System Approaches in Education

Lant Pritchett

May 14, 2018

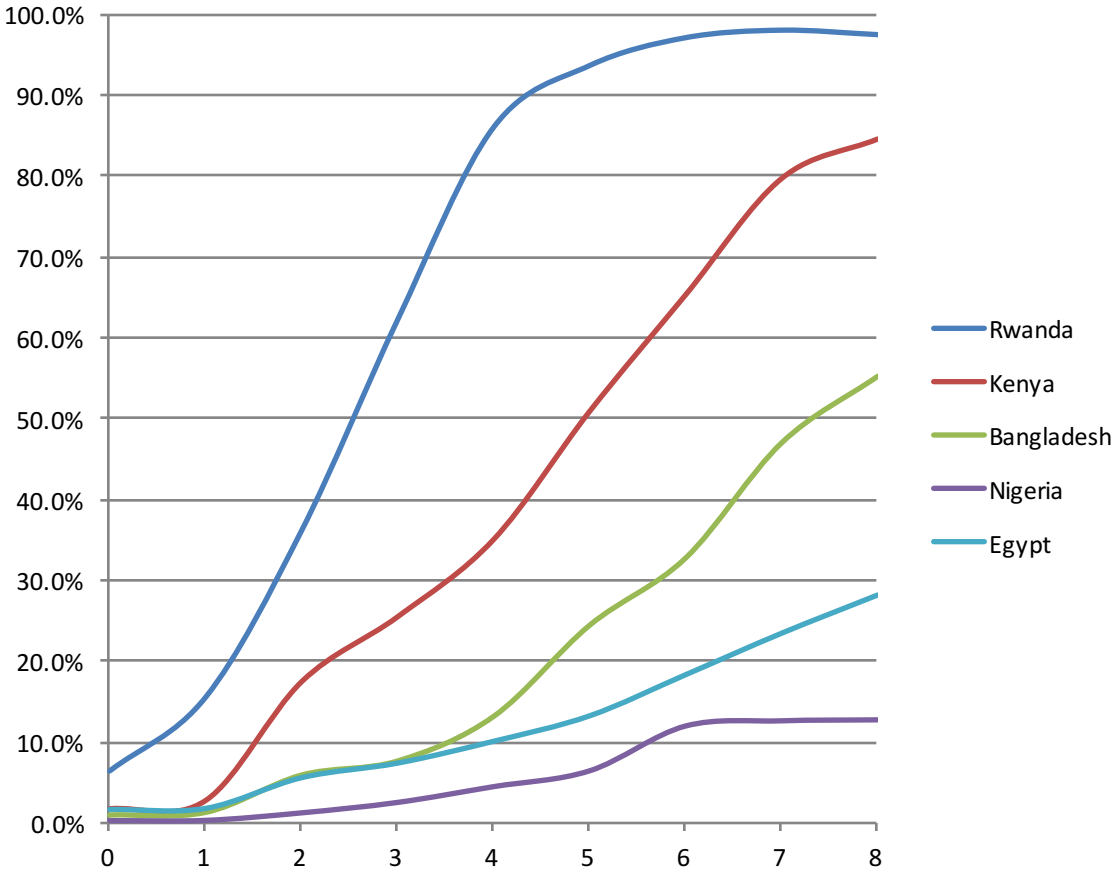
Systems of Public Service Delivery in Developing Countries

Outline

- Why we need one—big differences in performance, small differences/gains from observed inputs/interventions—literature has now proven lack of external validity
- A particular instance of a system approach: The RISE 5 by 4 accountability framework (adapted from WDR 2004)—used to illustrate the presence or lack of "coherence" in systems
- Top-down encompassing organization (spider) versus "performance pressured starfish" approaches (and organizational mismatch)

The vast difference in learning profiles and the vanishing gains from enrollment expansion

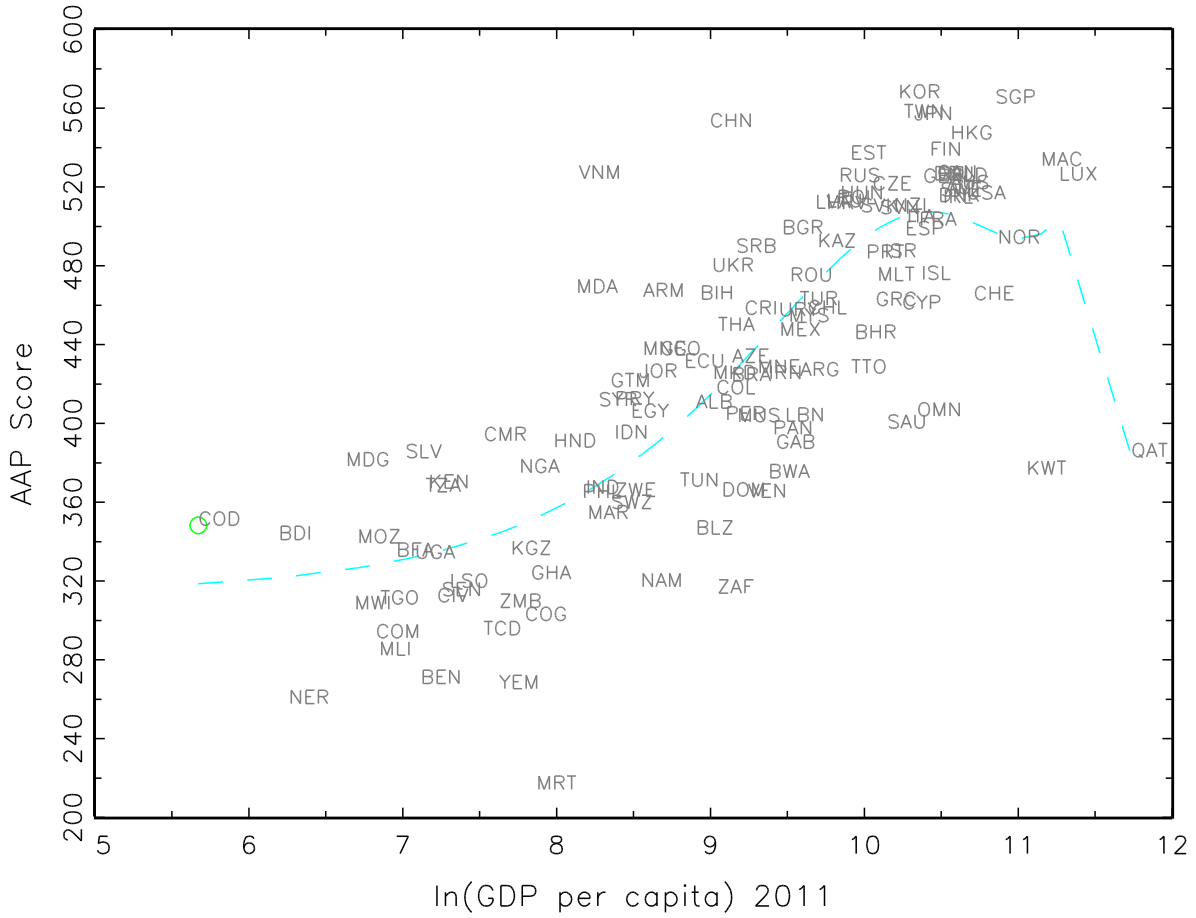
Percent of adult women who completed only the grade reported who can read a complete sentence (in their chosen language)



