New Insights on the Nonresponse Bias Analysis (NRBA) For the National Assessment of Educational Progress (NAEP)

Sabrina Zhang, Tom Krenzke

Agenda

- What is NAEP?

- Data

- Variable selection for NRBA

- Evaluation

- Nonresponse adjustment variables

- Conclusion

- Next steps

What is NAEP?

- Often called The Nation's Report Card

- Largest nationally representative assessment in the U.S.

- Sponsored by the National Center for Education Statistics (NCES)

Non-response Guidelines for Public and Private Schools

- NRBA needs to be done to ensure quality of NR adjustment

- NCES statistical standards call for a NRBA to be conducted for a sample with a response rate below 85 percent at any stage of sampling

- Nonresponse bias is a main concern due to low response rates

School Nonresponse Bias Analysis

- Three Parts
  - Effect of nonresponse: Original eligible sample vs. original respondents
  - Effect of substitution: Original eligible sample vs. respondents with substitutes
  - Effect of substitution and nonresponse adjustment: Original eligible sample vs. nonresponse adjusted respondents with substitutes

Student Nonresponse Bias Analysis

- Two Parts
  - Effect of nonresponse: Eligible students vs. responding students
  - Effect of nonresponse adjustment: Eligible students vs. nonresponse adjusted responding students

Objective

- Data source search for NRBA purposes
  - Find new data sources to enhance NRBA
  - Area, school and student characteristics
- New variables could be considered for NR adjustment

Data

X: Covariates

- 2022 Common Core of Data (CCD): school-level covariates for public schools

- 2022 Private School Universe Survey (PSS): school-level covariates for private schools

- 2021 5-year ACS: tract-level demographic covariates

- 2018 & 2019 Education Demographic and Geographic Estimates (EDGE): school-level Wage Index for Teachers

- 2022 Environmental Justice Index (EJI): tract-level Environmental and health related covariates in percentile ranks

Y: Achievement scores from ETS

- 2022 state representative samples for grade 4 & 8, math and reading subjects
  - 224,426 responding students in grade 4
  - 222,226 responding students in grade 8

Variable Selection for NRBA

Old Variables Used at School Level for Private Schools

- *Census region (4 levels: NE, S, MW, W)

- *Private school reporting subgroup (2 levels: Catholic, Non-Catholic)

- *Urban-centric locale (4 levels: city, suburban, town, rural)

- *School size class (3 levels: large, medium, small)

- #Mean enrollment averaged across students

- #Mean grade enrollment

- #Race/ethnicity percentages

*: Chi-square tests on distributions   #:t-tests on means

Old Variables Used at Student level for Public Schools

- *Gender (2 levels)

- *Race/ethnicity (7 levels)

- *Relative age (2 levels)

- *Free or reduced price lunch eligibility (3 levels)

- *Students with Disabilities (SD) status (2 levels)

- *English Learner (EL) status (2 levels)

*: Chi-square tests on distributions

New Variables Used at Both School and Student Levels

- 2021 5-year ACS

  - Tract-level demographic covariates

- 2018 & 2019 EDGE

  - Neighborhood poverty estimates

  - The Comparable Wage Index for Teachers (CWIFT)

- 2022 EJI

  - Environment (e.g., Percentile rank of proportion of tract's area within 1-mile buffer of airport)

