Survey Sampling Weights and Item Response Parameter Estimation

Spring 2014 Survey Methodology Colloquium Series

Simmons School of Education and Human Development

Center on Research & Evaluation

Paul Yovanoff, Ph.D.
Department of Teaching & Learning
pyovanoff@smu.edu
(541) 346-1495

Topics

- Measurement modeling of survey questionnaire responses is completed for various purposes, e.g., estimate respondent latent traits, estimate item parameters, or to improve survey instrumentation.

- Survey respondents are often sampled such that sampling weights are necessary analytically for statistical representation of sampled subpopulations.

- Are sampling weights important when modeling survey question item parameters?

Survey Item Response Modeling (briefly)

Why should we model survey item responses?

- By item design, an item response is generated by an underlying hypothetical process.

- Responses to items are used to estimate respondent characteristics.

- Therefore, item psychometrics should be known.

Survey Item Response Modeling (briefly)

The Dichotomous Response Model

$$P\left(X_{ij}\middle|\theta_{j},\beta_{i},\alpha_{i},\gamma_{i}\right)=\gamma_{i}+\left(1-\gamma_{i}\right)\frac{\exp\left[\alpha_{i}\left(\theta_{j}-\beta_{i}\right)\right]}{1+\exp\left[\alpha_{i}\left(\theta_{j}-\beta_{i}\right)\right]}$$

where,

X is the response to item i by respondent j

θ is the respondent j latent trait level

β is the item i difficulty

α is the item i discrimination (constrained to 1 for these items)

γ is the item i guessing (constrained to 0)
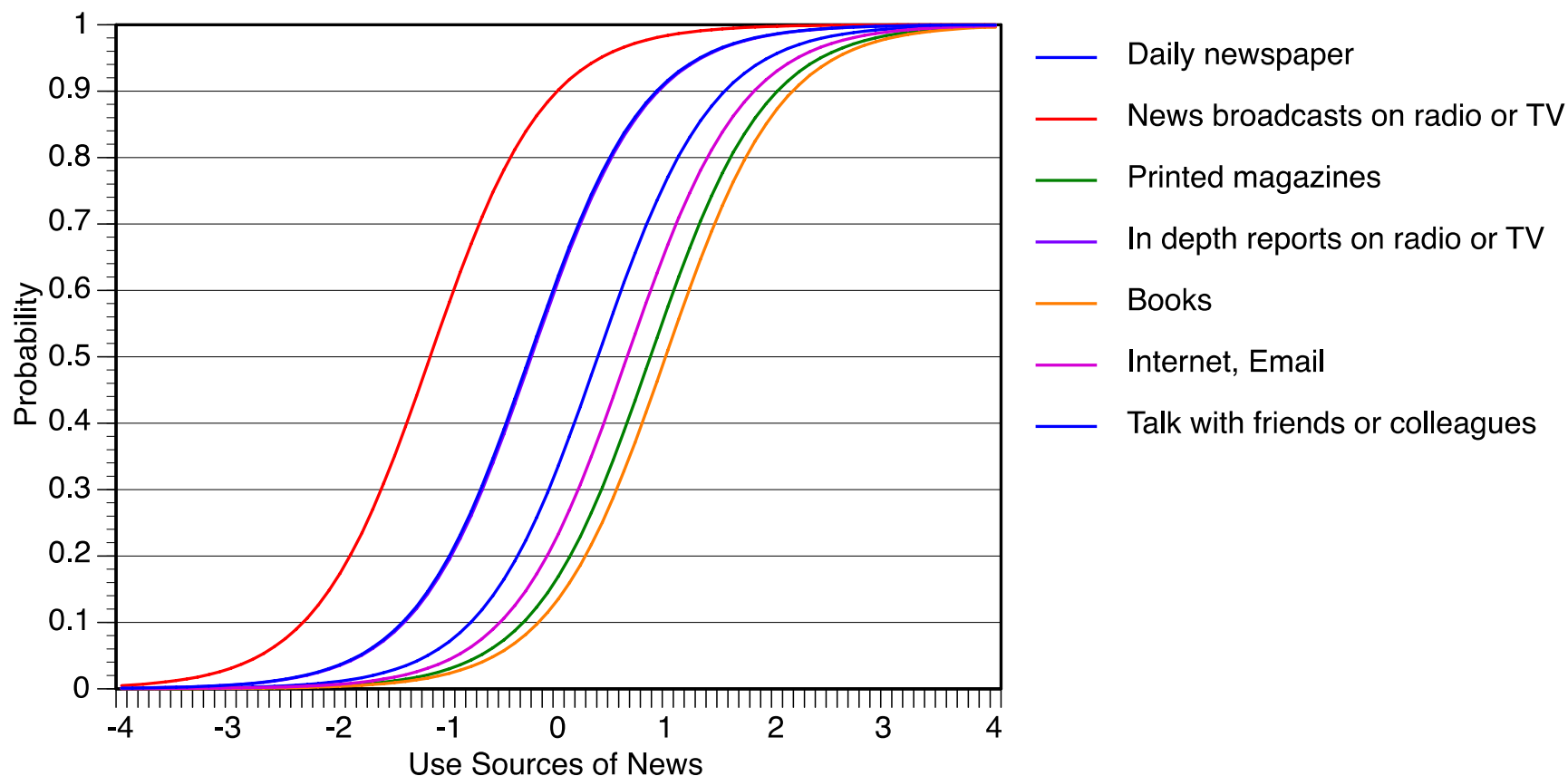
Survey Item Response Modeling (briefly)

