IEA Bruce H. Choppin Memorial Award at the doctoral level:

ADDRESSING AN OLD ISSUE FROM A NEW METHODOLOGICAL PERSPECTIVE: A PROPOSITION ON HOW TO DEAL WITH BIAS DUE TO MULTILEVEL MEASUREMENT ERROR IN THE ESTIMATION OF THE EFFECTS OF SCHOOL COMPOSITION

Ioulia Televantou

Acknowledgements…

Acknowledgements…

OUTLINE

- An overview of my thesis with emphasis on the studies involving use of IEA's Trends in International Mathematics and Science Study (TIMSS)

- Substantive and Methodological focus of my thesis

- For each study presented I draw on:
  - Empirical research methods / statistical methodology
  - Main results

- Discussion of findings in light of IEA's mission

  " (…) to provide high quality information on student achievement and educational contexts in which students achieve"
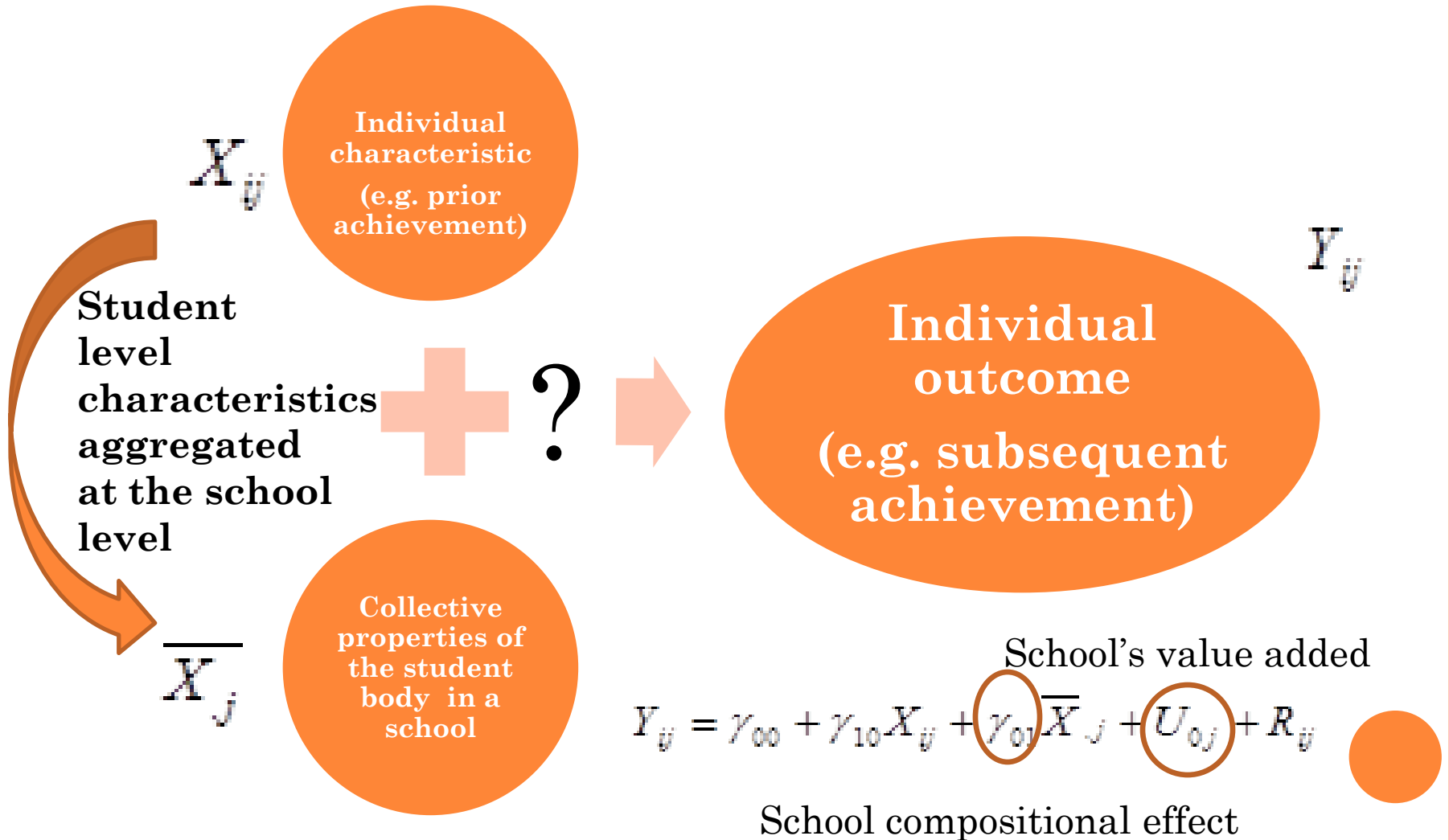 (TIMSS 2011 Assessment frameworks, Mullis, Martin, Ruddock, Sulivan and Preuschoff, 2009) :

