Exploration of NAEP-SSOCS linkage: Investigating the Association between Behavioral Incidents, School Behavioral Climate, and NAEP Mathematics Achievement

Saki Ikoma, Senior Researcher (AIR) | Paul Bailey, Senior Economist (AIR) |

Jizhi Zhang, Principal Researcher (AIR) | Kelly Guinn, Senior Research Scientist (ICF) |

Markus Broer, Managing Researcher (AIR)

FCSM| October 2022

Please do not cite or distribute without permission of the presenting authors. This project has been funded at least in part with Federal funds from the U.S. Department of Education under contract number ED-IES-12-D-0002/0002. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Copyright © 2021 American Institutes for Research®. All rights reserved.

Agenda

I. Contextual aspect

- Motivation of this study
- Data Linkage between NAEP and SSOCS
- Analytical approach
- Major takeaways on the criminological contexts

II. Technical aspect

- Underestimation of regression coefficients
- Plausible values and EdSurvey Dire Package
- Comparison of results between the original NAEP PVs and the Dire PVs
- Final thoughts

AIR®
Advancing Evidence.
Improving Lives.

I. Contextual aspect

Motivation of this study

- Importance of school safety
- Criminological contexts in education
- Ongoing efforts on NCES survey linkage
- NAEP linking to the School Survey on Crime and Safety (SSOCS)

Criminological contexts in education

- School behavioral environment and behavioral climate
  - Peguero & Bracy 2015; Ruiz et al. 2018
- Violence in school and surrounding school
  - Sharkey 2010; Mccoy, Raver, & Sharkey 2015; Carlson & Cohen 2015; Burdick-Will 2016
- School level spillover effects onto student level
  - Hagan & Foster 2012; Lacoe & Steinberg 2019

Data linkage using NCES surveys

- Linking NAEP to the Schools and Staffing Survey (Kaufman 1996)
- NAEP to the Trends in International Mathematics and Science Study (Johnson 1998)
- NCES surveys to state and federal administrative data (Gamoran 2016; Soldner 2017)
- Creating an ID crosswalk to link the Civil Rights Data Collection to other NCES data collections (Sable and Miller 2020)
- Commissioned by NCES, AIR has examined the overlap samples of NAEP and the High School Longitudinal Study of 2009 as well as other datasets linked to schools.

Why NAEP and SSOCS?

There is valuable information in SSOCS that is not included in NAEP, and this information can shed light on important and policy relevant associations between school level crime and behavioral indicators and student achievement.

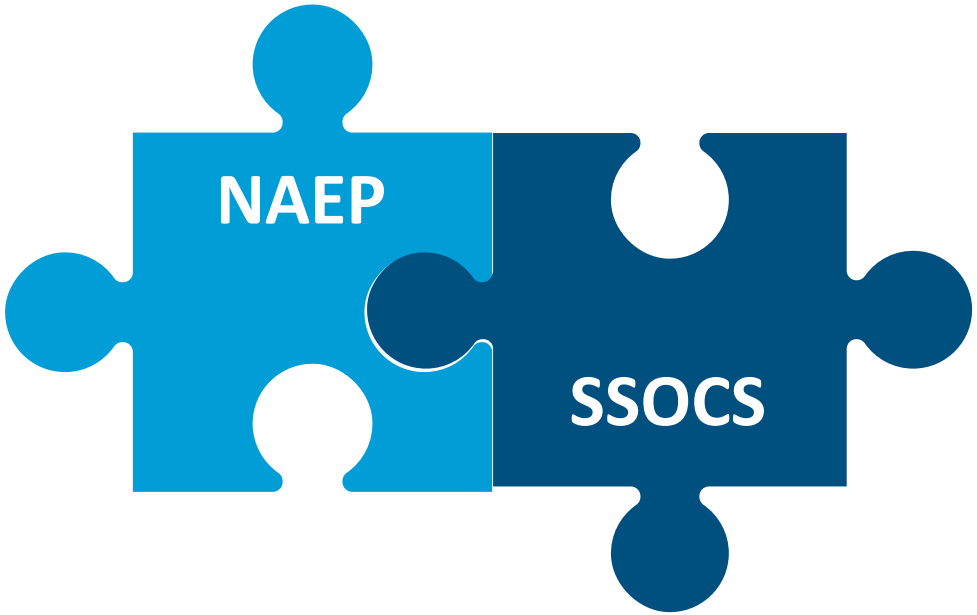
NAEP
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