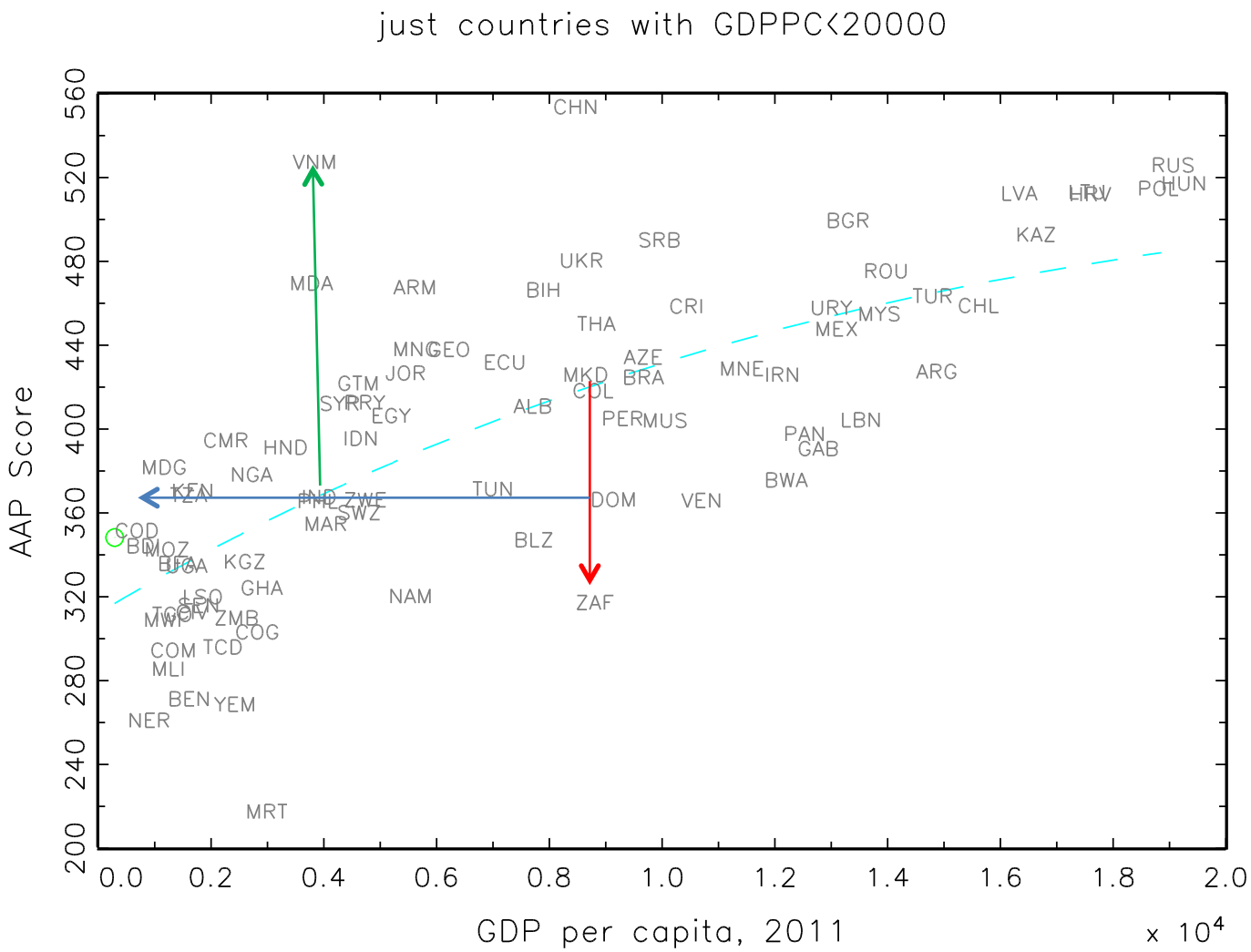
Source: Oye, Pritchett and Sandefur, 2016

Whether a woman who completed six grades of schooling can read a sentence ranges from over 90 percent in Rwanda to only 10 in Nigeria (median across all 50 countries is 50 percent!)

While learning is correlated with GDP per capita, there are massive variations



Vietnam is 161 points above its predicted performance (above the UK on PISA), whereas South Africa is 104 points lower. Dominican Republic (DOM) has scores lower than Tanzania or Kenya—at four times higher income



Need for a “system” approach

- The large differences across countries in the learning performance of the typical student cannot be explained by standard “thin” measures of quality (e.g. expenditure per pupil, class size, schooling of teachers, etc.)
- The improvement over time is an improvement of around 1 point per annum…so a country 100 points behind the US/UK/OECD would expect convergence (optimistically) in 100 years (and in many countries observed improvement is zero (e.g. Indonesia) or negative (e.g. India)

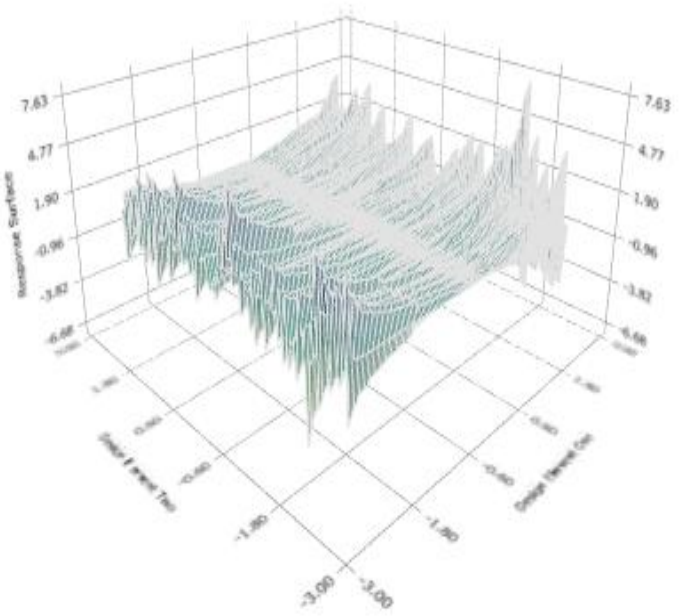
The (second)need for system approach: The “what works” approach clearly doesn’t work (and was silly from the get-go)

- Development economics has been invaded by a virus—the presumption that rigorously estimating the impact of specific programmatic interventions and aggregating those into compendia of “what works” through “systematic reviews” would be an important contributor to better outcomes—and the education sector has been particularly susceptible (as it produces easily measurable outcomes (enrollment, learning) of individualized ‘treatments’ and hence can be statistically “powered up”)
- The results to date there is neither external (e.g. Glewwe and Muralidharan) nor construct validity (Pritchett) of the empirical results—nor can there be any hope of such as the idea of “systematic reviews” to “resolve” the existing literature is logically incoherent (Pritchett and Sandefur)
- Moreover, the idea that the knowledge of the “response surface” over the design space was the key constraint to performance was known to be pretty silly even before all this started.

A quick build of the policy/program/project Trinity

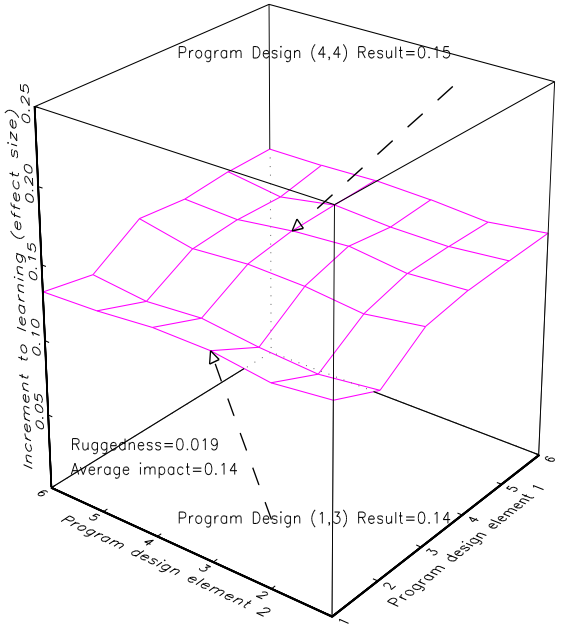
- Technically correct (response surface of outputs/outcomes over a p/p/p design space
- Administratively feasible (capability surface over the design space)
- Politically supportable (what can be politically adopted and sustained)

A response surface or fitness function is the mapping from the design space to an outcome of interest