A "Methodological-Substantive Synergy" (Marsh and Hau, 2007)

- Substantive focus:
  - School Compositional Effects
- Methodological focus:
  - Explore the impact of measurement error in estimation
  - Demonstrate models that control for potential bias due to measurement error and/or sampling error

School Compositional Effects

X_{ij}

Individual characteristic (e.g. prior achievement)

Student level characteristics aggregated at the school level

$\overline{X}_j$

Collective properties of the student body in a school

$+$ $?$ $\rightarrow$

Y_{ij}

Individual outcome (e.g. subsequent achievement)

School's value added

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}\overline{X}_j + U_{0j} + R_{ij}$$

School compositional effect
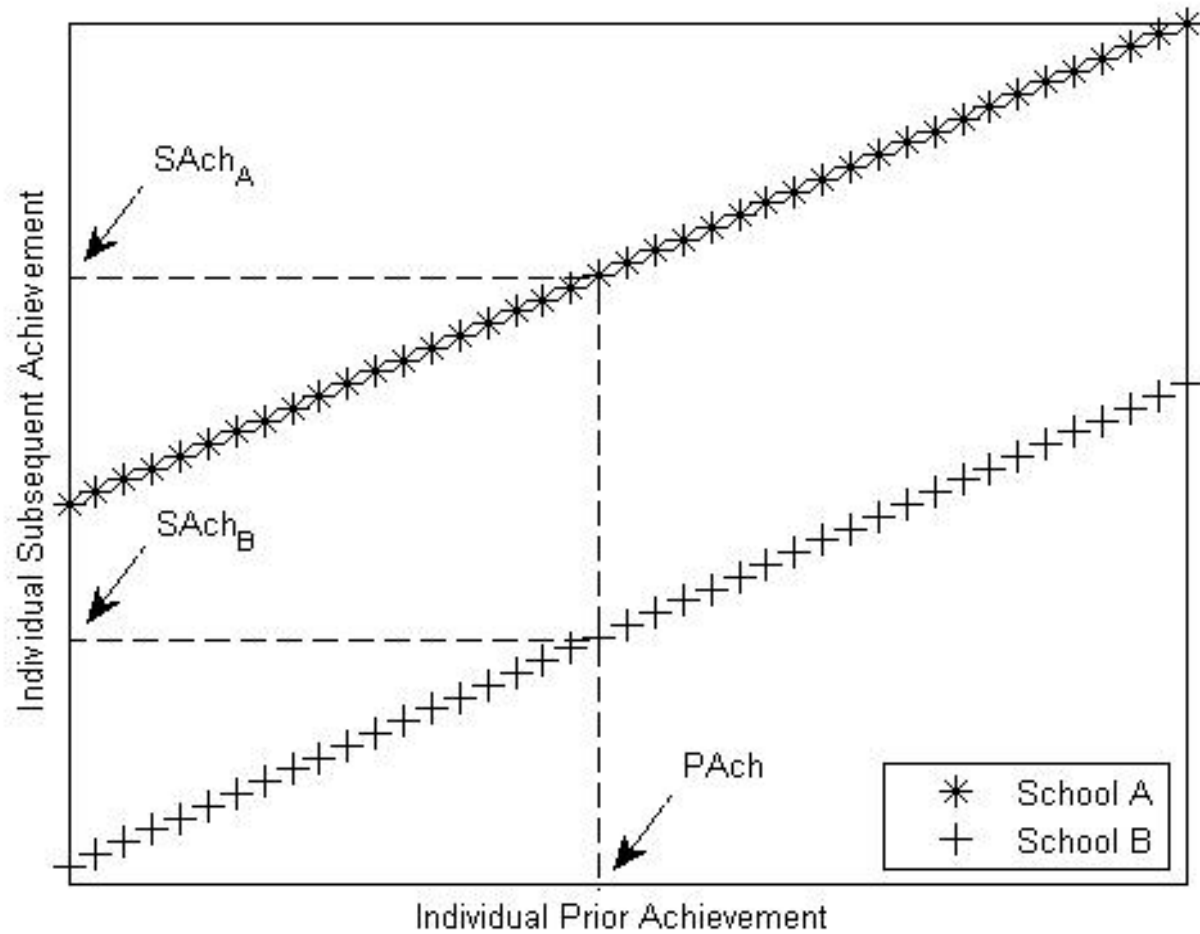
WHAT IS THE MAGNITUDE AND DIRECTION OF SCHOOL COMPOSITIONAL EFFECTS?