School Survey On Crime & Safety

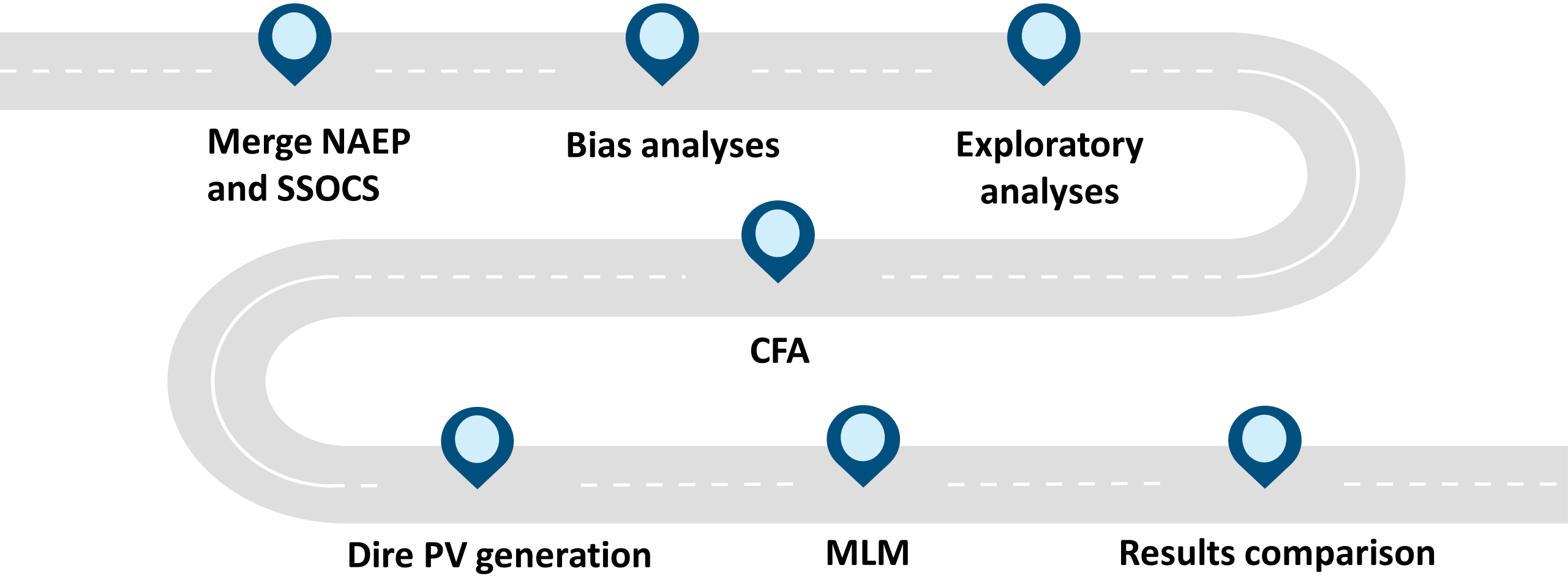
AIR®

Data Linkage between NAEP and SSOCS

- Analytic sample description
  - *NAEP 2017 G8 Mathematics & SSOCS 2017-18*
  - *≈ 330 schools*
  - *≈ 8,470 students*
- Bias analysis
  - *distributions, t-tests*
- Choosing weights
  - *NAEP weights*



Analytical Approach



Major takeaways regarding the criminological contexts

- After taking into account student and school level characteristics,
  - **no violent incidents within the school were significantly associated with student's higher NAEP mathematics performance**.
  - However, the MLM results also revealed that **the school climate index** was **not significantly associated** with student's mathematics performance.

AIR®
Advancing Evidence.
Improving Lives.

II. Technical aspect

Underestimation of regression coefficients

- When regression analyses are conducted with plausible values (PVs) as an outcome and explanatory variables from external sources, the regression coefficients will be underestimated due to the fact that the external variables are not included to produce the plausible values as regressors (Mislevy, Beaton, Sheehan, and Kaplan 1992; von Davier, Gonzalez, and Mislevy 2009; Wu 2015; Rubin 1987)

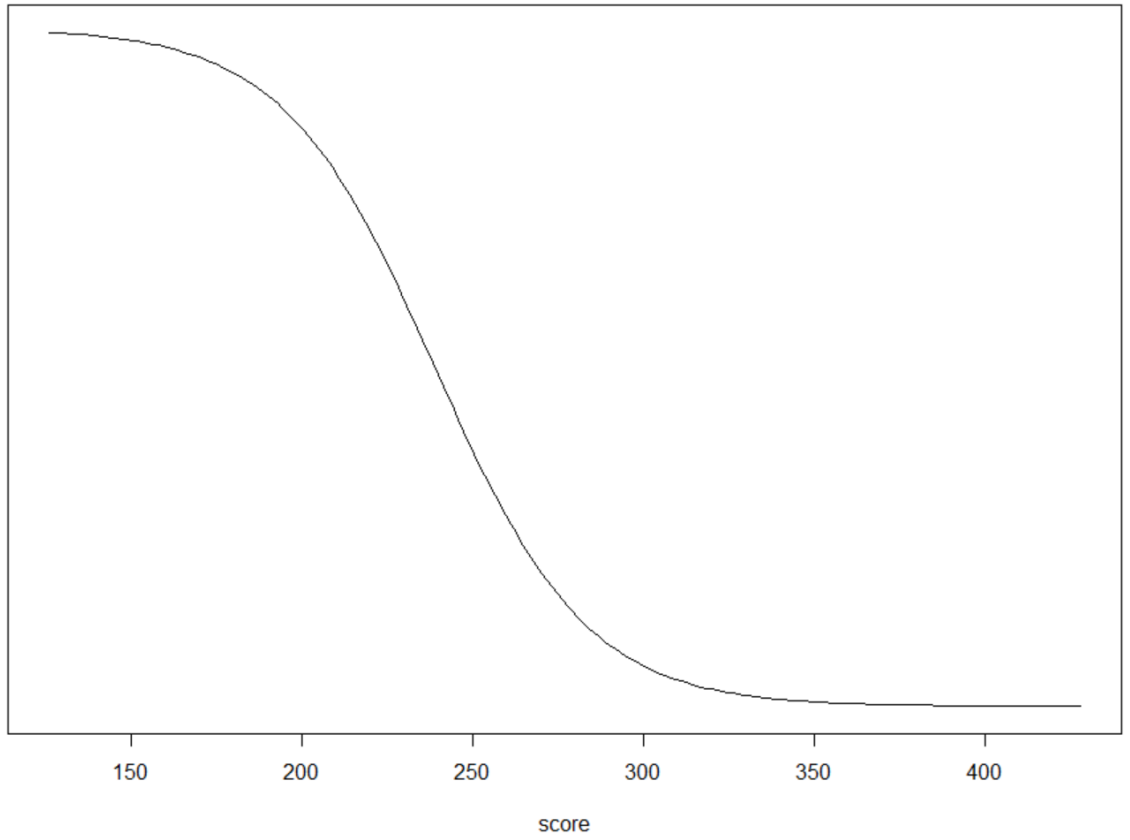
- The Marginal Maximum Likelihood (MML) model allows us to incorporate additional covariates in the model

