Published in final edited form as:
Open Fam Stud J. 2011 ; 4(Suppl 1-M9): 81–88.

Differential Effects for Sexual Risk Behavior: An Application of Finite Mixture Regression

Stephanie T. Lanza[1,2], **Kari C. Kugler**[1], and **Charu Mathur**[1]
[1]The Methodology Center, The Pennsylvania State University
[2]The College of Health and Human Development, The Pennsylvania State University

Abstract

Understanding the multiple factors that place individuals at risk for sexual risk behavior is critical for developing effective intervention programs. Regression-based methods are commonly used to estimate the average effects of risk factors, however such results can be difficult to translate to prevention implications at the individual level. Although differential effects can be examined to some extent by including interaction terms, as risk factors and moderators are added to the model interpretation can become difficult. The current study presents finite mixture regression as an alternative approach, where population subgroups are identified based on the pattern of associations between multiple risk factors and sexual risk behavior. Data from participants in the National Longitudinal Study on Adolescent Health were used to explore the effects of five adolescent risk factors (early sexual debut, heavy episodic drinking, school connectedness, positive consequences of having sex, and negative consequences of having sex) on the total number of sexual partners in adulthood. Four latent classes were identified on the basis of the Poisson regression parameter estimates. Gender, race, and grade were included as predictors of latent class membership. Results suggest that prevention programs focused on mediating these particular risk factors may be most effective for adolescents who are at lower risk for later engaging in risky sexual behaviour; however, for the subgroup of adolescents who go on to have the most sexual partners, the evidence is less conclusive and warrants further study.

Over the past two decades a wealth of literature has been published on the risk and protective factors associated with high-risk sexual behavior among youth (Buhi & Goodson, 2007; DiClemente et al., 2008). The factors are vast and include individual characteristics (e.g., attitudes towards sex) and behaviors (e.g., alcohol use), social interactions (e.g., family and peer connections), and environmental influences (e.g., community resources; Kirby, 2007). The identification of these factors has been critical in designing and implementing effective risk-reduction programs (DiClemente et al., 2008); however, as noted in a recent review article, there is room for improvement, especially in terms of conducting more rigorous methodological assessments (Buhi & Goodson, 2007). Using data from a nationally representative longitudinal study, we demonstrate how using a more advanced statistical technique can simultaneously (a) identify population subgroups based on the relative importance of risk factors for lifetime sexual partners among young adults and (b) identify likely members of those subgroups so that relevant preventive interventions might be designed for those individuals. While the absolute number of sexual partners a person has in his/her lifetime in is not a risk in and of itself, it has become a widely used indicator of risky sexual behavior because many young people inconsistently or incorrectly use condoms

Correspondence may be sent to: Stephanie T. Lanza, 204 E. Calder Way, Suite 400, The Methodology Center, The Pennsylvania State University, State College, PA 16801, 814-865-7095, SLanza@psu.edu.

(Santelli, Brener, Lowry, Bhatt & Zabin, 1998), which increases the risk for unintended pregnancies and transmission of sexually transmitted infections, including HIV.

The risk factors utilized in the study were inspired by a social cognitive framework, which posits that behavior is the result of interactions between individual cognitions, the environment, and personal behavior (Bandura, 1986). Within the cognitive domain, we investigated motivations for engaging in sexual intercourse, specifically positive and negative consequences of having intercourse. A recent review article found motivations to engage in sex as one of the most stable predictors of sexual behavior outcomes (Buhi & Goodson, 2007), including number of sexual partners (Kan, Cheng, Landale & McHale, 2010). The environmental factor considered in the study was school connectedness. Although there is mounting evidence that a youth's perception of relationships to people at school is protective against sexual risk behaviors, the evidence is inconclusive for predicting number of sexual partners (Markham et al., 2010), thus warranting further exploration. Finally, there is substantial evidence that other behaviors such as early use of alcohol (DiClemente et al.; 2008; Santelli, et al., 1998) and early sexual debut (Sandfort et al. , 2008) predict multiple sexual partners later in life; one study found that the relationship between alcohol use and multiple sexual partners was mediated by early sexual debut (Strachman, Impett, Henson & Pentz, 2009).