- On going debates prevalent in school effectiveness research since the emergence of the field

- The issue of phantom compositional effects arising due to the inadequacies of the statistical procedures used

PHANTOM COMPOSITIONAL EFFECTS OCCUR DUE TO…

- **Omitted Variable Bias:**
  Not all relevant pupil level
  variables controlled for

- **Measurement Error Bias:**
  Unreliability in
  individual level variables
  on which aggregation is based

Correcting for Multilevel Measurement Error
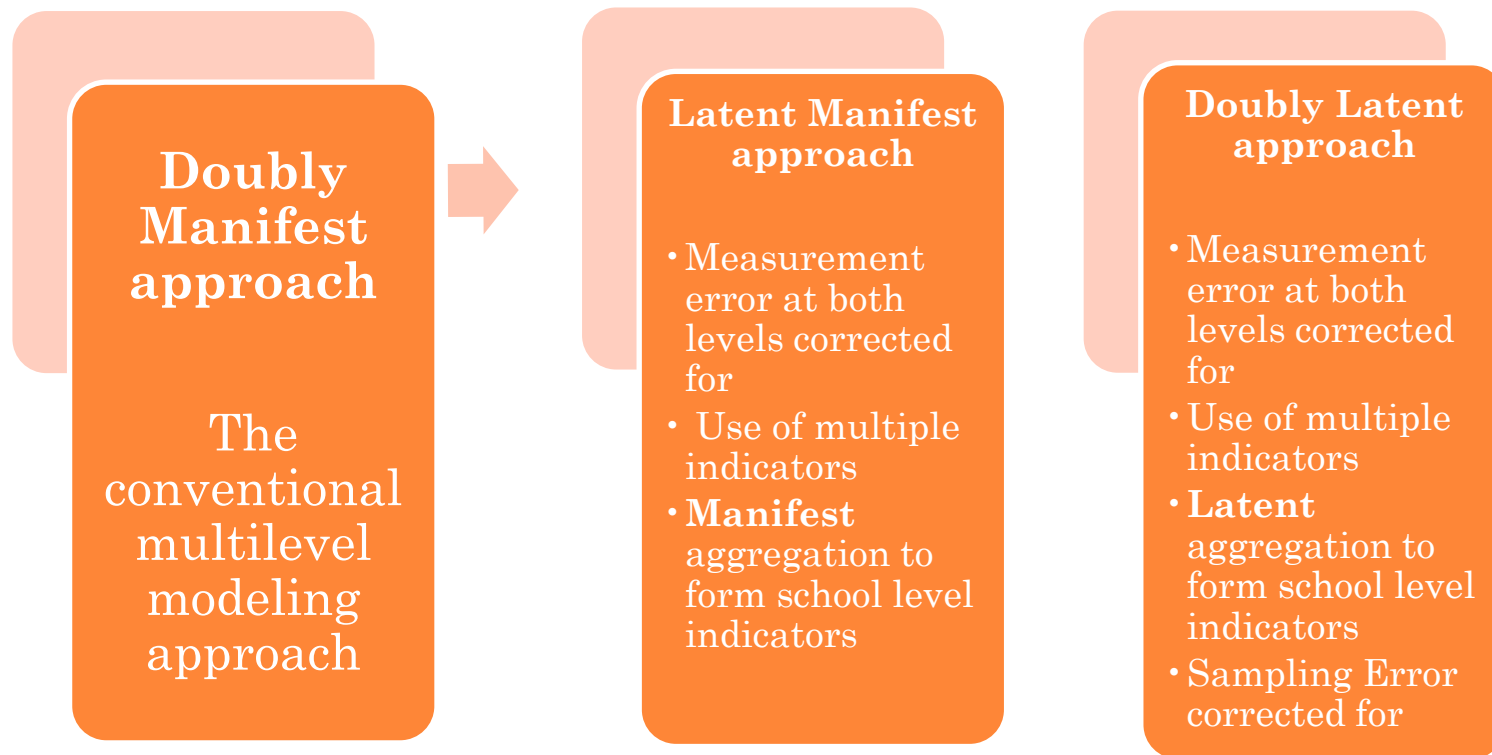
- **Level 1:**
  - **Measurement Error:** It results from the unreliability of the set of indicators used to assess students' achievement

