Researching Learning in Low-Income Countries

World Bank 2013
Jishnu Das
World Bank

Overview

- Growing interesting in learning
- Three questions
- What (I think) we want to learn about learning
  - Why?
- What is going on right now?
- Can we do better; if so, how?

What we want to learn about learning

- How are countries doing **over time** and **in comparison with one another**
- What interventions can help to improve learning
- “Deep” research on how children learn and the determinants of learning

Data requirements

- How are countries doing: Standardized, comparable data on many countries for many years
  - Example: Data in the United States and France allow us to look at learning since the 1960s
  - Example: Data from TIMSS, PISA allow us to compare different countries at any given point in time, and the same country over time
- **Broad Consensus** that this is required

Data Requirements (II)

- What interventions work
  - Example: Many examples of learning interventions and their impact on test-scores
    - Typically data collected by the authors of the study
  - Want to move to system where intervention is designed through research but data are obtained through administrative means
    - Example: United States, Chile
- Will argue that this is *far* more efficient

Data Requirements (III)

- Deep research on how children learn and the determinants of learning
- New research from a handful of countries shows that we are just starting to scratch the surface on how children are learning
  - And these “learning trajectories” can be very different from one place to the other
- Spend some time on this
  - Data are new

How are children learning?

	2003	2004	2005	2006	2011
	Mean	Mean	Mean	Mean	Mean
Math 1	-0.060	0.146	0.850	1.182	1.634
English 1	-0.091	0.152	0.691	1.046	1.851
Urdu 1	-0.133	0.173	0.828	1.261	2.191
N cohort 1	12,110	12,815	12,113	10,064	865
Math 2			-0.372	0.283	1.375
English 2			-0.031	0.497	1.569
Urdu 2			-0.277	0.422	1.841
N cohort 2			11,860	12,741	190

- Sourced from LEAPS study in Pakistan

How are children learning?

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 5 Drop outs
	Mean	Mean	Mean	Mean	Mean	Mean
Math	0.021	0.391	1.022	1.315	1.981	1.579
English	-0.085	0.180	0.730	1.046	2.225	
Urdu	-0.063	0.353	0.955	1.417	2.621	
N	500	500	500	500	500	309

- Balanced Panel Sourced from LEAPS study in Pakistan

How are children learning?

	Year 1	Year 2	Year 3	Year 4	Year 5
	Prop correct	Prop correct	Prop correct	Prop correct	Prop correct
Total kids	500	500	500	500	500
English					
Eng 12: Match picture with word, Banana	0.63	0.73	0.822	0.852	0.942
Eng 19: Fill missing letter for picture, Flag	0.27	0.276	0.46	0.53	0.778
Eng 30: Fill missing word in sentence	0.246	0.302	0.426	0.566	0.786
Eng 43: Construct sentence with word 'deep'	0.008	0.018	0.028	0.104	0.294
Eng 44: Construct sentence with word 'play'	0.008	0.024	0.106	0.212	0.382
Math					
Math 1: Count number of moons, write number	0.64	0.67	0.812	0.746	0.962
Math 9: Add 3 + 4	0.91	0.922	0.962	0.938	0.964
Math 12: Multiply 4 x 5	0.6	0.666	0.748	0.806	0.912
Math 25: Add 678 + 923	0.546	0.566	0.712	0.738	0.846
Math 27: Subtract 98 - 55	0.718	0.756	0.854	0.848	0.878
Math 30: Multiply 32 x 4	0.492	0.604	0.7	0.74	0.852
Math 32: Divide 384 / 6	0.164	0.234	0.438	0.552	0.672
Math 34: Cost of necklace, simple algebra	0.106	0.156	0.284	0.28	0.514
Math 39: Convert 7/3 into mixed fractions	0.02	0.056	0.068	0.148	0.168
Urdu					
Urdu 3: Match picture with word, Book	0.74	0.824	0.926	0.948	0.982
Urdu 5: Match picture with word, House	0.546	0.566	0.674	0.744	0.988
Urdu 12: Combine letters into word	0.362	0.468	0.54	0.618	0.784
Urdu 20: Antonyms, Khushk	0.388	0.5	0.65	0.642	0.752
Urdu 36: Complete passage for grammar	0.29	0.36	0.564	0.704	0.792

- Balanced Panel Sourced from LEAPS study in Pakistan

How are children learning?