World Values Survey (dichotomous response)

People use different sources to learn what is going on in their country and the world. For each of the following sources, please indicate whether you used it last week or did not use it last week to obtain information.

	No	Yes
a. Daily newspaper	1	2
b. News broadcasts on radio or TV	1	2
c. Printed magazines	1	2
d. In depth reports on radio or TV	1	2
e. Books	1	2
f. Internet, Email	1	2
g. Talk with friends or colleagues	1	2

Survey Item Response Modeling (briefly)

Survey Item Response Modeling (briefly)

Polytomous Response Model

Graded Response Model (Samejima, 1997)

$$P^*_{x_{ik}}(\theta_j) = P\left(X_{ik} \geq x_{ik} \mid \theta_j\right) = \frac{e^{Da_i\left(\theta_j - b_{x_{ik}}\right)}}{1 + e^{Da_i\left(\theta_j - b_{x_{ik}}\right)}}$$

for m categories there are m-1 category thresholds

where,

$$P\left(X_{ik} = x_{ik} \mid \theta_j\right) = P^*_{x_{ij}}(\theta_i) - P^*_{x_{ij}+1}(\theta_i)$$

for person j
responding in category k of item i

Survey Item Response Modeling (briefly)
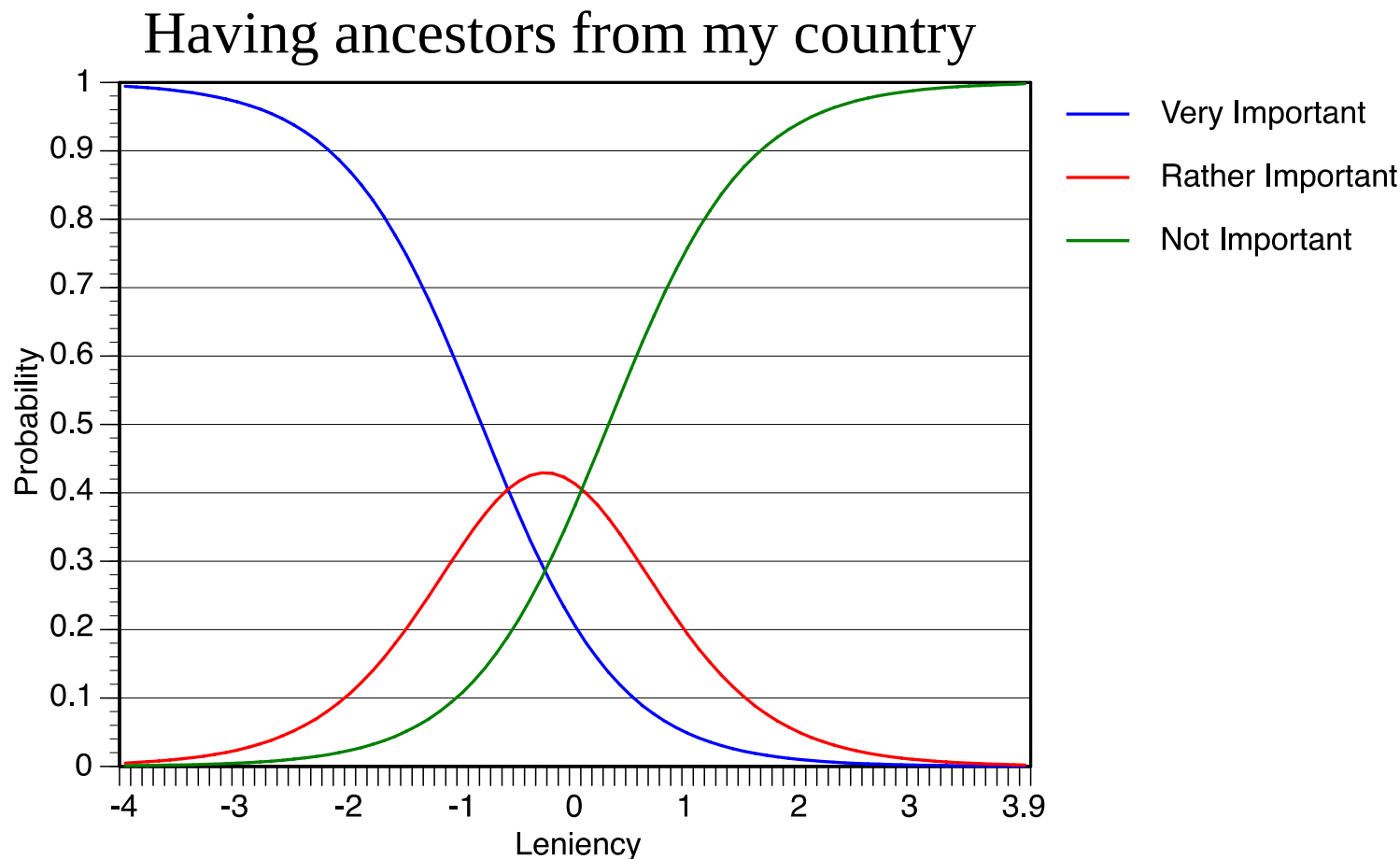
## World Values Survey (polytomous response)