- **Level 2:**
  - **Measurement error** aggregated at level 2
  - **Sampling Error:** Only a finite subset of students sampled from each school

Investigating School Compositional Effects Using Models From Marsh, Lüdtke et al. (2009) Multilevel Structural Equation Modeling Framework

Doubly Manifest approach

The conventional multilevel modeling approach

Latent Manifest approach

- Measurement error at both levels corrected for
- Use of multiple indicators
- **Manifest** aggregation to form school level indicators

Doubly Latent approach

- Measurement error at both levels corrected for
- Use of multiple indicators
- **Latent** aggregation to form school level indicators
- Sampling Error corrected for

PHANTOM COMPOSITIONAL EFFECTS: EVIDENCE FROM CYPRUS

- Data:
  - Mathematics achievement tests constructed based on the released items of the Trends in International Mathematics and Science Study (TIMSS) 2007
  - 1694 students in 59 schools (academic year 2010 - 2011)
  - "Dynamics of Educational Effectiveness Research" project (www. ucy.ac.cy/esf)

- **Criterion:** individual mathematics achievement at the end of year four
- **Predictors:** individual and school average mathematics achievement at the beginning of year four
- What is the school compositional effect of school average achievement in the beginning of year four on students' individual achievement at the end of year four?

**Application of the 2x2 taxonomy of models to the data from Cyprus**

	$ES_{\beta_{within}}$ (s.e.)	$ES_{\beta_{com}}$ (s.e.)
Doubly Manifest	1.224 (.034)	.137 (.058)
Latent Manifest	1.401 (.094)	-.018 (.109)
Doubly Latent	1.467 (.095)	-.034 (.141)

INTEGRATING MULTILEVEL STRUCTURAL EQUATION MODELS WITH THE REGRESSION DISCONTINUITY APPROACH: AN APPLICATION TO TIMSS-95

- The RD approach allows the estimation of the **absolute effect of schooling**; distinguishing it from the effect of chronological age
- **Absolute effect of schooling:** the effect of an extra year of schooling on students' outcomes
- **Strict Requirements:**
  - Needs to be applied to data from educational systems in which students begin formal education strictly based on an age cut-off (e.g. England)
  - Data from two consecutive year groups

