Classification Error in Measuring Sexual Victimization among Inmates: The National Inmate Survey

Marcus E. Berzofsky[1], Susan Edwards[1], Paul P. Biemer[1]
[1]RTI International, 3040 Cornwallis Rd., RTP, NC 27709

Abstract
All survey estimates are subject to measurement error – classification error in the case of categorical outcomes. This is especially true of sensitive outcomes such as sexual victimization. The National Inmate Survey (NIS), sponsored by the U.S. Bureau of Justice Statistics, is a nationally representative survey of inmates in prisons and jails which measures two types of sexual victimization – inmate-on-inmate and staff sexual misconduct with an inmate. This paper builds on the research of Berzofsky, Biemer, and Kalsbeek (2014) to present the results of a latent class analysis (LCA) designed to assess the measurement error in each type of sexual victimization. LCA uses multiple indicators of a construct embedded in the survey instrument to estimate the false positive and false negative probabilities in each indicator. Due to the rare nature of sexual victimization among inmates, our analysis combines data from the 2007-08 NIS and the 2009 NIS in order to achieve adequate precision in the results. One issue with LCA is how missing data for indicators and grouping variables are taken into account. Traditionally, if either type of variable was missing the model would use listwise deletion and remove the case from the model. Newer software, such as LatentGold, incorporates full information maximum likelihood (FIML) for dependent variables to utilize all records. This paper assesses the impact that the inclusion of cases with missing data have on the LCA estimates. Using MAR adjustments for missing data, we found evidence that inmates who do not respond to all indicators are more likely to be victims and more likely to not provide truthful responses for the items they do answer.

Key Words: classification error, measurement error, National Inmate Survey (NIS), sexual victimization, rare event, sensitive event

1. Background

1.1 Background

The National Inmate Survey (NIS) is the first nationally representative survey of adult inmates designed to measure the prevalence rate of sexual victimization among inmates in prison and local jails. Created through the Prison Rape Elimination Act of 2003 (PREA; P.L. 108-79) and sponsored by the Department of Justice, Bureau of Justice Statistics, the NIS attempts to measure the 12-month prevalence of sexual victimization among adult inmates that is perpetrated by either another inmate or a staff member both at the facility level and the national level.

The NIS has been conducted three times. The first NIS was conducted in 2007-08 (NIS-1); the second in 2009 (NIS-2); and the third in 2011-12 (NIS3; Beck, 2014). Table 1 present the number of facilities that participated and the number of inmates interviewed by facility type and NIS cycle (Beck and Harrison, 2007; Beck and Harrison, 2008; Beck, Harrison, Berzofsky, Krebs, and Caspar, 2010; Beck, Harrison, Berzofsky, Krebs and Caspar, 2013).

Table 1: Number of Facilities Participating and Inmates Interviewed by Facility type and NIS cycle

	Prisons		Local Jails	
Cycle	Number of Facilities	Number of Interviews	Number of Facilities	Number of Interviews
NIS-1	146	23,398	282	40,419
NIS-2	167	29,954	310	45,126
NIS-3	233	38,251	358	52,926

Each cycle of the NIS utilized a two-stage sample design. In the first stage a stratified sample of facilities was selected with probability proportionate to size (PPS) based on the number of inmates held in the facility. In the second stage a simple random sample of inmates was selected from a roster of inmates provided by the facility just prior to data collection. The response rates in each NIS cycle were 72%, 71%, and 60% in NIS-1, NIS-2, and NIS-3, respectively (Beck and Harrison, 2007; Beck, et. al., 2010; Beck, et. al., 2013).

1.2 Study Motivation

Previous studies of sexual victimization among inmates in prison or jail were methodologically flawed in that they were not probability based or did not utilize weights to take into account differences between respondents and nonrespondents (see, for example, Struckman-Johnson and Struckman-Johnson, 2000; Gaes & Goldberg, 2004; Tewksbury, 1989; and Hensley, Tewksbury, and Castle, 2003) and, therefore, not generalizable to all inmates. Because no prior nationally representative estimates of sexual victimization existed, the study designers were concerned about validating the NIS findings without a basis for comparison.

