

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

Stephanie T. Lanza<sup>\*1,2</sup>, Kari C. Kugler<sup>1</sup> and Charu Mathur<sup>1</sup>

<sup>1</sup>The Methodology Center, The Pennsylvania State University, University Park, PA 16802, USA

<sup>2</sup>The College of Health and Human Development, The Pennsylvania State University, University Park, PA 16802, USA

**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, interpretation can become difficult as risk factors and moderators are added to the model. The current study presents finite mixture regression as an alternative approach, in which population subgroups are identified based on the pattern of associations between multiple risk and protective 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 and protective factors (early sexual debut, past-year binge 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. Males and Black adolescents were significantly more likely to be in subgroups characterized by higher-risk behavior. Results suggest that prevention programs focused on mediating these particular risk factors may be most effective for adolescents at lower risk for later engaging in risky sexual behavior; however, for the subgroup of adolescents who go on to have the most sexual partners, the regression weights are significantly weaker, warranting further investigation into the most salient risk factors, and therefore targets for intervention programs, for high-risk adolescents.

**Keywords:** Adolescent risk factors, Lifetime sexual partners, HIV prevention, Mixture model.

## INTRODUCTION

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 [1, 2]. 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; [3]). The identification of these factors has been critical in designing and implementing effective risk-reduction programs [2]; however, as noted in a recent review article, there is room for improvement, especially in terms of conducting more rigorous methodological assessments [1]. Furthermore, the lack (until recently) of longitudinal data sets and advanced statistical methods has precluded the ability to test the underpinnings of behavioral theory (i.e., that adolescent cognitions, behaviors, and environments influence behaviors well into adulthood). 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 and protective 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 [4], which increases the risk for unintended pregnancies and transmission of sexually transmitted infections, including HIV.

The risk and protective factors used in the study were inspired by a social cognitive framework, which posits that behavior is the result of interactions between individual cognitions, personal behavior, and the environment [5]. Within the cognitive domain, we investigated motivations for engaging in sexual intercourse, specifically perceived 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 [1], including number of sexual partners [6]. 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 [7], thus warranting further exploration. Finally, there is substantial evidence that other behaviors such as early use of alcohol [2, 4] and early sexual debut [8] 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

\*Address correspondence to this author at the Methodology Center, The Pennsylvania State University, 204 E. Calder Way, Suite 400, State College, PA 16801, USA; Tel: 814-865-7095; Fax: 814-863-0000; E-mail: SLanza@psu.edu

debut [9]. The environmental factor considered in the study was school connectedness.

## METHODOLOGICAL APPROACHES TO MODELING RISK FACTORS

Most of the literature to date relating risk and protective factors to sexual risk behavior is based on regression models, which 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 adolescents' perceived positive consequences of having intercourse on lifetime sexual partners at age 30, including a term for the interaction between gender and perceived positive consequences permits an examination of whether that effect 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 [10-12] which can be used to identify population subgroups of individuals for whom certain risk and protective factors are most salient. If targeted in an intervention program, these factors may be most effective at reducing the problem behavior. Finite mixture regression is similar in spirit to the more familiar latent class analysis [13], which is used to identify latent subgroups on the basis of responses to a set of categorical items, usually from one time point. The difference is that indicators in LCA are actual observed variables, but the indicators in finite mixture regression 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 in which a scale for perceived 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 of perceived positive consequences of intercourse 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, in which the subgroups differ in terms of their intercept (i.e., mean number of partners), the effect of perceived 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 latent class 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 latent classes. However, we can identify potential targets for the intervention by including directly

measured characteristics (e.g., age, gender, race) as predictors of latent class membership. Then, the class-specific regression models can inform how to tailor the intervention program to the different latent classes based on the most salient predictors.

## CURRENT STUDY

The current study presents an empirical demonstration of finite mixture regression, in which the total number of sexual partners in adulthood is predicted by early sexual debut, past-year binge drinking, school connectedness, perceived positive consequences of having intercourse, and perceived negative consequences of having intercourse. Gender, race, and grade are examined as predictors of latent class membership. 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.