academic achievement

Y_{20}

Higher Grade

Lower Grade

cut off

chronological age
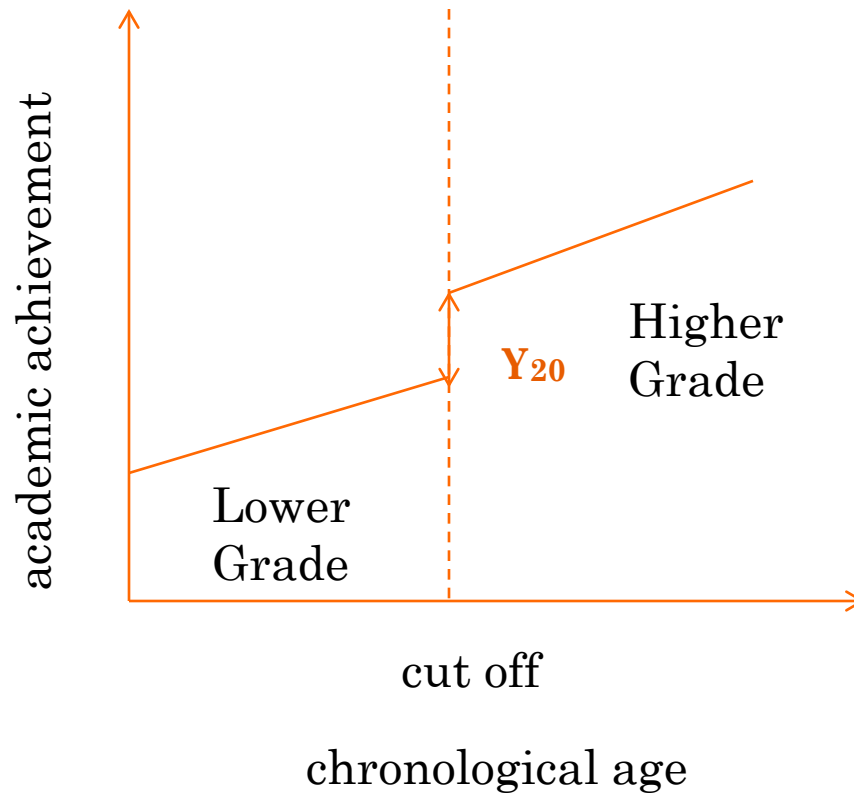
Added year scores
Vs
Value added scores

$$y_{ij} = Y_{00} + Y_{10}age_{ij} + (Y_{20} + U_{1j})grade_{ij} + U_{0j} + \varepsilon_{Ij}$$

DATA

- Mathematics achievement data from the "Third International Mathematics and Science Study" (TIMSS 95)
- Primary school: Cross-sectional data from Grade 4 and Grade 5
- Response variable in the RD models: Students' mathematics achievement
- RD approach implemented in the M-SEM framework
- Content areas in Mathematics were used to group items together and form multiple indicators

RESEARCH HYPOTHESES / RESEARCH QUESTIONS

- Measurement error in the mathematics scores of the students
  - **not** leading to bias on the unstandardized regression discontinuity estimates
  - only leading to larger standard errors of the estimated effects

My hypothesis was verified

- Do school-level variables of school composition (school average achievement) have a significant effect on schools' added year scores (the measure of effectiveness in RD approach)?

No significant effects of school average achievement were found

Summary of the Findings

The findings of the analyses presented suggest that school composition, as quantified using school average achievement, has:

- No significant effect on students' mathematics achievement (schools' value added scores), after controlling for measurement error using contextual value added models from the 2x2 taxonomy

  (study using data from Cyprus)

- No significant effect on schools' added-year scores – the measure of school effectiveness in regression discontinuity designs both with and without adjustments for measurement error

  (study using TIMSS 1995)

Methodological Implications

- Whenever possible educational researchers should consider employing the partial and full correction approaches from the Marsh et al. (2009) 2x2 taxonomy to eliminate measurement and sampling error bias in the estimates of the compositional effects that they retrieve in their analysis

- This is especially the case when using item-level large-scale data (e.g. TIMSS, PIRLS)

- Only when employing appropriate statistical methodology the potential for a more reliable assessment of the school compositional effect will be possible

Thank you!