Moreover, given the sensitive nature of the survey, BJS was concerned about any estimate being highly susceptible to measurement error. In general, research has found that sensitive events such as crime victimization, drug use, or unemployment status have high false negative rates (i.e., the respondent indicates the outcome of interest did not occur when it truly did) and negligible false positive rates (i.e., the respondent is unlikely to indicate the event occurred when it did not) (See, for example, Berzofsky (2011) for analysis of crime victimization; Biemer and Wiesen (2002) for analysis of drug use; and Biemer and Bushery (2001) for analysis of unemployment status). However, for the NIS, there was equal concern over false negative and false positive errors. For example, if an inmate feared retribution from his/her perpetrator then he/she may be reluctant to indicate a victimization occurred (i.e., a false negative response).On the other hand, an inmate may think that falsely reporting a victimization will somehow bring harm to the facility and thus may indicate a victimization occurred even when one did not (i.e., a false positive response).

Because of these concerns, study designers incorporated items in the NIS instrument that would allow classification errors to be estimated via latent class analysis (LCA). LCA is a modeling technique that does not require a gold standard estimate to estimate classification error. Rather LCA uses multiple indicators for the latent construct of interest to simultaneously model the structural components (i.e., latent true values) and measurement component (i.e., classification error parameters rates; Biemer, 2011). LCA does have one key assumption: local independence. Local independence holds when the probability of response for one indicator conditioned on the latent construct does not dependent on the response to any other indicator. Written in terms of probabilities, local independence for indicators A and B, and a latent variable X, occurs when

$$P(AB|X) = P(A|X)P(B|X)$$

Local dependence can occur when one of the following three conditions is not met (Berzofsky, Biemer, and Kalsbeek, 2008):

- *Univocality*. all indicators are fully correlated with the latent construct,
- *Group homogeneity*. within a group of persons, the classification error rates are equal, and
- *Zero behavioral correlation*. the correlation between any pair of indicators conditioned on the latent variable is zero.

The NIS has two latent constructs of interest: an inmate's status of being victimized by another inmate (inmate-on-inmate victimization) and an inmate's status of being victimized by a staff member (staff sexual misconduct). Inmate-on-inmate victimization is defined as a sexual contact or act that is experienced under force or pressure by another inmate. Staff sexual misconduct is defined as sexual contact or act with a facility staff member that is wanted or unwanted.

1.3 Study Considerations

On a national level sexual victimization among inmates is a rare event (see, Beck, et. al, 2013). Rare events are often difficult to model using LCA because of sparseness in the data causing weak identifiability and model convergence issues (Berzofsky and Biemer, 2012). Table 2 shows the number of victimizations, unweighted and weighted, and prevalence for inmate-on-inmate victimization in prisons by gender of inmate, and survey cycle. As can be seen, the inmate-on-inmate victimization rate is below 8% for both males and females in all cycles.

Because of the rare nature of victimization two issues needed to be considered prior to conducting the analysis. First, to increase the number of true positives in the analysis data set, it was necessary to combine cycles to promote model convergence. As can be inferred from Table 2, victimization rates within a gender are not statistically different across cycles. Therefore, we believe that combining data across cycles will not introduce undesirable heterogeneity. Second, we also assessed whether data across gender can be combined. Note from Table 2 that the prevalence rates between males and females are statistically different in each cycle. Thus, the structural portion of the LCA model (i.e., the portion of the model that estimates the prevalence) should include gender to account for these significant differences.

Table 2: Number and Prevalence of Inmate-on-Inmate Victimization in Prisons by Gender of Inmate and NIS Cycle

	Male Inmates				Female Inmates			
NIS Cycle	**Unwgt victims**	**Unwgt non-victims**	**Wgt %**	**Wgt % SE**	**Unwgt victims**	**Unwgt non-victims**	**Wgt %**	**Wgt % SE**
NIS-1	327	19,193	1.59	0.14	145	1,934	6.55	0.77
NIS-2	370	22,214	1.70	0.21	291	4,995	5.33	0.59
NIS-3	475	28,842	1.40	0.13	470	5,856	7.60	0.84

1.4 Study Goals

The primary goal of this study is to assess the classification error rates of the measures for inmate-on-inmate sexual victimization and staff sexual misconduct in both prisons and jails. This paper focuses on inmate-on-inmate victimization in prisons.