## METHOD

### Participants

Participants were drawn from the National Longitudinal Study of Adolescent Health [14]. 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 was representative of US schools with respect to region of country, urbanicity, school size, school type, and ethnicity. Of the 15,701 respondents at Wave 4, 14,800 were available for a longitudinal analysis linking risk factors at Wave 1 to behavior at Wave 4 and had longitudinal weights appropriate for analysis from Waves 1 to 4. The current study was limited to high school students (i.e. adolescents who were in ninth through twelfth grade at Wave 1) who had complete data on all variables used in the models ( $N = 10,562$ ; unweighted median age of 16; 52.8% female; 61.5% White, 20.7% Black, 17.8% other). 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, assessed at Wave 4 when participants were adults, ranged from zero to 900 with a weighted mean of 14.2 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 = 13.3; see Fig. 1). It is important to note that in this study we included both same-sex and opposite-sex partners in this count; thus, we expect the mean number of partners reported here to be slightly higher than those reported in many other studies.

**Risk and Protective Factors.** Five risk and protective factors for sexual risk behavior, 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

## Distribution of Lifetime Partners

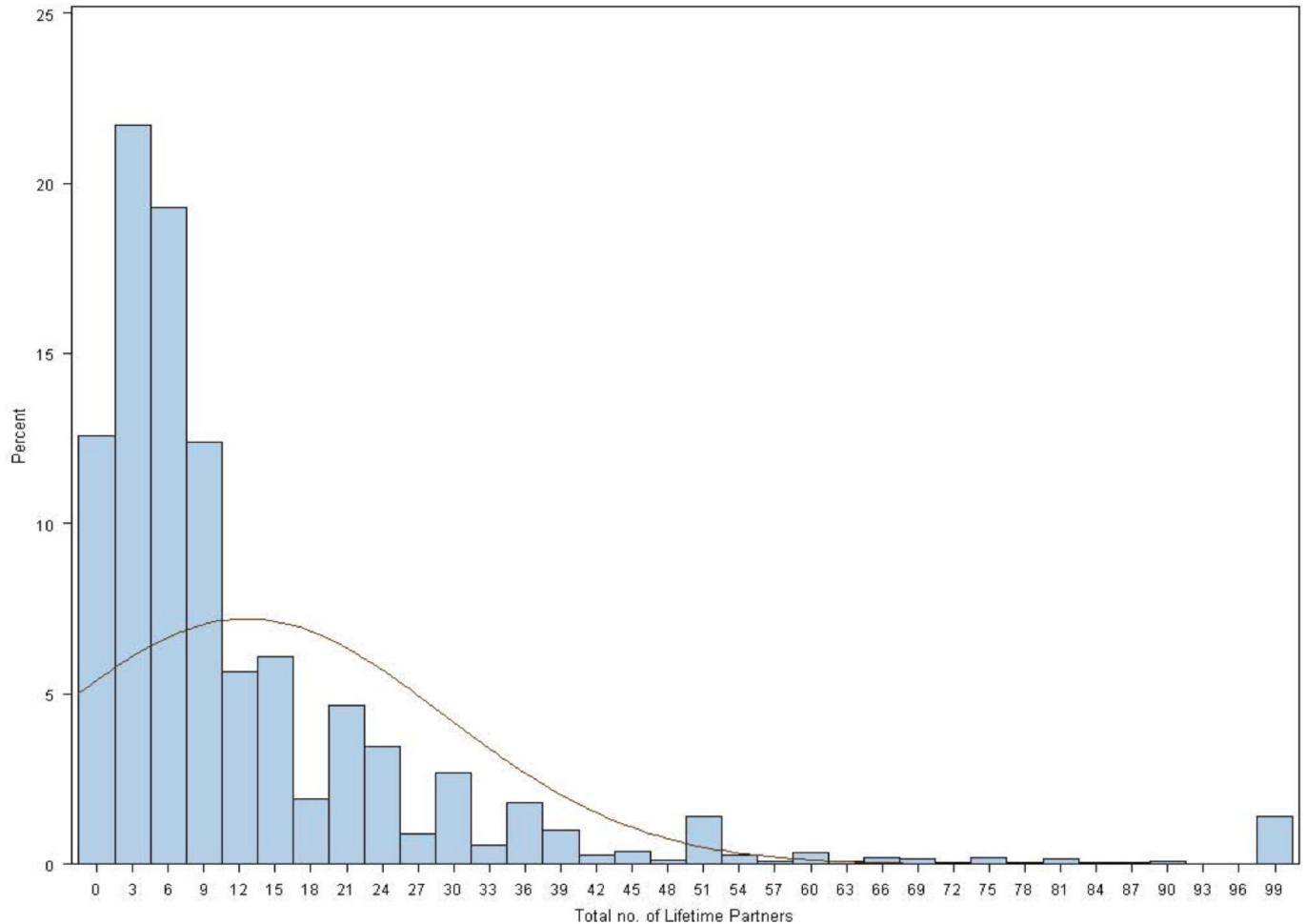


Fig. (1). Distribution of Lifetime Partners.

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. *Past-year binge 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 five items coded on five-point scales ranging from 1 (*strongly agree*) to 5 (*strongly disagree*): you feel close to people at your school, you feel like you are part of your school, you are happy to be at your school, the teachers at your school treat students fairly, and you feel safe in your school. The scale was standardized, with high scores reflecting more risk ( $\alpha = .75$ ). A factor analysis of a set of items measuring motivations to engage in risky behaviors suggested two unique scales. *Perceived 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 ( $\alpha = .84$ ). *Perceived 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 ( $\alpha = .82$ ). 