In your opinion, how important should the following be as requirements for somebody seeking citizenship of your country?

Specify for each requirement if you consider it as very important, rather important or not important
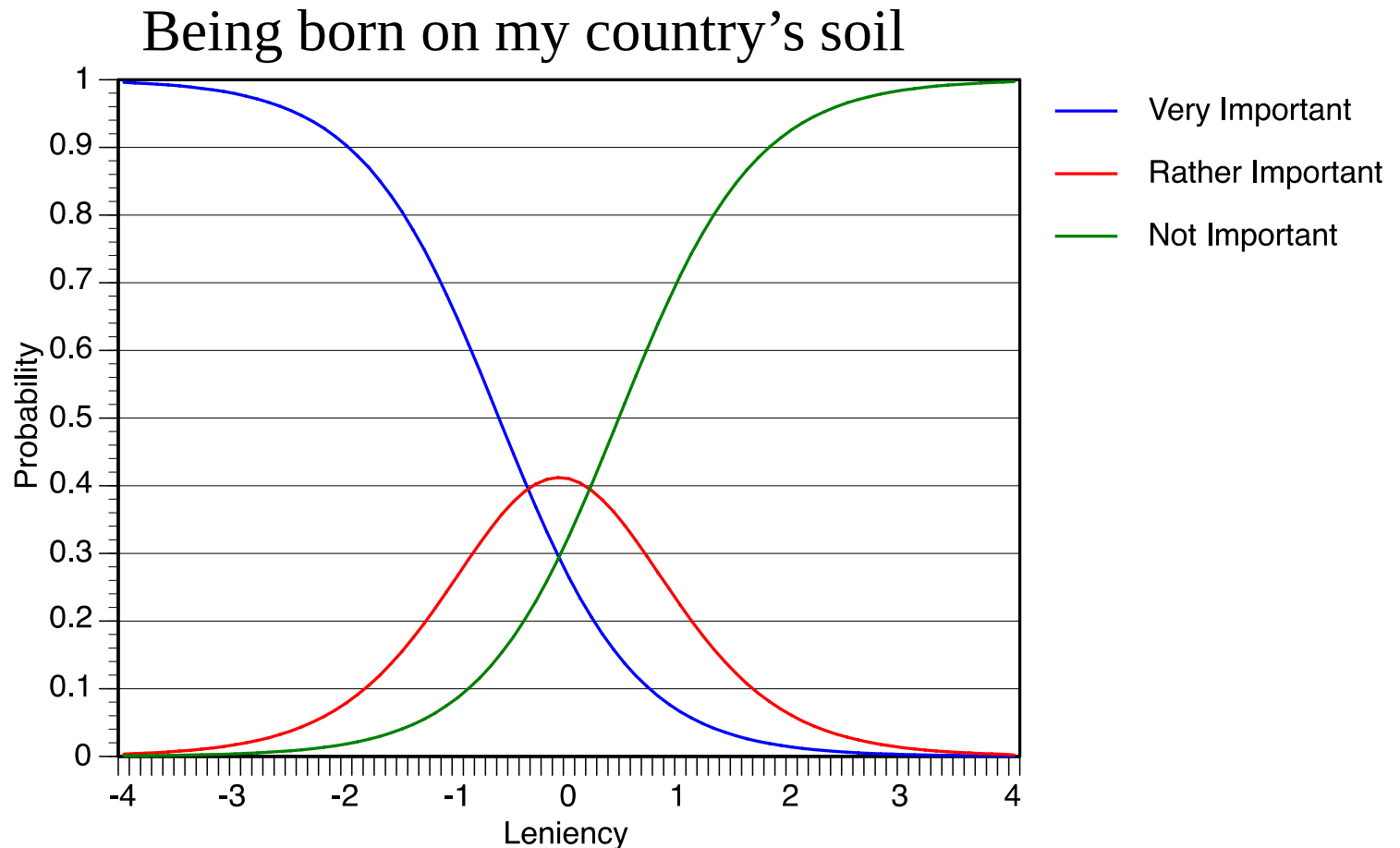
	Very Important	Rather Important	Not Important
a. Having ancestors from my country	1	2	3
b. Being born on my country's soil	1	2	3
c. Adopting the customs of my country	1	2	3
d. Abiding by my country's laws	1	2	3