  - Health

  - Demographic

Variable Selection Criteria for NRBA

- Only consider variables with missing rate <5%

- Check correlations between the achievement scores vs. new variables

  - Continuous covariates used Pearson t-test
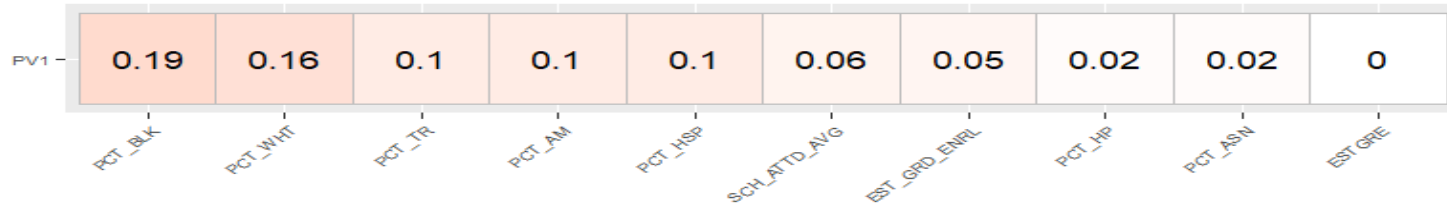
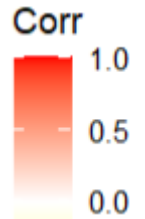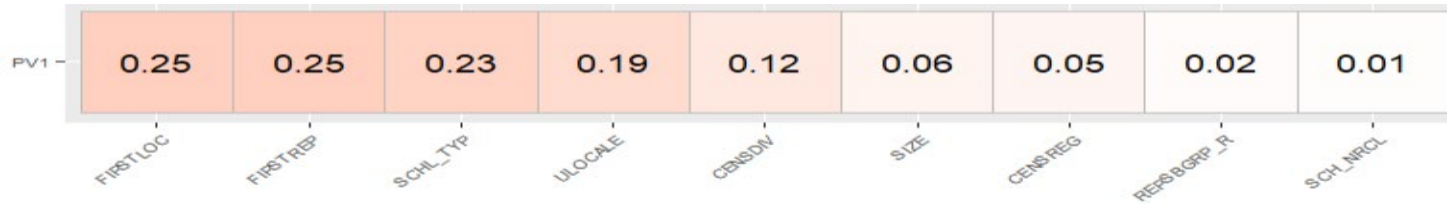 - Categorical covariates used ANOVA

Evaluation

Correlation Results at School Level for Grade 4 Reading Private

Y vs. cont. variable

	PCT_BLK	PCT_WHT	PCT_TR	PCT_AM	PCT_HSP	SCH_ATTD_AVG	EST_GRD_ENRL	PCT_HP	PCT_ASN	EST_GRE
PV1	0.19	0.16	0.1	0.1	0.1	0.06	0.05	0.02	0.02	0

Y vs. cat. variable

	FIRSTLOC	FIRSTREP	SCHL_TYP	ULOCALE	CENSDIV	SIZE	CENSREG	REPSBGRP_R	SCH_NRCL
PV1	0.25	0.25	0.23	0.19	0.12	0.06	0.05	0.02	0.01

Corr
1.0
0.5
0.0

Y vs. area-level variable

	SNAP_HHs	PctSNAP	PctDisabled	PctDisabled_18to64	PctPubAsst	PctNComPlmb	PctEducBchPlus	PctNoPhone	PctEducGHS	NumBelow18Pov_0to17
PV1	0.21	0.2	0.2	0.2	0.19	0.18	0.17	0.17	0.16	0.16
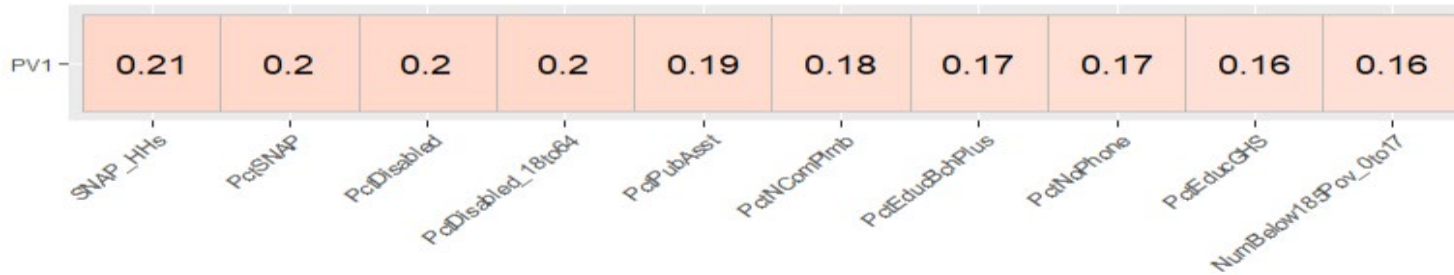
Summary – Final Variables Used in School-Level NRBA Private