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. Weighted proportions of these three characteristics are *gender* (49.4% female, 50.6% male), *race* (reduced to three categories: 16.1% Black, 59.7% White, and 11.9% other racial groups), and *grade in school* at Wave 1 (26.2% in Grade 9, 24.9% in Grade 10, 23.5% in Grade 11, and 25.3% 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 in which 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, past-year binge drinking, school connectedness, perceived positive consequences, and perceived 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 and protective factors, providing a basis for comparison between a standard regression model and the proposed finite mixture regression model.

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 and protective 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 model explains variability in lifetime sexual partners [15]. 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 [15] 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, in which the number of sexual partners

in adulthood was predicted from the five risk and protective factors assessed during high school. The amount of variance explained in the outcome by these five predictors was 10%. The intercept, 2.33, 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 perceived positive and negative consequences, the log of the expected number of partners in adulthood was 2.33. Exponentiating this value corresponds to  $e^{2.33} = 10.28$  mean lifetime partners for such adolescents.

The remaining coefficients reflect the difference in the log 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 log of expected number of partners was 0.43 units higher for those who had an early sexual debut. This implies that, holding other predictors constant, adolescents engaging in early intercourse have  $e^{(2.33+0.43)} = 15.80$  mean lifetime partners. In other words, the main effect of early sexual debut was  $e^{0.43} = 1.54$  times more partners (i.e., 54% more partners). Similarly, past-year binge drinking corresponded to  $e^{0.30} = 1.35$  times more partners; one standard deviation higher on the school connectedness risk factor corresponded to  $e^{0.04} = 1.04$  times more partners; one standard deviation higher on perceived positive consequences corresponded to  $e^{0.19} = 1.21$  times more partners; and perceived negative consequences corresponded to  $e^{-0.21} = 0.81$  times as many partners.