Survey Item Response Modeling (briefly)

Having ancestors from my country

Survey Item Response Modeling (briefly)

Being born on my country's soil

Survey Item Response Modeling (briefly)

Adopting the customs of my country

Survey Item Response Modeling (briefly)

Abiding by my country's laws

Survey Item Response Modeling (briefly)

Sample Invariant Parameter Estimates

- Item response model parameters are expected to be sample invariant with respect to trait level.

- Samples that vary with respect to trait level are expected to provide equivalent parameter estimates.

Question. Are parameter estimates invariant with respect to group identity, e.g., sex, ethnicity?

Survey Sampling Weights (briefly)

What are sampling weights?

- Survey data are often collected using a nonrandom sampling procedure, e.g., cluster sampling, stratified sampling, multistage sampling; subpopulations may have been over- or under-sampled.

- With known probability of being included (inclusion probabilities), sampling weights can be computed

- Sampling weights are the inverse of the inclusion probability of being selected given the sampling design.

Survey Sampling Weights (briefly)

Why use sampling weights?

- They provide unbiased parameter estimates

- They provide relatively accurate standard errors

Survey Sampling Weights (briefly)

Why don't we use sampling weights?.

- Weights are often unavailable.

- Not all computing software enables use of weights.

- Complex methods may be needed to properly use survey weights, e.g., multilevel regression modeling, computation of normalized weights rather than raw weights.

Item Response Modeling and Sampling Weights

2007 National Household Education Surveys Program Parent and Family Involvement in Education Survey (http://nces.ed.gov/nhes/surveytopics.asp)

Complex Sampling Design

Subsampling of telephone numbers

2-phase stratification by minority

Phase 1—over sampling areas with high % of Black or Hispanic

Phase 2—within minority stratum mailable

Item Response Modeling and Sampling Weights

2007 National Household Education Surveys Program Parent and Family Involvement in Education Survey (http://nces.ed.gov/nhes/surveytopics.asp)

Case Level Sampling Weights

80 replicates (80 subsamples)

Full sample weight (average of the 80 replicates)

Item Response Modeling and Sampling Weights

Methods

- 10,681 Children Attending Public or Private School
  - Unweighted (13.3% in private school, N=1,687)
  - Weighted (11.1% in private school)

- Item Response Modeling. Satisfaction with School Features; Graded response model parameter estimation

- Test of Measurement Invariance using SEM with unweighted and weighted samples

Item Response Modeling and Sampling Weights

2007 Parent and Family Involvement in Education Survey

Would you say that you are very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied with . . .

	Very Dissatisfied	Somewhat Dissatisfied	Somewhat Satisfied	Very Satisfied
a. School	1	2	3	4
b. Teachers	1	2	3	4
c. Academic Standards	1	2	3	4
d. Order and Discipline	1	2	3	4
e. Staff/Parent Interaction	1	2	3	4

Item Response Modeling and Sampling Weights

Odds That Respondents from Private Schools are Satisfied Compared to Respondents from Public Schools

School Feature	Unweighted		Weighted	
	Odds Ratio	CI	Odds Ratio	CI
School	5.13	1.75, 8.79	5.61	5.57, 5.66
Teacher	2.80	1.75, 4.47	3.03	3.01, 3.05
Standards	4.94	2.83, 8.64	4.49	4.45, 4.52
Discipline	3.91	2.53, 6.04	4.08	4.05, 4.11
Communication	3.57	2.39, 5.33	4.08	4.05, 4.10

Clearly, respondents from private schools are more satisfied.

Item Response Modeling and Sampling Weights

Unweighted Data

Model Information Fit Statistics for Model Comparison

Model	AIC	BIC	A-BIC
Null	42754.512	42865.820	42811.800
School	42740.661	42871.611	42808.058
Teacher	42734.264	42865.214	42801.661
Standards	42731.756	42862.707	42799.153
Discipline	42746.041	42876.991	42813.438
Communication	42752.954	42883.904	42820.351

best fitting model

Item Response Modeling and Sampling Weights

Weighted Data

Model Information Fit Statistics for Model Comparison

Model	AIC	BIC	A-BIC
Null	43111.932	43223.240	43169.220
School	43102.391	43233.342	43169.789
Teacher	43088.161	43219.111	43155.558
Standards	43085.894	43216.844	43153.291
Discipline	43098.215	43229.166	43165.612
Communication	43113.756	43244.706	43181.153

