr'} = 0$ and letting t vary from 1 to T in equation 1.2 (instead of $t = 2, \ldots, T$), namely a *multivariate* population-averaged model for cluster-correlated longitudinal data. This combines the model presented by Preisser et al. (31) for longitudinal clustered data and the model presented by Das et al. (13) for cross-sectional models for multivariate outcomes. In this case, equation 1.2 is also extended to at least include additional POR parameters for wave 1—that is,

$$\alpha_1 I(j = j', \ t = t' = 1) + \sum_{\substack{\text{behavior} \\ \text{pairs}(m,l)}} \alpha_{1,ml} I\big(j = j', \ t = t' = 1, \ r = m, \ r' = l\big).$$

This was the approach taken for model 1 (Table 2). If the data allow, additional terms could be added to quantify correlations of the same behavior across time within an individual—for example, $\alpha_m I(j = j', \ r = r' = m)$, with $m = 1, \ldots, R$.

In all cases of the model, the intercept β_{ort} is specific to behavior r at wave t. However, the interpretation of this coefficient differs depending on the special case of the model being used. For the transition models, the intercept is the log odds of the probability of onset of behavior r, given that no behaviors were reported at the previous wave, for a person with average covariate values. For the multivariate population-averaged model, the intercept is the log odds of the prevalence of behavior r at wave t.

APPENDIX 2

SAS Code for Clustered Univariate and Multivariate Transition Models

Below we provide a description of the data management and analytic steps necessary to fit the clustered univariate and multivariate transition models described in this paper, along with sample SAS code (SAS Institute Inc., Cary, North Carolina) demonstrating each of the steps.

The original Project on Human Development in Chicago Neighborhoods data set contained 1 record for each participant ($n = 1{,}517$), including variables from each of the 3 waves of data collection. We first convert this data set to a stacked data set, with 1 record for each participant-wave of data (1,517 participants $\times$ 3 study waves = 4,551 records). In this example, 3 risk behaviors are of interest, represented by dichotomous (1/0) variables indicating their presence or absence at each wave (A represents antisocial behavior, B represents substance use, and C represents problems with the police). Since we are interested in risk transitions, the presence or absence of the risk behavior at the previous wave is also included in the data set (PREV_A, PREV_B, PREV_C). Finally, the data set contains demographic characteristics (COHORT, AGE, SEX, SES, RACE), characteristics of the participant's neighborhood (DISADV), an indicator for study wave (WAVE), and identification numbers for each participant (ID) and neighborhood cluster (NC) included in the study. Note that the demographic and neighborhood characteristics may be time-invariant or time-varying.

```
/*
 * Convert original data set to stacked data set with 1 record for each participant-wave
 */
data stacked;
       set flat;
       A = A1; PREV_A = .;
       B = B1; PREV_B = .;
       C = C1; PREV_C = .;
       DISADV = DISADV90; AGE = AGE1; WAVE = 1;
       output;
       A = A2; PREV_A = A1;
       B = B2; PREV_B = B1;
       C = C2; PREV_C = C1;
       DISADV = DISADV90; AGE = AGE2; WAVE = 2;
       output;
       A = A3; PREV_A = A2;
       B = B3; PREV_B = B2;
       C = C3; PREV_C = C2;
       DISADV = DISADV00; AGE = AGE3; WAVE = 3;
       output;
       keep  ID  NC  A  PREV_A  B PREV_B  C PREV_C
             COHORT AGE SEX SES RACE DISADV WAVE;
run;
```

For multivariate models, which will jointly investigate multiple behaviors simultaneously, an additional data step is necessary. This converts the stacked data set created above, which has 1 record for each participant-wave, to a stacked data set with 1 record for each participant-wave-behavior (in this example, 1,517 participants × 3 study waves × 3 behaviors = 13,653 records, assuming there was no attrition). In this example, the variable BEHAVIOR contains the value (1/0) of the respective behavior variable, and the BEHAV indicator identifies which behavior (1 = A, 2 = B, 3 = C) it reflects. Note that for univariate models, which are fitted for each behavior separately, this step is not necessary.

```
/*
 * Further convert stacked data set to include 1 record for each participant-wave-behavior
 */
data stacked2;
       set stacked;
       BEHAVIOR = A; BEHAV = 1; output;
       BEHAVIOR = B; BEHAV = 2; output;
       BEHAVIOR = C; BEHAV = 3; output;
run;
```