Secondarily, assessing how to incorporate partially complete data records (i.e. records with item nonresponse) and the impact that those records had on model estimates was important due to concerns that non-respondents are likely victims. Therefore, techniques needed to be identified that allowed missing data to be included in the LCA model.

Based on these goals, for this paper, there were three main research questions:

1. Do the model estimates of victimization prevalence change when cases with missing data are included in the model? If so, to what degree?
2. Do the model estimates of classification error change when cases with missing data are included? If so, to what degree?
3. Is the estimate of inmate-on-inmate victimization biased due to classification error? If so, to what degree?

2. Methods

2.1 Latent Class Analysis in the NIS

The NIS incorporated five indicators for inmate-on-inmate victimization in order to assess classification error. Table 3 lists the indicators and their question characteristics. Indicator 1 is the official indicator used in all published NIS reports while indicators 2 – 5 are only used for LCA (see, Berzofsky, Biemer, and Kalsbeek (2014) for detailed definition of each indicator). However, in NIS-3 Indicator 4 and Indicator 5 were removed from the instrument to reduce survey length. Therefore, our analysis only used NIS-1 and NIS-2 data. Furthermore, the results in this paper are based on data from prison inmates only. An LCA based upon all three NIS data sets is still in progress.

Table 3: Question characteristics by NIS Indicator

Question characteristic	**Indicator 1**	**Indicator 2**	**Indicator 3**	**Indicator 4**	**Indicator 5**
Multiple indicators	X				
Single indicator		X	X	X	X
Dichotomous	X	X		X	
Recency			X		X
Sexual acts and sexual touching	X	X	X		
Sexual acts only				X	X

Our paper used the methods described in Berzofsky, Biemer, and Kalsbeek (2014) to fit the LCA model. Namely, we used the following three step process.

1. Determine if any indicators are bivocal (i.e., indicators that are not fully correlated with the latent construct of interest). If any indicators are, incorporate a second latent variable into model.
2. Determine best grouping variables to account for group heterogeneity.
3. Determine if there is any behavioral correlation among the indicators. If there is, incorporate direct effects (i.e., the interaction of two correlated indicators) in the model.

2.1.1 Determining Bivocal Indicators

As described above, the latent construct for inmate-on-inmate victimization includes any sexual contact or act that is forced or pressured upon the inmate by another inmate. As illustrated in Table 3, Indicator 4 and Indicator 5 only address sexual acts. Therefore, while highly correlated with the latent construct of interest, these two indicators are bivocal. To account for this model assumption failure, a second latent construct was incorporated into the model.

2.1.2 Determining Best Grouping Variables

The NIS has 13 grouping variables which include 10 demographic and criminal history variables and three paradata variables. In order to account for the NIS's complex survey design (i.e., stratification and clustering within PSU), LatentGold version 4.5 software (Magidson and Vermunt, 2005), which utilizes pseudo maximum likelihood (Pfeffermann, 1993), was used. Because of sparseness, all 14 potential grouping variables (listed in Table 4) cannot fit into a single model where backwards selection is used. Therefore, a forward selection approach was used. Using the BIC to determine the better model, the following steps were used in the model selection:

1. Fit all single variable models (i.e., full models with grouping variable in structural and measurement components); select best based on BIC.
2. Fixing the best grouping variable from (1), fit all possible two variable models (main effects model that does not include interaction between grouping variables); select best based on BIC.
3. If BIC of (2) is substantially greater than (1) then continue to add third grouping variable.
4. Continue process until best model with *k* variables has a negligible better BIC than the *k*-1 model. At this point the *k* variable model is the best model.
5. Using Wald statistic add variables to structural component only to determine best structural model.

2.1.3 Determining Direct Effects

Using the model developed in Section 2.1.2, the final step in the modeling process is to determine if there is any behavioral correlation between the indicators. Using LatentGold, this can be done by assessing the bivariate residuals. If the bivariate residual for any pair of indicators is greater than 3.84 (the chi-square critical value with one degrees of freedom and the 95% confidence level) a direct effect should be included (Vermunt and Magidson, 2005). Direct effects were added to the model in a stepwise fashion whereby the direct effects were added one at a time beginning with the direct effect with the largest residual. Direct effects were added until all bivariate residuals had a value less than 3.84.