Why use Multiple Imputation on Test Scores

- To measure an individual student's score we use a high reliability test, usually the variation of any student's estimate is dominated by their ability, not noise
- These tests have too few items to be reliable for individual students
- If we applied the same test score we use for individual student assessments to the NAEP test form, we would get biased results (Mislevy, Beaton, Sheehan, and Kaplan 1992)

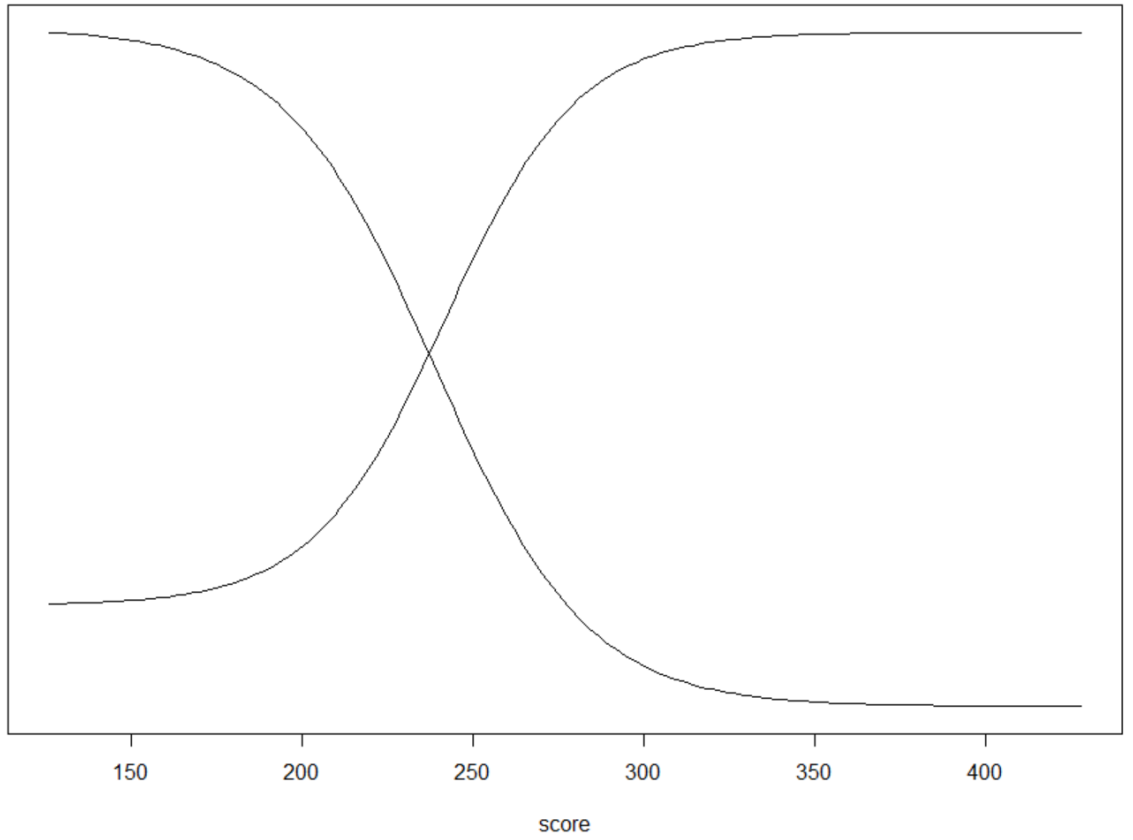
Forming Plausible values: the likelihood

- This item the student got wrong, their likelihood is higher in lower scores.



