
Critical Issues in Growth Models using Performance on Annual State Tests

Joseph Stevens
University of Oregon
and
Ann Schulte
North Carolina State University
Arizona State University (as of 7/1/2012)

Presented at the CCSSO National Conference on Student Assessment, CCSSO,
Minneapolis, MN, June, 2012.

© Stevens, 2012



1

Contact Information:

Joseph Stevens, Ph.D.
University of Oregon
Eugene, OR 97403
(541) 346-2445
stevensj@uoregon.edu

Ann Schulte, Ph.D.
North Carolina State University
Ann_Schulte@ncsu.edu

Presentation available at:
NCAASE web site: <http://www.ncaase.com/>

This work was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R32C110004 awarded to the University of Oregon. The opinions expressed are those of the authors and do not necessarily represent views of the Institute or the U.S. Department of Education.



2

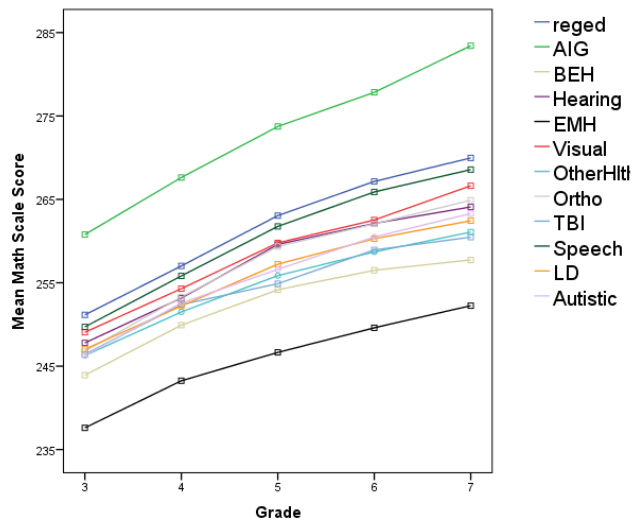
Brief Background

- NCAASE is studying achievement growth of students with and without disabilities in four partner states (AZ, NC, OR, PA)
- NCAASE will study formative (EasyCBM), interim (NWEA MAP), and summative assessments
- Study of student's individual growth trajectories
- Study of school effects
- In this presentation we share some preliminary results of analyses on first cohort of data from North Carolina
- Describe additional issues and some future directions

NCAASE Cornerstone Study: North Carolina

- What are special challenges in modeling growth?
 - Attrition, mobility, retention
 - Correctly modeling functional form
 - Other issues (see Stevens, Zvoch, & Biancarosa 2012 NCME paper)
- Basic questions about the population of students with disabilities and achievement have yet to be answered
 - How stable is identification/exceptionality category?
 - 81% of published growth studies omit SPED in analysis, 94% do not distinguish exceptionality categories
- Information needed for accountability models that include SWDs, capture SWD achievement and growth, and schools' performance with these subgroups

Observed Means Over Time (descriptive)



5

HLM Growth Models: Evolution of Analytic Sample

- Original cohort sample grades 3-8, 2001-2006, N=103,123
- Exclusions:
 - Off sequence cases N=8,315 (8.1%)
 - Missing ethnicity (N=14) or sex codes (N=3); dropped “other” ethnicity (N=30)
 - Missing exceptionality codes (N=255) or exceptionality group less than 50 (Blind, Multi-handicapped, Trainable MH, TBI, Visual, Severe MH; N=288)
 - Restricted to grades 3-6 math (same edition)
 - Missing all math scores or missing demographic codes (N=2,445; 2.4%)
 - Final analytic sample, N=92,028 (89%)

Measurement Occasions

- Grades 3-6, “forward matched”
- All cases with one or more math scores retained
- Distribution of scores over occasions:

Only one score:	6.2%
Two scores:	4.8%
Three scores:	5.2%
All four scores:	83.8%

Missing Data, Attrition, and Mobility

- Missing data can undermine accuracy of growth estimates and may occur for many reasons (e.g., illness, drop out, mobility)
- Analysis of only the stable, non-mobile students in the same school for two or three years yields biased results (see Lockwood et al., 2006; Zvoch & Stevens, 2005)
- Percentage of students matched over 3-4 years ranges from less than 30% to about 85% (Lockwood, et al, 2006; McCall, Kingsbury, & Olson, 2004)
- Stevens (2005) study of middle school students ($N \sim 24,000$ per grade) matched 85% over two years, 81% over three years, and 75% over four years