	English					
(6038 children)	Eng 12	Eng 18	Eng 19	Eng 30	Eng 43	Eng 44
(0, 0, 0, 0)	0.042	0.064	0.275	0.207	0.861	0.716
(0, 0, 0, 1)	0.044	0.049	0.127	0.130	0.086	0.143
(0, 0, 1, 1)	0.067	0.052	0.117	0.099	0.012	0.049
(0, 1, 1, 1)	0.153	0.104	0.067	0.085	0.003	0.008
(1, 0, 1, 1)	0.043	0.053	0.051	0.045	0.001	0.005
(1, 1, 1, 1)	0.510	0.543	0.133	0.070	0.000	0.002

- Balanced Panel Sourced from LEAPS study in Pakistan

How are children learning: India

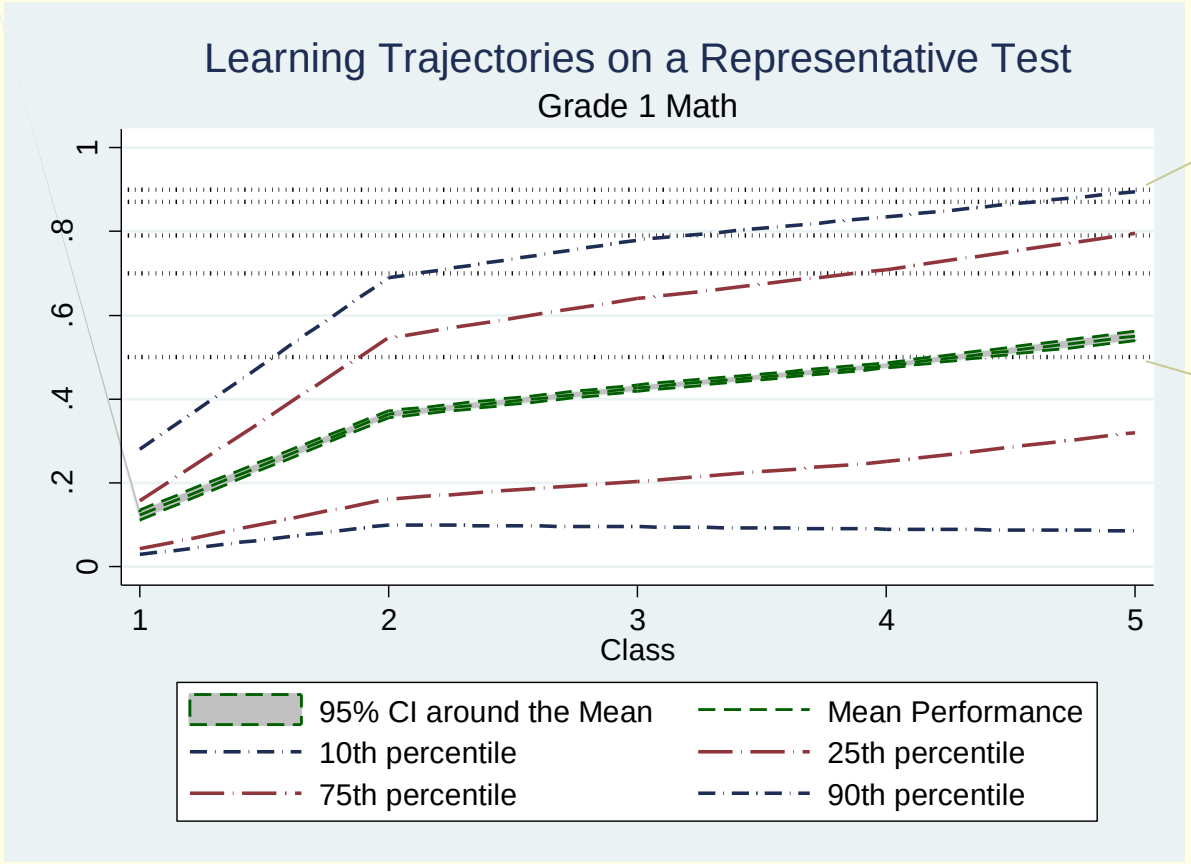
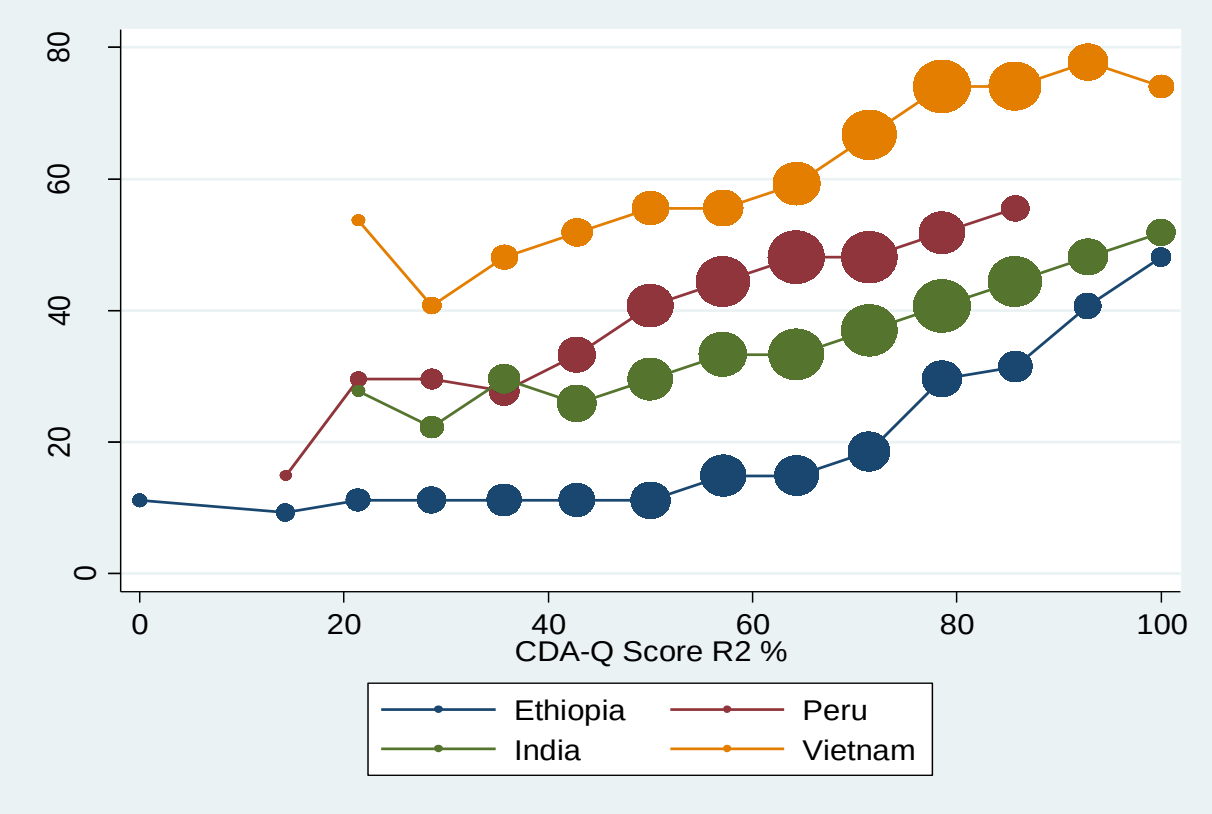


