How Subgroup Characteristics Affects Equating Methods' Academic Growth Detection

Ou Zhang
Research and Evaluation Methodology Program, College of Education
University of Florida
E-mail: zhango@ufl.edu, zhangou888@gmail.com

Charles DePascale, PH.D.,
Center for Assessment (NCIEA)
Dover, NH 03821-0351

M. David Miller, PH.D.,
Research and Evaluation Methodology Program, College of Education
University of Florida

UF | UNIVERSITY of FLORIDA

Introduction

The relationship of Academic growth and subgroup performance with NCLB (2001)

- Academic growth → AYP

  Academic growth is usually assessed by comparing the performance of students on the standardized test across years.

- Subgroup performance difference and academic growth → AYP

Theoretical Framework

- Scale Transformation and Equating Method to Detect Academic Growth in Previous Research
  - Moment equating methods (e.g. Keller, et al, 2007)
  - Test characteristic curve methods (TCC, e.g. Hanson & Béguin, 2002)
  - Fixed common item parameter methods (FCIP, Paek & Young, 2005)
  - Concurrent calibration (e.g. Kim & Kolen, 2006)

  Stocking-Lord Test Characteristic Curve method → best performance

- IRT models for Mixed Format Test
  - Grade Response Model (GRM, Samejima, 1969;1996)
  - 2PL model

Purpose of Study

1. To what extent differential academic growths are captured through common equating designs and IRT procedures, as different levels of growth occur in various size subgroups.

2. To determine whether the equating approach recovers the person ability estimates, as population distribution suppression, or inflation is found when subgroup growth occurs across years.

3. To investigate the robustness of the IRT estimation and equating methods in population achievement level classification as the subgroup growths vary.

Methods

- **Test Form**

  Item Parameter Source: New England Common Assessment Program (NECAP) 2008 8th Grade Math test

  ITEMS: total : 47 items (64 points) including 38 dichotomous items and 9 polytomous items (five 3-category items and four 5-category items)

-

- **Classification cut-scores**

  Ability (theta) scale

  Raw score cuts: 19, 28, and 48 (NECAP math 2008)

Methods (cont.)

- **Simulation design**

  This research investigates academic growth detection and ability estimate recovery in 3 different conditions:

  1. Subgroup proportion of the population (0.05, 0.1, 0.25, and 0.5)
  2. Subgroup mean growth (0.25, 0.5, 0.75, and 1.0)
  3. Population distribution change (no changes-normal distributed, mean shift, skewness and kurtosis change)

Therefore, three factors are completely crossed:

4 (subgroup ratio) × 5(subgroup mean growth) × 2(population distribution change)

20 conditions, 2 population distribution changes, Total population: N=20,000 (M=0, SD=1), 100 iterations

Methods (cont.)

- Condition Table:

Subgroup mean and SD Change						
		Mean Growth Level				
		0	**0.25**	**0.5**	**0.75**	**1**
Subgroup: Population Ratio	**0.05**	1	2	3	4	5
	0.1	6	7	8	9	10
	0.25	11	12	13	14	15
	0.5	16	17	18	19	20

Methods (cont.)

- Software: R.12.2.1
- R packages : 'ltm'(Rizopoulos, 2006) and 'plink' (Weeks, 2010)
- IRT models and Equating Method

  Grade Response Model and 2PL model

  IRT true score equating via Stocking-Lord

- **Criteria**

  Measure of Growth (mean differences of subgroup, majority group, total population)

  Under-Classification and Over-Classification

  Under-Classification coefficient → AYP (negatively impact)

Results

- Ability Estimates Mean Difference

Subgroup Proportion of Population		Subgroup Ability Expected Growth			
0.05 (1000:19000)	Sub group expected mean growth	**0.25**	**0.50**	**0.75**	**1.00**
	Subgroup estimated mean growth	0.1212	0.2401	0.3914	0.5461
	Majority group estimated mean changes	-0.0095	-0.0172	-0.0299	-0.0406
0.1 (2000:18000)	Sub group expected mean growth	**0.25**	**0.50**	**0.75**	**1.00**
	Subgroup estimated mean growth	0.0896	0.2164	0.3181	0.4829
	Majority group estimated mean changes	-0.0170	-0.0301	-0.0531	-0.0783
0.25 (5000: 15000)	Sub group expected mean growth	**0.25**	**0.50**	**0.75**	**1.00**
	Subgroup estimated mean growth	0.0626	0.1221	0.1976	0.2879
	Majority group estimated mean changes	-0.0289	-0.0670	-0.1074	-0.1658
0.5 (10000:10000)	Sub group expected mean growth	**0.25**	**0.50**	**0.75**	**1.00**
	Subgroup estimated mean growth	0.0221	0.0395	0.0578	0.0803
	Majority group estimated mean changes	-0.0367	-0.0806	-0.1478	-0.2118

Results (cont.)