Missing Data, Attrition, and Mobility

- Uncertain how attrition will bias results for school accountability
 - More temporary “attendance-type” attrition may be essentially random and not bias estimates (may inflate *SE*’s and precision of estimation)
 - More permanent enrollment and mobility changes likely to introduce bias because mobility rates and patterns correlated with student socio-demographics
- Researchers often fail to report details on missing data, attrition, and mobility; list-wise deletion most common
- Luo & Kwok (2012) simulation study found that, unless cross-classified models used:
 - School estimates are biased
 - Extent of bias a function of the size and pattern of mobility
 - Spurious results obtained even when the mobility rate was relatively low

Table 1
Student Demographic Characteristics by Analytic Sample (from Zvoch & Stevens, 2005)

<i>Student Characteristic</i>	<i>Accountability Sample (N = 3,334)</i>		<i>Complete Cohort Sample (N = 6,098)</i>	
	<i>Frequency</i>	<i>Percent</i>	<i>Frequency</i>	<i>Percent</i>
Female	1,710	51.3	3,016	49.5
Non-Anglo	1,797	53.9	3,536	58.0
English Language Learner	397	11.9	1,121	18.4
Free Lunch Recipient	1,170	35.1	2,628	43.1
Special Education	101	3.0	1,092	17.9

Stability Across Three Years in NC

	Spec ed 3 rd gr	Gen ed 3 rd gr
First time in 3 rd grade in 2001 (n)	14,380	88,429
In a NC school following 2 years (n, %)	12,731 88.5	79,841 94.9
Same school (%, all subsequent figures based on n of students present all 3 years)	63.7	69.4
Retained (%)	14.0	7.4
Same sped status (Y/N) across yrs (%)	76.3	95.0
Same ec category across years (%)	64.6	---
Took EOG all three yrs		
Reading (%)	69.7	98.3
Math (%)	76.3	98.4
Use/nonuse of accommodations consistent across years (%)	61.0	92.2

HLM Models

- Grades 3-6 Math within test edition
- Two level models with measurement occasions at level 1 nested within students at level 2
- Unconditional model followed by conditional models

Two-level HLM Full Model, Grades 3-6

Fixed Effect	Coefficient	SE	<i>t</i>	<i>df</i>	<i>p</i>
Intercept, γ_{00}	261.500853	0.0509	5134.062	91863	<0.001
LEP, γ_{01}	0.349517	0.1434	2.438	91863	0.015
Parent Ed., γ_{02}	1.429273	0.0165	86.627	91863	<0.001
Sex, γ_{03}	-0.317504	0.0422	-7.531	91863	<0.001
Title I, γ_{04}	-0.141558	0.0125	-11.286	91863	<0.001
FRL, γ_{05}	-1.717237	0.0264	-65.085	91863	<0.001
Accom., γ_{06}	-1.798118	0.1163	-15.466	91863	<0.001
AIG, γ_{07}	7.480157	0.0700	106.902	91863	<0.001
BEH, γ_{08}	-5.259667	0.2912	-18.060	91863	<0.001
Hearing, γ_{09}	-3.220139	0.5584	-5.767	91863	<0.001
EMH, γ_{010}	-11.795806	0.1697	-69.505	91863	<0.001
Other, γ_{011}	-5.416424	0.2300	-23.546	91863	<0.001
Speech, γ_{012}	-1.211880	0.1457	-8.315	91863	<0.001
SLD, γ_{013}	-3.475170	0.1387	-25.061	91863	<0.001
Autistic, γ_{014}	-7.523085	0.7219	-10.422	91863	<0.001