2.2 Accounting for Missing Data

In LCA models, as in any regression model, missing values can occur in the independent variables (i.e., grouping variables) and the dependent variable(s) (i.e., indicator variables). Without any treatment, records with a missing value at either level (dependent or independent) are removed from the analysis (i.e., listwise deletion). In order to circumvent this problem, most statistical software for LCA, including LatentGold, utilizes Fuchs method of full information maximum likelihood (FIML; Fuchs, 1982). Under this version of FIML, when an independent variable is missing, the case is dropped from the analysis. Therefore, independent variables must be imputed prior to incorporating FIML to use all available data.

2.2.1 Grouping variables

As noted in the previous section, there are 14 grouping variables in the NIS survey instrument. Table 4 presents the level of each of the grouping variables and their level of missingness. As can be seen, no individual characteristic has greater than 4% of their data missing; however, across all variables 11.3% of the data are missing.

Table 4: Number of percent of data missing by grouping variable

Variable	Missing (n)	Missing (%)	Variable	Missing (n)	Missing (%)
Any	**5,896**	**11.27**	Sexual orientation	1,712	3.27
Marital status	311	0.59	Trouble inmate	1,001	1.91
Controlling offense	2,059	3.93	Inmate problems with survey	772	1.48
Education	84	0.16	Inmate dishonest on survey	1,258	2.40
Age Category	0	0.00	Int. perceived distractions	1,226	2.34
Race	453	0.87	Int. perceived misunderstanding	70	0.13
Time since admission	4	0.01	Int. perceived inmate upset	90	0.17

When indicated, LatentGold 4.5 will incorporate a mean imputation procedure to impute missing data for grouping variables included in the model (Vermunt and Magidson, 2005). The mean imputation procedure uses the distribution among respondents to determine the classification probabilities for the missing cases. This procedure was used in this analysis.

2.2.2 Indicator variables

Due to the sensitive nature of the NIS, it is not surprising that respondents did not respond to all of the indicator variables. Table 5 presents the number of cases missing by indicator and NIS cycle. As seen in the table, across all five indicators, around 8-9% of cases have at least one missing indicator. After imputing for the grouping variables, FIML was used to include all cases in the dataset. FIML assumes that dependent variables are missing at random (MAR). However, if victimization is correlated to an inmate not responding then the data are not missing at random (NMAR). While NMAR bias can be reduced by MAR adjustments, it is possible that not all of the bias can be corrected.

Furthermore, Indicator 1 has the highest level of missing data. This is likely because it is a composite variable based on 7 individual questions asking about specific sexual contacts or acts. When creating the composite variable if any of the individual items were not answered (i.e., don't know or refuse answer provided) then the composite variable was set to missing.

Table 5: Number of Cases with a Missing Indicator by Indicator and NIS Cycle

		Number of Respondents with Missing Indicator					
	Total	Indicator 1	Indicator 2	Indicator 3	Indicator 4	Indicator 5	Cross-5
NIS-1	22,868	1,269	303	977	390	1,042	2,126
NIS-2	29,458	1,588	64	711	170	780	2,468

3. Assessing Impact of Excluding Missing Data

In order to compare a model that excluded missing data to a model that included the missing data (as described in Section 2.2), two parallel models were fit. In other words, the process described in Section 2.1 was conducted with listwise deletion (i.e., missing data were excluded) and with the full dataset.

3.1 LCA Model Excluding Missing Cases

The final model when missing data were excluded included 46,198 cases (88.3% of all cases). The model found four grouping variables to be significant in the measurement component of the latent class model. Namely, in the order they were identified,

- sexual orientation (heterosexual vs. non-heterosexual),
- inmate written-up or spend a night in administrative segregation (yes/no),
- inmate indicating having problems taking the survey (yes/no), and
- inmate indicated responding honestly to all questions (yes/no)

were identified as the best set of grouping variables. Furthermore, in the structural component, the model included gender and the interaction between gender and sexual orientation to account for the structural differences in the estimates by gender (see Table 2). In terms of direct effects, four direct effects were needed based on the bivariate residuals. In all cases, the direct effects involved the interaction between one of the univocal indicators (i.e., Indicator 1, Indicator 2, or Indicator 3) and the bivocal indicators (i.e., Indicator 4 or Indicator 5).