Methodological approaches to modeling risk factors

Most of the literature to date relating risk factors to sexual risk behavior has been based on regression models. These methods are appropriate for estimating the average effects of predictors in a population. Any exploration of differential effects across population subgroups is traditionally done by adding interaction terms. For example, when estimating the effect of perceived positive consequences of having intercourse during adolescence on lifetime sexual partners at age 30, inclusion of a term for the interaction between gender and positive consequences permits an examination of whether the effect of positive consequences on number of partners is different for males and females. Although such moderation analyses can be highly informative, a model considering multiple predictors and multiple moderating variables simultaneously can be very difficult to interpret. As an example, a model with three risk factors and the same number of moderators would require three main effects for the risk factors, three main effects for the moderators, and nine two-way interaction terms.

An alternative is finite mixture regression (McLachlan & Peel, 2000; Wedel & DeSarbo, 2002; Wedel & Kamakura, 2001), which can be used to identify population subgroups of individuals for whom certain risk factors are most salient and, if targeted in an intervention program, may be most effective at reducing the problem behavior. Finite mixture regression is similar in spirit to the more familiar latent class analysis (Collins & Lanza, 2010), which is used to identify latent subgroups on the basis of responses to a set of categorical items, usually from one time point. What is different, however, is that unlike in latent class analysis, where the indicators are actual observed variables, in finite mixture regression the indicators are regression coefficients.

Conceptually, finite mixture regression is a statistical model that posits a single regression model for a population (similar to standard regression) and posits that two or more unobserved (latent) subgroups exist in the population such that those subgroups differ in their regression coefficients. Consider again the example where a scale for positive consequences of having intercourse, measured during adolescence, is used to predict the total number of sexual partners at age 30. Standard linear regression analysis would yield the expected effect that positive consequences of intercourse have on the number of partners, *on* *average in the whole population*. In contrast, finite mixture regression analysis may suggest two or more subgroups (i.e., latent classes) in the population, where the subgroups differ in terms of their intercept (i.e., mean number of partners), the effect of positive consequences on number of partners, or both. Parameter estimates include the intercept and slope within each subgroup, as well as the proportion of the population in each subgroup.

Because the latent class variable is unobserved, the actual subgroup membership of individuals in a sample is unknown. Rather, each individual has a (typically nonzero) probability of membership in each latent class. Based on this, it may appear to be impossible to direct intervention resources to the unobserved subgroups. However, identifying who to target actually can be achieved by including directly measured characteristics (e.g., age, gender, race) as predictors of latent class membership. Then, the set of characteristics that are associated with latent class membership, where the classes are defined by both the level on the outcome (i.e., intercept) and the effects of multiple predictors on that outcome, may be used to target intervention resources to certain individuals and tailor the intervention program to those respective subgroups based on the most salient predictors suggested within the class-specific regression models.

The current study presents an empirical demonstration of finite mixture regression, where the total number of sexual partners in adulthood is predicted by early sexual debut, heavy episodic drinking, school connectedness, positive consequences of having intercourse, and negative consequences of having intercourse. Emphasis is given to a comparison between the prevention implications suggested by a standard Poisson regression analysis and those suggested by a finite mixture regression analysis.

Methods

Participants

Participants were drawn from the National Longitudinal Study of Adolescent Health (Add Health; Harris et al., 2009). A sample of 80 high schools and 52 middle schools in the US was selected with unequal probability of selection. Incorporating systematic sampling methods and implicit stratification into the Add Health study design ensured this sample is representative of US schools with respect to region of country, urbanicity, school size, school type, and ethnicity. The current study included $N = 6{,}042$ adolescents who were in ninth through twelfth grade at Wave 1 (median age of 16; 54.1% female; 70.9% White, 17.6% Black, 11.5% other), collected in 1994–1995, and had complete data on the variables used in the models. Data on individual characteristics and risk factors were drawn from Wave 1, and data on total number of sexual partners were drawn from Wave 4, when participants were approximately 27 – 30 years old.

