

# **Annual Value-added School Effects from Various Models and their Inter-temporal Variability: Evidence from China**

Pai Peng, Eckhard Klieme

*German Institute for International Educational Research (DIPF)*

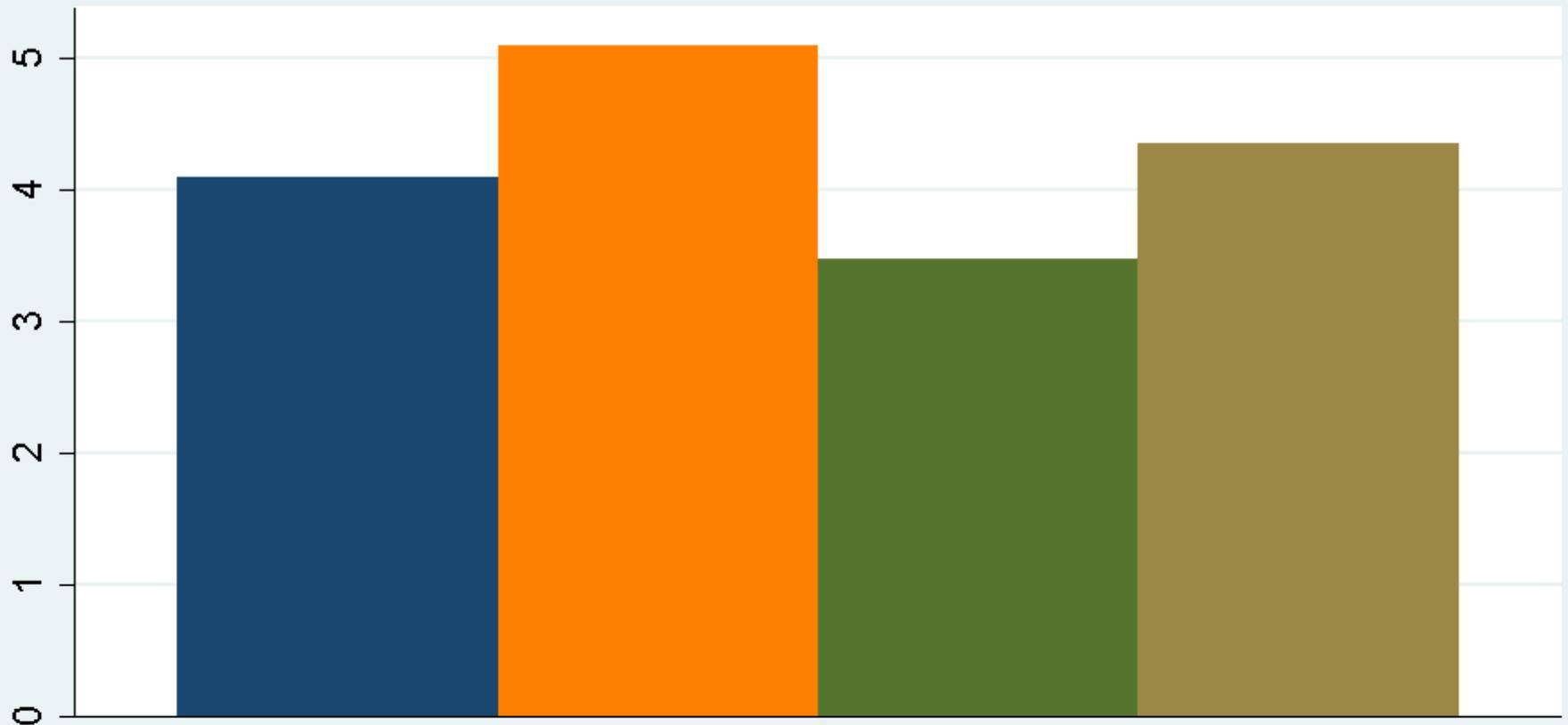
*Rome, 04.10.2012*

# Background

- Growing interest in using value-added around the world (*Braun, 2005; OECD, 2008; Thomas, 2010*)
- In China: Student raw score in HSEE is the dominant school evaluation criterion.
- Adverse effect of raw score measure:
  - Widespread school choice behavior
  - Mathew Effect of schools
  - Good/bad practice in teaching not being identified
  - Lack of morale in schools with low-achieving intake

# Teachers' views of evaluating methods

To what extent do you think are the following ways reasonable to evaluate your teaching?



- on student achievement of my class in critical exams (HSEE)
- on student achievement progress relative to baseline attainment
- on observations of my class teaching from extern experts
- on observations/grading from students in my class

# Research Questions

- What are the **conceptual differences** of some commonly used value-added models and their specifications?
- In Chinese contexts, are estimated annual value-added school effects **consistent across models**?
- To what extent are estimated annual value-added school effects **stable from grade to grade (year to year)**?