Table 6 presents the model diagnostics for the best one-, two-, three-, four, and five-variable models as well as the final model that included direct effects and additional structural component variables. The model diagnostics include the number of model parameters, the design effect, the BIC, and the dissimilarity index. In the table *X* represents the latent construct of sexual contact or sexual acts, *Y* represents the latent construct of sexual acts only; *A*, *B*, *C*, *D*, and E represent Indicator 1 through Indicator 5, respectively; *G* represents sexual orientation, *H* represents an inmate having discipline problems, *I* represents the inmate indicating having a problem taking the survey, *J* represents an inmate indicating responding honestly to all questions, and *K* represents gender.

Table 6: Model Selection Diagnostics by Model Iteration

#	Model	Number of Parameters	Design Effect	BIC	Dissimilarity Index
1	{XG YG AXG BXG CXG DYG EYG}	12	1.6773	15875.68	0.0472
2	{XG XH YG YH AXG AXH BXG BXH CXG CXH DYG DYH EYG EYH}	18	1.6110	15425.17	0.0463
3	{XG XH XI YG YH YI AXG AXH AXI BXG BXH BXI CXG CXH CXI DYG DYH DYI EYG EYH EYI}	30	1.5510	15198.22	0.0459
4	{XG XH XI XJ YG YH YI YJ AXG AXH AXI AXJ BXG BXH BXI BXJ CXG CXH CXI CXJ DYG DYH DYI DYJ EYG EYH EYI EYJ}	36	1.5850	15000.81	0.0454
5	{XG XH XI XJ XK YG YH YI YJ YK AXG AXH AXI AXJ AXK BXG BXH BXI BXJ BXK CXG CXH CXI CXJ CXK DYG DYH DYI DYJ DYK EYG EYH EYI EYJ DYK}	52	1.4954	14949.68	0.0453
6	{XG XH XI XJ XK YG YH YI YJ YK AXG AXH AXI AXJ AXK BXG BXH BXI BXJ BXK CXG CXH CXI CXJ CXK DYG DYH DYI DYJ DYK EYG EYH EYI EYJ DYK}	56	1.8389	14748.14	0.0454

3.2 LCA Model Including Missing Cases

The final model with all cases included contained all 52,319 respondents in NIS-1 and NIS-2 combined. This model found the same significant grouping variables. Furthermore, the direct effects were the same as well. In other words, the inclusion of missing data in the NIS did not alter the model building process. Table 7 presents the model diagnostics for the models that included the missing cases from the model selection process

Table 7: Model Selection Diagnostics by Model Iteration

#	Model	Number of Parameters	Design Effect	BIC	Dissimilarity Index
1	{XG YG AXG BXG CXG DYG EYG}	12	1.6992	21125.05	0.0534
2	{XG XH YG YH AXG AXH BXG BXH CXG CXH DYG DYH EYG EYH}	18	1.5972	20538.11	0.0526
3	{XG XH XJ YG YH YJ AXG AXH AXJ BXG BXH BXJ CXG CXH CXJ DYG DYH DYJ EYG EYH EYJ}	24	1.5995	20151.68	0.0519
4	{XG XH XI XJ YG YH YI YJ AXG AXH AXI AXJ BXG BXH BXI BXJ CXG CXH CXI CXJ DYG DYH DYI DYJ EYG EYH EYI EYJ}	36	1.5543	19882.32	0.0515
5	{XG XH XI XJ XK YG YH YI YJ YK AXG AXH AXI AXJ AXK BXG BXH BXI BXJ BXK CXG CXH CXI CXJ CXK DYG DYH DYI DYJ DYK EYG EYH EYI EYJ DYK}	52	1.4829	19789.90	0.0513
6	{XG XH XI XJ XK YG YH YI YJ YK AXG AXH AXI AXJ AXK BXG BXH BXI BXJ BXK CXG CXH CXI CXJ CXK DYG DYH DYI DYJ DYK EYG EYH EYI EYJ DYK}	56	1.6421	19567.07	0.0178