Measures

Sexual risk behavior—Total number of sexual partners, irrespective of sexual orientation, was assessed at Wave 4, when participants were age 27–30. The reported number of partners ranged from zero to 900, with a mean of 13.7 partners. For computational reasons, extreme high outliers on the count variable, which comprised less than one percent of the total sample, were truncated so that the maximum number of partners was coded as 100 (range 0 to 100, mean = 12.8)

Risk factors—Five risk factors for sexual risk behavior, which were assessed at Wave 1 when participants were in Grades 9 through 12, were included in the finite mixture regression models. *Early sexual debut* was coded 1 if the adolescent reported ever having sexual intercourse before ninth grade. Specifically, adolescents were coded as having early sex if those in Grade 9 reported having first intercourse during or before 1994, those in Grade 10 during or before 1993, those in Grade 11 during or before 1992, and those in Grade 12 during or before 1991. *Heavy episodic drinking* was coded 1 if the adolescent reported drinking five or more alcoholic drinks in a row at least one time in the past 12 months, and coded 0 if they reported not engaging in this behavior. *School connectedness* was assessed by taking the mean of two items: "You feel close to people at your school," and "You feel like you are part of your school," which were coded on five-point scales ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). The scale was standardized, and high scores reflect more risk (α = .77). A factor analysis of a set of items measuring motivations to engage in risky behaviors suggested two unique scales. *Positive consequences* reflected three perceived positive outcomes for having sexual intercourse: friends would respect you more, it would relax you, it would make you more attractive to the opposite sex. The standardized mean of the three five-point scales (from strongly disagree to strongly agree) was coded so that high scores, i.e. more positive consequences for having intercourse, correspond to more risk (α = .85). *Negative consequences* reflected three perceived negative outcomes for having sexual intercourse: your partner would lose respect for you, you would feel guilty, and it would upset your mother. The standardized mean of the three five-point scales (from strongly disagree to strongly agree) was coded so that high scores correspond to greater protection (α = .83). The five factors described above were included in the regression models as predictors of sexual risk behavior.

Individual characteristics—Three individual characteristics were included to demonstrate how profiles can be obtained for the individuals who comprise each latent class. These were *gender* (54.1% female, 45.9% male), *race* (reduced to three categories: 17.6% Black, 70.9% White, and 11.5% other racial groups), and *grade in school* at Wave 1 (26.4% in Grade 9, 25.4% in Grade 10, 25.7% in Grade 11, and 22.6% in Grade 12). These variables were included as predictors of latent class membership, as opposed to direct predictors of sexual risk behavior.

Analytic strategy

We first employed a standard linear regression model where the outcome variable (number of sexual partners by adulthood) was specified to be a count variable; thus, log-linear Poisson regression was used. That is, the natural log of the expected number of sexual partners was modelled as a function of the predictors. Early sexual debut, heavy episodic drinking, school connectedness, positive consequences, and negative consequences were included as numeric predictors of the number of sexual partners. Results from this model yielded the population average expected effects of the five risk factors, providing a basis of comparison between a standard regression model and the proposed finite mixture regression model. The profile of individual characteristics for the standard approach is simply reflected by the overall distribution on gender, race, and grade in school, described above.

Next, we conducted finite mixture regression models with the number of sexual partners again specified to be an outcome to be treated as a count variable, with the five risk factors as predictors. In these models, log-linear Poisson regression was used for the within-class regression models. Models with one through seven latent classes were compared; the one-class model is equivalent to the standard regression model described above. Model selection for the finite mixture regression analysis was conducted using a combination of interpretability, the BIC, and the pseudo R^2, which is a heuristic measure of how well the set of predictors explains variability in lifetime sexual partners (see Vermunt & Magidson, 2005). Gender, race, and grade in school were included to predict membership in each latent class, allowing us to determine whether these individual characteristics were predictive of certain regression patterns, and also to profile the resultant latent classes according to these characteristics. Latent Gold (Vermunt & Magidson, 2005) was used to fit the standard Poisson regression and the finite mixture regression models.

Results

Standard Poisson regression

Table 1 shows results from the standard Poisson regression analysis, where number of sexual partners in adulthood was predicted from the five risk factors assessed during high school. The amount of variance explained in the outcome by these five predictors was 9%. The intercept, 2.32, is the Poisson regression estimate when all predictors in the model are evaluated at zero. For adolescents with no early sexual debut, no heavy drinking, and average school connectedness and positive and negative consequences, the log of the expected number of partners in adulthood was 2.32. By exponentiating this value, we estimated that this intercept corresponded to $e^{2.32} = 10.18$ lifetime partners.