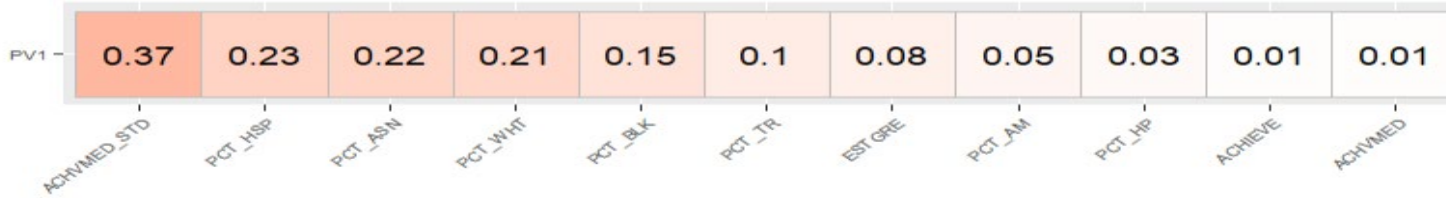| Variable Name	Label	Source	Level	Correlation
PCT_BLK	Percent black	PSS	School	-0.21
ULOCALE	Urban-centric locale code	PSS	Area	0.19
SCHL_TYP	Six level private school types	PSS	School	0.19
PCT_WHT	Percent white	PSS	School	0.18
Division	Census division	PSS	Area	0.17
PCTEDUCBCHPLUS	Percent college degree or more ages 25 and over	ACS	Area	0.15
PCTFAM100POV	Percent families below poverty level	ACS	Area	-0.15
PCTFAM200POV	Percent families below 200% poverty level	ACS	Area	-0.11
PCT_HSP	Percent Hispanic	PSS	School	-0.10
PCT_ASN	Percent Asian	PSS	School	0.09
CENSREG	Census region	PSS	Area	0.07
ESTGRE	Estimated enrollment used in school sampling	PSS	School	0.06
PCT_TR	Percent two or more races	PSS	School	0.05
SchoolType	Catholic and Non-Catholic	PSS	School	0.04
PCT_HP	Percent Hawaiian/Pacific Islander	PSS	School	0.03
PCT_AM	Percent American Indian/Alaskan	PSS	School	0.00

"Correlation" is the median correlation with achievement scores across grades and subjects

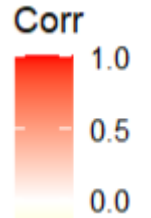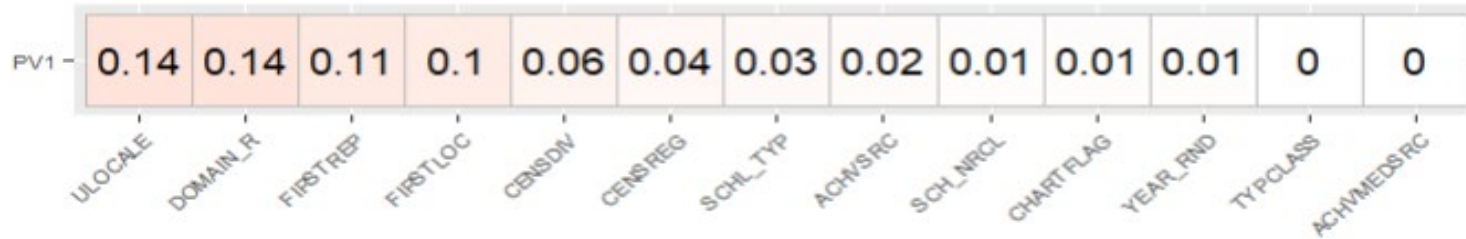# Summary at School Level – 2022 NRBA Results with Added New Covariates for Grade 4 Private