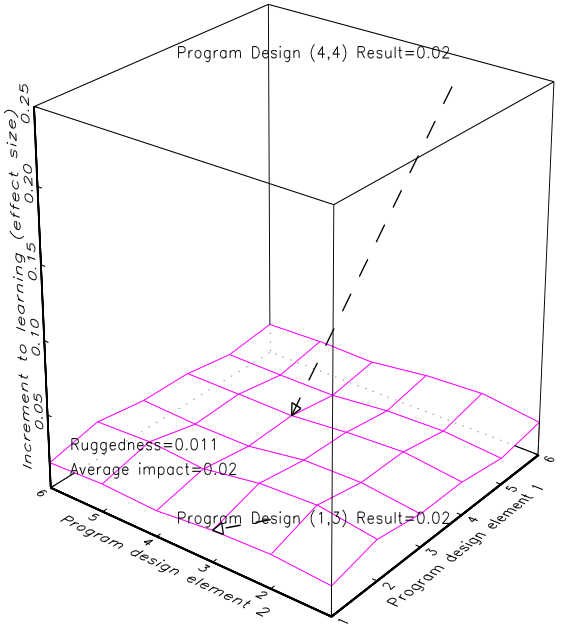


“Pure” external validity

Response surface in context A—design doesn’t matter much, all works

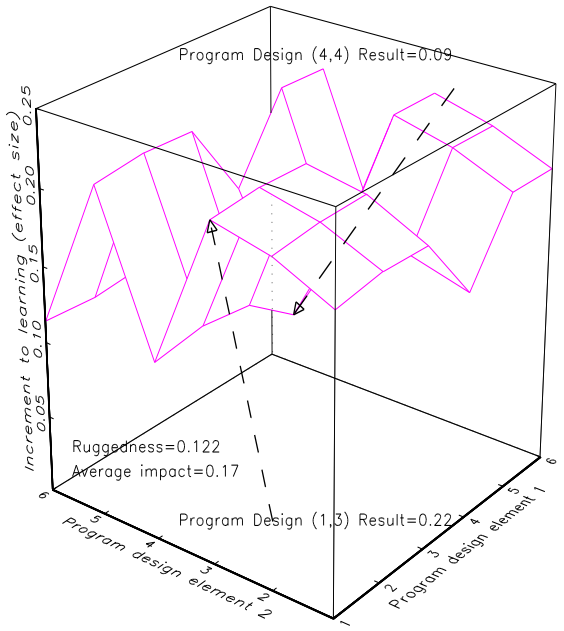


Response surface in context B—design doesn’t matter much, nothing works

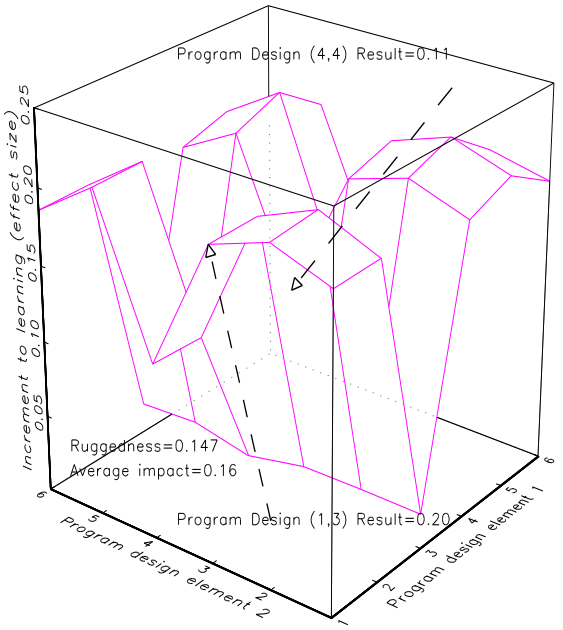


Construct validity: Rugged fitness functions imply different designs produce different results

One "class" of program ("textbook provision")

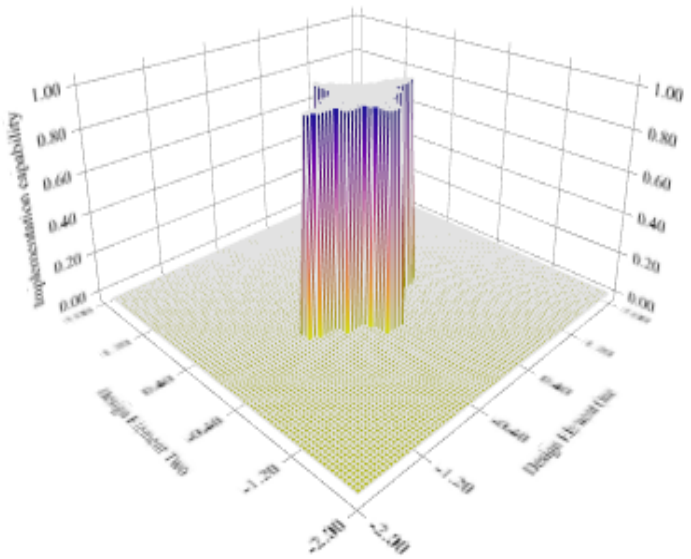


A different class of program ("teacher training")

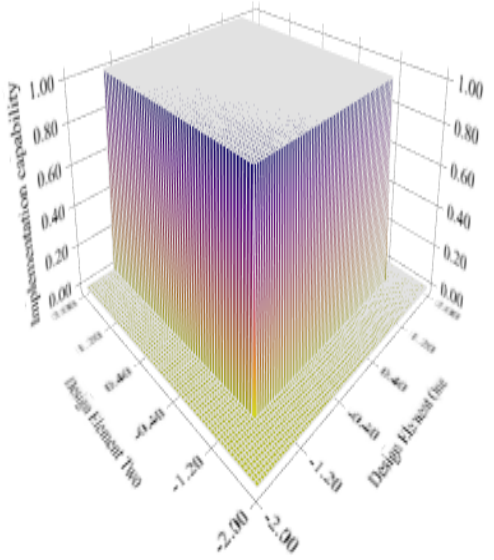


Mappings of organization capability to replicate a policy/program/project with fidelity over the design space