Figure 6. Learning Trajectories for Grade 1 Appropriate Material, Cohort 5 fractional polynomial fit

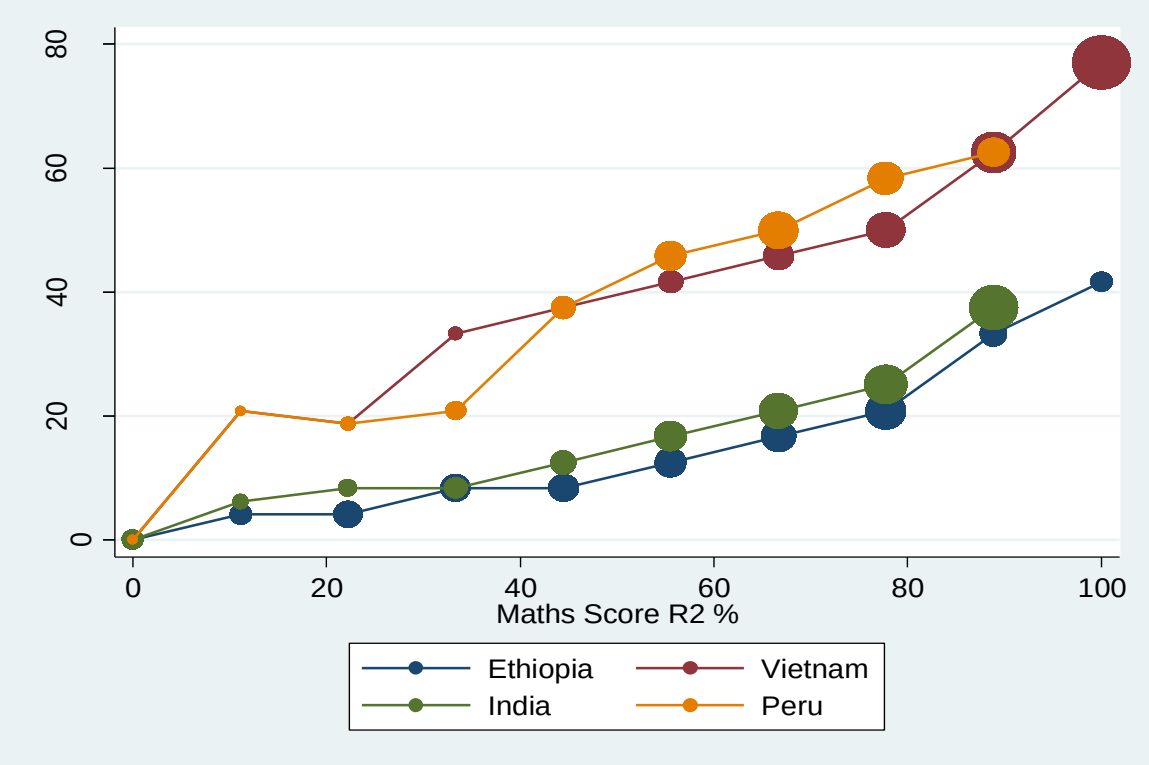
- Sourced from: Muralidharan and Zieleniak

How are children learning: Many countries



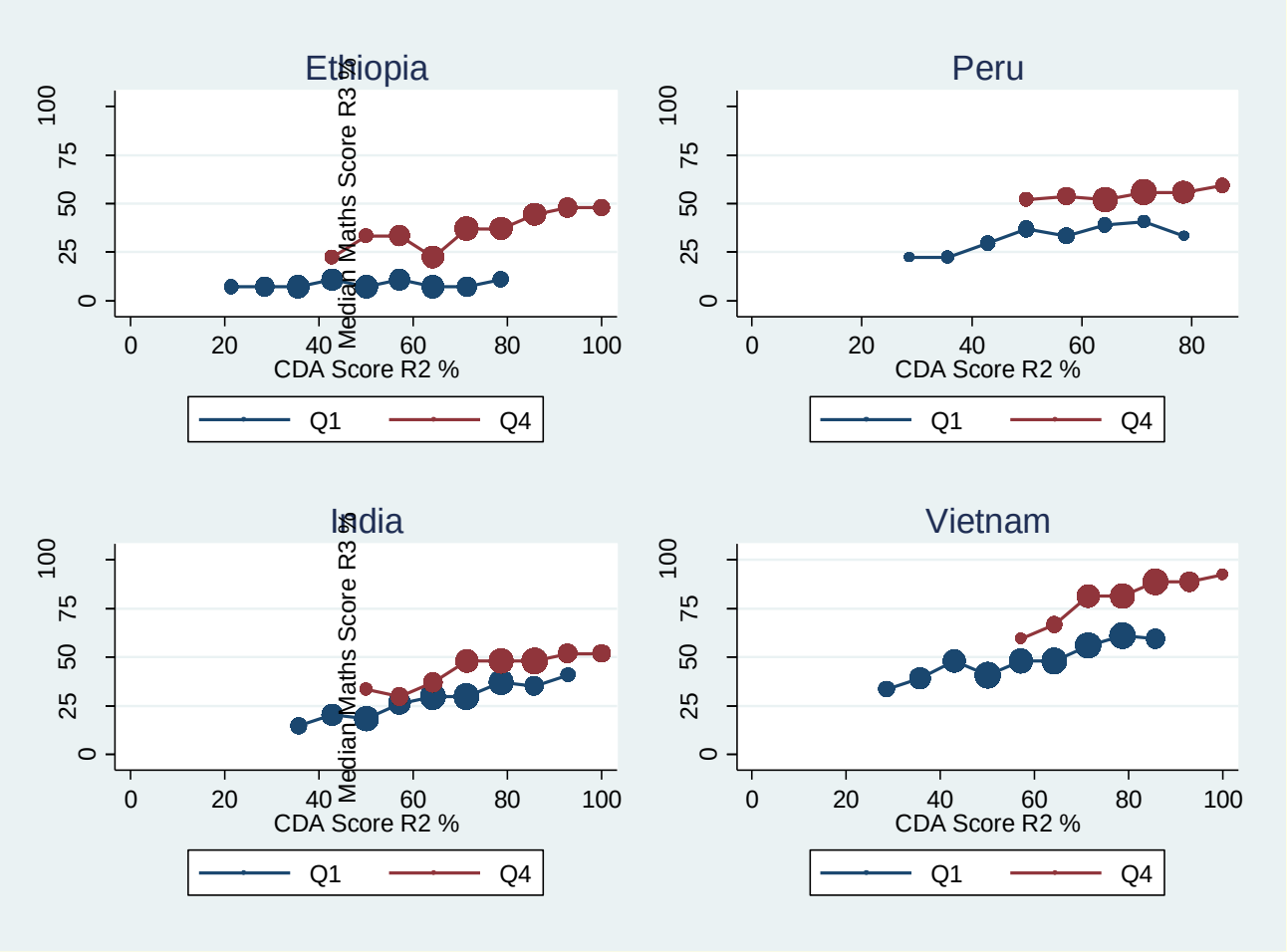
- Sourced from Rolleston, James and Aurino, Young Lives Study, Ages 5-8