best fitting model

Item Response Modeling and Sampling Weights

Satisfaction w/Teachers Parameters by School Type

Parameter	Public		Private	
	Est	SE	Est	SE
unweighted				
Category 1 – 2+	-6.931	0.150	-6.056	0.452
Category 2 – 3+	-4.762	0.094	-4.541	0.266
Category 3 - 4	-1.130	0.060	-0.928	0.131
weighted				
Category 1 – 2+	-6.959	0.203	-6.105	0.534
Category 2 – 3+	-4.731	0.135	-4.461	0.308
Category 3 - 4	-1.096	0.086	-0.996	0.167

Item Response Modeling and Sampling Weights
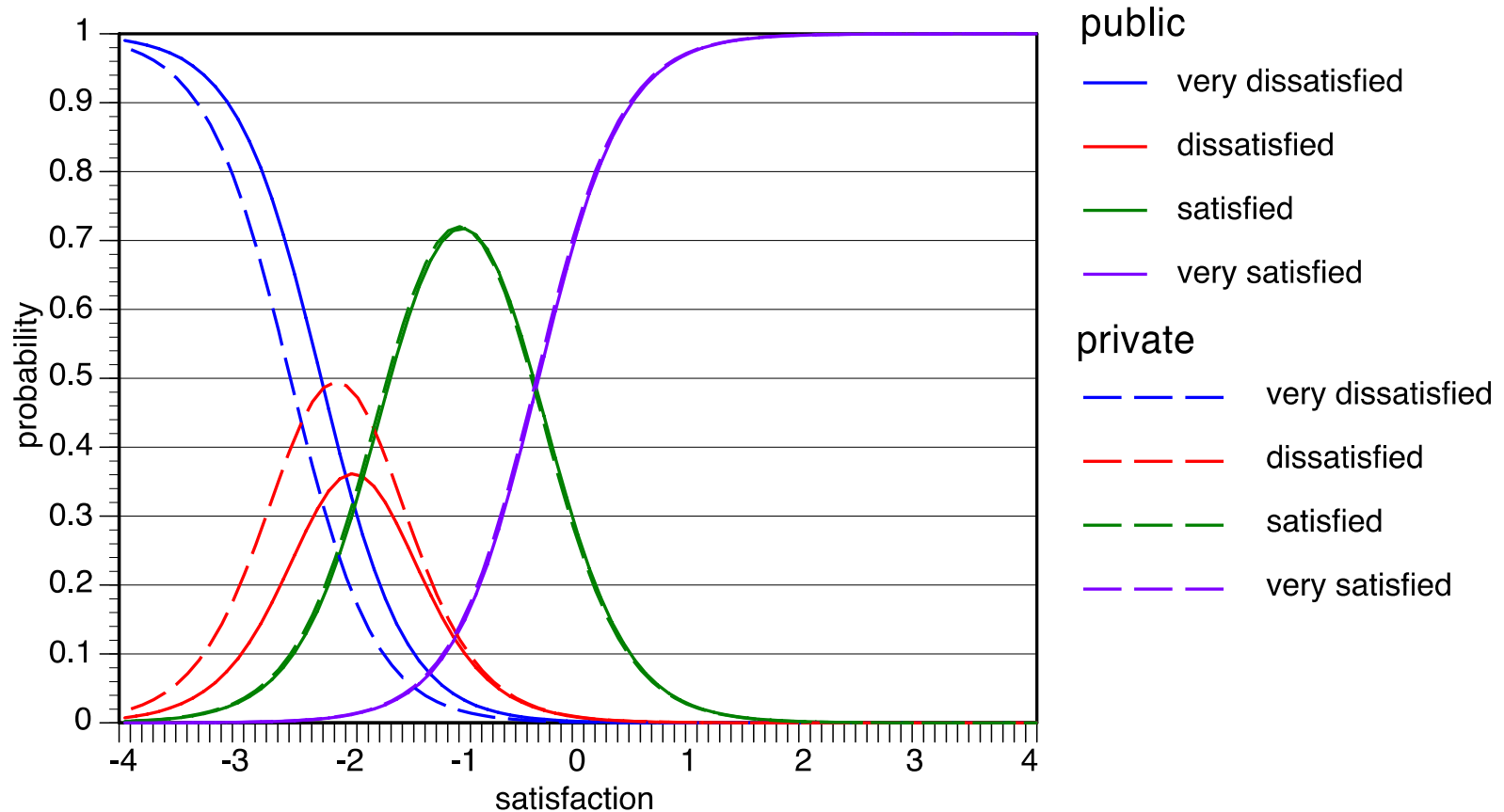
Satisfaction w/Standards Parameters by School Type