NRBA Step	Analysis	Characteristics with nonresponse bias
1	Original sample vs. responding sample (without substitutes)	Census region, Catholic/Non-Catholic, size class, Census division, school type, Mean enrollment averaged across students, Mean grade enrollment, Percent Hispanic, Percentage college degree or more ages 25 and over
2	Original sample vs. responding sample (with substitutes)	Census region, Catholic/Non-Catholic, size class, school type, Mean enrollment averaged across students, Mean grade enrollment, Percent White, Percent Hispanic, Percentage two or more races
3	Original sample vs. responding sample (with substitutes and nonresponse adjustments)	Census region, Census division, school type, Mean enrollment averaged across students, Percentage college degree or more ages 25 and over

Y vs. cont. variable

	ACHVMED_STD	PCT_HSP	PCT_ASN	PCT_WHT	PCT_BLK	PCT_TR	EST_GRE	PCT_AM	PCT_HP	ACHIEVE	ACHVMED
PV1	0.37	0.23	0.22	0.21	0.15	0.1	0.08	0.05	0.03	0.01	0.01

Y vs. cat. variable

	ULOCALE	DOMAIN_R	FIRSTREP	FIRSTLOC	CENSDIV	CENSREG	SCHL_TYP	ACHVSRC	SCH_NRCL	CHARTFLAG	YEAR_RND	TYPCLASS	ACHVMEDSRC
PV1	0.14	0.14	0.11	0.1	0.06	0.04	0.03	0.02	0.01	0.01	0.01	0	0

Corr
1.0
0.5
0.0

Y vs. area-level variable

	PctMgmtBusSciArt	WClsk	PctEducBchPlus	MedHHInc	PctEducGHS	EducAvgYrs	Pct200Pov	PctRecvIDR	PctFam200Pov	Pct150Pov
PV1	0.33	0.33	0.32	0.31	0.31	0.31	0.31	0.3	0.29	0.29

Summary – Final Variables Used in **Student**-Level NRBA **Public**

Variable Name	Label	Source	Level	Correlation
ACHVMED_STD	Standardized achievement data	CCD	School	0.38
SCH_RACE	Race from school records	CCD	School	0.35
SD	Students with disabilities status	CCD	School	0.35
NSLP	National school lunch program	CCD	School	0.35
PCTEDUCBCHPLUS	Percent college degree or more ages 25 and over	ACS	Area	0.30
EL	English learner status	CCD	School	0.27
PCTFAM200POV	Percent families below 200% poverty level	ACS	Area	-0.27
PCTFAM100POV	Percent families below poverty level	ACS	Area	-0.22
SEX	Sex	CCD	School	0.17
REL_AGE	Relative age	CCD	School	0.10

"Correlation" is the median correlation with achievement scores across grades and subjects

Summary at Student Level – 2022 NRBA Results with Added New Covariates for Grade 8 Reading Public

Domain	Analysis	Characteristics with nonresponse bias
State A	Student base weights	Sex, Race, SD, EL, Standardized achievement data, Percentage college degree or more ages 25 and over
State A	Nonresponse adjusted student weights	EL
State B	Student base weights	Race, SD, Percentage college degree or more ages 25 and over, Percentage families below 100% poverty level, Percentage families below 200% poverty level
State B	Nonresponse adjusted student weights	Sex, Race, SD, Percentage college degree or more ages 25 and over
State C	Student base weights	Sex, Free lunch, SD, EL, Standardized achievement data, Percentage college degree or more ages 25 and over, Percentage families below 100% poverty level, Percentage families below 200% poverty level
State C	Nonresponse adjusted student weights	SD, Percent families below 100% poverty level
State D	Student base weights	Relative age, SD, Standardized achievement data
State D	Nonresponse adjusted student weights	None