### Finite Mixture Regression

In contrast to the standard Poisson regression analysis, in finite mixture regression a complete set of regression coefficients, including intercept and slopes, is estimated within each latent class. In addition, latent class membership

**Table 1. Parameter Estimates from the Standard Poisson Regression and 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	0.52	0.32	0.13	0.03
Mean number of partners	13.32	4.17	13.35	33.42	78.31
<b>Proportion of Variance Explained in Outcome</b>					
Overall	0.10	0.94			
Within latent class	n/a	0.38	0.61	0.52	0.34
<b>Regression Coefficients (Standard Errors)</b>					
Intercept	2.33 (.02)*	1.09 (.04)‡	2.29 (.04)	3.29 (.03)	4.26 (.04)
Early sexual debut	0.43 (.04)*	0.47 (.05)‡	0.38 (.05)	0.30 (.05)	0.11 (.05)
Past-year binge drinking	0.30 (.04)*	0.48 (.06)‡	0.45 (.06)	0.32 (.06)	0.17 (.04)
Low school connectedness	0.04 (.02)*	0.06 (.02)	0.05 (.03)	0.01 (.03)	0.00 (.03)
Perceived positive consequences	0.19 (.02)*	0.14 (.04)‡	0.14 (.04)	0.09 (.03)	0.02 (.02)
Perceived negative consequences	-0.21 (.02)*	-0.16 (.03)	-0.16 (.03)	-0.12 (.04)	-0.09 (.03)

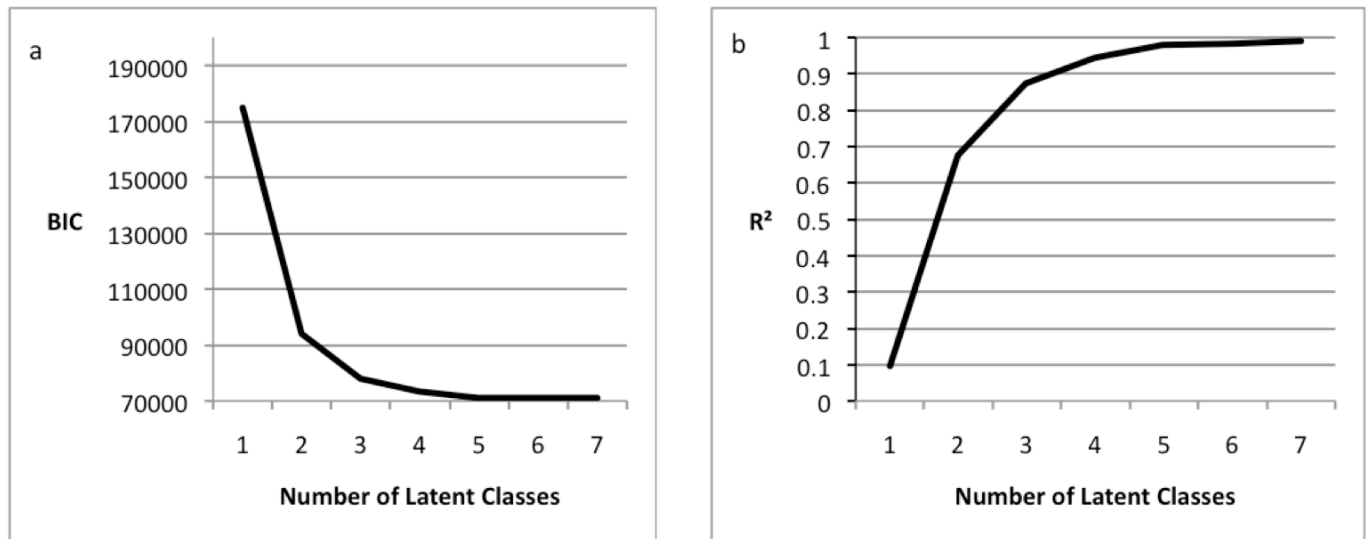
\* Coefficient is significant ( $p < .05$ ) in standard Poisson regression model and in finite mixture regression model.

‡ Coefficient varies significantly across latent classes ( $p < .05$ ).