# Various editions of value-added

- *Linear regression value-added models*

$$Y_{ij2} = a_0 + \alpha_1 Y_{ij1} + \sum \beta_{pij} X_{pij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

- *Fixed-effects value-added models*

$$Y_{ij2} = a_0 + \alpha_1 Y_{ij1} + \sum_p \beta_{pij} X_{pij} + \theta_j + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

- *Random-effects or multi-level value-added models*

$$Y_{ij2} = a_0 + \alpha_1 Y_{ij1} + \sum_p \beta_{pij} X_{pij} + \zeta_j + \varepsilon_{ij}$$
$$\zeta_j \sim N(0, \tau^2) \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

- *\* all models are based on single-cohort data*

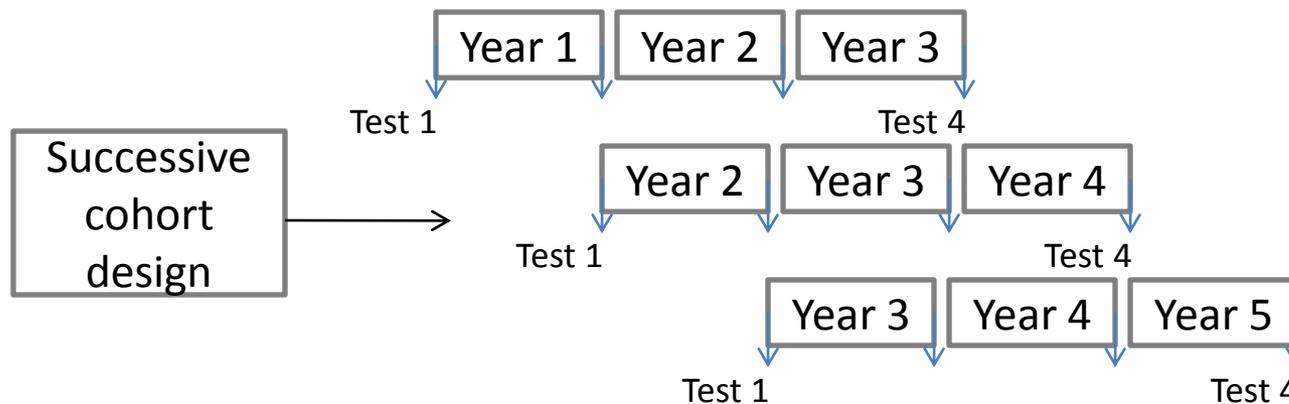
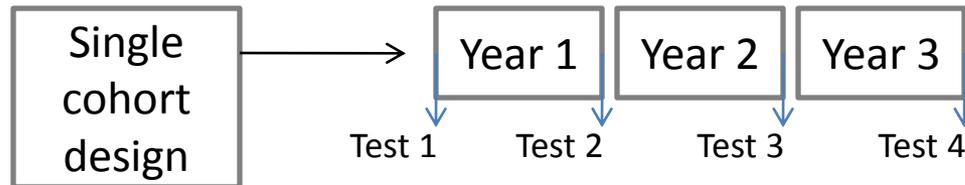
*\* not gain score*

# Different model specifications

- Control for what in the model?  
(school cannot control)
  - Prior attainment only
  - Prior attainment and student SES (Type A)
  - Prior attainment, student SES and school aggregated variable (e.g. school mean prior attainment, school mean SES) (Type B)

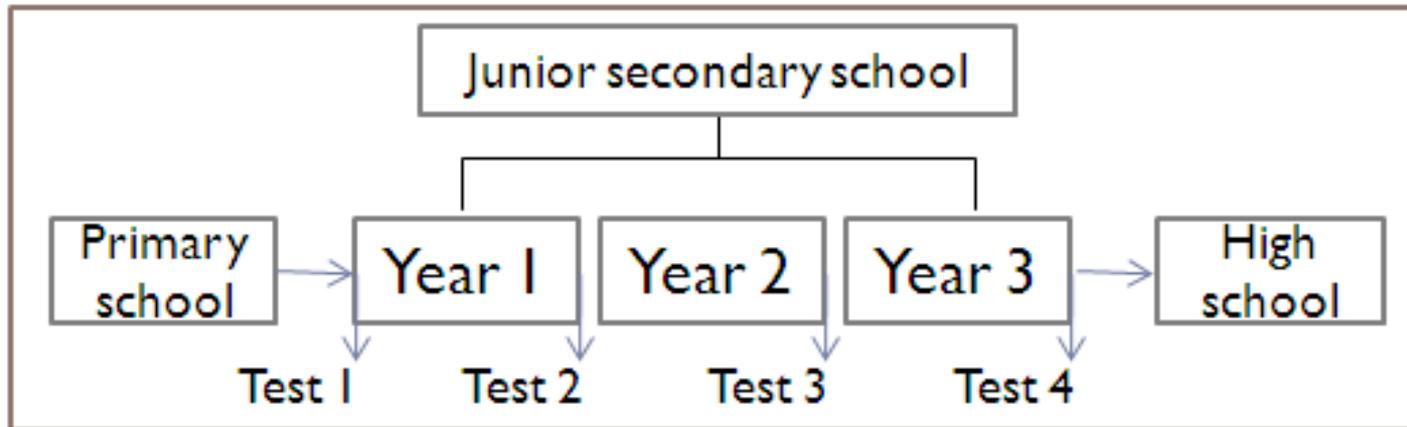
# Two types of stability of school effects

