

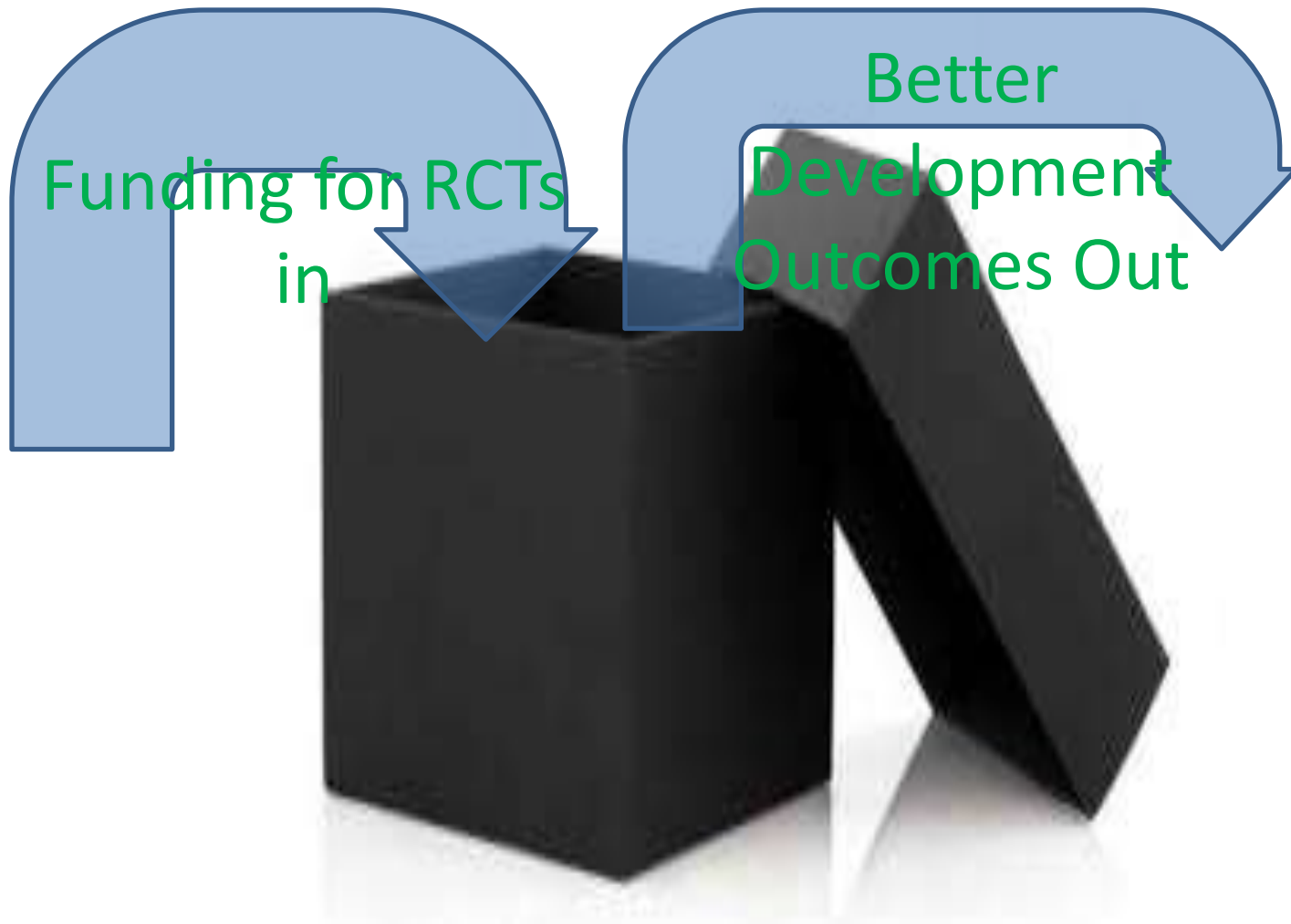
The Debate about RCTs in  
Development is Over:  
From RCT as IIE to MeE/SMART/etc.  
(Minimalist version)

Lant Pritchett

February 26<sup>th</sup>, 2018

LIEP

The important claims were about how to get published in economic journals but about *impact on development outcomes*—what was in the black box?