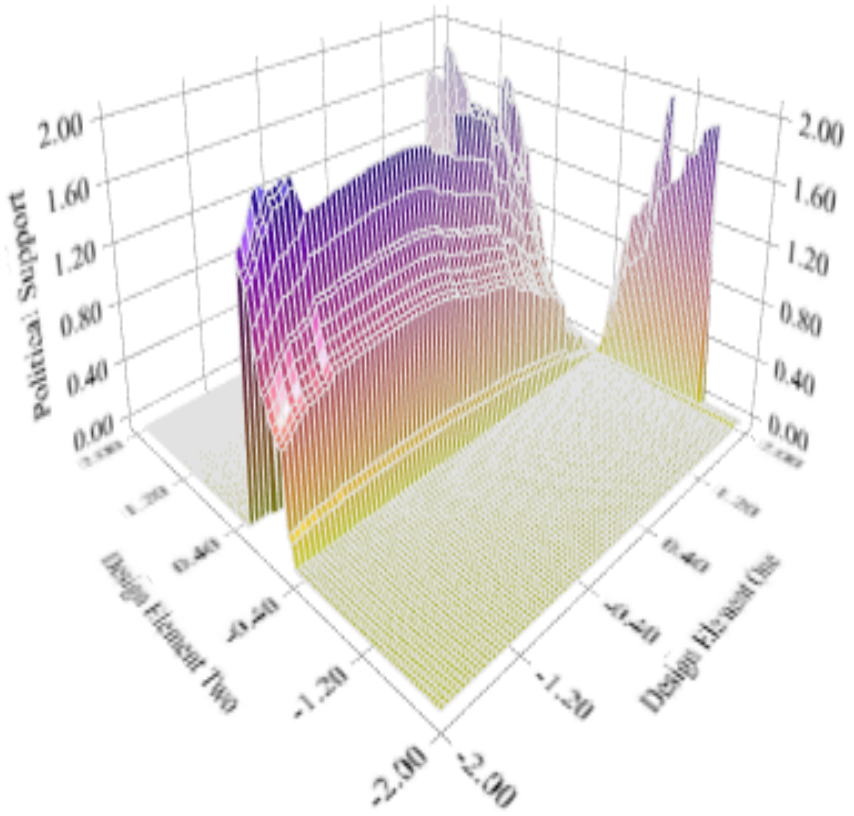
Limited implementation capability



Lots of implementation capability



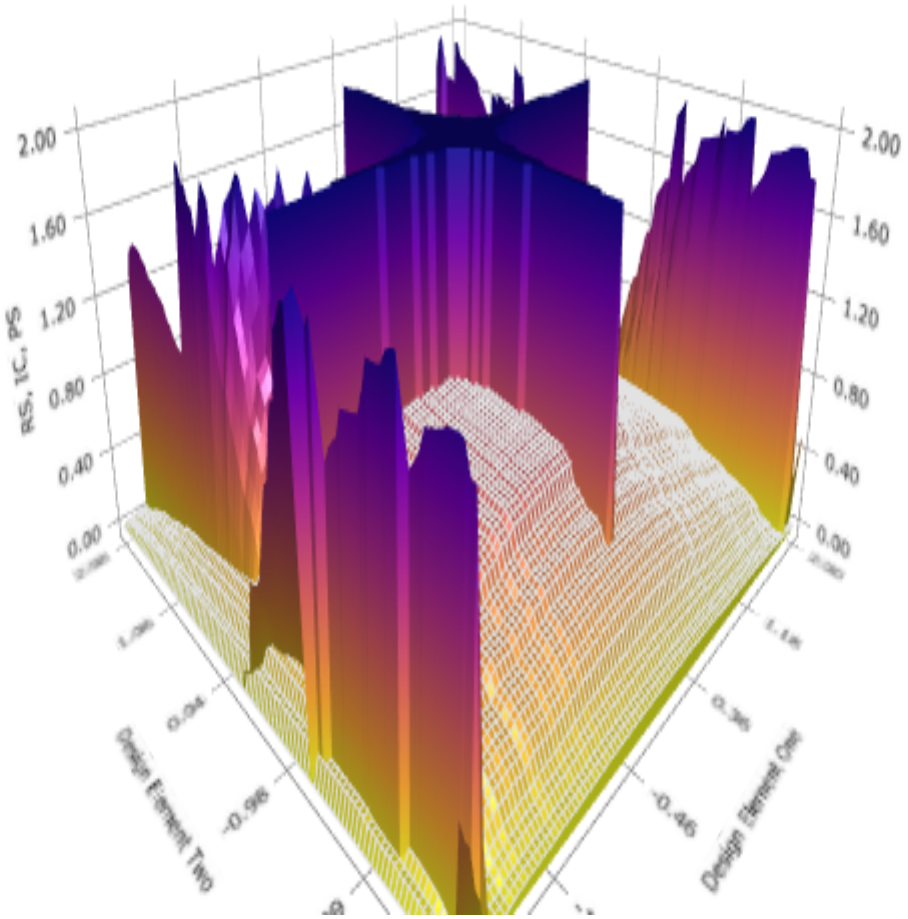
Political support surface



In order to increase well-being a Policy/Program/Project has to meet the Trinity

- *Instrumentally correct:* the design has to be such that, if it were implemented with fidelity it would lead to higher levels of well-being for the intended beneficiaries.
- *Administratively feasible:* The responsible organization has to be able to implement with reasonable fidelity the P/P/P with the resources made available to it.
- *Politically supportable:* One has to create and sustain a political coalition with sufficient power to authorize the P/P/P

Response Surface with Implementation and Politics



The "RCT as IIE" or *randomista 1.0* logframe for development impact has six *necessary* steps and (at least) five of the six are false

The knowledge about the response surface over P/P/P acquired through RCTs ….	
…can be generated about highly consequential actions	False. National development is a four fold transformation at *ontologically* aggregate process and individuated interventions are second order.
…leads to feasible large scale interventions	False. Efficacy of P/P/P is mostly limited by low organizational capability for implementation not knowledge of the response surface.
…either is in regions of political support and/or changes political support sufficient to authorize action	False. RCT knowledge has no special traction on political decision making.
… is of sufficient construct validity to guide action	False. Response surfaces are rugged over super high dimensional design spaces.
…is of sufficient external validity to be "amortized" and made cost effective	False. The external validity of RCT evidence is in many/most key instances is l
…is superior to other evaluation methods.	True.

	2018: Debate over. *Every* point to non-RCT advocates.
Topics important for development	National Development leads to better well being. National development is ontologically a social process (markets, politics, organizations, institutions). RCTs have focused on topics that account for roughly zero of the observed variation in human development outcomes.
Organizational capability and learning	Organizations doing any non-logistical activity (and most even of those) cannot be beaten into doing better by evidence from "independent" outsiders.
Political economy	There is massive evidence that governments do not implement many many many projects/proposals/programs that are cost effective and do spend budget on items known to be not cost effective. The NAP model of a benign SWF planner hampered by lack of rigorous evidence on effectiveness whose behavior evidence from an RCT will change is complete wacky nonsense.
Construct validity	RCTs examine an instance (or small numbers of treatment arms) which, in a rugged response surface over a high dimensional design space reveals next to nothing. Simple iterative/adaptive methods dominate RCTs in finding good policy designs.
External validity	External validity (a) logically incoherent when existing evidence has variance, (b) RCTs worse predictors of impact than OLS, (c) reviews show massive variance. If experiments were the hallmark of science alchemists would win Nobel prizes.





The “what works” approach is not commensurate to the transformational task at hand

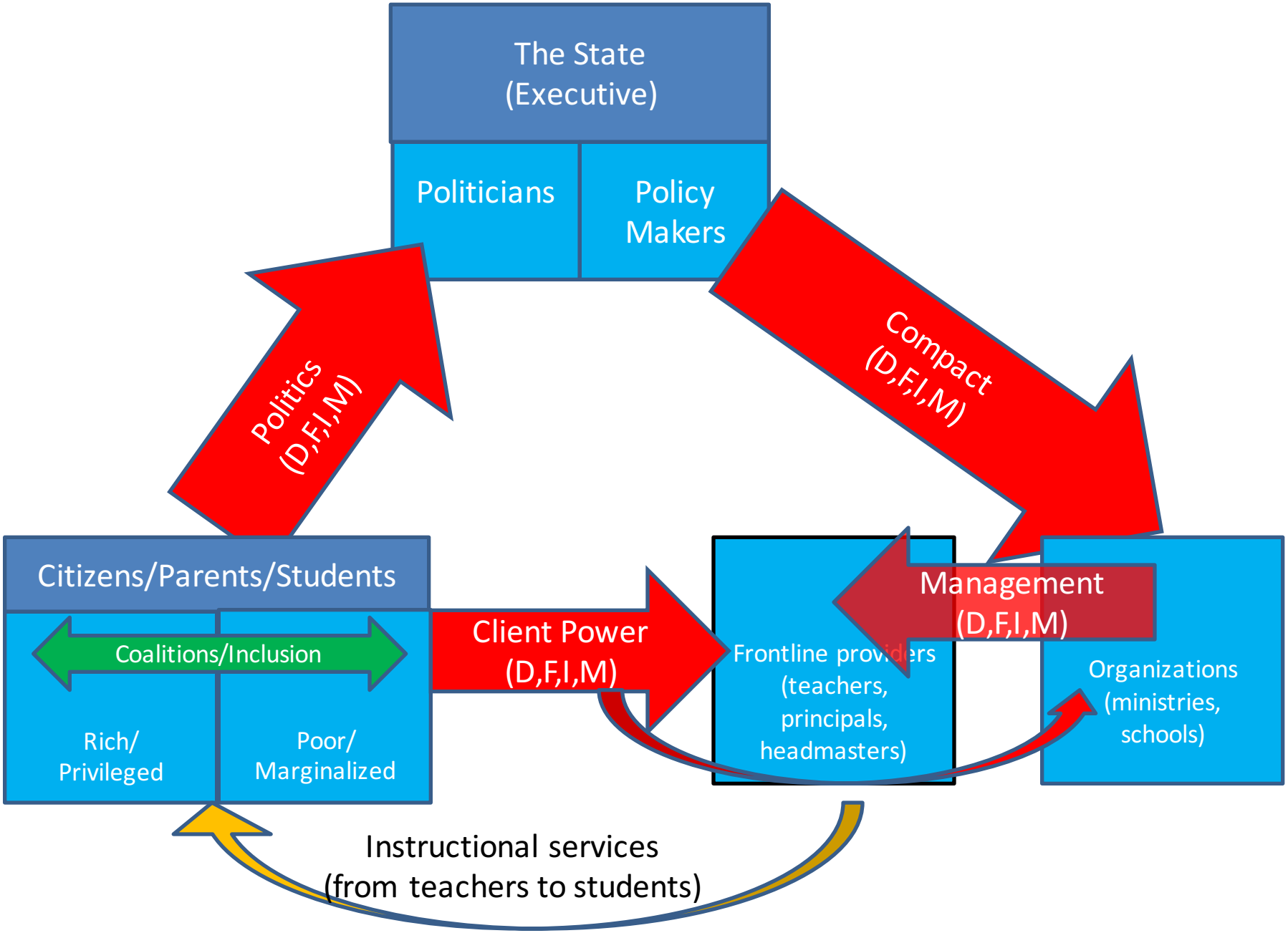
- The existing empirical literature identifies many programmatic efforts as having zero impact and even those with positive impact are, on this scale, on the order of 10 point gains for “treatment”—so the agenda that the 100 to 200 point gap in learning performance is to be erased by a sequence of 10 to 20 large impact programs seems…goofy.