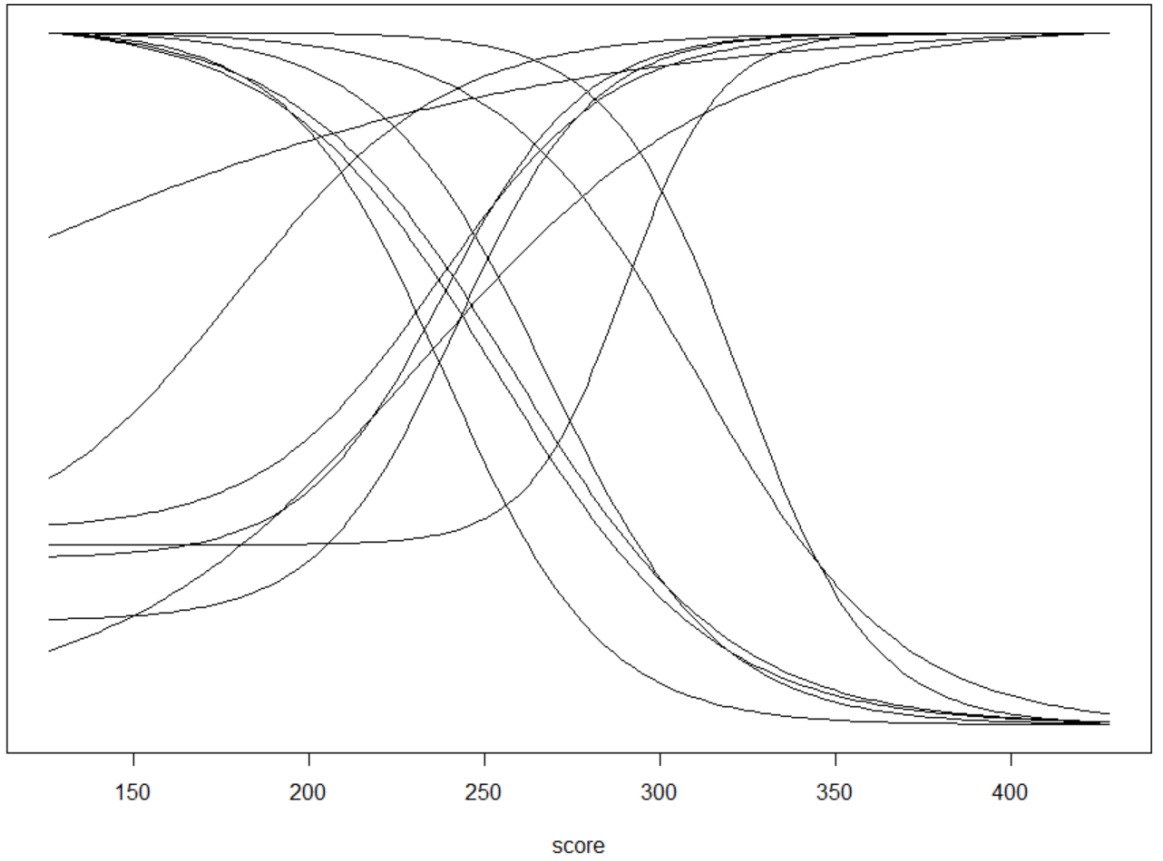
Latent Regression: the Likelihood

- A second item the student got right and so is more likely for higher scores
- The model allows “guessing” and so it does not go to zero at low scores



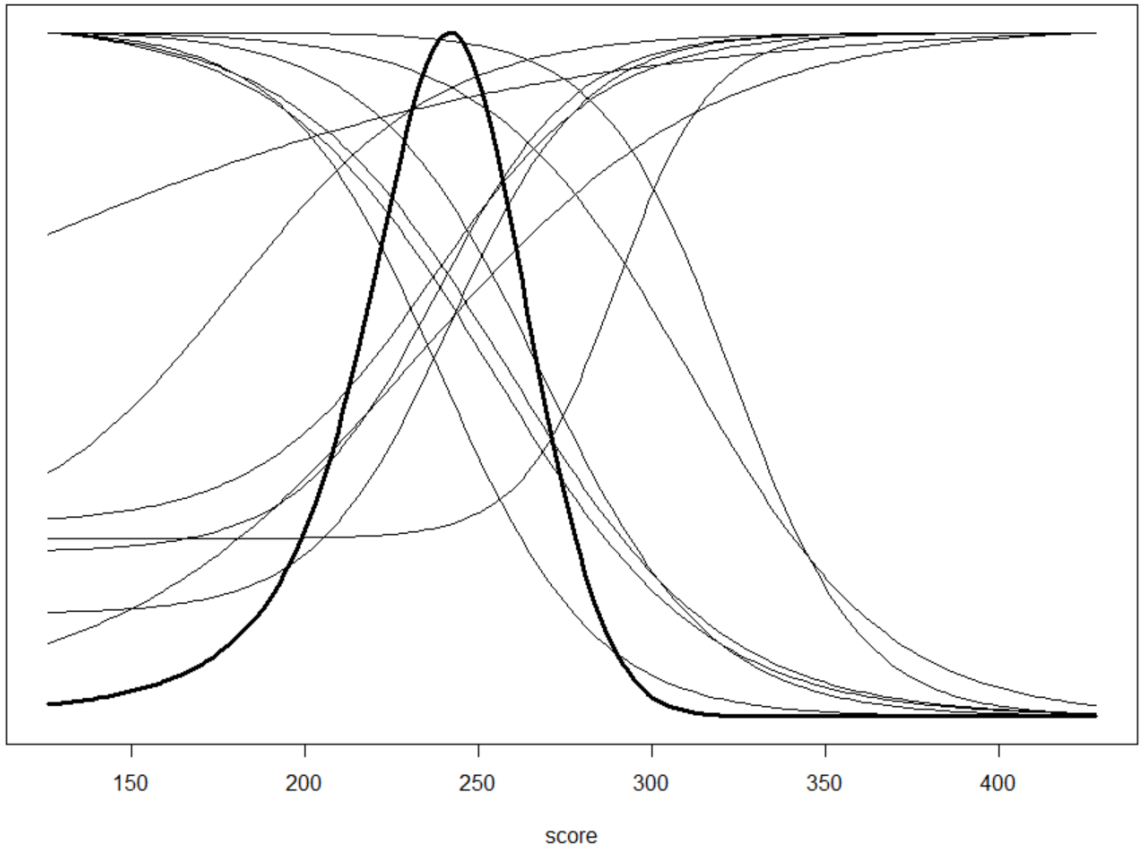
Latent Regression: the Likelihood

- The student took 13 dichotomous items, this is all their likelihoods



Latent Regression: the Likelihood

- Combining them we get a skewed normal likelihood surface
- If we wanted to assign the respondent a score we could take the peak or the expectation of this likelihood
  - We do not want to assign the respondent a score