School-level Nonresponse Adjustment

Characteristics Used for School Nonresponse Adjustment for Private Schools

- School type (Catholic, non-Catholic)

- Census region

- Urban-centric locale (city, suburban, town, rural)

- Minority stratification

Evaluation for NR Adjustment Variables at School Level for Private Schools

#	Model Description	Adj. R^2 for grade 4	Adj. R^2 for grade 8
1	Y = SchoolType*Region*URBSTRA*MINSTRA	0.03	0.03
2	Y = SchoolType*Region*URBSTRA*PCTEDUBCHPLUS	0.06	0.06
3	Y = SchoolType*Region*URBSTRA*PCT_WHT	0.02	0.05
4	Y = SchoolType*Region*MINSTRA*PCTEDUBCHPLUS	0.06	0.06
5	Y = SchoolType*Region*MINSTRA*PCT_WHT	0.02	0.05
6	Y = SchoolType*Region*PCT_WHT*PCTEDUBCHPLUS	0.05	0.07
7	Y = SchoolType*Division*URBSTRA*MINSTRA	0.04	0.04
8	Y = SchoolType*Division*URBSTRA*PCTEDUBCHPLUS	0.06	0.07
9	Y = SchoolType*Division*URBSTRA*PCT_WHT	0.03	0.06
10	Y = SchoolType*Division*MINSTRA*PCTEDUBCHPLUS	0.06	0.08
11	Y = SchoolType*Division*MINSTRA*PCT_WHT	0.03	0.06
12	Y = SchoolType*Division*PCT_WHT*PCTEDUBCHPLUS	0.06	0.08
13	Y = SchoolType*URBSTRA*MINSTRA*PCTEDUBCHPLUS	0.07	0.07
14	Y = SchoolType*URBSTRA*MINSTRA*PCT_WHT	0.03	0.05
15	Y = SchoolType*URBSTRA*PCT_WHT*PCTEDUBCHPLUS	0.06	0.07
16	Y = SchoolType*MINSTRA*PCT_WHT*PCTEDUBCHPLUS	0.06	0.07
17	Y = SchoolType*Division*URBSTRA	0.03	0.03
18	Y = SchoolType*Division*MINSTRA	0.02	0.04
19	Y = SchoolType*Division*PCTEDUBCHPLUS	0.05	0.06
20	Y = SchoolType*Division*PCT_WHT	0.02	0.06

Conclusion

NRBA evaluation with added covariates at school and student levels

- School-level NRBA: Private school samples for grade 4 and 8 had less than an 85 percent response rate

- Student-level NRBA: Several states in the public-school samples for grade 4 and 8 had less than an 85 percent response rate

- Summary

  - Grade 4 & 8

    - Decrease in number of characteristics with significant differences
    - Possibly significant nonresponse bias after the NR adjustment

Conclusion

Evaluation for NR adjustment variables at the school level

- Suggest to add percent college degree or more ages 25 and over for the grade 4 private school NR adjustment

- Suggest to add census division, percent college degree or more ages 25 and over and percent white categorical variable for the grade 8 private school NR adjustment

Next Steps

- Student-level NR adjustment evaluation

- Decide on whether to add new variables to the NR adjustment

Thank you

SabrinaZhang@westat.com
TomKrenzke@westat.com

westat.com

Photos are for illustrative purposes only. All persons depicted, unless otherwise stated, are models.

References

- Peytcheva, E., and Groves, R.M. (2009). Using Variation in Response Rates of Demographic Subgroups as Evidence of Nonresponse Bias in Survey Estimates. *Journal of Official Statistics*, **25**, 193-201.

- Rao, J.N.K., and Scott, A.J. (1984). On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data.

- Schneider, B., Krenzke, T., and Gamble, L. (2023, July). Prediction-based methods for assessing nonresponse bias and the effectiveness or nonresponse adjustments in an international study. European Survey Research Association Conference, Milan, Italy.