Parameter	Public		Private	
	Est	SE	Est	SE
unweighted				
Category 1 – 2+	-6.276	0.130	-6.927	0.615
Category 2 – 3+	-4.460	0.089	-5.140	0.323
Category 3 - 4	-0.984	0.059	-1.731	0.147
weighted				
Category 1 – 2+	-6.287	0.171	-7.755	0.629
Category 2 – 3+	-4.360	0.121	-4.816	0.464
Category 3 - 4	-0.887	0.091	-1.754	0.189

Item Response Modeling and Sampling Weights
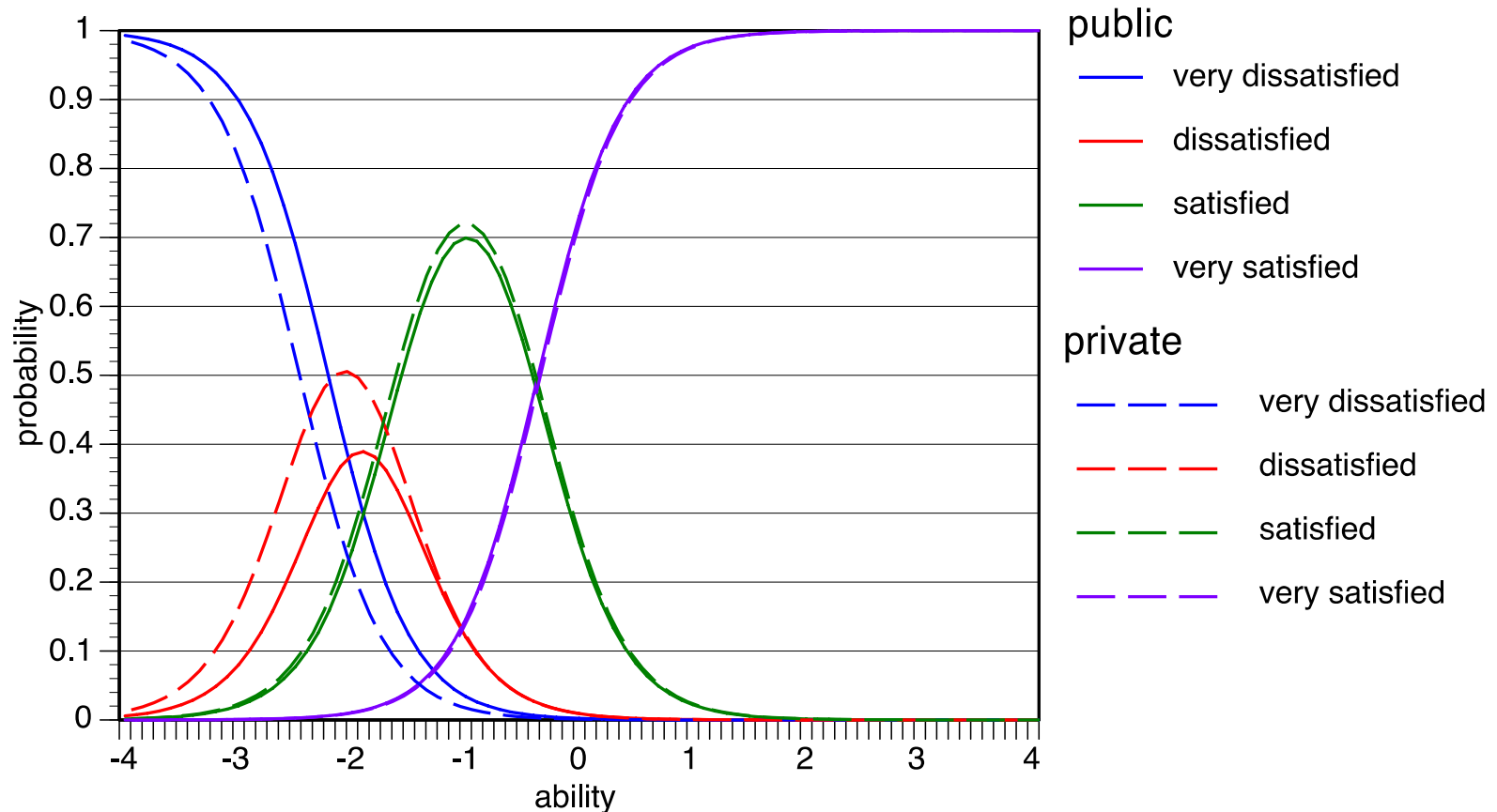
Satisfaction with Teachers (unweighted)