- One of the most important issues in school effectiveness research (*Scheerens & Bosker, 1997*)
- **Year-to-year & Cohort-to-cohort**



# Methodology

- Single-cohort longitudinal design



- Tests:
  - curriculum-based (Chinese, Maths, English)
  - Achievement: total score, standardized
  - Reliability (*Cronbach's alpha*=0.72-0.92)
  - Externally designed and graded
- All 2012 cohort of students in all 25 junior secondary schools in one LEA

LEA: Local Education Authority

# Methods

- Matrix of models

---

	Prior attainment Only	Type-A	Type-B
Linear Regression Model (1)	1PO	1A	
Fixed-effect Model (2)	2PO	2A	
Random-effect Model (3)	3PO	3A	3B

---

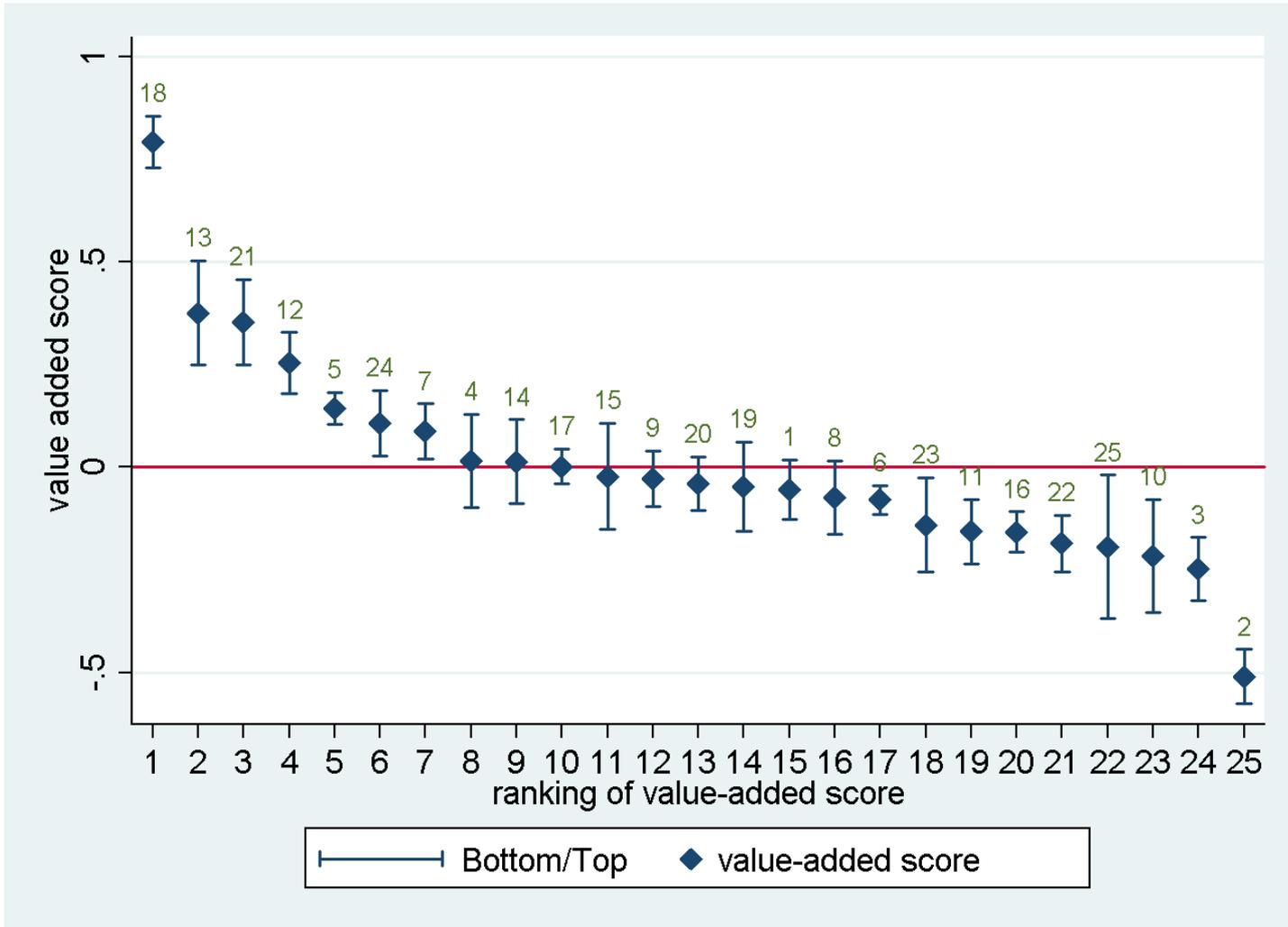


Table 4: Parameter Estimates of Seven Value-added Models for Each Year

		Prior score	Boy	Age in month	SES	Mean prior score	Mean SES
1PO	year 1	0.93(0.01)***					
	year 2	0.93(0.01)***					
	year 3	0.93(0.01)***					
2PO	year 1	0.96(0.01)***					
	year 2	0.92(0.01)***					