The remaining coefficients reflect the difference in the logs of expected number of partners corresponding to a one-unit change in the predictor variable, given that the other predictor variables in the model are held constant. For example, the first coefficient compares adolescents who have and have not engaged in early sexual intercourse, given that the other variables are held constant in the model. The difference in the logs of expected number of partners was expected to be 0.45 units higher for those who had an early sexual debut. This corresponds to a main effect of $e^{0.45} = 1.57$ times more partners for those with early sexual debut (compared to those who did not) holding all other predictors constant. Similarly, heavy episodic drinking corresponded to $e^{0.27} = 1.31$ times more partners; one standard deviation higher on school connectedness corresponded to $e^{0.02} = 1.02$ times more partners; one standard deviation higher on positive consequences corresponded to $e^{0.21} = 1.23$ times more partners; and negative consequences corresponded to $e^{-0.23} = 0.79$ times as many partners.

Finite mixture regression

In contrast to the standard Poisson regression analysis, in finite mixture regression there is a complete set of regression coefficients, including intercept and slopes, estimated within each latent class. In addition, latent class membership probabilities are estimated that indicate the relative class sizes. Models with one through seven latent classes were compared. Table 2 shows information relevant to selecting the number of latent classes. In particular, the BIC suggested that the model with seven latent classes had the most optimal balance between fit and parsimony among the models considered. The BIC, however, indicated only incremental improvements when moving from a four-class model to more complex ones (see Figure 1, Panel a). Classification error is a criterion that indicates the certainty with which we can predict the individuals' latent class membership given the model and their responses to the observed variables; values closer to zero are preferred, although some amount of error is to be expected for models with two or more latent classes. Perhaps more convincing is the overall R^2, or amount of variability in the number of sexual partners explained by the model. Standard Poisson regression (i.e. the one-class model) explained only 9% of the variance. Figure 1, Panel b plots the overall R^2 for the models under consideration. Based on the information presented in Table 2 and Figure 1, and a careful inspection of the parameter estimates from competing models, we selected the four-class model to describe heterogeneity both in sexual risk behavior in adulthood and in the effects of risk factors predicting that behavior.

Table 1 shows the proportion of individuals in each latent class, the class-specific mean lifetime sexual partners, and the class-specific regression coefficients based on the finite mixture regression model. Classes are ordered according to their relative size (largest to smallest), which corresponds to an increasing average lifetime sexual partners. Latent Class 1 comprised 53% of the sample, with mean lifetime partners equal to 4.12. Latent Class 2 included 32% of the sample, with 13.09 lifetime partners on average. Similarly, Latent Classes 3 comprised 13% of the sample, with 32.15 mean lifetime partners, and Latent Class 4 included just 3% of the sample, with 79.91 lifetime partners on average. As with the standard Poisson regression, class-specific intercept coefficients from the mixture regression model were exponentiated to yield mean lifetime partners in adulthood among adolescents with no early sexual debut, no heavy drinking, and average school connectedness and positive and negative consequences. For example, in Latent Class 1 the intercept corresponded to $e^{1.11}$ = 3.03 lifetime partners. The corresponding values for Latent Classes 2, 3, and 4 were 9.87, 26.84, and 73.70 lifetime partners, respectively.

All five predictors were significantly related to lifetime sexual partners overall (that is, combining across latent classes). The regression coefficients varied significantly ($p < .0001$) across the latent classes for the following four predictors: early sexual debut, heavy episodic drinking, positive consequences, and negative consequences. In every case, the effects were strongest in Latent Class 1, and weakened across the latent classes as the mean number of partners increased. In other words, the known risk factors for lifetime sexual partners had effects in the anticipated direction for every latent class, however, the effects were significantly stronger in the lower-risk classes. For example, in Latent Class 1, adolescents with early sexual debut were expected to have $e^{0.53}$ = 1.70 times more partners than those without early debut. In contrast, the effect of early sexual debut was significantly weaker in Latent Class 4, corresponding to $e^{0.09}$ = 1.09 times more partners. The effect of school connectedness, while significant overall, did not vary significantly across the latent classes.