**Table 2. A Comparison of Competing Finite Mixture Regression Models**

Number of Latent Classes	Log-Likelihood	AIC	BIC	Number of Parameters	Degrees of Freedom	Entropy	Overall R <sup>2</sup>
1	-87402.8	174817.6	174860.9	6	10171	1.000	.096
2	-46975.9	93977.8	94071.8	13	10164	.953	.675
3	-38798.4	77636.8	77781.4	20	10157	.903	.872
<b>4</b>	<b>-36473.7</b>	<b>73001.3</b>	<b>73196.5</b>	<b>27</b>	<b>10150</b>	<b>.826</b>	<b>.941</b>
5	-35680.6	71429.1	71674.9	34	10143	.782	.978
6	-35426.0	70934.0	71230.4	41	10136	.683	.983
7	-35314.1	70724.2	71071.1	48	10129	.664	.990

Note. Bold font signifies selected model.

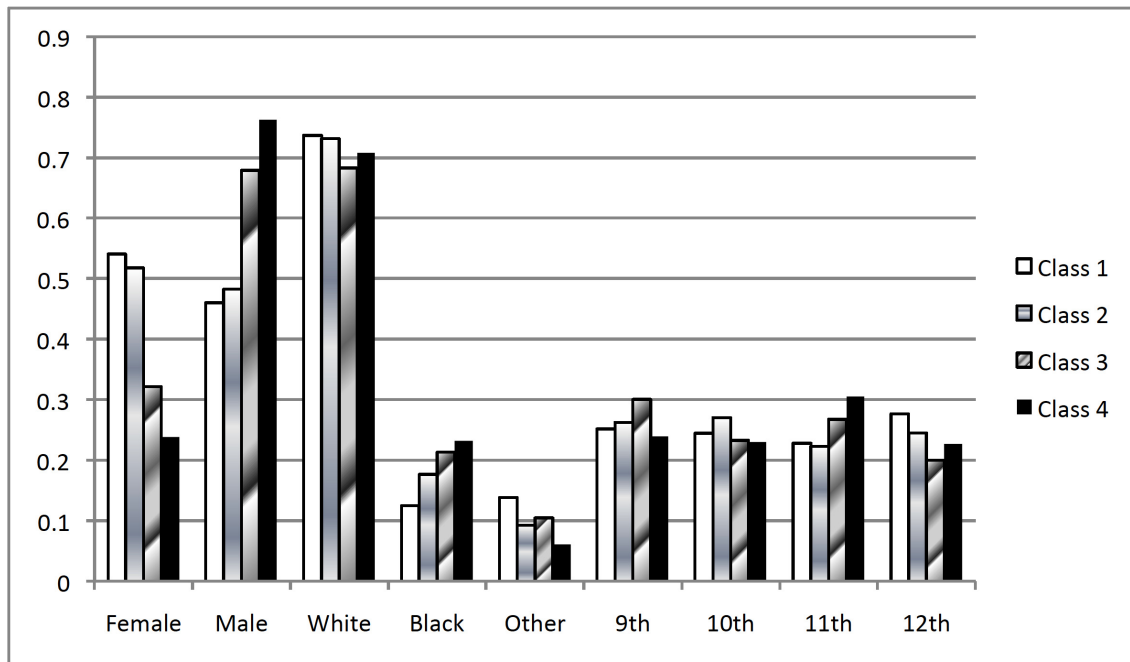


**Fig. (2).** Plots for Models With One Through Seven Latent Classes: (a) BIC and (b) Total Amount of Variance in Number of Sexual Partners Explained by Model.

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. Although the information criteria suggested that the model with seven latent classes had the most optimal balance between fit and parsimony among the models considered, there were only small, incremental improvements when moving from a four-class model to more complex ones (see Fig. 2, Panel a for a plot of the BIC values). Entropy 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 one are preferred, although some amount of classification 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 10% of the variance. Fig. 2, Panel b plots the overall  $R^2$  for the models under consideration. Based on the information presented in Table 2 and Fig. (2) and a careful inspection of the parameter estimates from competing models, we selected the four-class model to describe

heterogeneity both in adult sexual risk behavior 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, variance explained by the regression model within each latent class, 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 increasing mean lifetime sexual partners. Latent Class 1 comprised 52% of the sample, with 4.17 mean lifetime partners. Latent Class 2 included 32% of the sample, with 13.35 mean lifetime partners. Similarly, Latent Classes 3 comprised 13% of the sample, with 33.42 mean lifetime partners, and Latent Class 4 (adults with many partners) included just 3% of the sample, with 78.31 mean lifetime partners. 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  $e^{1.09} = 2.97$  lifetime partners. The corresponding values for Latent