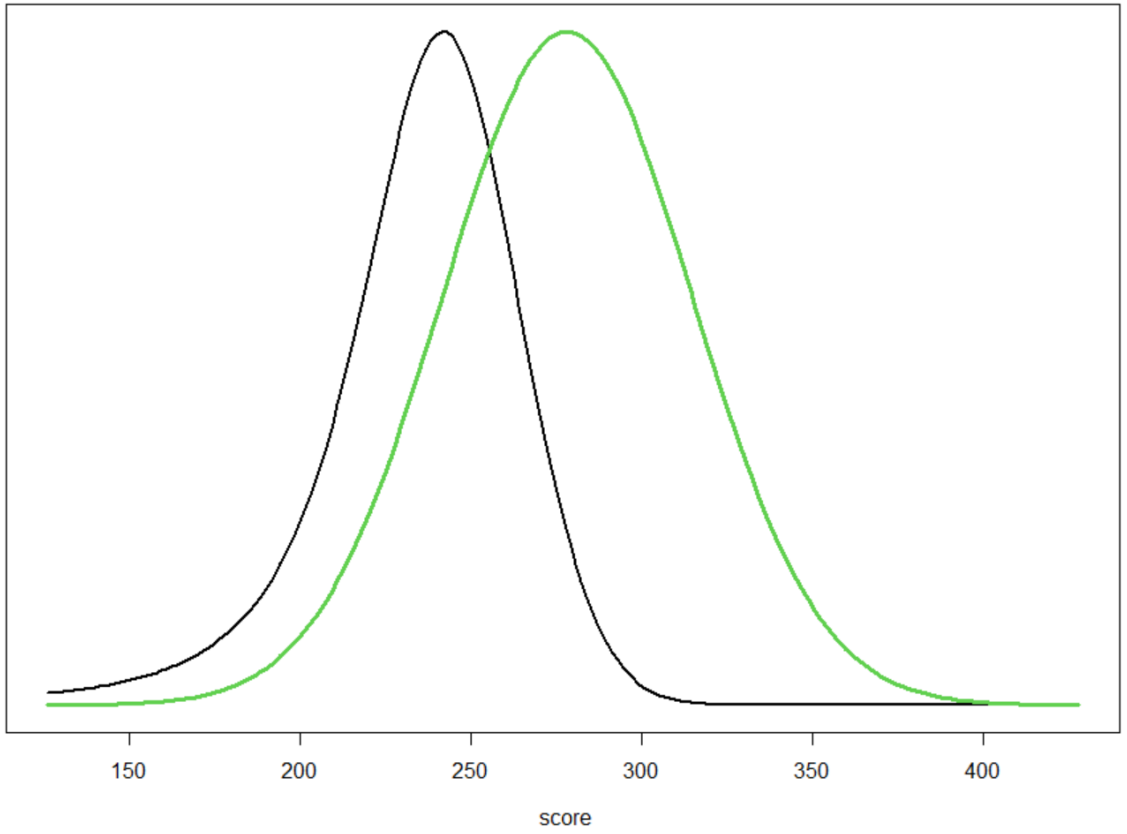
Latent Regression: the Mean Model

- The green curve is the population model or mean model

$$\hat{y} = X\beta + \epsilon$$

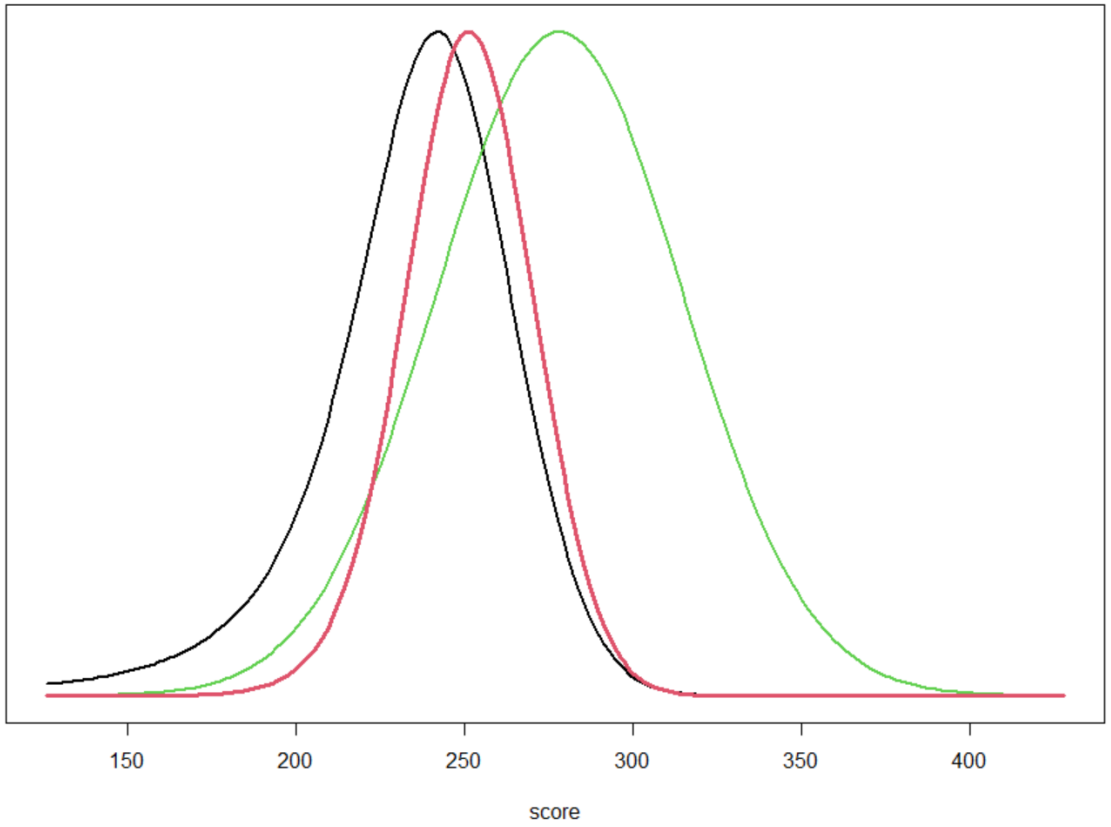
$$\epsilon \sim N(0, \hat{\sigma})$$

- Fit parameters (β) can move it to the left/right, or ($\hat{\sigma}$) make the distribution wider or narrower



Latent Regression: Final Model

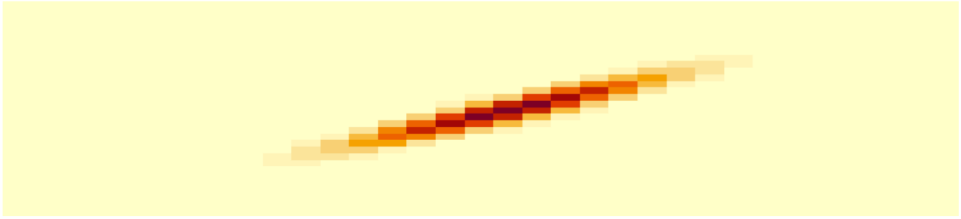
- The student's final likelihood is the product of the mean model and the likelihood
- The mean model is fit by maximizing the "likelihood" of this product
- This is analogous to a regression where the outcome is not precisely measured, but we still want accurate coefficients