We then sort this data set by neighborhood cluster ID (variable NC), participant ID (variable ID), and study wave (variable WAVE). A unique identifier is then created for each participant within each neighborhood cluster (KID), ranging sequentially from 1 to the maximum number of participants in the respective neighborhood cluster. In this example, the largest neighborhood cluster contains 39 participants. Finally, we create a unique identifier for each observation within each neighborhood cluster (Y), ranging sequentially from 1 to the maximum number of distinct observations in the respective neighborhood cluster. This corresponds to 39 participants × 9 observations per participant = 351 in this example, since each participant has 9 distinct observations of interest (i.e., 1 for each of the 3 behaviors at each of the 3 study waves). To create this unique identifier, the following formula can be used: $Y = \{[(\text{KID} - 1) \times (\text{number of observations per participant})] + [(\text{WAVE} - 1) \times (\text{number of behaviors of interest at each wave})] + \text{BEHAV}\}$, assuming that these variables have been coded as specified above. Note that for univariate models, the BEHAV term can be left out of the formula to create the unique identifier Y for each observation within each neighborhood.

```
/*
 * Create unique identifier for each participant, and each observation, within each neighborhood
 */
proc sort data = stacked2;
by NC ID WAVE;
run;
data stacked3;
       set stacked2;
       by NC ID WAVE;
       if first.NC then KID = 0;
       if first.ID then KID + 1;
run;
data stacked4;
       set stacked3;
       Y = (KID - 1)*9 + (WAVE - 1)*3 + BEHAV;
run;
```

To employ the alternating logistic regression algorithm in SAS, a z-matrix must be specified. This data set contains 1 record for each possible pair of observations within 1 neighborhood cluster. The z-matrix is then applied to each neighborhood cluster during the analysis (i.e., this is a replicated, or ‘‘zrep,’’ matrix, rather than a fully specified z-matrix). The number of records in the z-matrix data set is equal to [(the maximum number of observations within any 1 neighborhood cluster) × (that maximum number – 1)] divided by 2. In this example, the maximum number of observations in any 1 neighborhood cluster = 39 participants × 3 study waves × 3 behaviors = 351. Thus, the number of records in the zrep matrix = [(351 × 350)/2] = 61,425. The zrep-matrix contains variables representing pairs of participants (KID1, KID2), behaviors (BEHAV1, BEHAV2, BEHAV3), and study waves (WAVE1, WAVE2, WAVE3). The formulas used to calculate the values of these variables can be generalized as follows, where X represents the first (1) or second (2) member of the pair of observations: KIDX = {int(YX/[number of observations per participant]) + 1}. Here, the ‘‘int’’ function in SAS returns an integer after removing any decimal portion of the result, producing values of 1–40 when, as in this example, 39 participants are present in the largest neighborhood cluster and have 9 observations each; BEHAVX = {mod(YX, number of behaviors of interest)}, where the ‘‘mod’’ function in SAS returns the remainder after division, producing values of 0, 1, and 2 when 3 behaviors are of interest; and WAVEX = {int((mod(YX, number of observations per participant) + 2)/3)}, producing values of 0, 1, and 2 when 3 study waves are of interest. If only 2 waves of data are of interest, WAVEX can be simplified to mod(YX,2). Note that in this example, values of 0 for the ‘‘BEHAV’’ and ‘‘WAVE’’ variables must be recoded to 3, and values of 40 for the ‘‘KID’’ variables must be reduced to 39. Also note that $Y1$ will range from 1 to 350 in this example, while $Y2$ will range from 2 to 351.

```
/*
 * Creating the z-matrix to be applied to each neighborhood cluster
 */
data zrep;
       do m = 1 to 350;
            do n = m+1 to 351;
            Y1 = m;
            Y2 = n;
            KID1 = int(Y1/9)+1;
            KID2 = int(Y2/9)+1;
            BEHAV1 = mod(Y1,3); if BEHAV1 = 0 then BEHAV1 = 3;
            BEHAV2 = mod(Y2,3); if BEHAV2 = 0 then BEHAV2 = 3;
            WAVE1 = int((mod(Y1,9)+2)/3);
            if WAVE1 = 0 then do; WAVE1 = 3; KID1 = KID1 - 1; end;
            WAVE2 = int((mod(Y2,9)+2)/3);
            if WAVE2 = 0 then do; WAVE2 = 3; KID2 = KID2 - 1; end;
            output;
            end;
      end;
run;
```