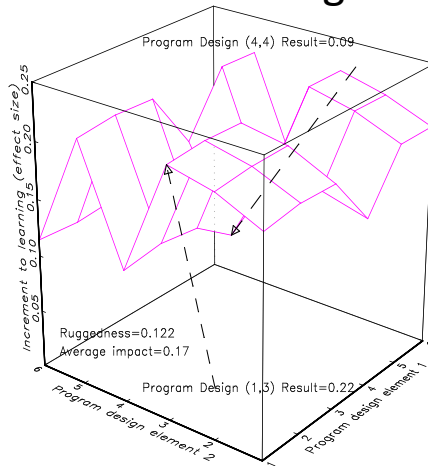
# In order to increase well-being a Policy/Program/Project has to meet the Trinity

- *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

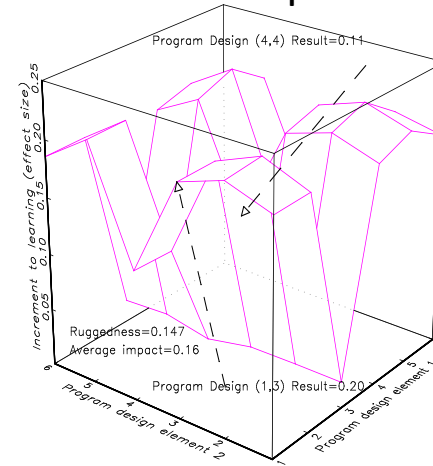
# Response surfaces with context (external validity) and construct validity (ruggedness)

Context A

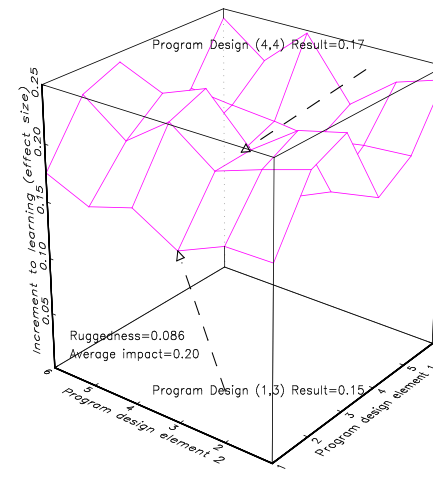
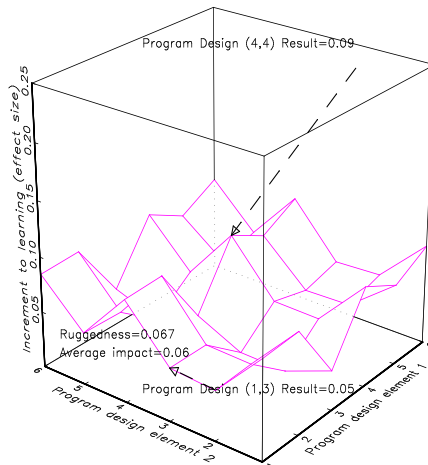
“Teacher training”



“Textbook provision”