Four examples (of many possible) of zero impact from more inputs into failing systems

- Indonesia *doubled* teacher salaries and a rigorous evaluation show exactly zero impact on learning.
- India increased federal spending by *ten fold* and overall per pupil expenditure *tripled* and yet a decade of ASER assessments (and other sources) show learning getting worse
- A rigorous evaluation of reducing class sizes in Kenya by hiring contract teachers shows learning improvements when implemented by an NGO—but the exact same program had zero impact when implemented by the MoE.
- Additional textbooks had zero impact unless they were accompanied by changes in teacher incentives.

World Development Report 2004 was a first cut at a system perspective

- Four primary principal-agent relationships of accountability among broad types of actors (*politics:* citizens to the state, *compact:* state to organizational providers, *management:* organizations to front-line workers, *client power: citizens to* front-line/organizational providers)
- Each relationship of accountability has four design elements: *delegation, finance, information, motivation* and performance of agents is endogenous (determined by those)


The State
(Executive)
Politicians
Policy
Makers
Politics
(D,F,I,M)
Compact
(D,F,I,M)
Citizens/Parents/Students
Coalitions/Inclusion
Rich/
Privileged
Poor/
Marginalized
Client Power
(D,F,I,M)
Frontline providers
(teachers,
principals,
headmasters)
Management
(D,F,I,M)
Organizations
(ministries,
schools)
Instructional services
(from teachers to students)

What is 'Accountability'? – Demystifying the Elements of the Accountability Relations

There are Four Features to Any Accountability Relationship and Performance is Endogenous

Feature	What	Example 1: Buying a Sandwich	Example 2: Going to a Doctor
Delegation	**You give a task to the accountable 'agent'**	• You ask for a sandwich	• You go to the doctor to be treated
Financing	**You give the 'agent' the money to do the task**	• You pay for the sandwich	• You pay the doctor for the treatment
Performing	**The 'agent' does the assigned task**	• The sandwich is made for you	• The doctor treats you to try cure your ailment
Informing	**You find out how well the 'agent' has done the work**	• You eat the sandwich which informs you of its quality	• You see if you are feeling better – you assess the performance of the doctor
Motivation	**You reward good performance and discourage bad performance**	• You choose whether to buy a sandwich from the seller the next time, affecting his profits	• You go to him next time (if he was good) or choose to go somewhere else if not

This is a system approach

- Specifies *actors* in the system (individuals or organizations or collections of individuals)
- Specifies connections between the actors in structured way (principals to agents) with specific conceptual dimensions of what flows between actors in a "design elements" of the principal-agent relationship
- Choices of agents are endogenous to the structure of the design elements of the agency relationship

[A plea about what graphs mean—it would be nice if the graphical elements had clear, one to one, conceptual meaning (or at least ontological status)—in most "system" diagrams there are arrows connecting things that don't mean the same thing (or anything at all?)]

Research Conjectures based on a system approach

- Education systems were coherent around enrollment objectives but incoherent around learning objectives.

- Accelerating progress on learning objectives will require systemic reforms to achieve coherence around learning.

What is a coherent system?

- System is a specification of the elements (e.g. actors/agents) and of the ways in which the actors are connected (e.g. feedback loops)

- Coherence is that the pieces of the system fit together towards a common purpose

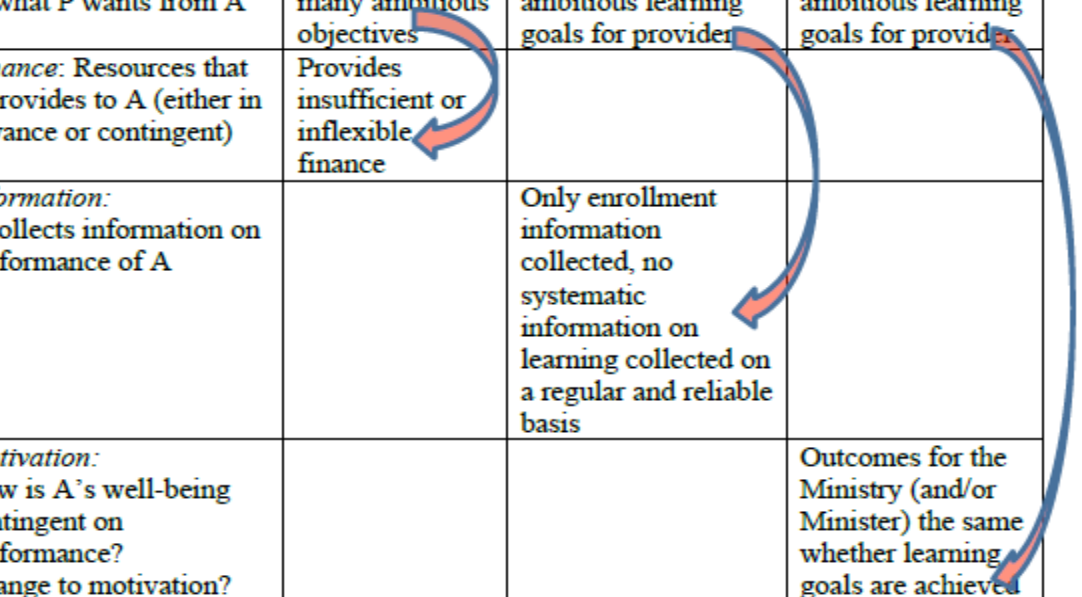
Table 3: Four by four diagnostic for systems of basic education				
Four design elements of each relationship of accountability (Principal (P) to Agent (A))	Principal-agent relationships			
	Politics: Citizens to "the state"/politicians (many P to one A)	*Compact:* "The state" to organizations (one P to one A or one P to many A with non-state providers)	*Management:* Organizations to front-line providers (one P to many A)	*Voice/ Client power:* Service recipients (parents/children) direct to FLP/Organizations (many P to one A)
Delegation: Specification of what P wants from A				
Finance: Resources that P provides to A (either in advance or contingent)				
Information: P collects information on performance of A				
Motivation: How is A's well-being contingent on performance? Change to motivation? - Intrinsic - Extrinsic - Exit (force out)				
Performance of agent (endogenous)				

Three types of incoherence in systems: *Within* a relationship of accountability

- Within a single relationship of accountability incoherence between the elements
  - Examples in the *Management* relationship between say a *Ministry* and *Headmasters* and *Teachers*
    - *Incoherence of delegation and magnitude and structure of finance:* goals are given without adequate and adequate autonomy over the use of resources to accomplish the task
    - *Incoherence of delegation and information:* Goals are set but no regular, reliable, repeated measurement of progress on goals
    - *Incoherence of delegation and motivation:* Goals are set but there is no connection between teacher performance assessment and structure of compensation and the goals.