The distribution of individuals across the four latent classes varied significantly across gender, race, and grade. Figure 2 provides the prevalence of the four latent classes across all levels of the individual characteristics considered in the model. Each set of bars in the figure sums to 1.0. For example, among all White adolescents (regardless of their gender or grade) the distribution across latent classes exactly mirrored the overall distribution shown in Table 1, specifically 53%, 31%, 12%, and 3% for Latent Classes 1 through 4, respectively. This distribution was considerably different, however, for the other racial groups. Among Black adolescents, the prevalence of Latent Classes 1 and 2 were nearly equal (43% and 37%, respectively), and the prevalence of Latent Class 3 was higher, at 16%, than for White adolescents. In contrast, among adolescents from other racial groups, the distribution was shifted toward lower risk, with the prevalence of Latent Class 1 equal to 63% and just 1% in Latent Class 4.

Discussion

A variety of regression modeling approaches are available for examining the associations between multiple factors related to sexual risk behavior. For example, the most straightforward approach is to simply include the risk factors as predictors of the outcome. The standard analysis reported in Table 1 provides a demonstration of what conclusions may be drawn from such an approach. That is, the population average effect of an individual risk factor on the outcome is estimated, holding all other predictors constant. For example, early sexual debut conferred the strongest risk for having more sexual partners by adulthood. Specifically, that level of risk was estimated to be 1.57 times for the overall population. To include a moderator such as race, as defined in this study to include three categories, ten interaction terms must be specified: early sexual debut by White, early sexual debut by Black, heavy episodic drinking by White, and so on. While such an approach could help to identify differential effects of risk factors (for example, whether the effect of early sexual debut varies across race groups), it would be difficult to consolidate the implications from so many coefficients in order to inform development of prevention interventions for various subgroups.

By moving to a finite mixture regression framework, heterogeneity in the regression estimates, including the intercept and the five slopes, was explained by a latent subgroup variable. In other words, rather than having three observed moderators (gender, race, and grade), this model incorporates a single latent moderator. This model allows us to further unpack these associations in the population, providing information about the processes involved in developing sexual risk behavior. Specifically, this approach allowed us to determine which risk factors play an important role in predicting sexual risk behavior within each of the latent subgroups. Then, by including individual characteristics as predictors of the latent class variable, we were able to obtain profiles of the individuals who comprise each latent class. This approach is much more holistic than a standard regression model that includes a large number of interaction terms, as it paints a picture of key subgroups of individuals in terms of their group-specific level of risk and processes leading to the risky behavior.

A primary finding based on the finite mixture regression model was that the strength of association between four of the five risk factors (all but school connectedness) and lifetime sexual partners systematically weakened across the latent classes. That is, with an increasing average number of partners, the absolute size of effects diminished. While exploring this phenomenon goes beyond the current study, we propose here several possible explanations. A more substantive hypothesis is that the factors that predict the number of partners is less well understood for the highest-risk individuals (i.e., Latent Class 4). It is possible that other predictors such as psychopathology may play a stronger role in this subgroup than that played by the risk factors known to play an important role in the overall sample. A more methodological hypothesis stems from the fact that unreliability in measurement attenuates correlations. It is possible that reported number of partners becomes substantially less reliable as the number increases. For example, an individual who has had, say, six partners can most likely remember them all by name, whereas an individual who has had dozens may not be able to report the exact number accurately. To the extent that measurement error increases as the number of lifetime partners increases, it is possible that measurement error is attenuating the effects of the predictors within the higher-risk classes (Brown & Sinclair, 1999).

Limitations and future directions

Several limitations and areas for future research merit mentioning. First, while the overall amount of variance explained by the four-class regression model was extremely high (94.2%), the amount explained within each latent class was somewhat lower, ranging from 27.7% for Class 4 and 62.3% for Class 2. The within-class regression model could be improved by including other known predictors of sexual risk behavior such as parental monitoring (Markham et al., 2010) or depressive tendencies (Lehrer, Shrier, Gortmaker, & Buka, 2006). Second, the current analysis did not incorporate survey weights. This issue will be important to consider in future finite mixture regression analyses relying on data such as these, collected using a complex survey design, to draw inference. Third, as found in Kan et al. (2010), there is variability in the acquisition of new partners over a lifetime and, therefore, different predictors might be highlighted at different times of life (early adolescence, late adolescence, early adulthood, late adulthood). The current study could be extended by examining the outcome over time, possibly with time-specific effects of the risk factors. Finally, as previously mentioned, our outcome of lifetime number of sexual partners is not in and of itself a risk. Our understanding of factors related to high-risk sexual behavior among young adults would be enhanced by investigating more nuanced outcomes that incorporate partner characteristics (e.g., HIV/STI status) or distinguish between concurrent or serial relationships and whether or not condoms were used consistently with all partners.