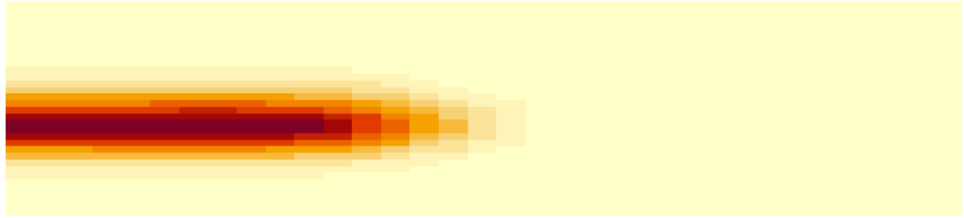
Latent Regression: 2-Dimensions

- For multiple dimensions we must estimate the joint density across the dimensions.
- Here a student's likelihood is plotted for two constructs to estimate the correlation between the two

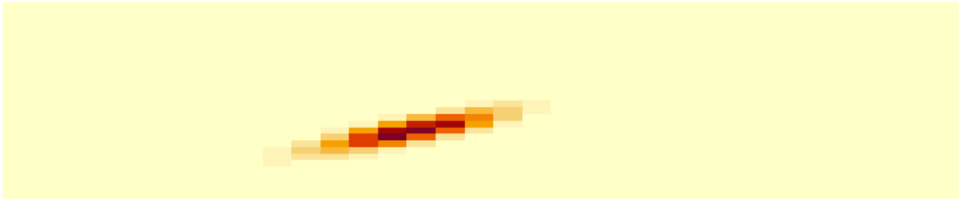
Conditioning model



Likelihood



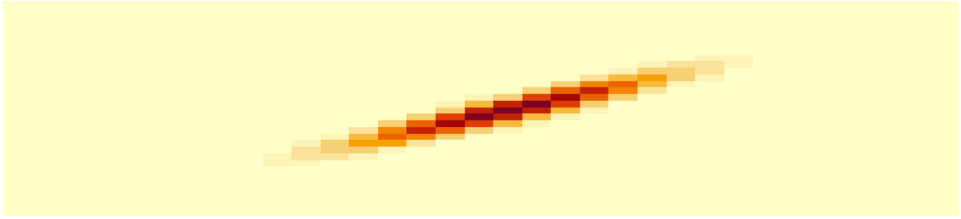
Posterior distribution



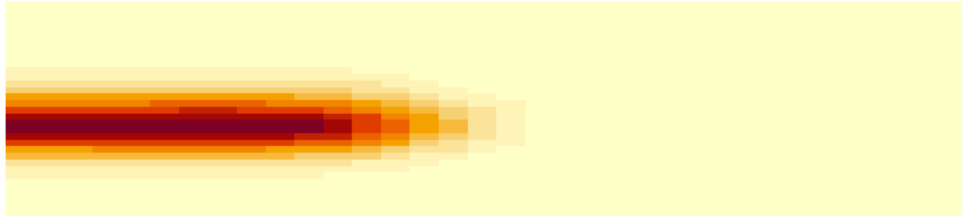
Latent Regression: n-Dimensions

- Because a multivariate normal distribution, when marginalized to two variables, is bivariate normal the same procedure used for 2-dimensions generalizes completely to any number of dimensions—only bivariate correlations need to be calculated

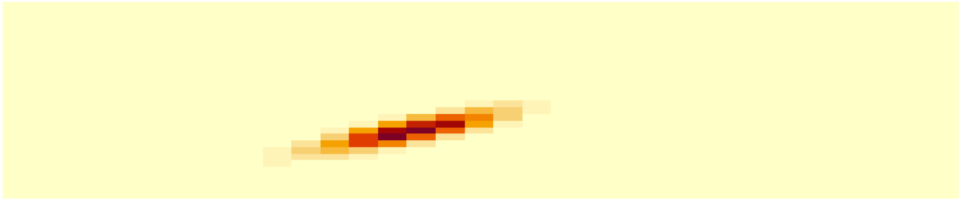
Conditioning model



Likelihood



Posterior distribution



EdSurvey Dire Package

- EdSurvey is an R package developed by NCES and AIR to help users download, understand available covariates for, run common analyses on, and extract data for additional analyses
- Dire adds the ability to use the type of latent regression described above to EdSurvey, but it can also work with other data

Plausible Values

- EdSurvey, with Dire, can use a latent regression to draw plausible values
  - This is multiple imputation for the latent variable or test score
- This is typically used by data owners before distribution
  - This prevents users from having to use latent regression
  - Also prevents data owners from having to share all the conditioning variables
- In addition, this allows repeated use of a conditioning model
  - A conditioning model is biased if a variable of interest is not included, but not if unused variables are not included

Selected comparison results for multiple regressions

	Original NAEP PVs			Dire PVs		
	Coeff. (SE)		*p*-value	Coeff. (SE)		*p*-value
Intercepts	322.86	(2.59)	0.00	322.81	(2.12)	0.00
Residual variances	981.18	(24.94)	0.00	807.11	(19.11)	0.00
R-squared	0.34			0.42		
Student characteristics						
Gender						
Female	-4.28	(1.17)	0.00	-4.01	(1.03)	0.00
Race/ethnicity						
Black	**-23.27**	**(1.99)**	**0.00**	**-25.49**	**(1.65)**	**0.00**
Hispanic	**-12.37**	**(1.58)**	**0.00**	**-12.58**	**(1.47)**	**0.00**
Asian	**10.93**	**(3.23)**	**0.00**	**13.03**	**(2.39)**	**0.00**
Other	-9.21	(2.33)	0.00	-7.30	(2.24)	0.00
Student with disabilities (SD)						
Yes, SD	**-37.45**	**(1.88)**	**0.00**	**-37.63**	**(1.65)**	**0.00**
English learner (EL)						
Yes, EL	**-22.58**	**(2.85)**	**0.00**	**-27.43**	**(2.44)**	**0.00**
NSLP Eligibility						
Eligible	**-16.56**	**(1.30)**	**0.00**	**-18.62**	**(1.18)**	**0.00**