- Over-classification and Under-classification (Example)

	Estimation Class 1	Estimation Class 2	Estimation Class 3	Estimation Class 4	Classification based on ability
Ability Class 1	5471	1049	0	0	**6520**
Ability Class 2	0	3290	19	0	**3309**
Ability Class 3	0	256	7164	0	**7420**
Ability Class 4	0	0	862	1889	**2751**
Classification based on Estimation	**5471**	**4595**	**8045**	**1889**	**20000**

Results (cont.)

- Over-classification and Under-classification (Total Population)

Subgroup Proportion of Population		Subgroup Ability Expected Growth				
0.05 (1000:19000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0534	0.0459	0.03975	0.03025	0.02825
	Under-estimation Proportion	0.0559	0.06475	0.07075	0.08665	0.08275
0.1 (2000:18000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.04585	0.03925	0.0336	0.0141	0.0058
	Under-estimation Proportion	0.07655	0.0961	0.09375	0.11495	0.1377
0.25 (5000: 15000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.04005	0.02745	0.0019	0.00005	0.0000
	Under-estimation Proportion	0.1263	0.1446	0.19025	0.2386	0.3439
0.5 (10000:10000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.02825	0.006	0.0000	0.0000	0.0000
	Under-estimation Proportion	0.1836	0.22905	0.3131	0.44025	0.5276

Results (cont.)

- Over-classification and Under-classification (Subgroup)

Subgroup Proportion of Population		Subgroup Ability Expected Growth				
0.05 (1000:19000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0450	0.0440	0.0430	0.0290	0.0290
	Under-estimation Proportion	0.0120	0.0220	0.0330	0.0640	0.0680
0.1 (2000:18000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0390	0.0430	0.0355	0.0120	0.0080
	Under-estimation Proportion	0.0200	0.0305	0.0475	0.0660	0.1070
0.25 (5000: 15000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0460	0.0312	0.0032	0.0002	0.0000
	Under-estimation Proportion	0.0288	0.0492	0.1012	0.1504	0.2830
0.5 (10000:10000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0480	0.0102	0.0000	0.0000	0.0000
	Under-estimation Proportion	0.0593	0.1109	0.2150	0.3736	0.4780

Results (cont.)

- Over-classification and Under-classification (Majority group)

Subgroup Proportion of Population		Subgroup Ability Expected Growth				
0.05 (1000:19000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.0534	0.0459	0.03975	0.03025	0.02825
	Under-estimation Proportion	0.0559	0.06475	0.07075	0.08665	0.08275
0.1 (2000:18000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.04585	0.03925	0.0336	0.0141	0.0058
	Under-estimation Proportion	0.07655	0.0961	0.09375	0.11495	0.1377
0.25 (5000: 15000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.04005	0.02745	0.0019	0.00005	0.0000
	Under-estimation Proportion	0.1263	0.1446	0.19025	0.2386	0.3439
0.5 (10000:10000)	Sub group expected mean growth	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**
	Over-estimation Proportion	0.02825	0.006	0.0000	0.0000	0.0000
	Under-estimation Proportion	0.1836	0.22905	0.3131	0.44025	0.5276

Findings from previous 3 tables

1. The under-classification and over-classification occur in the null condition.
2. As the subgroup growth increases, the over-classification proportion decreases accordingly.
3. As the subgroup/total population ratio increases, the over-classification proportion decreases as well.
4. As the subgroup growth increases, the under-classification proportion increases accordingly.
5. As the subgroup/total population ratio increases, the under-classification proportion increases as well.