Once the z-matrix is created, we create a number of variables identifying pairs of observations that may be highly correlated (e.g., observations from participants residing in the same neighborhood cluster, observations within the same participant, observations for the same behavior by the same participant at different study waves). These variables are referred to as ''clustering parameters'' throughout the text. A full list of all of the clustering parameters considered in this analysis is available from the authors upon request. The creation of the clustering parameters utilized in the analyses (intraneighborhood correlation, SAMENC; baseline intraparticipant correlation between a reference pair of behaviors at waves 1, 2, and 3, SAMEKID1 SAMEKID2, and SAMEKID3, respectively; wave-specific correlation between substance use and problems with the police, BCWAVE1, BCWAVE2, and BCWAVE3, respectively; and wave-specific correlation between antisocial behavior and problems with the police, ACWAVE1, ACWAVE2, and ACWAVE3, respectively) is demonstrated below.

```
/*
 * Create clustering parameters for pairs of observations in the z-matrix
 */
data zrep2;
        set zrep;
        * intraneighborhood correlation;
        * note that since the z-matrix will be applied to each neighborhood cluster;
        * all of the observations in the z-matrix will belong to the same neighborhood;
        SAMENC = 1;
        * individual-level correlation;
        if (KID1 = KID2) and (WAVE1 = WAVE2 = 1) then SAMEKID1 = 1; else
                SAMEKID1 = 0;
        if (KID1 = KID2) and (WAVE1 = WAVE2 = 2) then SAMEKID2 = 1; else
                SAMEKID2 = 0;
        if (KID1 = KID2) and (WAVE1 = WAVE2 = 3) then SAMEKID3 = 1; else
                SAMEKID3 = 0;
        * differential correlation between substance use (B) and problems with the police
(C) at wave 2;
        if ((KID1 = KID2) and (WAVE1 = WAVE2 = 2) and (BEHAV1 = 2) and (BEHAV2 = 3))
                then BCWAVE2 = 1; else BCWAVE2 = 0;
* Other differentials constructed similarly;
        keep Y1 Y2  SAMENC   SAMEKID1  SAMEKID2  SAMEKID3
                             BCWAVE1   BCWAVE2   BCWAVE3
                             ACWAVE1   ACWAVE2   ACWAVE3;
run;
```

Model 1 in Table 2 of the paper is a 3-wave multivariate population-averaged model assessing the relation between individual- and neighborhood-level covariates and each of the 3 risk behaviors of interest. We use the following SAS code to fit this model. In this example, ''repeated subject = NC'' reflects the clustering of participants within neighborhoods; ''withinsubject = Y'' reflects that the variable Y is a unique identifier for each distinct observation within each ''subject'' (NC, in this case); the ''logor = zrep'' option indicates that the alternating logistic regression algorithm will be used, with the zrep-matrix specified in the ''zdata'' option; the clustering parameters are specified in the ''zrow'' option; and the indicators for each pair of observations in the zrep-matrix are specified in the ''ypair'' option. This model gives estimates for each of the 3 risk behaviors of interest by using interactions between the ''BEHAV'' variable and each of the covariates.

```
/*
 * model 1
 */
proc genmod data = stacked4 descending;
class NC ID WAVE BEHAV RACE Y;
model BEHAVIOR = BEHAV*WAVE BEHAV*COHORT BEHAV*SEX BEHAV*SES
       BEHAV*RACE BEHAV*DISADV
       /dist = bin noint noscale scale = 1 type3;
repeated subject = NC /withinsubject = Y logor = zrep zdata = zrep2 zrow = (SAMENC
       SAMEKID1 BCWAVE1 ACWAVE1 SAMEKID2 BCWAVE2 ACWAVE2
SAMEKID3 BCWAVE3 ACWAVE3) ypair = (Y1 Y2) ecovb;
run;
```

Model 2 in Table 3 of the paper is a first-order clustered multivariate transition model, from which the transition probabilities for each behavior can be estimated for participants who did and did not engage in that same behavior at the previous wave. In this example, we create variables to indicate whether a participant engaged in each behavior at the current and previous time points (e.g., A2A1 = 1 indicates that a young person met criteria for antisocial behavior at wave 2 as well as at wave 1, while C3C2 = 1 indicates that a young person reported problems with the police at both waves 2 and 3).

Since we are only interested in the outcomes at waves 2 and 3, the wave 1 records can be removed from the stacked data set containing records for each participant-wave-behavior (“stacked2”), leaving a total of 1,517 participants × 2 study waves × 3 behaviors = 9,102 records in this data set. Note that the unique identifier *Y* can be specified using the following formula in this case: {[(KID – 1) × 6] + [(WAVE – 2) × 3] + BEHAV}. “Wave – 2” is used in this case because only records corresponding to waves 2 and 3 are of interest, since these records also contain the status of the risk behavior at the previous wave. Note also that the *z*-matrix in this example will have a total of 27,621 records because the maximum number of observations in any given neighborhood cluster is 39 participants × 3 behaviors × 2 study waves = 234, and [(234 × 233)/2] = 27,621.

```
/*
 * model 2
 */
proc genmod data = stacked4_waves23 descending;
class NC ID WAVE BEHAV RACE Y;
model BEHAVIOR = BEHAV*WAVE A2A1 A2B1 A2C1 A3A2 A3B2 A3C2 B2B1 B2A1 B2C1 B3B2 B3A1 B3C1 C2C1
C2A1 C2B1 C3C2 C3A2 C3B2 COHORT SEX SES RACE/dist = bin noint noscale scale = 1 type3;
repeated subject = NC /withinsubject = Y logor = zrep zdata = zrep27621
      zrow = (SAMENC SAMEKID2 BCWAVE2 ACWAVE2 SAMEKID3
              BCWAVE3 ACWAVE3) SAMENC SAMEKID2 BCWAVE2
              ACWAVE2 BCWAVE3 SAMEKID3 ACWAVE3 SAMENC SAMEKID
              BC_ANYWAVE) ypair = (Y1 Y2) ecovb;
output out = pmodel2 prob = P1;
run;
```