Acknowledgments

This study was supported by Award Number P50-DA010075 from the National Institute on Drug Abuse. The authors wish to thank colleagues at the Methodology Center at Penn State for helpful feedback on an early draft of this manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse or the National Institutes of Health. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

References

Bandura, A. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.

Brown NR, Sinclair RC. Estimating number of lifetime sexual partners: Men and women do it differently. The Journal of Sex Research. 1999; 36:292–297.

Buhi ER, Goodson P. Predictors of adolescent sexual behavior and intention: A theory-guided systematic review. Journal of Adolescent Health. 2007; 40:4–21. [PubMed: 17185201]

Collins, LM.; Lanza, ST. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. New York: Wiley; 2010.

DiClemente RJ, Crittenden CP, Rose E, Sales JM, Wingood GM, Crosby RA, Salazar LF. Psychosocial predictors of HIV-associated sexual behaviors and the efficacy of prevention interventions in adolescents at-risk for HIV infection: What works and what doesn't? Psychosomatic Medicine. 2008; 70:598–605. [PubMed: 18541908]

Harris KM, Halpern CT, Whitsel E, Hussey J, Tabor J, Entzel P, Udry JR. The National Longitudinal Study of Adolescent Health: Research Design. 2009 Retrieved from http://www.cpc.unc.edu/projects/addhealth/design.

Kan ML, Cheng YA, Landale NS, McHale SM. Longitudinal predictors of change in number of sexual partners across adolescence and early adulthood. Journal of Adolescent Health. 2010; 46:25–31. [PubMed: 20123254]

Kaplan D. Finite mixture dynamic regression modeling of panel data with implications for dynamic response analysis. Journal of Educational and Behavioral Statistics. 2005; 30:169–187.

Kirby, D. Emerging Answers: Research Findings on Programs to Reduce Teen Pregnancy and Sexually Transmitted Diseases. Washington, D.C.: The National Campaign to Prevent Teen Pregnancy; 2007.

Lehrer JA, Shrier LA, Gortmaker S, Buka S. Depressive symptoms as a longitudinal predictor of sexual risk behaviours among US middle and high school students. Pediatrics. 2006; 118:189–200. [PubMed: 16818565]

Markham CM, Lormand D, Gloppen KM, Peskin MF, Flores B, Low B, House LD. Connectedness as a predictor of sexual and reproductive health outcomes for youth. Journal of Adolescent Health. 2010; 46:S23–S41. [PubMed: 20172458]

McLachlan, GJ.; Peel, D. Finite Mixture Models. New York: Wiley; 2000.

Sandfordt TGM, Orr M, Hirsch JS, Santelli J. Long-term health correlates of timing of sexual debut: Results from a national US study. American Journal of Public Health. 2008; 98:155–161. [PubMed: 18048793]

Santelli JS, Brener ND, Lowry R, Bhatt A, Zabin LS. Multiple sexual partners among U.S. adolescents and young adults. Family Planning Perspectives. 1998; 30:271–275. [PubMed: 9859017]

Strachman A, Impett EA, Henson JM, Pentz MA. Early alcohol use and sexual experiences by emerging adulthood: A 10-year longitudinal investigation. Journal of Adolescent Health. 2009; 45:478–482. [PubMed: 19837354]

Vermunt, JK.; Magidson, J. Technical Guide for Latent GOLD 4.0: Basic and Advanced. Belmont, MA: Statistical Innovations Inc.; 2005.

Wedel M, DeSarbo WS. A mixture likelihood approach for generalized linear models. Journal of Classification. 2002; 12:21–55.

Wedel, M.; Kamakura, WA. Market segmentation: Conceptual and methodological foundations. 2nd ed.. Boston, MA: Klewer Academic; 2001.