Discussion

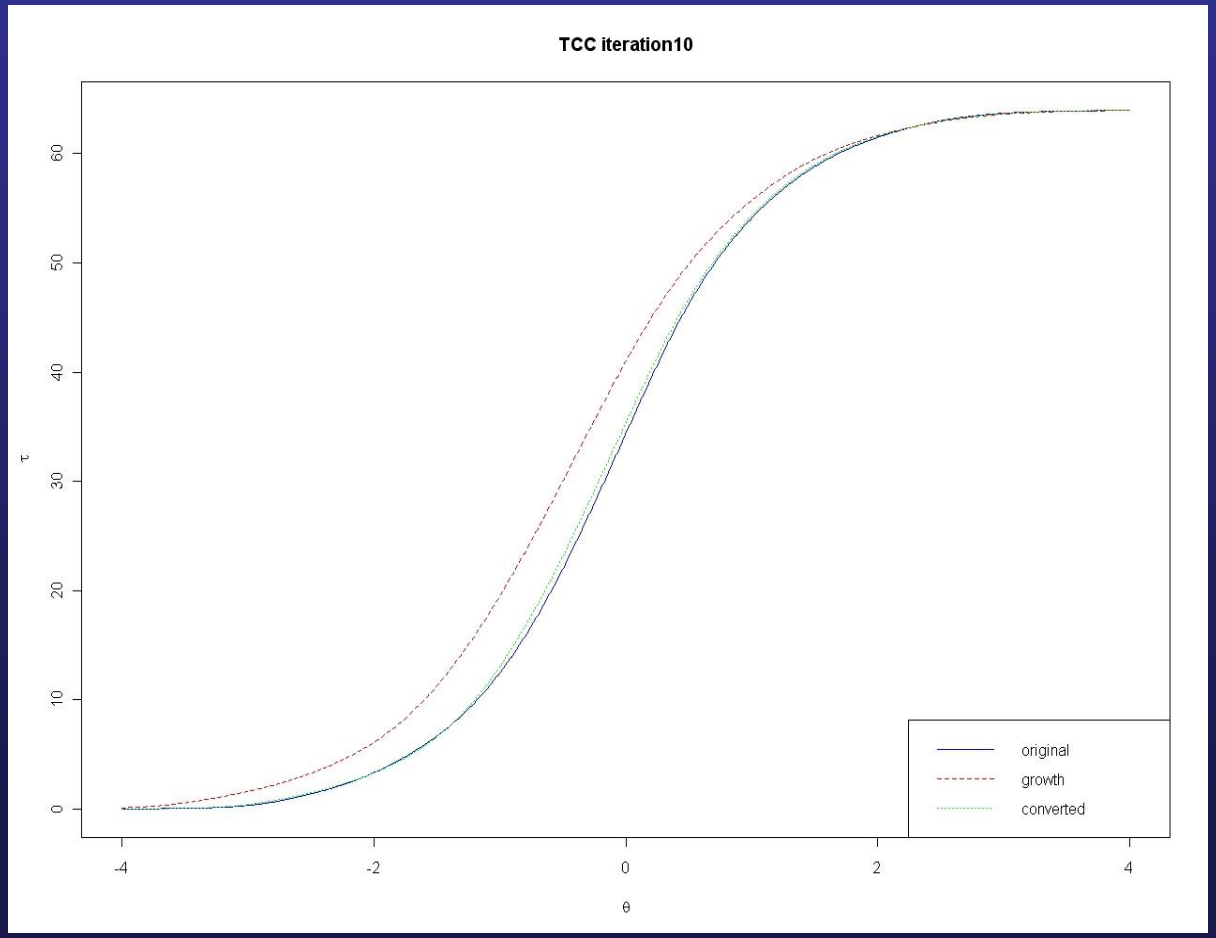
- The results suggest that the size of the subgroup population (i.e. large subgroup/total population ratio) affects the performance of IRT estimation and equating design most, compared with other factors.

  This phenomenon indicates that the non-normal characteristics of the total population distribution negatively affect the performance of default IRT estimation (i.e. normally distributed population distribution assumption is hold) even before the equating approach is applied.

- The size of the subgroup population influences the over-classification and under-classification most.