How are children learning: Many countries



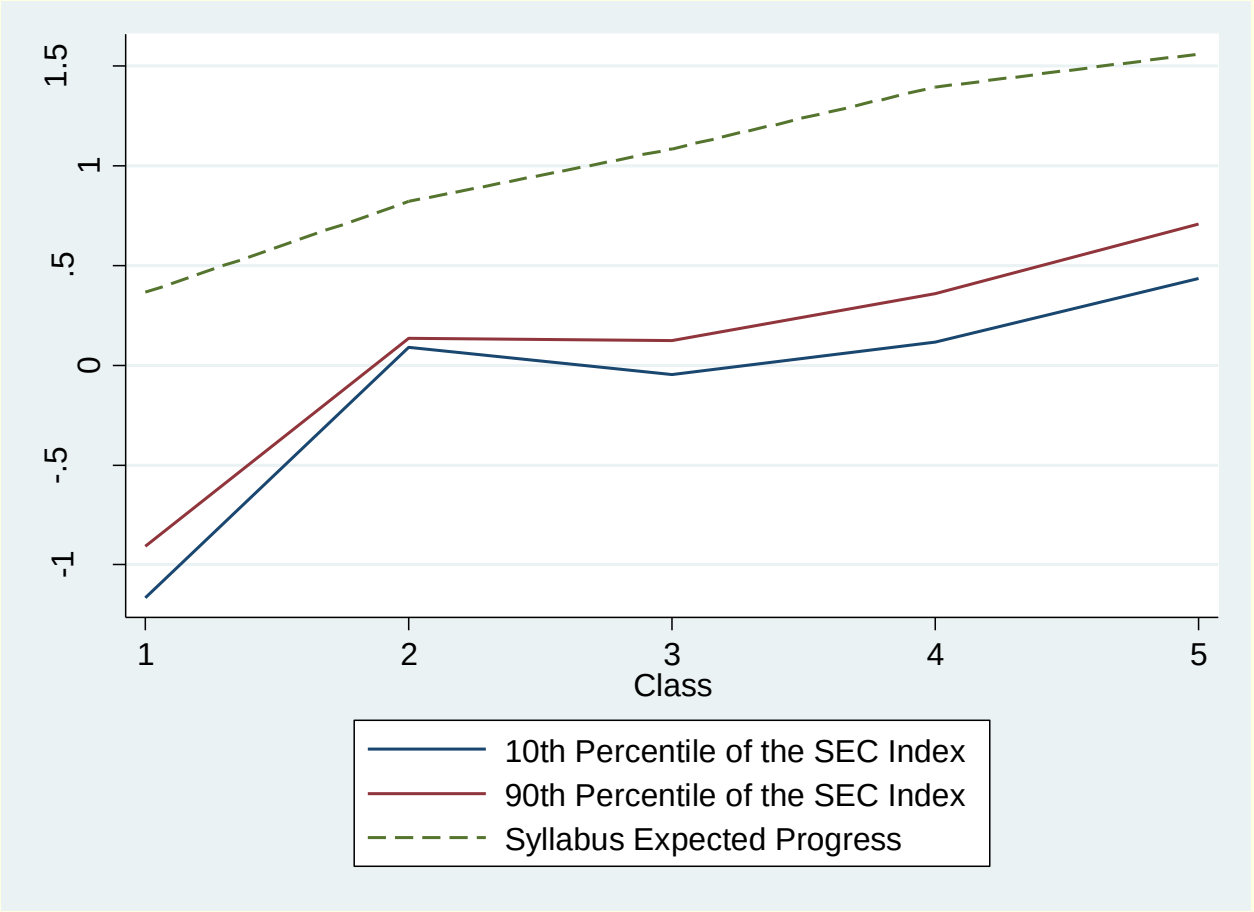
- Sourced from Rolleston, James and Aurino, Young Lives Study, Ages 12-15