Item Response Modeling and Sampling Weights
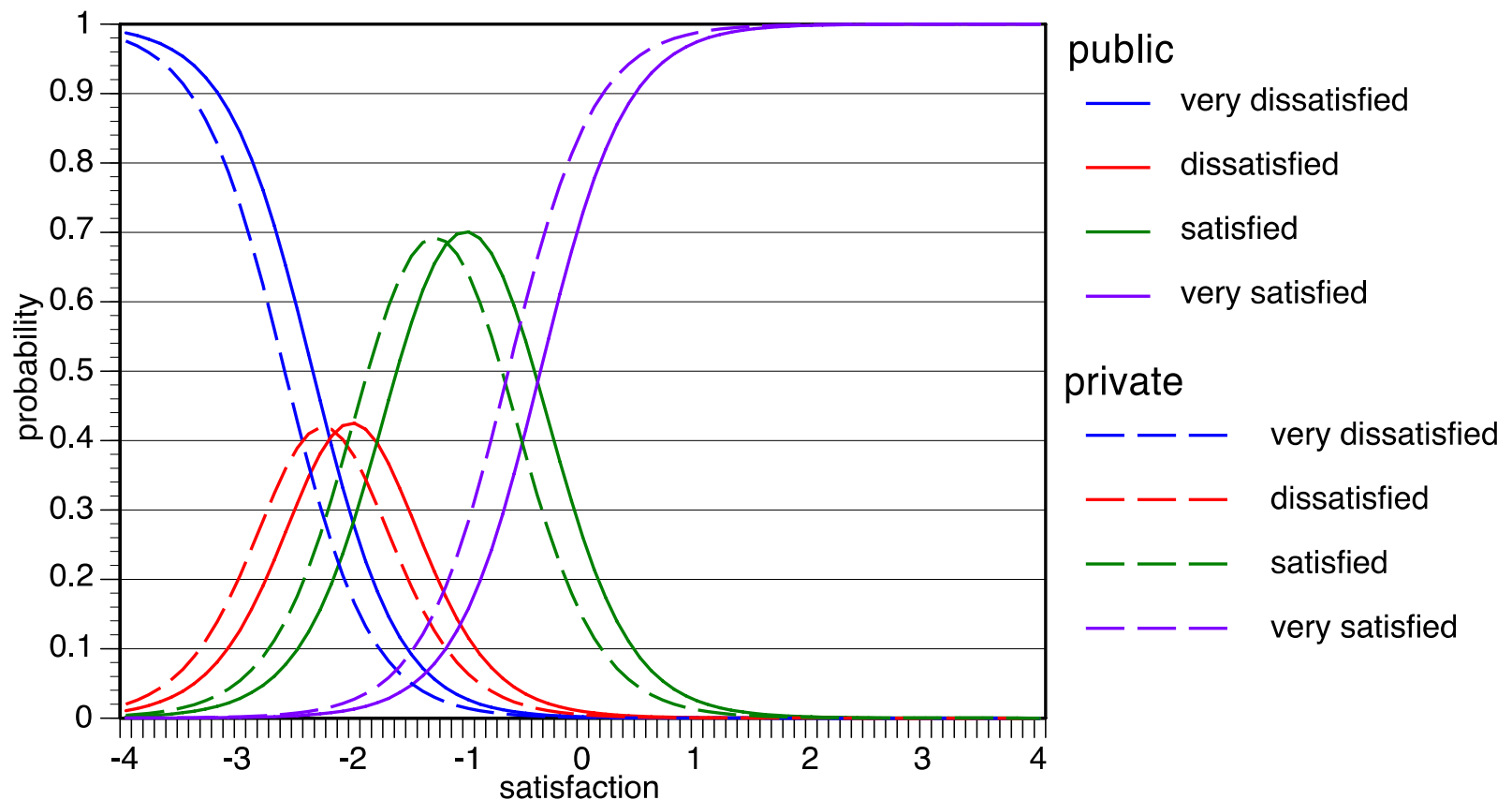
Satisfaction with Teachers (weighted)