Table 5: Illustration of potential incoherence *within* a single relationship of accountability, illustrated with *compact* (between executive apparatus of the state and organizational providers, e.g. between a Ministry of Finance and Ministry of Education)

Four design elements of each relationship of accountability (Principal (P) to Agent (A))	*Compact:* "The state" to "organizational providers (e,g. one Principal (e.g. Ministry of Finance) to one Agent (e.g. Ministry of Education) or one Principal to many Agents with non-state providers (e.g. state resources follows the student to schools)		
	Delegation to Finance incoherence	Delegation to information incoherence	Delegation to motivation incoherence
Delegation: Specification of what P wants from A	Delegation lists many ambitious objectives	Delegation lists ambitious learning goals for provider	Delegation lists ambitious learning goals for provider
Finance: Resources that P provides to A (either in advance or contingent)	Provides insufficient or inflexible finance		
Information: P collects information on performance of A		Only enrollment information collected, no systematic information on learning collected on a regular and reliable basis	
Motivation: How is A's well-being contingent on performance? Change to motivation? - Intrinsic - Extrinsic - Exit (force out)			Outcomes for the Ministry (and/or Minister) the same whether learning goals are achieved or not. Outcomes depend on budget utilization and process compliance.
Performance of agent (endogenous)	Cannot perform as delegation specifies. Weak *compact* accountability.	Performance of agent cannot be reliably assessed. Weak *compact* accountability.	No motivation for agent to perform well. Weak *compact* accountability.



Second type is *Incoherence* between same element across relationships

- Example: The *information* collected and used is different in each of the relationships.
- The information used in *management* (often "thin" information about *logistics*) is different from parent/child information about their own experience (*client power)* is different from how the state manages the ministry (*compact)* and all of these are different from the information that is (or is made) salient politically (*politics).*

Table 6: Illustration of incoherence in the same element of accountability across different relationships: Example of *information*

Four design elements of each relationship of accountability (Principal (P) to Agent (A))	Principal-agent relationships			
	Politics: Citizens to "the state"/politicians (many P to one A)	*Compact:* "The state" to organizations (one P to one A or one P to many A with non-state providers)	*Management:* Organizations to front-line providers (one P to many A)	*Voice/ Client power:* Service recipients (parents/children) direct to FLP/Organizations (many P to one A)
Information: P collects information on performance of A	Citizens know their own child's experience, but there is typically only aggregate (national/state/locality) information about enrollments, budgets, and inputs, not learning or learning progress. This often channels citizen pressure for "better" schools into these measured characteristics as politically salient.	Between the executive apparatus of the state and the organizational providers (typically Ministries of Education) the main apparatus and hence information is the budget allocation (both aggregate and across categories of expenditure (e.g. wages versus other) and program. To the extent "performance" elements are measured they tend to be measures of inputs or outputs, rarely outcomes and more rarely still learning outcomes.	Particularly in public sector organizational providers the information collected on teachers is mostly bureaucratic process compliance based on official internal systems. Teacher attendance is perhaps measured (though often not well), teacher participation in trainings, teacher reports on compliance with programmatic activities, and reports on measures of enrollment and (perhaps) student attendance. Very little information on teacher performance of any kind.	Parents/students know their daily experience with schooling and hence have "thick" information on aspects of teaching (e.g. is the teacher present, is class time boring, is the teacher kind or mean) and at least intuitive information on progress (e.g. does the child understand the lessons, is she/he able to do homework). But parents/students often lack any concrete, comparable, objective metric of their own child's position or progress or of the child's school versus others (particularly on a "value added" basis that adjusts learning outcomes).

Third type of *incoherence* is between entire relationships

- Teachers are caught between the accountability relationship to their employer (e.g. Ministry) and the accountability relationship to the students/parents they work with every day. All parts of this can be incoherent—*delegation* is different, *financing* is different (e.g. exclusively from Ministry), *information* is different (locally "thick" versus bureaucratic "thin"), *motivation* is different

Table 7: Illustrating incoherence between two different relationships of accountability affecting the same "agents" (teachers and headmasters)

Four design elements of each relationship of accountability (Principal (P) to Agent (A))	Principal-agent relationships	
	Management: Organizational providers (e.g. Ministry of Education) to public sector teachers/headmasters	*Voice/ Client power:* Parents/students to teachers/headmasters
Delegation: Specification of what P wants from A	Desired objectives for teachers designated as process compliance, e.g. teach in this school in this classroom these materials, not outputs or outcome performance	Parents want teachers to advance the interests of their children and to treat their children well.
Finance: Resources that P provides to A (either in advance or contingent)	Wages of teachers are fixed by teacher characteristics (whether related to learning or not)	Parents often provide little or no direct finance to teachers or school.
Information: P collects information on performance of A	Information on teacher performance based on official reports (e.g. attendance), process compliance and (perhaps) some supervision and (weak) performance assessments	Students (hence parents via students) have access to daily experiential observation on teacher behaviors and some knowledge about their own progress.
Motivation: How is A's well-being contingent on performance? Change to motivation? - Intrinsic - Extrinsic - Exit (force out)	Outcomes for teachers/headmasters based almost exclusively on seniority, cannot be fired, disciplined only with great difficulty, little extra reward for superior performance possible.	Parents/students would like to have to have only teachers who do well by their assessment of teacher performance.
Performance of agent (endogenous)	Cannot have two strong relationships of accountability that are themselves incoherent	

System coherence as organizing principle

- There are many ways to achieve “coherence” and this does not dictate any particular structure (as we have seen many structures succeed—from top-down authoritarian to “money follows the student”)
- Incoherence creates the possibility of lots and lots of action and effort and programs and spending and still no progress as systems are either coherent only around enrollment or worse, have exploited incoherence to introduce other drivers (e.g. political patronage)

	Starfish					Spiders
	Locality-level decentralization	Charter schools (only public-sector entrants)	Community-controlled schools	Private (for and not for profit entrants)	Pure markets for instruction (e.g., tutoring)	
Open?	Entry only by localities	Entry by designated organizations	Entry only by locally organized groups	Open entry	Completely open entry	Closed
Locally operated?	Mixed	Yes	Yes	Yes	Yes	No
Performance pressured?	Mixed	Mixed	Mixed	Yes	Depends on metric	Mixed
Professionally networked?	Regionally	Mixed	Mixed	Mixed	Weak	Hierarchy
Technically supported?			Yes		No	Yes
Flexibly Financed?	Mixed		Mixed	Yes	No financing	No flexibility

Global isomorphism and system mismatch: trying to make a government Weberian bureaucracy be "the system"

Table 1: How the importance of 'thick' information and economies of scale affect expected organization size

		Extent to which successful creation of value in the activity relies on application by front-line workers of specialized knowledge to difficult to externally observe features of the particular case	
		Thick	Thin
Extent of economies of scale or scope	Small	'Practices'—small organizations, often owned by professionals as sole proprietors or partners Examples: dentists, architects, lawyers, medical specialists	
	Large	'Franchises'—large organizations that reap economies of co-ordination in some areas (e.g. marketing) while relying on small units for 'thick' aspects of operation Examples: fast food, budget hotels, armies	'Bureaucracies'—large organizations, owned by large anonymous shareholders or non-profits, nearly all workers on salary Examples: postal services, railroads, automobile producers

Source: Author's compilation.