	year 3	0.90(0.01)***					
3PO	year 1	0.96(0.01)***					
	year 2	0.92(0.01)***					
	year 3	0.90(0.01)***					
1A	year 1	0.91(0.01)***	-0.02(0.01)	-0.00(0.00)*	0.03(0.01)***		
	year 2	0.91(0.01)***	-0.07(0.01)***	-0.00(0.00)	0.01(0.01)		
	year 3	0.90(0.01)***	-0.01(0.01)	-0.00(0.00)***	0.03(0.01)***		
2A	year 1	0.94(0.01)***	-0.01(0.01)	-0.00(0.00)*	0.01(0.01)		
	year 2	0.92(0.01)***	-0.07(0.01)***	-0.00(0.00)	0.02(0.01)*		
	year 3	0.90(0.01)***	-0.01(0.01)	-0.00(0.00)***	0.01(0.01)		
3A	year 1	0.94(0.01)***	-0.01(0.01)	-0.00(0.00)*	0.01(0.01)		
	year 2	0.92(0.01)***	-0.07(0.01)***	-0.00(0.00)	0.02(0.01)		
	year 3	0.90(0.01)***	-0.01(0.01)	-0.00(0.00)***	0.01(0.01)		
3B	year 1	0.94(0.01)***	-0.01(0.01)	-0.00(0.00)*	0.01(0.01)	-0.37(0.12)***	0.28(0.10)**
	year 2	0.92(0.01)***	-0.07(0.01)***	-0.00(0.00)	0.02(0.01)*	0.07(0.10)	-0.06(0.07)
	year 3	0.90(0.01)***	-0.01(0.01)	-0.00(0.00)***	0.01(0.01)	-0.04(0.09)	0.06(0.07)

Note: Standard errors in parentheses.