How are children learning: SEC



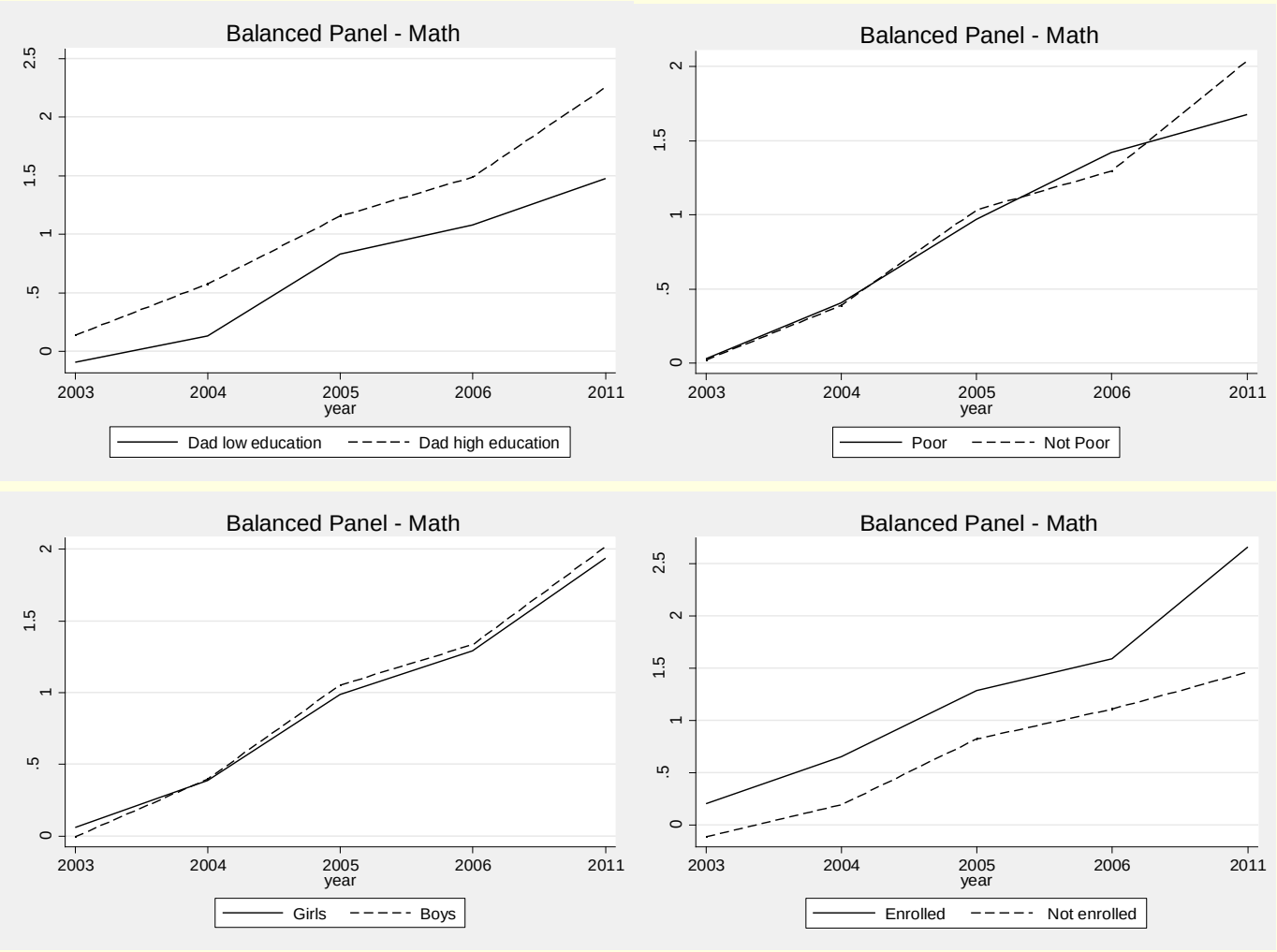
- Sourced from Rolleston, James and Aurino, Young Lives Study, Ages 5-8