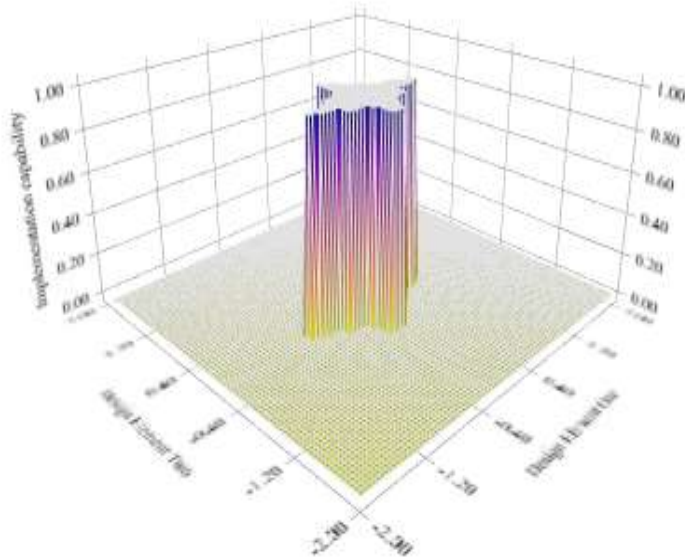


Context B

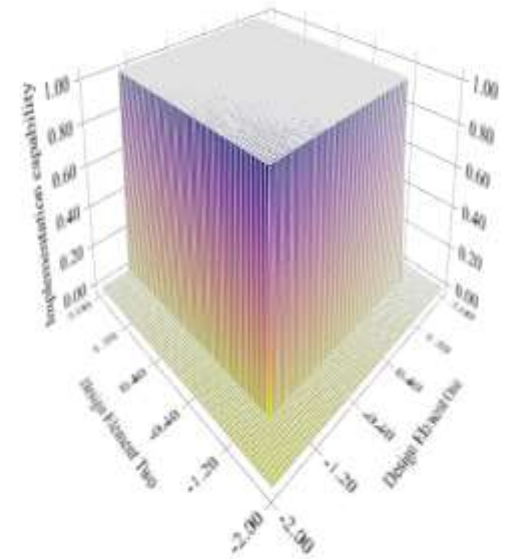


# Mappings of organization capability to replicate a policy/program/project with fidelity over the design space

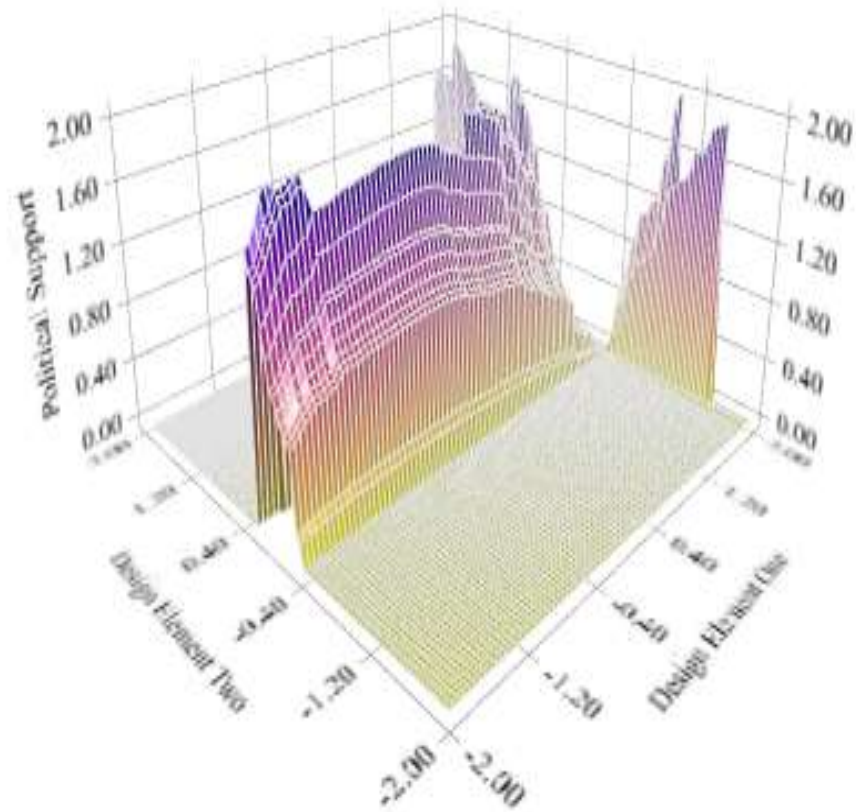
Limited implementation capability



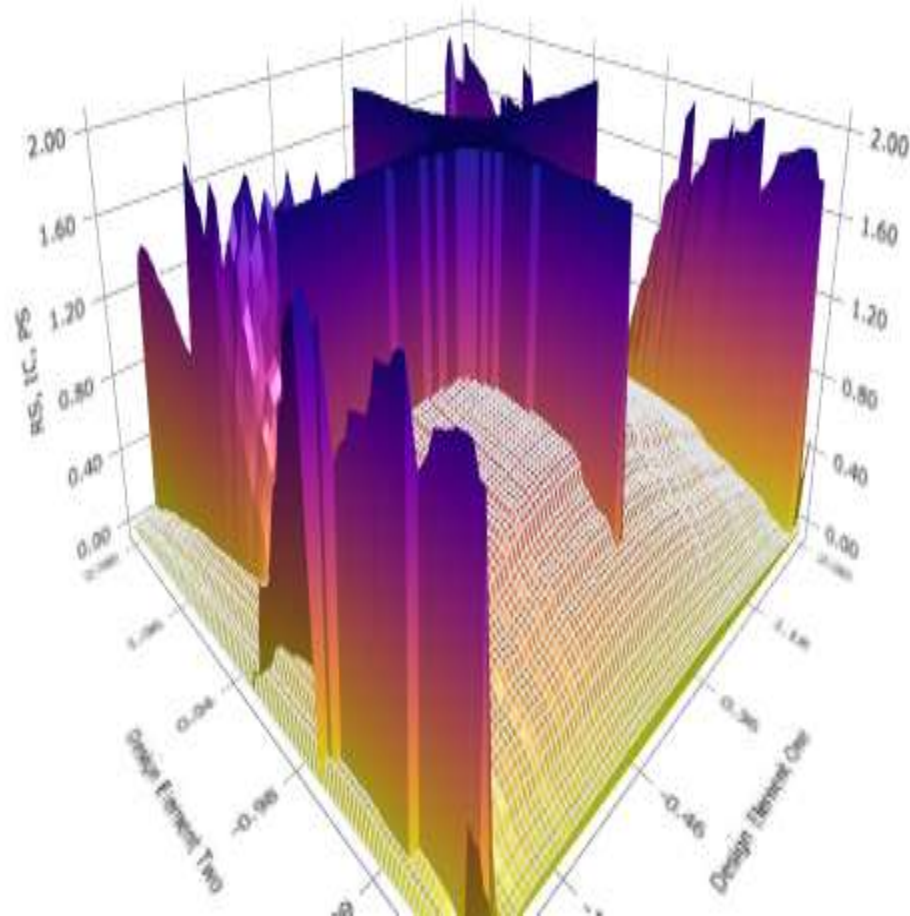
Lots of implementation capability



## Political support surface



# Response Surface with Implementation and Politics



# The first generation *randomistas* claim

Significantly more funding of rigorous *independent impact evaluations* using techniques of randomized control trials will lead, not just to more academic papers with firmer results, but to actual significant improvements in the development process (policies, programs, projects) that will lead to higher human well-being.



# Put another way

The kinds and types of knowledge that can, in principle and in practice, be generated by applying RCT techniques via independent impact evaluations to development projects/programs/policies are a key binding constraint on development practice (e.g. has a very high Lagrangian) and hence greater investments in RCTs will lead *pari passu* to significantly higher levels of human well-being cost effectively (relative to other available investments)

# The “RCT as IIE” or *randomista 1.0* logframe for development impact has six *necessary* steps and five of the six are false