13

Two-level HLM Full Model, Grades 3-6

Fixed Effect	Coefficient	SE	<i>t</i>	<i>df</i>	<i>p</i>
Slope, γ_{10}	2.734180	0.007771	351.842	91863	<0.001
LEP, γ_{11}	0.314132	0.023897	13.145	91863	<0.001
Parent Ed., γ_{12}	0.051177	0.002473	20.694	91863	<0.001
Sex, γ_{13}	0.076430	0.006451	11.848	91863	<0.001
Title I, γ_{14}	-0.016490	0.001896	-8.699	91863	<0.001
FRL, γ_{15}	-0.065537	0.004043	-16.210	91863	<0.001
Accom., γ_{16}	-0.114250	0.018838	-6.065	91863	<0.001
AIG, γ_{17}	0.033028	0.012167	2.714	91863	0.007
BEH, γ_{18}	-0.314230	0.056090	-5.602	91863	<0.001
Hearing, γ_{19}	-0.101264	0.091303	-1.109	91863	0.267
EMH, γ_{110}	-0.488218	0.040101	-12.175	91863	<0.001
Other, γ_{111}	-0.439229	0.038324	-11.461	91863	<0.001
Speech, γ_{112}	0.048946	0.020651	2.370	91863	0.018
SLD, γ_{113}	-0.259771	0.022540	-11.525	91863	<0.001
Autistic, γ_{114}	-0.112583	0.123632	-0.911	91863	0.362

14

Two-level HLM Full Model, Grades 3-6

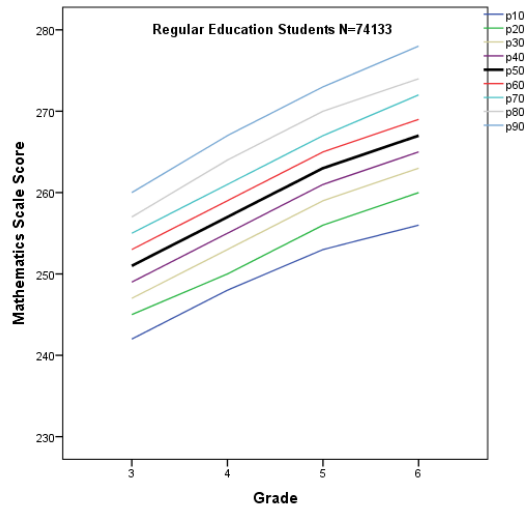
Fixed Effect	Coefficient	SE	<i>t</i>	<i>df</i>	<i>p</i>
Slope², γ_{20}	-0.203024	0.020108	-10.097	153002	<0.001
LEP, γ_{21}	-0.092911	0.059514	-1.561	153002	0.118
Parent Ed., γ_{22}	-0.029699	0.006446	-4.607	153002	<0.001
Sex, γ_{23}	0.067788	0.016722	4.054	153002	<0.001
Title I, γ_{24}	-0.018963	0.004942	-3.837	153002	<0.001
FRL, γ_{25}	-0.004707	0.010184	-0.462	153002	0.644
Accom., γ_{26}	0.077931	0.047022	1.657	153002	0.097
AIG, γ_{27}	-0.354281	0.034968	-10.131	153002	<0.001
BEH, γ_{28}	-0.345207	0.124969	-2.762	153002	0.006
Hearing, γ_{29}	-0.101650	0.244912	-0.415	153002	0.678
EMH, γ_{210}	-0.261297	0.118073	-2.213	153002	0.027
Other, γ_{211}	-0.072668	0.094208	-0.771	153002	0.440
Speech, γ_{212}	-0.029594	0.052433	-0.564	153002	0.572
SLD, γ_{213}	0.010125	0.055876	0.181	153002	0.856
Autistic, γ_{214}	0.047885	0.271605	0.176	153002	0.860

15

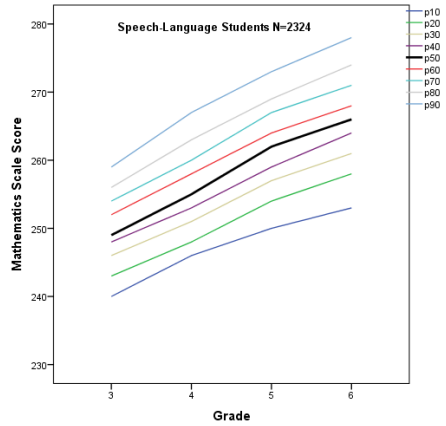
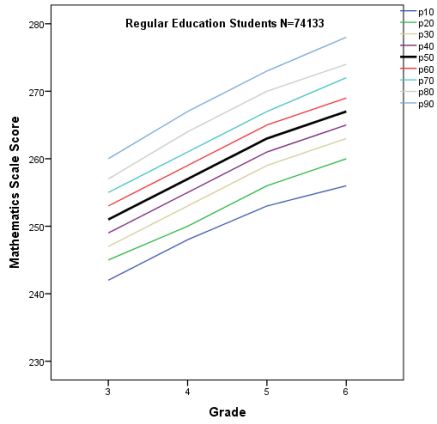
Population Growth Norms