How are children learning: SEC



- Sourced from: Muralidharan and Zieleniak

How are children learning: SEC



- Balanced Panel Sourced from LEAPS study in Pakistan

Child learning: Convergence?

Test Scores by Original Quintile

year	1	2	3	4	5
1	-2.12	-1.17	-0.75	-0.37	0.22
2	-1.43	-0.93	-0.55	-0.28	0.10
3	-0.85	-0.36	0.01	0.21	0.53
4	-0.75	-0.10	-0.02	0.25	0.64
5	-0.43	-0.17	-0.11	0.04	0.07

- Balanced Panel Sourced from LEAPS study in Pakistan

How are children learning

- We know that they are learning little
- But there are very large differences in how much they are learning and the correlations with SEC across countries
- Unique combination of household data and test-score data
- How are we going to learn more

What is going on right now?

- *Some* standardized testing through TIMSS and PISA
  - *Low-income countries conspicuous only by their absence and sporadic testing*
- Large number of evaluations (500 or more!)
  - *Different tests, quality uncertain, no validation, not comparable* except *on cost*
  - Which is uniformly large (think $200-$500,000 *per evaluation*)
    - The data from these evaluations are seldom used again
- Testing programs at national level in many countries

Is this the best way to do things?

- Poverty measurement at The World Bank
  - National measurement: Capacity building, training
  - Parallel research track on how to do things better
  - Public good provision
    - Issues in poverty measurement, how to do it better, methodology, accreditation through checks with best practice
    - International Price Comparison is a huge public good

How the poverty guys think about testing

- This is crazy (1)
  - At best, a duopoly, at worst localized monopolies on international testing
- This is crazy (2)
  - Donors and governments funding monopolies who do not share the IP that is built up (item bank, methodology, scoring)
  - Even scoring programs etc. not publicly available
- This is crazy (3)
  - *None* of the evaluation results are strictly comparable limiting our learning enormously
    - Although this does not stop researchers from comparing them
- This is crazy (4)
  - All this enterprise is very high cost and very hard to scale-up with the current model

Can we do better?

- United States
  - Testing by states and NAEP
  - Linking NAEP to TIMSS
  - Evaluations based on administrative data