NOTE: Race categories exclude persons of Hispanic origin. Asian includes students who are identified as Pacific Islander. Other includes students who are identified as either American Indian or Alaska Native or More than one race. NSLP = National School Lunch Program. Bold indicated that the magnitude of coefficients in the model using Dire PVs was larger than the model using the original NAEP PVs.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

Selected comparison results for multiple regressions (cont'd)

	Original NAEP PVs			Dire PVs		
	Coeff. (SE)		*p*-value	Coeff. (SE)		*p*-value
School characteristics						
School behavioral climate						
	-0.62	(0.78)	0.42	-0.47	(0.73)	0.52
Violent incidents (Ref: No violent incidents)						
Bottom 25th percentile	-11.47	(2.09)	0.00	-9.63	(1.79)	0.00
Middle percentile (25-50)	**-8.46**	**(2.04)**	**0.00**	**-9.68**	**(1.78)**	**0.00**
50th percentile or higher	-9.35	(1.98)	0.00	-8.26	(1.75)	0.00
Title I fund						
Yes, Title I received	**-6.05**	**(1.16)**	**0.00**	**-7.02**	**(1.04)**	**0.00**
School location (Ref: City)						
Town	**-6.00**	**(1.81)**	**0.00**	**-6.70**	**(1.59)**	**0.00**
Rural	-2.17	(1.73)	0.21	0.29	(1.52)	0.85
Suburban	-1.44	(1.53)	0.35	-0.72	(1.29)	0.58
Census region (Ref: Midwest)						
Northeast	**1.92**	**(1.96)**	**0.33**	**2.92**	**(1.85)**	**0.11**
South	-3.12	(1.29)	0.02	-1.68	(1.16)	0.15
West	-5.37	(1.83)	0.00	-3.71	(1.57)	0.02
Student/teaching-staff ratio (12-16)						
Less than 12	0.60	(2.24)	0.79	0.36	(2.02)	0.86
More than 16	**-1.41**	**(1.26)**	**0.27**	**-2.33**	**(1.14)**	**0.04**

NOTE: Bold indicated that the magnitude of coefficients in the model using Dire PVs was larger than the model using the original NAEP PVs.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

Final Thoughts

- These methods are general to other latent constructs, such as a compound SES measure, or a personality measure
- Learn more about EdSurvey on the AIR website
  - https://www.air.org/project/nces-data-r-project-edsurvey

Thank you

Saki Ikoma

Senior Researcher
sikoma@air.org

Paul Bailey

Senior Economist
pbailey@air.org

Notice of Trademark: “American Institutes for Research” and “AIR” are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.

Copyright © 2021 American Institutes for Research®. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on AIR.ORG.

References

Burdick-Will, J. (2016). Neighborhood Violent Crime and Academic Growth in Chicago: Lasting Effects of Early Exposure. *Social Forces*, *95*(1), 133–158.

Carlson, D., & Cowen, J. M. (2015). Student Neighborhoods, Schools, and Test Score Growth. *Sociology of Education*, *88*(1), 38–55.

Gamoran, A. (2016). Toward the Next Generation of U.S. National Longitudinal Surveys: Ideas From Researchers for the National Center for Education Statistics. *AERA Open 2*(2), 1–8.

Hagan, J., & Foster, H. (2012). Children of the American Prison Generation: Student and School Spillover Effects of Incarcerating Mothers: Children of the American Prison Generation. *Law & Society Review*, *46*(1), 37–69.

Johnson, E.G. (1998). *Linking the National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Study (TIMSS): A Technical Report*, NCES 98-499. U.S. Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office, September 1998.

Kaufman, P. (1996). Linking Student Data to SASS: Why, When, How. *Working Paper* No. 96-23, U.S. Department of Education, National Center for Education Statistics. Washington, D.C.

Lacoe, J., & Steinberg, M. P. (2019). Do Suspensions Affect Student Outcomes? *Educational Evaluation and Policy Analysis*, *41*(1), 34–62.

McCoy, D. C., Raver, C. C., & Sharkey, P. (2015). Children's Cognitive Performance and Selective Attention Following Recent Community Violence. *Journal of Health and Social Behavior*, *56*(1), 19–36.

Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, *29*(2), 133-161.

Peguero, A. A., & Bracy, N. L. (2015). School Order, Justice, and Education: Climate, Discipline Practices, and Dropping Out. *Journal of Research on Adolescence*, *25*(3), 412–426.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Ruiz, L. D., McMahon, S. D., & Jason, L. A.. (2018). The Role of Neighborhood Context and School Climate in School-Level Academic Achievement. *American Journal of Community Psychology*, *61*(3–4), 296–309.

Sable, J., & Miller, S., R. (2020, November 24). *Building bridges: Increasing the power of the Civil Rights Data Collection (CRDC) through data linking with an ID crosswalk*.

Sharkey, P. (2010). The acute effect of local homicides on children's cognitive performance. *Proceedings of the National Academy of Sciences*, *107*(26), 11733–11738.

Soldner, M. (2017). *Leveraging Administrative Data to Strengthen NCES Postsecondary Sample Surveys*. (NPEC 2017). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative.

von Davier, M., Gonzalez, E. J., & Mislevy, R. J. (2009). What are Plausible Values and Why are They Useful? In Matthias von Davier & Dirk Hastedt (eds.). *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments: Volume 2*. 2009, 9-36.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation, 31*(2-3), 114-128.