Table 3: Total enrollment and number of teachers in elementary education in selected Indian states, contrasted with other countries and with school districts in the USA

State/UT	Total enrollment in elementary education in government schools	Total government teachers
Bihar	20,519,815	347,330
Uttar Pradesh	19,585,396	509,508
West Bengal	13,256,933	449,724
Madhya Pradesh	10,221,216	268,471
Maharahstra	7,231,470	289,067
Rajasthan	7,155,509	266,505
Andhra Pradesh	6,175,060	348,221
Gujurat	5,982,181	206,203
Odisha	5,565,229	205,335
Jharkand	5,390,338	127,774
Karnataka	4,783,689	228,681
Tamil Nadu	4,226,225	149,868
Assam	4,174,185	145,935
Chattisgarh	3,789,376	161,268
Germany (total Primary, gov't and private)	2,912,938	
Punjab	2,193,899	110,284
Haryana	2,135,714	83,332
Delhi	1,742,738	44,523
Kerala	1,007,249	53,738
New York City Department of Education (All, K-12)	995,336	
Uttarkhand	907,931	44,643
Himachal Pradesh	695,417	17,776
Los Angeles Unified School District (all K-12)	667,273	
Sweden (all primary, gov't and private)	576,299	
Finland (all primary, gov't and private)	160,133	

Sources: DISE State Report Cards 2011 for enrollment and teachers in Indian states, UIS data for Germany and

Step 1: Unbundling

Question: What are the Key Functions and Activities in Primary Education?

Functions and Activities in Government Primary Schooling in Rural India

Function	Activity	Responsibility						
		Central	State	District	Block	Village		Service Provider (school)
						Gram Panchayat	User Groups	
Standards	Curriculum design Learning achievement standards							
Planning	Plans for physical expansion Plans for quality improvement							
Asset Creation	***Social Capital*** ***Physical Capital***							
Operation - Non Teacher	***Beneficiary Selection*** Choice of students for targeting programs Enrolment ***Recurrent*** Textbook choice/purchase Learning materials ***Maintenance*** Maintenance of school buildings/facilities Monitoring of school processes							
Operation - Teacher	Hiring Assignment Training Salary Supervision Dismissal							
Monitoring and Evaluation	Tests of learning achievement							

Output: A *Mutually Exclusive* and *Exhaustive* Classification of Primary Education into Functions and Activities

Step 4: Optimal Allocation based on *First Principles* Analysis

First Principles of *Public Finance*

Function	Public Finance First Principle			
	Economies of Scale	Externalities / System-wide Effects	Equity	Heterogeneity of Demand
Standards Setting				
Planning			-	
Asset Creation		-	-	
Operation - Non teacher		-	-	
Operation - Teacher		-	-	
Monitoring and Evaluation				

+

First Principles of *Accountability*

Function	Accountability First Principle		
	Discretionary?	Transaction Intensive?	Who Can Best Infer Performance (Technical or Local)?
Standards Setting	No	No	Technical
Planning	Somewhat	Somewhat	Bit Technical
Asset Creation	Yes	Yes	Local
Operation - Non teacher	Yes	Yes	Local
Operation - Teacher	Yes	Yes	Local / Technical
Monitoring and Evaluation	No	Yes	Technical

=

Functional Allocation in Primary Education – Based on *First Principles* Analysis

Function	Responsibility						
	Central Govt	State Govt	District	Block	Village: Gram Panchayat	Village: User Groups	Service Provider (school)
Standards Setting							
Planning							
Asset Creation			Support				
Operation - Non teacher			Support				
Operation - Teacher			Support				
Monitoring and Evaluation							

Key Messages

- States do **Standards Setting** and **Monitoring**
- PRIs assume responsibility for actual **Operation**
  - As much as possible as low as possible
  - Higher PRI tiers back-up on professionalism, technical

But our Analysis Avoids these Pitfalls by Suggesting Countervailing Forces to build accountability via decentralization

The Two Big Messages from Our Analysis

1

Strengthen Centre and State for *Standard Setting* and *Monitoring*

- **Consistent Standard Setting** is Critical To Provide the Guiding Framework for Local Governments to Operate and Manage
- **Uniform Monitoring** is Essential for Quality Control, Designing Rewards & Recognition Systems and Generating Credibility

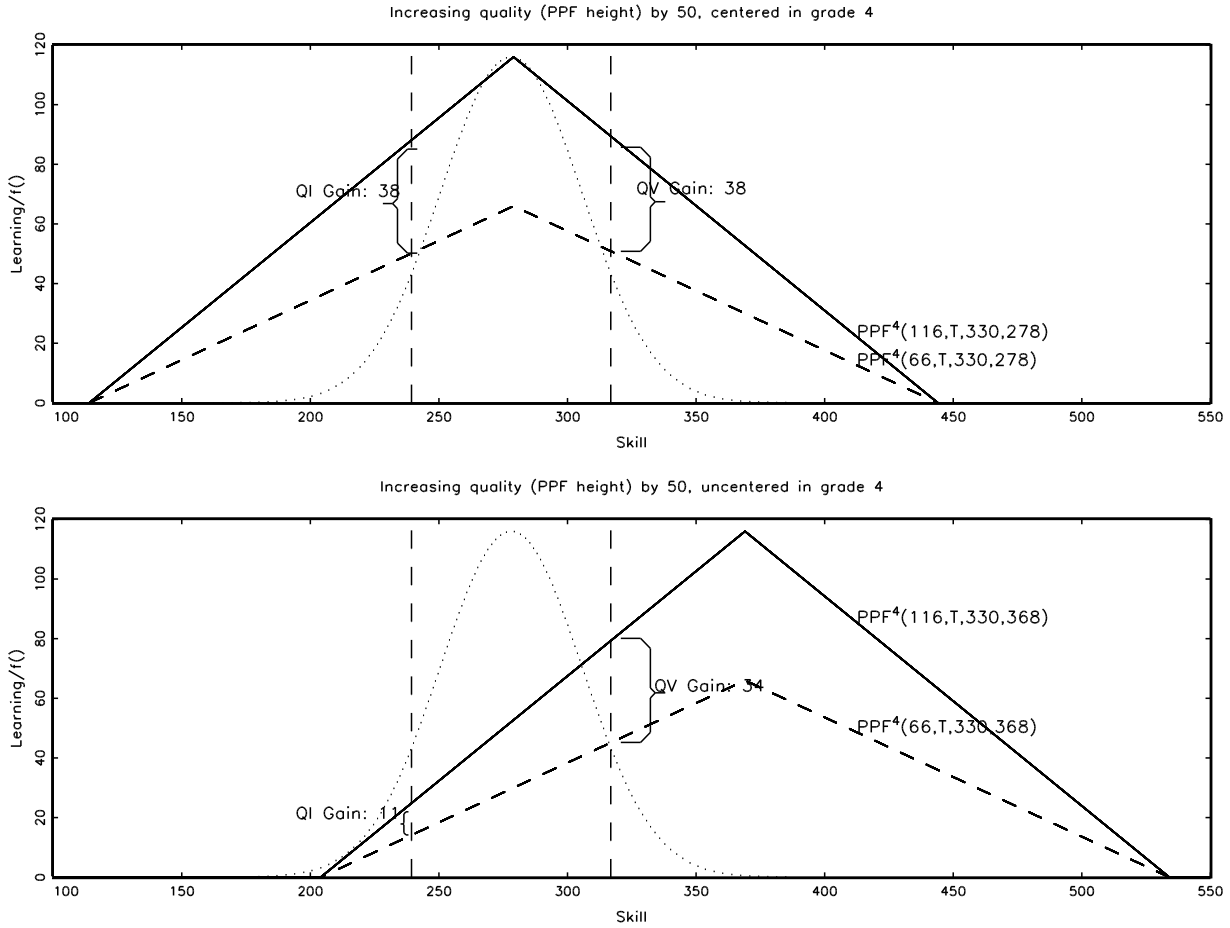
2

Greater *Operational* Responsibility to PRIs—the single biggest issue is teachers

Function	Responsibility						
	Central	State	District	Block	Village: Gram Panchayat	Village: User Groups	Service Provider (school)
Standards							
Planning							
Asset Creation							
Operation - Non Teacher							
Operation - Teacher							
Hiring							
Assignment							
Training							
Salary							
Supervision							
Dismissal		-					
Monitoring and Evaluation							