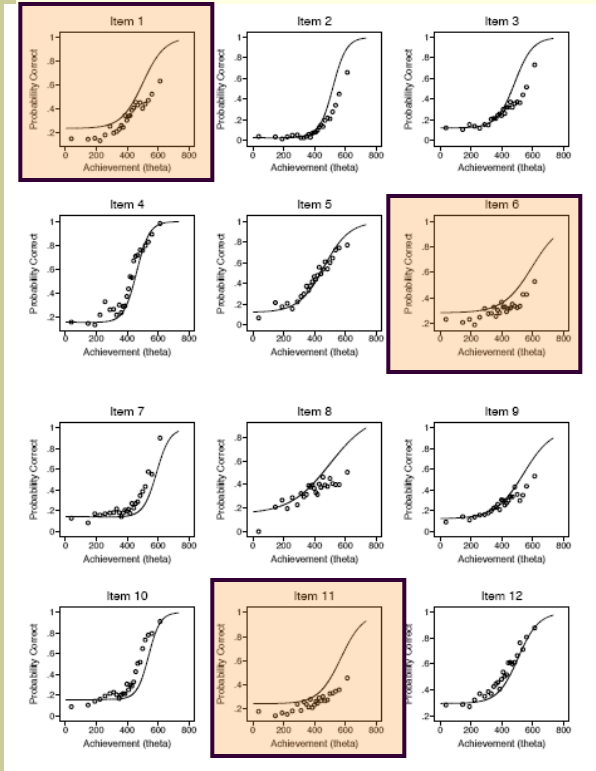
Can we do better: Break it up

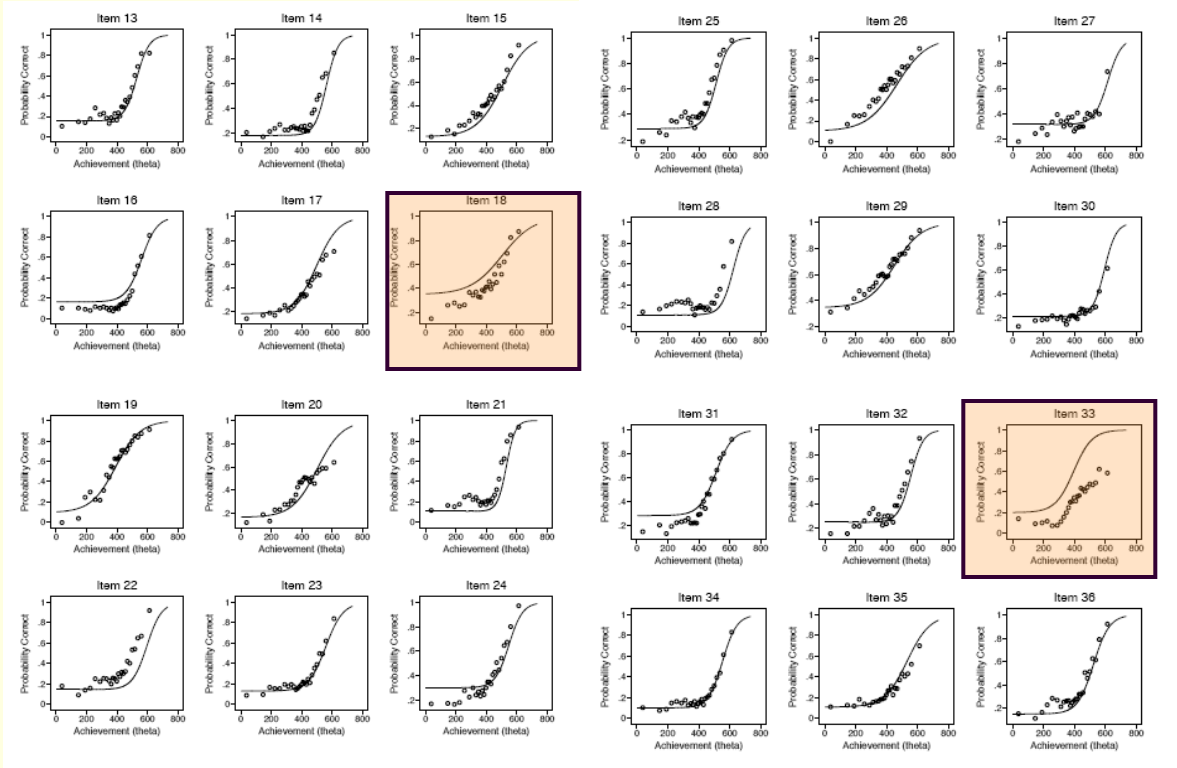
- Intellectual Property: This is a public good and should be publicly provided (Example: Calibrated item banks)
  - Donors are a good example
- Capacity building and implementation: Standard issues of how to structure (Government only, PPP)
  - In either case, monopolies are *not* the way forward
- Accreditation functions: Political economy is the hard part, implementation should not be

Can we do better: Example

- Testing, with linking to public/semi-public item bank
- Capacity building, technical training
- Research *in* low-income countries on methodology
  - Do we *really* need to do things the way that big firms do to get comparable results? What things need to be done the same way?
- Longitudinal studies in select countries that build up long-term panels for “deep dives”

Item Fits for all administered TIMSS items





Each figure plots the probability of getting a question right against the achievement level of the child. The solid lines are the expected responses based on the TIMSS item calibrations and the dots are the observed responses from the Indian test data. Poorly fitting items are shaded in light orange. Sourced from Das and Zajonc

Summary

- We are beginning to learn how children learn
  - They learn little
  - But the patterns are very different, both in overall learning and differences across households and children in different countries
- We also *think* we are learning about what works to improve learning
  - But this is caveated by very little interaction between serious measurement and evaluation of interventions by economists
- We know very little about internationally comparable country-level scores
  - The current model is problematic for scale-up