4. Results

4.1 Structural Components

Figure 1 presents the estimated structural component (error-free) estimates for inmate-on-inmate victimization for both latent variables included in the model. The model found that when missing data were excluded the estimate of the true victimization prevalence was 1.9%, but when cases with missing data were included the estimate increased to 2.1% - a 12.7% increase. This suggests that inmates excluded by list-wise deletion may be more likely to be victims of inmate-on-inmate sexual victimization. However, when the latent construct is restricted to only sexual acts except touching (e.g., penetrative sexual acts) the model that includes the missing produces estimates a somewhat lower estimate than when the missing data are excluded (0.95% vs. 1.01%). This indicates that the bias is towards an underreporting of sexual contact among those with missing data.

The published estimate of inmate-on-inmate victimization was 2.1% in NIS-1 and NIS-2 (Beck and Harrison, 2007; Beck, et. al., 2010). This indicates that the published estimate and the error-free estimate when missing data are included are nearly identical. This indicates that when the missing data are included, the false negative and false positive rates have a canceling effect on each other. However, when the missing cases are excluded the false positive rate has a greater influence on the error-free estimate leading to a lower estimated prevalence rate.

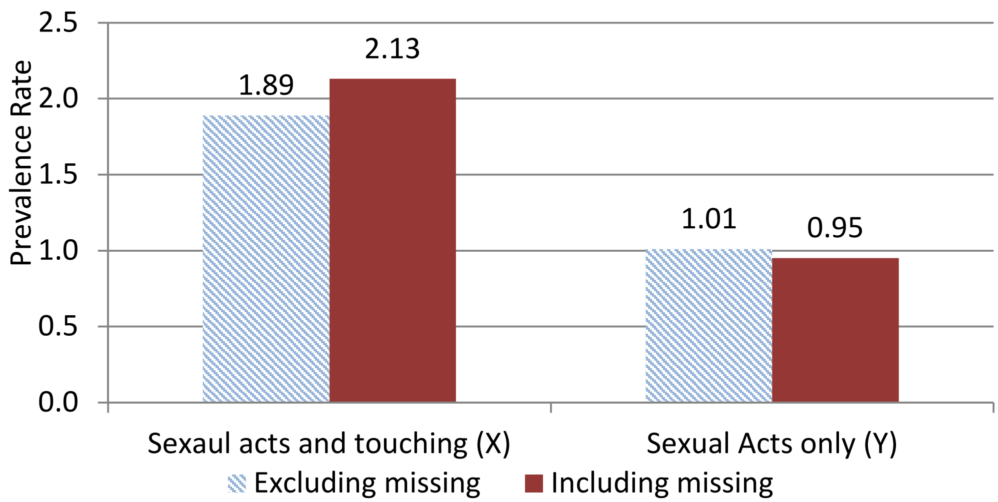

Figure 1: Estimated Structural Component (Error-free) Estimates by Latent Variable and Model Type (Excluding and Including Missing Data)

4.2 Measurement Component

Figure 2 presents the estimated false negative rates for each indicator as they relate to the main latent construct of interest (i.e., sexual contact and sexual acts by force or pressure from another inmate) by model type. Several findings are apparent from the figure.