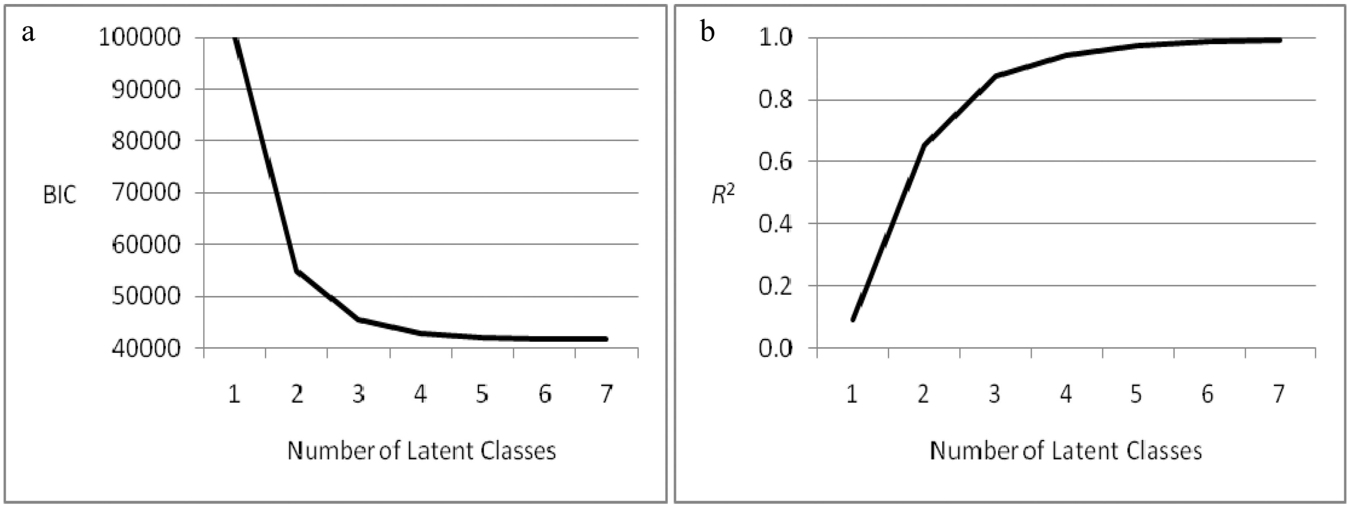

Figure 1.
Plots for Models With One Through Seven Latent Classes: (a) BIC and (b) Total Amount of Variance in Number of Sexual Partners Explained by Model