**Fig. (3).** Profiles of the Latent Classes: Characteristics of Individuals Comprising Each Latent Class.

Classes 2, 3, and 4 were 9.87, 26.84, and 70.81 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 ( $p < .05$ ) across the latent classes for the following three predictors: early sexual debut, past-year binge drinking, and perceived positive 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 and protective 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  $e^{0.47} = 1.60$  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  $e^{0.11} = 1.12$  times more partners. The effects of low school connectedness and perceived negative consequences, while significant overall, did not vary significantly across the latent classes.

Predictors were included in the four-class model so that the effects of covariates on latent class membership could be discerned. Note that the estimation of the mixture regression model and prediction of membership in the latent classes was done simultaneously. The distribution of individuals across the four latent classes varied significantly across gender, race, and grade. Fig. (3) provides information about the characteristics of individuals comprising each latent class. Each bar represents the proportion of individuals in a particular latent class with the corresponding characteristic. Among individuals in Latent Classes 1 and 2, the proportion of males and females were nearly equal, whereas among those in Latent Classes 3 and 4 there were substantially more males than females. The latent classes did not differ in terms

of the proportion of whites, but Latent Class 4 had twice as many Black individuals as Latent Class 1. In contrast, Latent Class 4 had half as many individuals from other racial groups compared to Latent Class 1.

## DISCUSSION

A variety of regression modeling approaches are available to examine the associations between multiple factors related to sexual risk behavior. For example, the most straightforward approach is to simply include the risk and protective factors as predictors of the outcome. The standard analysis reported in Table 1 provides a demonstration of conclusions that can be drawn from such an approach. That is, the population average effect of an individual risk or protective 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 early behavior increased the expected lifetime number of partners by 1.54 times. To include a moderator such as race (as operationalized in this study) would require ten interaction terms to be specified: early sexual debut by White, early sexual debut by Black, past-year binge drinking by White, and so on. While such an approach could help 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 the development of 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 comprised 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 group-specific processes leading to the risky behavior.

We identified four subgroups of adults, defined not only by their mean number of lifetime sexual partners, but also by the relative salience of the different risk and protective factors considered here. The predictive strength of three of these factors systematically and significantly weakened across the latent classes. One possible explanation is that there is a unique set of risk factors driving the lifetime number of partners among adults in Latent Class 4 (adults with many partners). For example, low self-esteem – which was not included as a risk factor in this study – may be an important risk factor for this subgroup but not for subgroups defined by lower risk. A more methodological hypothesis is 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 [16].

Interesting trends emerged in relation to the individual characteristics, providing some insight into how intervention programs might be tailored for particular population subgroups. The proportion of men comprising a latent class increased significantly across the classes, from 46% to 76% (see Fig. 3). This finding is consistent with data from the National Survey of Family Growth showing that males have higher numbers of lifetime partners than females [17]. In terms of racial groups, the proportion of White adults was quite stable across latent classes, while the proportion of Blacks in the latent classes increased and the proportion of Other race groups decreased. Again, this is consistent with the findings from the National Survey of Family Growth that found a higher percentage of African Americans reporting 15 or more partners in their lifetime than Whites or Hispanics (females: 11.3%, 8.9%, 4.4% and males: 30.0%, 20.9%, 19.1%, respectively) [17]. The unique contribution of this study compared to previous studies, however, is the holistic picture of the effect of multiple risk factors for the different subgroups. Finally, the lack of differences across grades suggests that interventions targeting high school students need not differentiate between age groups.