- Cross-sectional; use all students in the sample (before exclusions for HLM analysis, includes those with missing data; requires valid math score in year of interest)
- Typically involves smoothing (not done yet here)
- Advantages and disadvantages
- Sample size by exceptionality category:

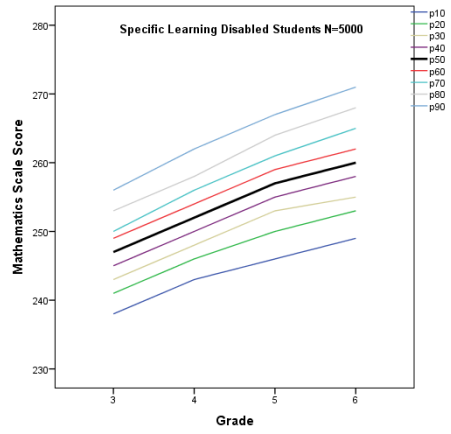
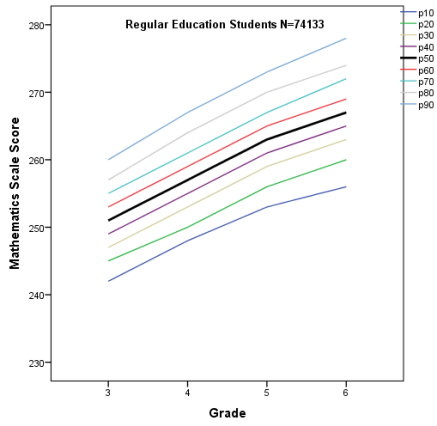
16



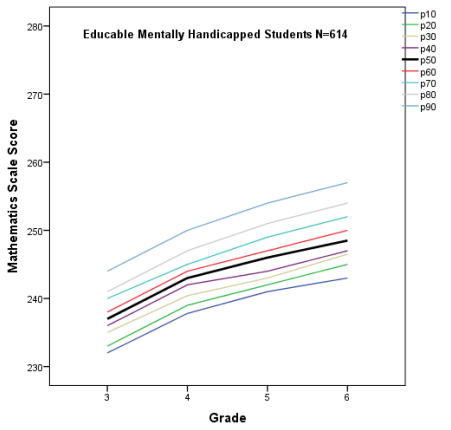
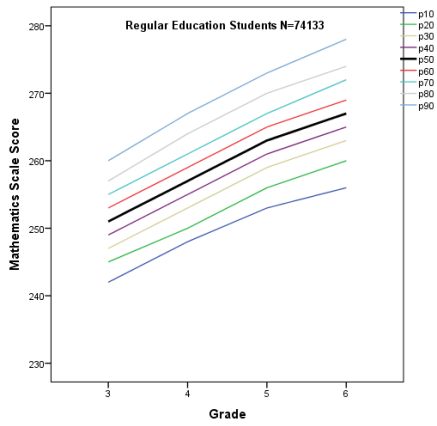
17



18



19



20

SWDs Present Special Difficulties for Student- and School-Level Inferences (status or growth)

- Identification as a SWD is not stable from year-to-year (Ysseldyke & Bielinski, 2002)
- Which assessment (general, alternate, modified) and which accommodations can vary across students and across years for each student
- More scores at the lower end of scale likely result in more error in scores
- Small cohort sizes for many schools, even when aggregated across grades



Special Education Prevalence Across 5 Years (Gr 3-8)

Group	Percent
Always in special education	8.8
Spec ed, but in general ed at least one year	10.4
Ever in special education	(19.2)
Never in special education	80.8
Cross sectional prevalence (Mean percent by year)	14.3

THE GRADE 3 LD DIASPORA

Yr	Gen ED					OHI					LD					EMH					BEH				
1	5,284																								
2	403					154					4,544					85					36				
3	299	14	85	1	2	4	111	36	2	1	410	85	3,927	48	28	3	0	22	59	1	2	2	3	0	29



Yr	Gen ED					OHI					LD					EMH					BEH				
3	3,927																								
4	423					60					3,347					43					20				
5	355	5	57	3	0	4	38	16	0	1	283	45	2,667	20	18	1	0	19	23	0	2	1	3	0	14

Current and Future Directions

- Additional cohorts of students
- Reading as well as math
- Estimation of school effects
- Comparison of alternative growth models
- Study of other issues and model variations
 - Mobility and cross-classification
 - Cohort stability