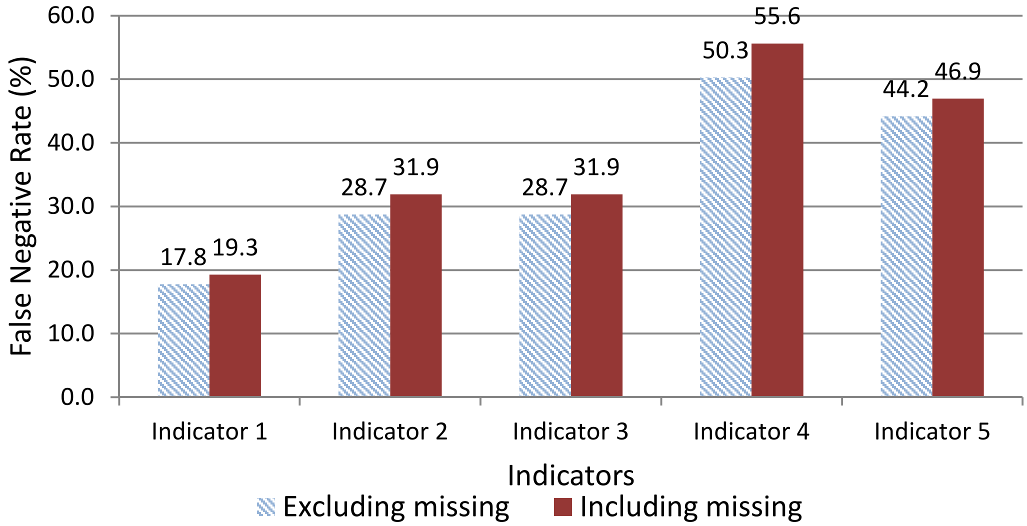

Figure 2: Estimated Measurement Component (False Negative Rate Only) Estimates by Latent Variable and Model Type (Excluding and Including Missing Data)

First, regardless of model type, the false negative rate for Indicator 1 is considerably lower than all other indicators. As noted earlier, Indicator 1 is the method used to create the published estimates of sexual victimization. This indicates that the multi-item approach asking about individual specific acts produces the most accurate estimate of inmate-on-inmate sexual victimization compared to the single item indicators used for the LCA.

Second, Indicator 1, Indicator 2, and Indicator 3 have significantly lower false negative rates compared to Indicator 4 and Indicator 5. This is likely a result of the fact that, as indicated in Table 3, Indicator 4 and Indicator 5 are bivocal with the latent construct of interest. Therefore, it is not surprising that the false negative rate is higher for these indicators because an inmate who was a victim of unwanted sexual contact by another inmate, but not a sexual act is accurate in responding 'no' to Indicator 4 and 'never' to Indicator 5 when, in fact, the inmate is a victim for the main latent construct of interest.

Third, for all indicators, the estimated false negative rate is larger in the model that includes the missing cases than the model that excludes the missing cases. The percent difference in the rates range from 1.5% to 5.3%. This suggests that the inmates who do not respond to all indicators are more likely to provide inaccurate answers. This suggests that missing is NMAR rather than MAR. Therefore, as discussed earlier, it may be that FIML is only partially adjusting for the bias in the classification error rates.

Furthermore, the estimated false positive rates are negligible under both model types (not shown). For example, for Indicator 1 the false positive rate is 0.26% and 0.22% when the missing cases are included and excluded, respectively. The false positive rates for the other four indicators are all smaller than the Indicator 1 rate under both model types. This suggests that the concern that inmates who were truly victims would indicate otherwise simply to indict the facility did not materialize.

5. Conclusions

For inmate-on-inmate victimization in the first two cycles of the NIS the LCA found that the published estimates are nearly identical to the error free estimates and that the false negative rate is smallest for Indicator 1. These results indicate two important findings: (1) when missing data are included in the model, the false negative rate and false positive rate have a canceling effect leading to estimates similar to the published estimates, and (2) by having the smallest false negative rate, fewer inmates indicated they were not victimized when they truly were compared to any other indicator.

Furthermore, when cases are missing either a grouping variable or an indicator variable, the estimated prevalence rate is larger for inmate-on-inmate victimization. Moreover, the inclusion of missing data does not alter the LCA model in terms of the model selection process. However, the models do produce different classification error rate, namely, the false negative rates are larger when the missing data are included.

In order to confirm that these findings hold more generally, this analysis should be repeated for staff sexual misconduct among prison inmates as well as inmate-on-inmate victimization and staff sexual misconduct among jail inmates. In addition, future work should determine how best to incorporate NIS-3 data that only has three indicators rather than the five in NIS-1 and NIS-2. Additionally, more work needs to be done to test the difference between estimates from models that exclude missing data are statistically different from models that include the missing data. Furthermore, as noted, LatentGold utilizes a mean imputation for the grouping variables. The impact of using other imputation procedures on model estimates needs to be assessed.

Acknowledgements

The authors would like to thank BJS for sponsoring this research and comments provided by Allen Beck. However, we would like to note that the views expressed in this paper are those of the authors only and do not reflect the views or position of BJS or the Department of Justice.