## LIMITATIONS OF THE CURRENT STUDY 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%), the amount explained within each latent class was somewhat lower, ranging from 34% for Class 4 to 61% for Class 2. The within-class regression model could be improved by including other known predictors of sexual risk behavior such as parental monitoring [7] or depressive tendencies [18]. Second, it has been shown that there is variability in the acquisition of new partners over a lifetime and, therefore, different predictors might be highlighted at different times of life (e.g., late adolescence, early adulthood, adulthood) [6, 19]. 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. Thus, merely reducing the lifetime number of sexual partners might not provide substantial public health benefit if safe sex is not practiced. Despite these limitations, however, the results from the current study provide important information for guiding tailored preventive interventions, and also for adding to our set of statistical tools for answering more holistic, process-oriented research questions, particularly that of the development of sexual risk behavior.

## ACKNOWLEDGEMENTS

This study was supported by Award Number P50-DA010075 from the National Institute on Drug Abuse. The authors wish to thank Amanda Applegate and other 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.

## CONFLICT OF INTEREST

None declared.

## REFERENCES

- [1] Buhi ER, Goodson P. Predictors of adolescent sexual behavior and intention: A theory-guided systematic review. *J Adolescent Health* 2007; 40: 4-21.
- [2] DiClemente RJ, Crittenden CP, Rose E, *et al.* 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? *Psychosom Med* 2008; 70: 598-605.
- [3] 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.
- [4] 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: 271-5.
- [5] Bandura A. *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall 1986.
- [6] Kan ML, Cheng YA, Landale NS, McHale SM. Longitudinal predictors of change in number of sexual partners across adolescence and early adulthood. *J Adolescent Health* 2010; 46: 25-31.
- [7] Markham CM, Lormand D, Gloppen, KM, *et al*, LD(2010) Connectedness as a predictor of sexual and reproductive health outcomes for youth. *J Adolescent Health* 2010; 46: S23-S41.
- [8] Sandfordt TGM, Orr M, Hirsch JS, Santelli, J. Long-term health correlates of timing of sexual debut: Results from a national US study. *Am J Public Health* 2008; 98: 155-161.
- [9] Strachman, A, Impett, EA, Henson, JM, Pentz, MA. Early alcohol use and sexual experiences by emerging adulthood: A 10-year longitudinal investigation. *J Adolescent Health* 2009; 45: 478-82.
- [10] McLachlan GJ, Peel, D. *Finite Mixture Models.* New York: Wiley 2000.
- [11] Wedel M, DeSarbo WS. A mixture likelihood approach for generalized linear models. *J Classif* 1995; 12: 21-55.
- [12] Wedel M, Kamakura WA. *Market segmentation: Conceptual and methodological foundations (2<sup>nd</sup> ed.).* Boston, MA: Kluwer Academic 2001.
- [13] Collins LM, Lanza ST. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences.* New York: Wiley 2010.
- [14] Harris KM, Halpern CT, Whitsel E, *et al.* The national longitudinal study of adolescent health: Research design. 2009. Retrieved from <http://www.cpc.unc.edu/projects/addhealth/design>.
- [15] Vermunt JK, Magidson J. *Technical guide for Latent GOLD 4.0: Basic and advanced.* Belmont, MA: Statistical Innovations Inc 2005.
- [16] Brown NR, Sinclair RC. Estimating number of lifetime sexual partners: Men and women do it differently. *J Sex Res* 1999; 36: 292-7.
- [17] Chandra A, Mosher WD, Copen C, Sionean C. *Sexual behavior, sexual attraction, and sexual identity in the United States: Data from the 2006-2008 National Survey of Family Growth.* National health statistics reports, No. 36. Hyattsville, MD: National Center for Health Statistics 2011.
- [18] Lehrer JA, Shrier LA, Gortmaker S, Buka S. Depressive symptoms as a longitudinal predictor of sexual risk behaviors among US middle and high school students. *Pediatrics* 2006; 118: 189-200.
- [19] Murphy DA, Brecht M, Herbeck DM, Huang D. Trajectories of HIV risk behavior from age 15 to 25 in the National Longitudinal Survey of Youth Sample. *J Youth Adolescence* 2009; 38: 1226-39.

Received: April 6, 2011

Revised: June 30, 2011

Accepted: June 30, 2011

© Lanza *et al.*; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.