\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001

# Consistency of value-added across models

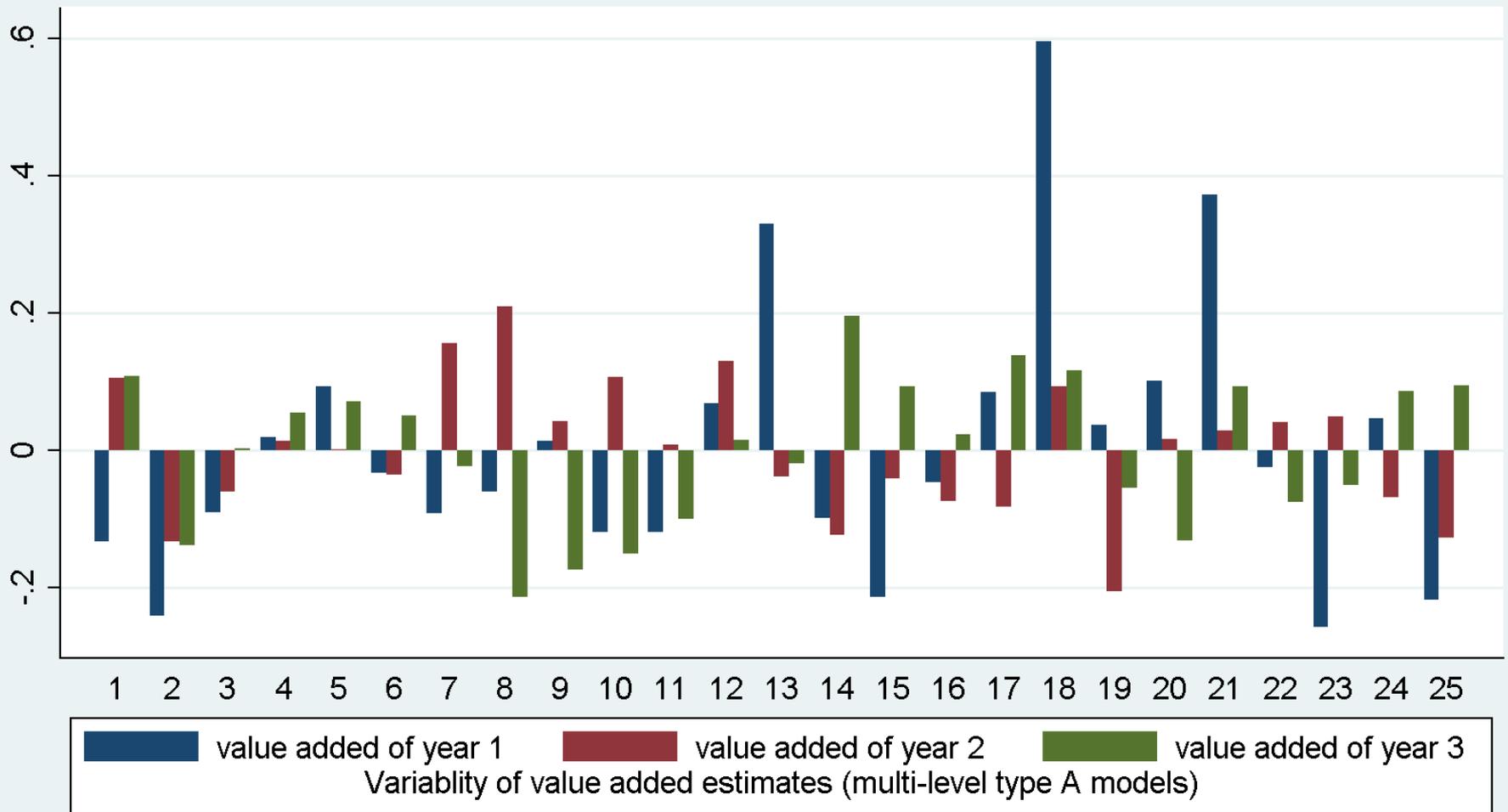
**Table 5: Correlations of school's value-added across models**

		1PO	2PO	3PO	1A	2A	3A	3B
Year 1	1PO							
	2PO	1.00						
	3PO	1.00	1.00					
	1A	0.99	0.99	0.99				
	2A	1.00	1.00	1.00	1.00			
	3A	1.00	1.00	1.00	1.00	1.00		
	3B	0.85	0.84	0.83	0.88	0.86	0.85	
Year 2	1PO							
	2PO	1.00						
	3PO	0.99	0.99					
	1A	0.99	0.99	0.99				
	2A	0.99	0.99	0.98	1.00			
	3A	0.98	0.98	0.99	0.99	0.99		
	3B	0.97	0.97	0.98	0.98	0.98	0.99	
Year 3	1PO							
	2PO	1.00						
	3PO	0.99	0.99					
	1A	0.98	0.98	0.98				
	2A	0.99	0.99	0.98	0.99			
	3A	0.98	0.98	0.99	0.99	1.00		
	3B	0.95	0.95	0.96	0.99	0.98	0.98	

# Stability across years

**Table 6: Correlations of Value-added Score Across 3 year**

	year1-year2	year2-year3	year1-year3
1PO	0.18	-0.35	0.15
2PO	0.18	-0.33	0.15
3PO	0.18	-0.27	0.19
1A	0.13	-0.39	0.20
2A	0.13	-0.39	0.20
3A	0.13	-0.35	0.23
3B	0.12	-0.38	0.19



# Conclusions and discussions

- Schools' value-added is highly consistent across different models.
  - The simpler, the better?
  - Transparency (Confidence interval; super population)
  - Local affair/no national assessment
  - Time-consuming to collect student background information
- Annual value-added of schools is not stable across years.
  - not proper for high-stake accountability
  - useful for school self-evaluation/improvement
  - Dilemma (timely feedback VS. accurate estimate)



Thank you very much!  
Vielen Dank!  
谢谢！

Presenter: Pai Peng

German Institute for International Educational Research  
peng@dipf.de