Descriptive statistics of variables of interest

Student characteristics	
Gender	
Male	50.1 (0.77)
Female	49.9 (0.77)
Race/ethnicity	
White	46.5 (2.29)
Black	12.8 (1.27)
Hispanic	27.9 (2.24)
Asian	5.5 (0.88)
Other	7.3 (0.54)
Student with disabilities (SD)	
Yes, SD	12.0 (0.67)
No, not SD	88.0 (0.67)
English learner (EL)	
Yes, EL	5.7 (0.83)
No, not EL	94.0 (0.83)
Eligibility for the National School Lunch Program (NSLP)	
Eligible	46.7 (2.16)
Not eligible	51.6 (2.07)
Information not available	1.7 (0.75)

NOTE: $n \approx 8{,}470$. Estimates are at student level. Race categories exclude persons of Hispanic origin. Asian includes students who are identified as Pacific Islander. Other includes students who are identified as either American Indian or Alaska Native or More than one race. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

Descriptive statistics of variables of interest (cont'd)

Criminological contexts	
	Percent (SE)
Violent incidents	
No violent incidents	7.4 (1.54)
Bottom 25th percentile	25.2 (4.09)
Middle percentile (25-50)	22.8 (2.81)
50th percentile or higher	44.6 (4.05)
School characteristics	
Title I fund	
No	46.3 (3.71)
Yes	53.7 (3.71)
School location	
City	20.3 (2.55)
Suburb	53.2 (3.71)
Town	7.3 (1.34)
Rural	19.2 (3.23)
Census region	
Northeast	12.4 (1.93)
Midwest	20.9 (2.48)
South	39.0 (3.43)
West	27.7 (3.41)
Student/teaching-staff ratio	
Less than 12	5.9 (1.77)

NOTE: $n \approx 8{,}470$. Estimates are at student level. SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

SSOCS items for school behavioral climate

Disciplinary Problems and Actions

35. To the best of your knowledge, how often do the following types of problems occur **at your school***?

Check one response on each line.

		Happens daily	Happens at least once a week	Happens at least once a month	Happens on occasion	Never happens
a. Student racial/ethnic tensions	374	1	2	3	4	5
b. Student **bullying***	376	1	2	3	4	5
c. Student **sexual harassment*** of other students	378	1	2	3	4	5
d. Student **harassment*** of other students based on **sexual orientation***	381	1	2	3	4	5
e. Student **harassment*** of other students based on **gender identity***	383	1	2	3	4	5
f. Student **harassment*** of other students based on religion	385	1	2	3	4	5
g. Student **harassment*** of other students based on disability (e.g., physical, mental, and learning disabilities)	387	1	2	3	4	5
h. Widespread disorder in classrooms	382	1	2	3	4	5
i. Student verbal abuse of teachers	380	1	2	3	4	5
j. Student acts of disrespect for teachers other than verbal abuse	384	1	2	3	4	5
k. **Gang*** activities	386	1	2	3	4	5

36. To the best of your knowledge, thinking about problems that can occur anywhere (both **at your school*** and away from school), how often do the following occur?

Check one response on each line.

		Happens daily	Happens at least once a week	Happens at least once a month	Happens on occasion	Never happens
a. **Cyberbullying*** among students who attend your school	389	1	2	3	4	5
b. School environment is affected by **cyberbullying***	391	1	2	3	4	5
c. Staff resources are used to deal with **cyberbullying***	393	1	2	3	4	5

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

Results of school behavioral climate CFA

Mode fit indices	
RMSEA	0.10
CFI	0.94
TLI	0.91
Item	**Standardized factor loading**
Racial tension	0.644
Bullying	0.561
Sexual harassment	0.757
General disorder	0.463
Verbal abuse of teachers by students	0.561
Disrespect of teachers by students	0.561
Cyber bullying	0.561
Gang activity	0.561

NOTE: The survey asked, "To the best of your knowledge, how often do the following types of problems occur at your school?". Five response categories ranging from never happens, happens on occasion, happens at least once a month, happens at least once a week to happens daily. The estimates of Cronbach's alpha are 0.76.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

MLM Results

Intercept	Estimate (SE)
Between school	307.35 (4.92)
Residual variance	**Estimate (SE)**
Within school	770.47 (26.67)
Between school	119.17 (18.41)

R-squared	
Within school	0.30
Between school	0.46

DV: Dire PVs	Coeff. (SE)	*p*-value
Within school		
Gender		
Female	-5.27 (1.15)	0.00
Race/ethnicity		
Black	-22.38 (2.84)	0.00
Hispanic	-11.08 (2.37)	0.00
Asian	10.69 (4.04)	0.00
Other	-9.01 (2.78)	0.00
Student with disabilities (SD)		
Yes, SD	-38.84 (2.18)	0.00
English learner (EL)		
Yes, EL	-24.85 (3.55)	0.00
NSLP Eligibility		
Eligible	-14.81 (1.71)	0.00

NOTE: Race categories exclude persons of Hispanic origin. Asian includes students who are identified as Pacific Islander. Other includes students who are identified as either American Indian or Alaska Native or More than one race. NSLP = National School Lunch Program.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.

MLM Results (cont'd)

DV: Dire PVs	Coeff. (SE)	*p*-value
Between school		
School behavioral climate		
	-1.33 (1.51)	0.19
Violent incidents (Ref: No violent incidents)		
Bottom 25th percentile	-14.64 (4.54)	0.00
Middle percentile (25-50)	-13.15 (4.34)	0.00
50th percentile or higher	-17.77 (4.25)	0.00
Title I fund		
Yes, Title I received	-15.65 (2.27)	0.00
School location (Ref: City)		
Town	-3.69 (3.32)	0.13
Rural	7.38 (3.26)	0.01
Suburban	3.52 (2.87)	0.11
Census region (Ref: Midwest)		
Northeast	3.53 (3.33)	0.14
South	-7.13 (2.66)	0.00
West	-5.55 (3.71)	0.07
Student/teaching-staff ratio (Ref: 12-16)		
Less than 12	-2.69 (3.99)	0.25
More than 16	-2.47 (2.53)	0.16

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2017 Grade 8 Mathematics Assessment. U.S. Department of Education, National Center for Education Statistics, 2017–18 School Survey on Crime and Safety (SSOCS), 2018.