Item Response Modeling and Sampling Weights
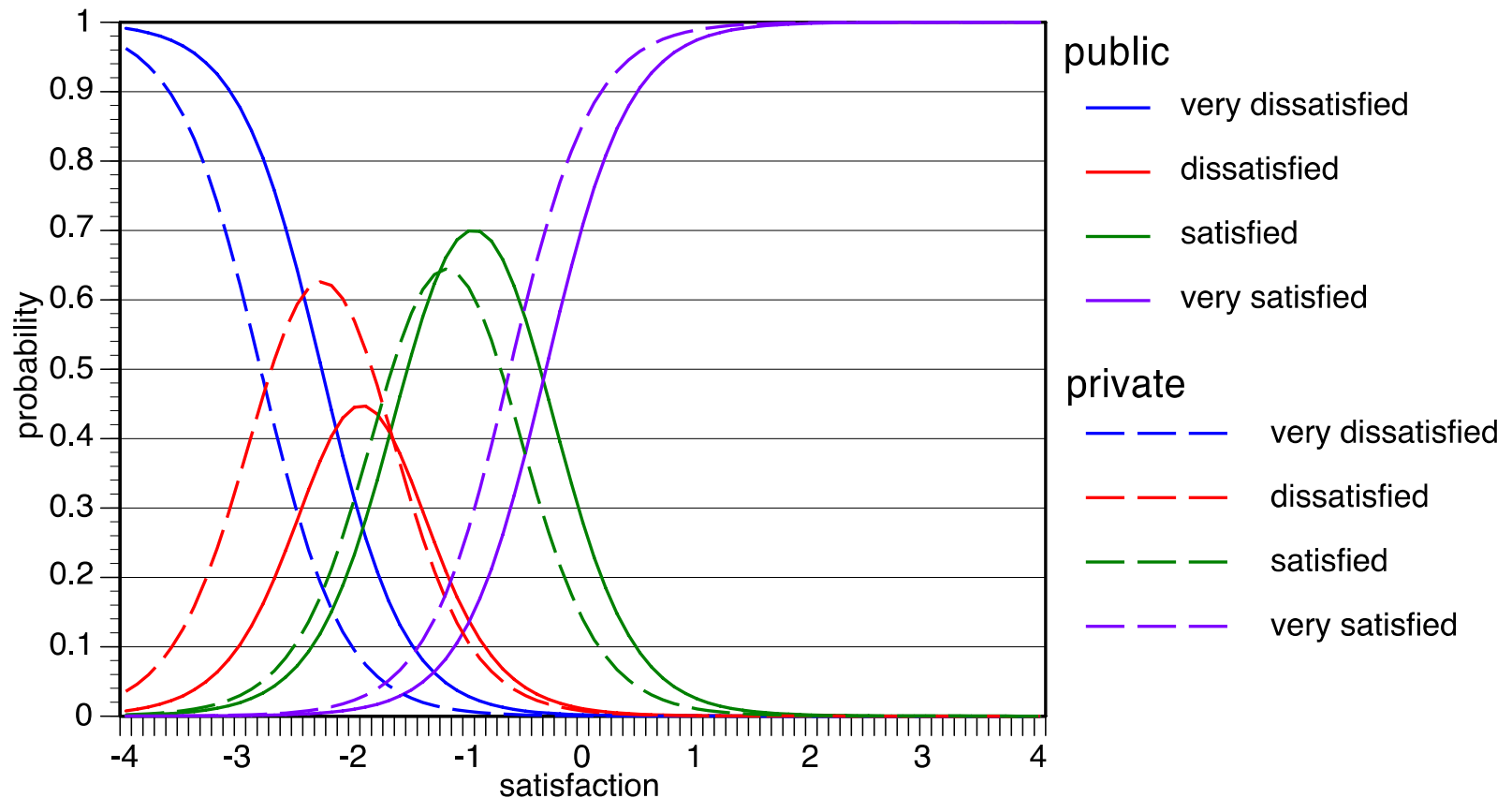
Satisfaction with Standards (unweighted)

Item Response Modeling and Sampling Weights

Satisfaction with Standards (weighted)

Item Response Modeling and Sampling Weights

Question. Are survey sampling weights relevant when estimating item response parameters?

- Parameter estimates are affected by weights, they will theoretically be less biased.

- Parameter standard errors tend to be affected by weights; they tend to increase.

- Whether estimating parameters or testing for measurement invariance, use sampling weights when possible.

Item Response Modeling and Sampling Weights

Question. Are survey sampling weights relevant when estimating item response parameters?

- If the weights change representation of subsamples and if the items are biased with respect to subsamples, then the estimates are biased in both cases. The weighted data provides a better basis for testing for sample invariance.

- Be sure to use software that accommodates weights

References and Resources

Dayton, C. M. (2003). Model comparison using information measures. Journal of Modern Applied Statistical Methods, 2, 281-292.

Embretson, S.E., & Reise, S.P. (2000). Item response theory for psychologists. Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc., Publishers.

Hagedorn, M., Roth, S. B., O'Donnell, K., Smith, S., & Mulligan, G. (2008). National Household Education Surveys Program of 2007. Data File User's Manual, Volume 1. Study Overview and Methodology. National Center for Education Statistics.

Hahs-Vaughn, D. L. (2005). A primer for using and understanding weights with national datasets. Journal of Experimental Education, 73, 221-248.

Kaplan, D., & Ferguson, A. J. (1999). On the utilization of sample weights in latent variable models. Structural Equation Modeling: A Multidisciplinary Journal, 6, 305-321.

References and Resources

Millsap, R. E. (2011). Statistical approaches to measurement invariance. New York, New York: Routledge.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, 17.

Thomas, S. L., & Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling designs. Research in Higher Education, 42, 517-540.

National Center for Education Statistics. National Household Surveys. http://nces.ed.gov/nhes/surveytopics.asp

World Values Survey. http://www.worldvaluessurvey.org/