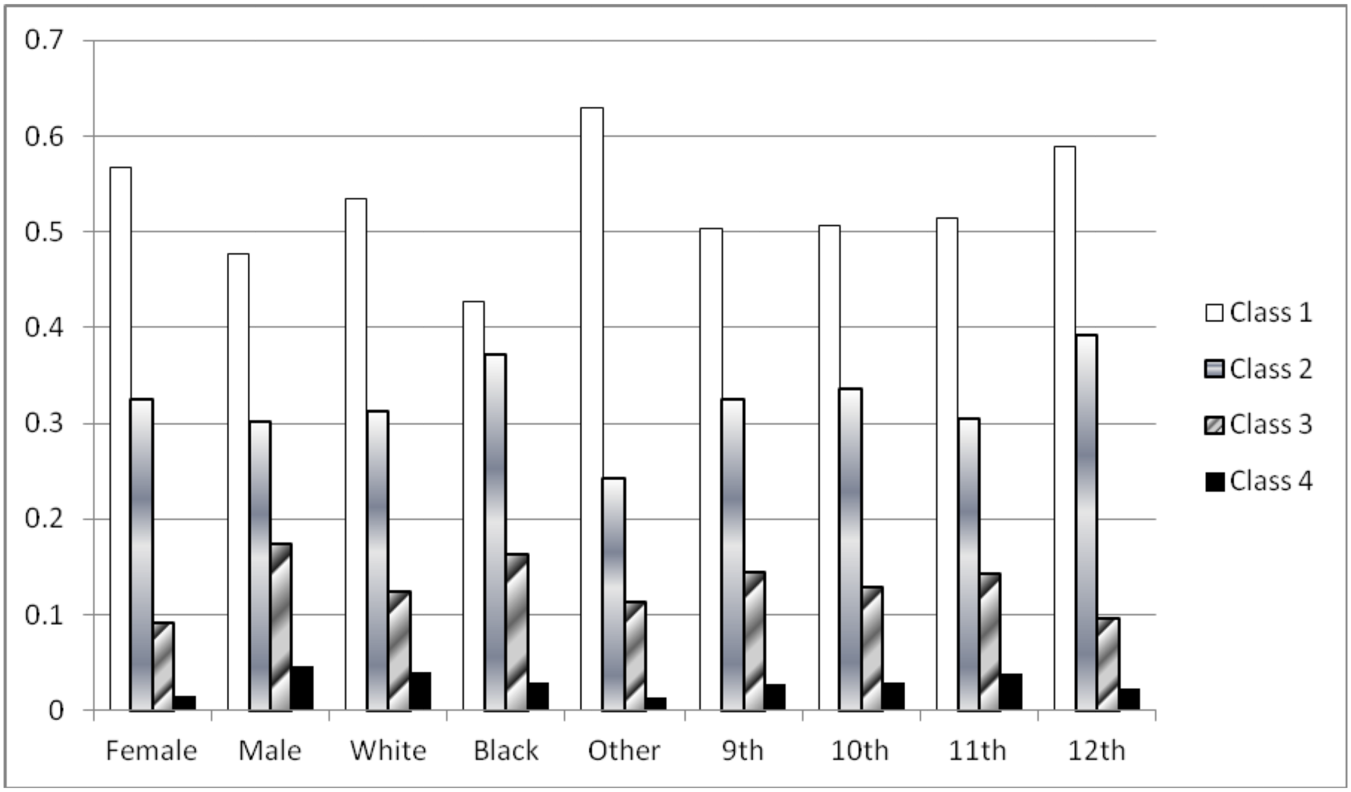

Figure 2.
Marginal Latent Class Prevalences Across Levels of Gender, Race, and Grade

Table 1

Parameter estimates from the standard Poisson regression and from the four-class mixture regression model

	Standard Poisson Regression	Finite Mixture Regression Model			
		Class 1	Class 2	Class 3	Class 4
Proportion in group	1.00	0.53	0.32	0.13	0.03
Mean number of partners	12.85	4.12	13.09	32.15	79.91
R^2					
Overall	0.09	0.94			
Within latent class	n/a	0.38	0.62	0.48	0.28
Intercept	2.32	1.11	2.29	3.29	4.30
Regression coefficients					
Early sexual debut	0.45*	0.53‡	0.47	0.32	0.09
Heavy episodic drinking	0.27*	0.43‡	0.41	0.26	0.16
School connectedness	0.02*	0.03	0.04	0.04	0.03
Positive consequences	0.21*	0.17‡	0.16	0.10	0.02
Negative consequences	−0.23*	−0.18‡	−0.16	−0.11	−0.06
Profile characteristics					
Female	n/a	0.58	0.56	0.38	0.29
Male	n/a	0.42	0.44	0.62	0.71
Black	n/a	0.14	0.21	0.22	0.24
White	n/a	0.72	0.70	0.68	0.71
Other race	n/a	0.14	0.09	0.10	0.05
Grade 9	n/a	0.25	0.27	0.30	0.25
Grade 10	n/a	0.25	0.27	0.25	0.25
Grade 11	n/a	0.25	0.25	0.28	0.33
Grade 12	n/a	0.25	0.21	0.17	0.17

* Predictor is significant ($p < .0001$) in standard Poisson regression model and in finite mixture regression model.

‡ Effect of predictor varies significantly across latent classes ($p < .0001$).

Table 2

A comparison of competing finite mixture regression models

Number of Latent Classes	Log-Likelihood	BIC	Number of Parameters	Degrees of Freedom	Classification Error	Overall R^2
1	−49964.3	99980.8	6	6036	.000	.094
2	−27390.2	54919.7	16	6026	.011	.655
3	−22662.7	45551.7	26	6016	.028	.874
4	**−21280.1**	**42873.6**	**36**	**6006**	**.082**	**.942**
5	−20815.2	42030.9	46	5996	.120	.975
6	−20682.8	41853.2	56	5986	.212	.985
7	−20634.8	41844.1	66	5976	.239	.990

Note. Bold font signifies selected model.