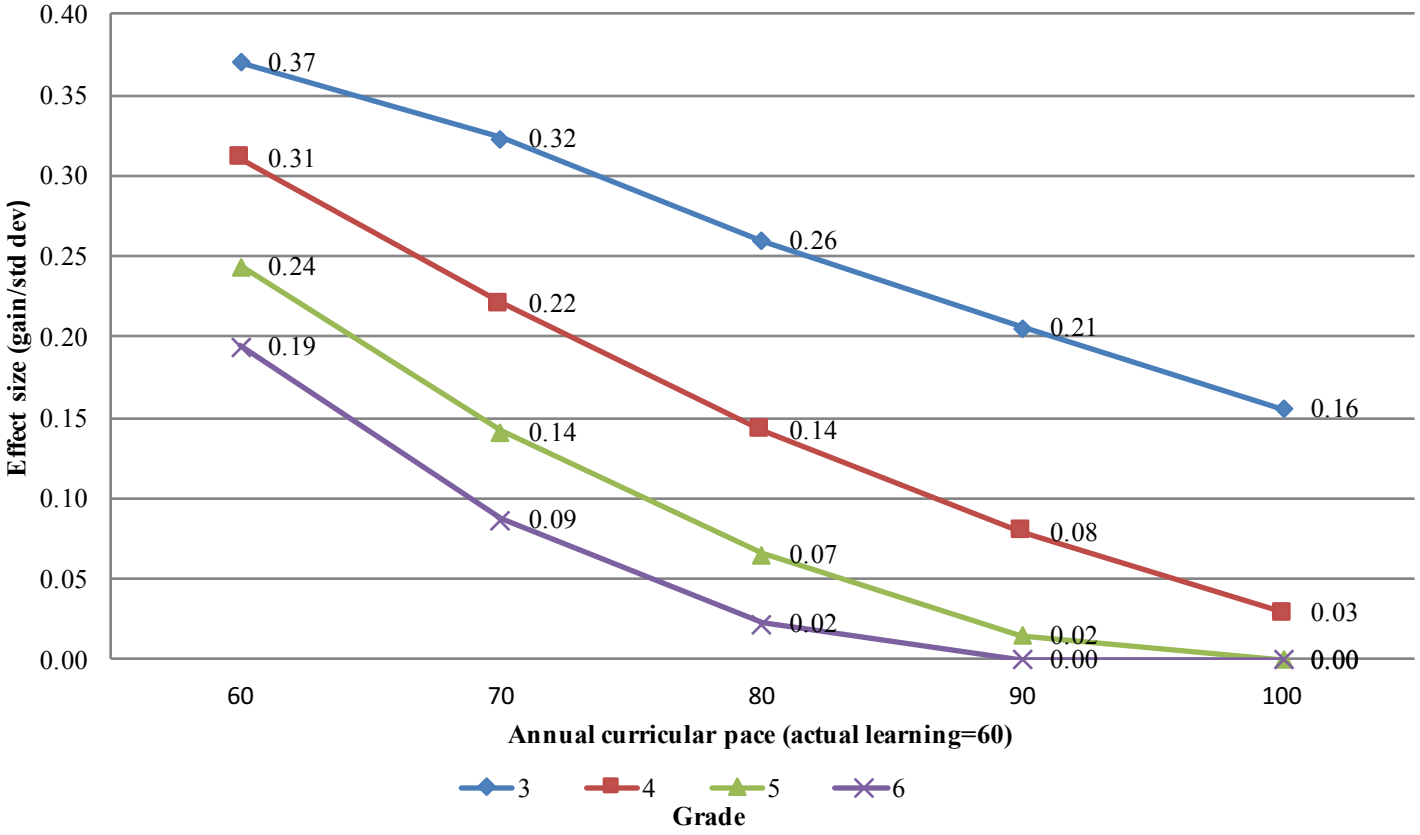
Slides for later

Example of a causal mechanism that produces different impacts across contexts: "uncentered" teaching

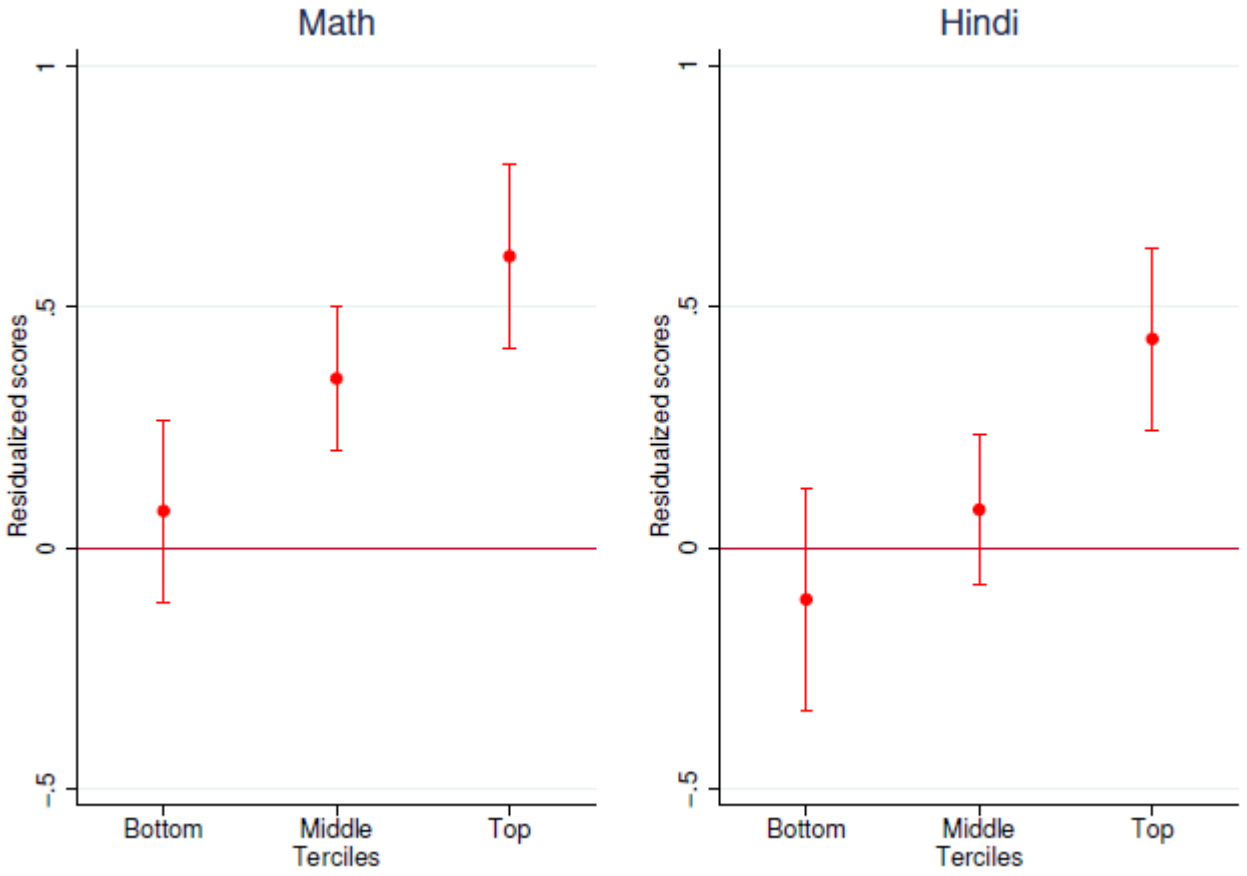


With uncentered learning any experiment uncovers a mix of increase in PPF (pedagogical production function) and curricular mismatch—the *exact same expansion* produces estimates from .37 to 0 effect sizes

Figure 15: The impact on student learning (in effect sizes) of increasing the PPF height by 10 in each grade varies by curricular pace



Bottom third of kids in Delhi schools grades 6 to 9 are not learning *anything*—control group value added score for bottom third is zero.



Source: Muralidharan et al (forthcoming—not to be cited yet)