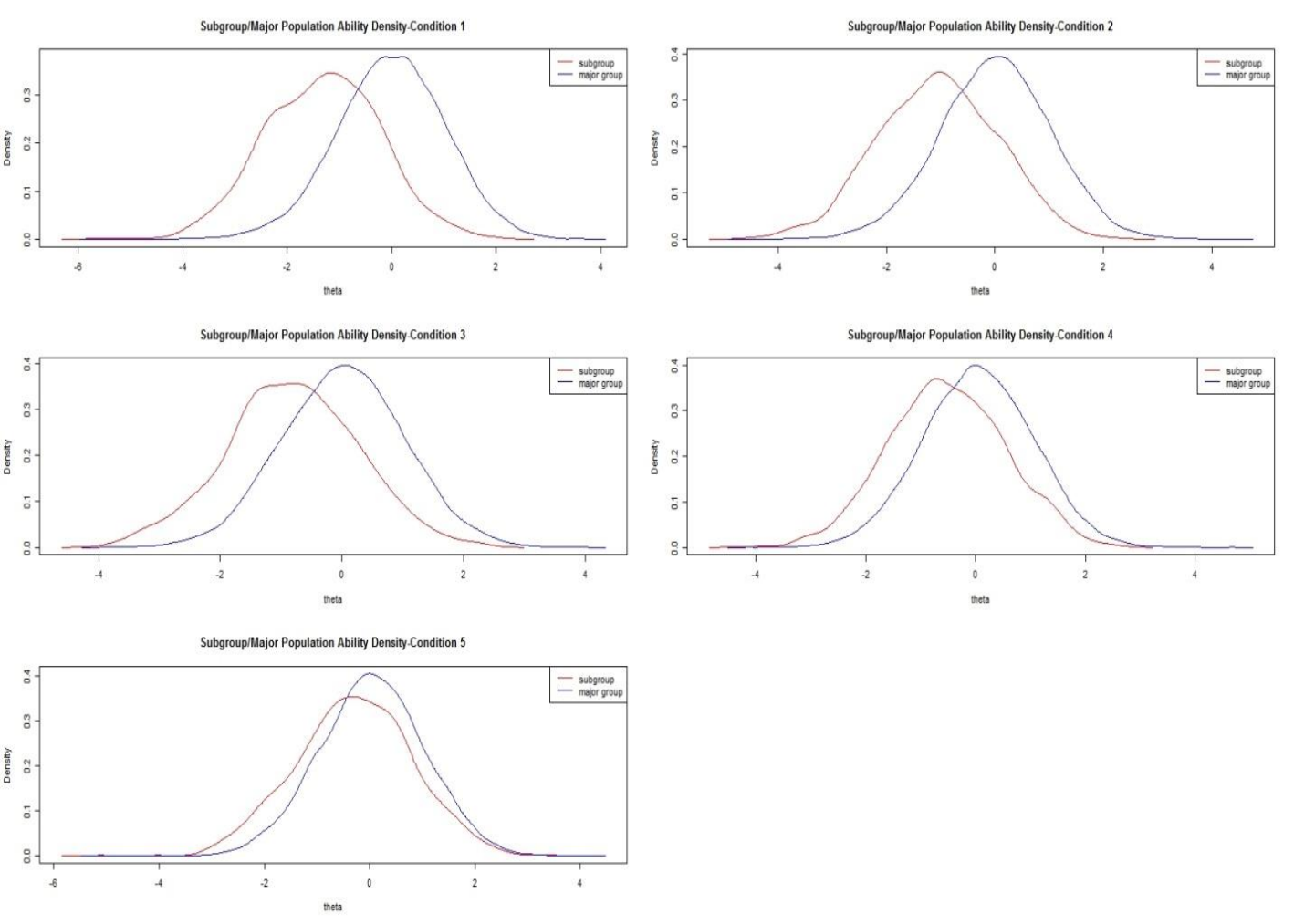
Discussion (cont.)

- Regardless of the negative effects from the non-normal characteristics of the total population distribution, true score equating method via Stocking-Lord scale linking approach did play a positive role in recovering the person ability estimates as subgroup growth occurs across years.



Limitation and Future Research Direction

- Multi-group mixed normal distribution simulation



Limitation and Future Research Direction (cont.)

- The weakness of the default IRT estimation when the non-normal population distribution characteristics exists
- The prior ability distribution is set as default normal distribution to match the circumstance as the usual procedure in state's large scale assessment.
- Future studies:
  posterior ability distribution updates (Paek & Young, 2005)
  nonparametric IRT approach (Sijtsma, 2002)

Conclusion

- Inappropriate subgroup population sample size selection may raise questions as to the appropriateness of the results of equating method academic growth detection.
- It is important to consider the size of subgroup population and the distribution of population in the academic growth detection analysis.

Thank you!

zhango@ufl.edu

UF | UNIVERSITY of FLORIDA