References

Beck, A.J. (2014). PREA Data Collection Activities, 2014. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 245694. Retrieved from http://www.bjs.gov/content/pub/pdf/pdca14.pdf on September 20, 2014.

Beck, A.J & Harrison, P.M (2007). Sexual Victimization in State and Federal Prisons as Reported by Inmates, 2007. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 219414. Retrieved from http://www.bjs.gov/content/pub/pdf/svsfpri07.pdf on September 20, 2014.

Beck, A.J & Harrison, P.M (2008). Sexual Victimization in Local Jails as Reported by Inmates, 2007. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 219414. Retrieved from http://www.bjs.gov/content/pub/pdf/svljri07.pdf on September 20, 2014.

Beck, A.J, Harrison, P.M, Berzofsky, M.E., Krebs, C., & Caspar, R. (2010). Sexual Victimization in Prisons and Jails as Reported by Inmates, 2008-09. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 231169. Retrieved from http://www.bjs.gov/content/pub/pdf/svpjri0809.pdf on September 20, 2014.

Beck, A.J, Berzofsky, M.E., Krebs, C., & Caspar, R. (2013). Sexual Victimization in Prisons and Jails as Reported by Inmates, 2011-12. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 241399. Retreived from http://www.bjs.gov/content/pub/pdf/svpjri1112.pdf on September 20, 2014.

Berzofsky, M.E. (2011). *Methods and Approaches for Evaluating the Validity of Latent Class Models with Applications.* Unpublished dissertation.

Berzofsky, M. E., & Biemer, P. P. (2012). Weak Identifiability in Latent Class Analysis. In *Proceedings of the ASA Section on Survey Methodology, Joint Statistical Meetings* (pp. 4346–4354), Alexandria, VA: American Statistical Association.

Berzofsky, M. E., Biemer, P. P., & Kalsbeek, W. D. (2008). A brief history of classification error models. In *Proceedings of the ASA Section on Survey Methodology, Joint Statistical Meetings* (pp. 3667–3673), Alexandria, VA: American Statistical Association.

Berzofsky, M. E., Biemer, P. P., & Kalsbeek, W. D. (2014). Local dependence in latent class analysis of rare and sensitive events. *Sociological Methods and Research, 43*(1), 137–170. doi:10.1177/0049124113506407

Biemer, P.P. (2011). *Latent Class Analysis of Survey Error*. Hoboken, NJ: John Wiley & Sons.

Biemer, P., & Bushery, J. (2001). Application of Markov latent class analysis to the CPS. *Survey Methodology, 26*(2), 136–152.

Biemer, P., & Wiesen, C. (2002). Latent class analysis of embedded repeated measurements: An application to the National Household Survey on Drug Abuse. *Journal of the Royal Statistical Society: Series A*, *165*(1), 97–119.

Fuchs, C. (1982). Maximum Likelihood Estimation and Model Selection in Contingency Tables with Missing Data. Journal of the American Statistical Association, 77(378), 270 – 278.

Gaes, G.G., & Goldberg, A.L. (2004). *Prison rape: A critical review of the literature.* Washington, DC: U.S. Department of Justice, National Institute of Justice. Retrieved 2/1/05 from http://www.nicic.org/Downloads/PDF/2004/019813.pdf.

Hensley, C., Tewksbury, R., & Castle, T. (2003). Characteristics of prison sexual assault targets in male Oklahoma correctional facilities. *Journal of Interpersonal Violence*, *18*(6), 595–607.

Pfeffermann, D. (1993). The role of sampling weights when modeling survey data. Internat Surv Rev. 61: 317–337.

Struckman-Johnson, C., & Struckman-Johnson, D. (2000). Sexual coercion rates in seven mid-western prison facilities for men. *The Prison Journal*, *80*(4), 379–390.

Tewksbury, R. (1989). Measures of sexual behavior in an Ohio prison. *Sociology and Social Research*, *74*(1), 34–39.

Vermunt, J. K., & Magidson, J. (2005). *Technical guide to Latent Gold 4.0: Basic and advanced*. Belmont, MA: Statistical Innovations.