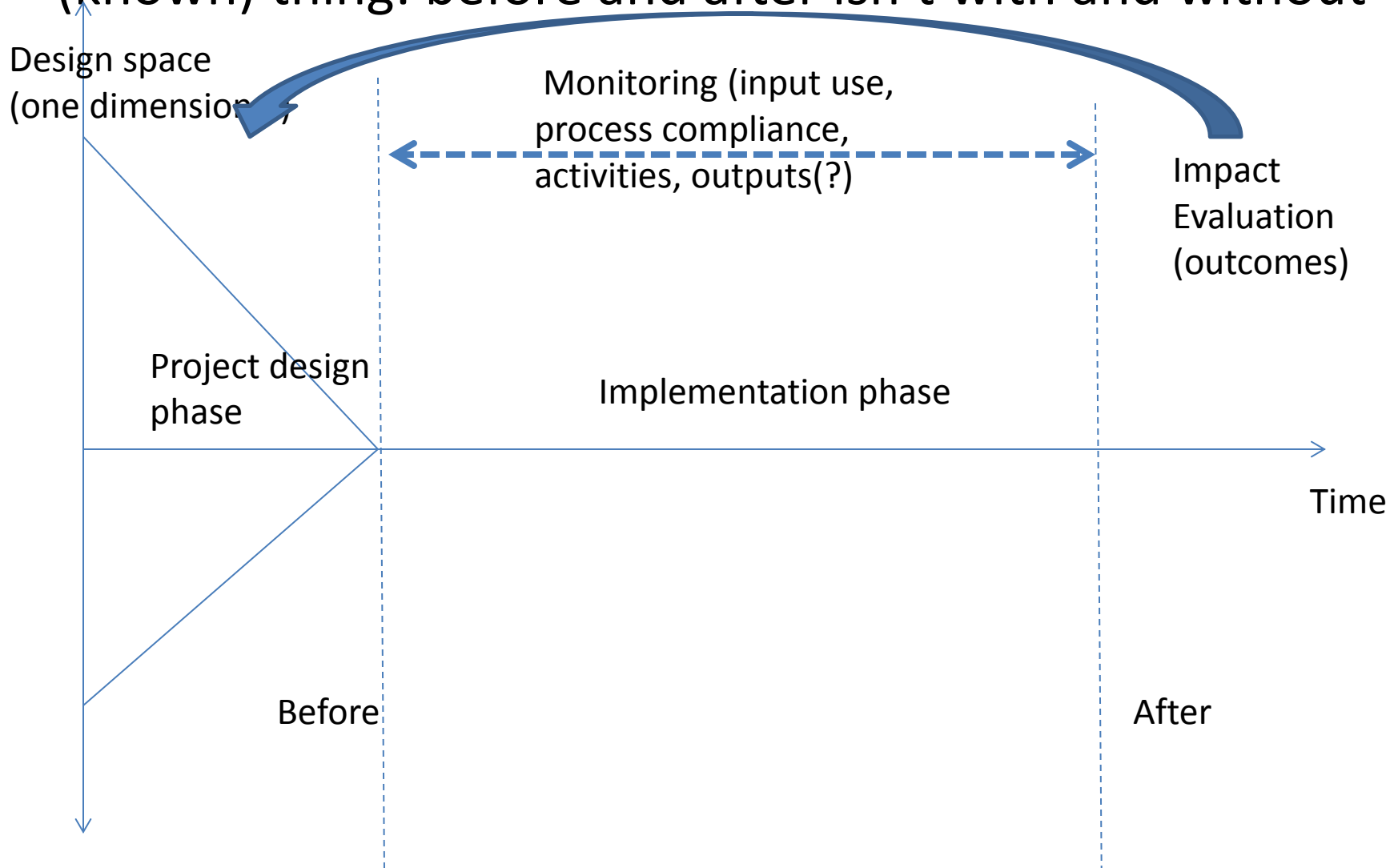
## The knowledge about the response surface over P/P/P acquired through RCTs ...

...can be generated about highly consequential actions	False. National development is a four fold transformation at <i>ontologically</i> 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 I
...is superior to other evaluation methods.	True.

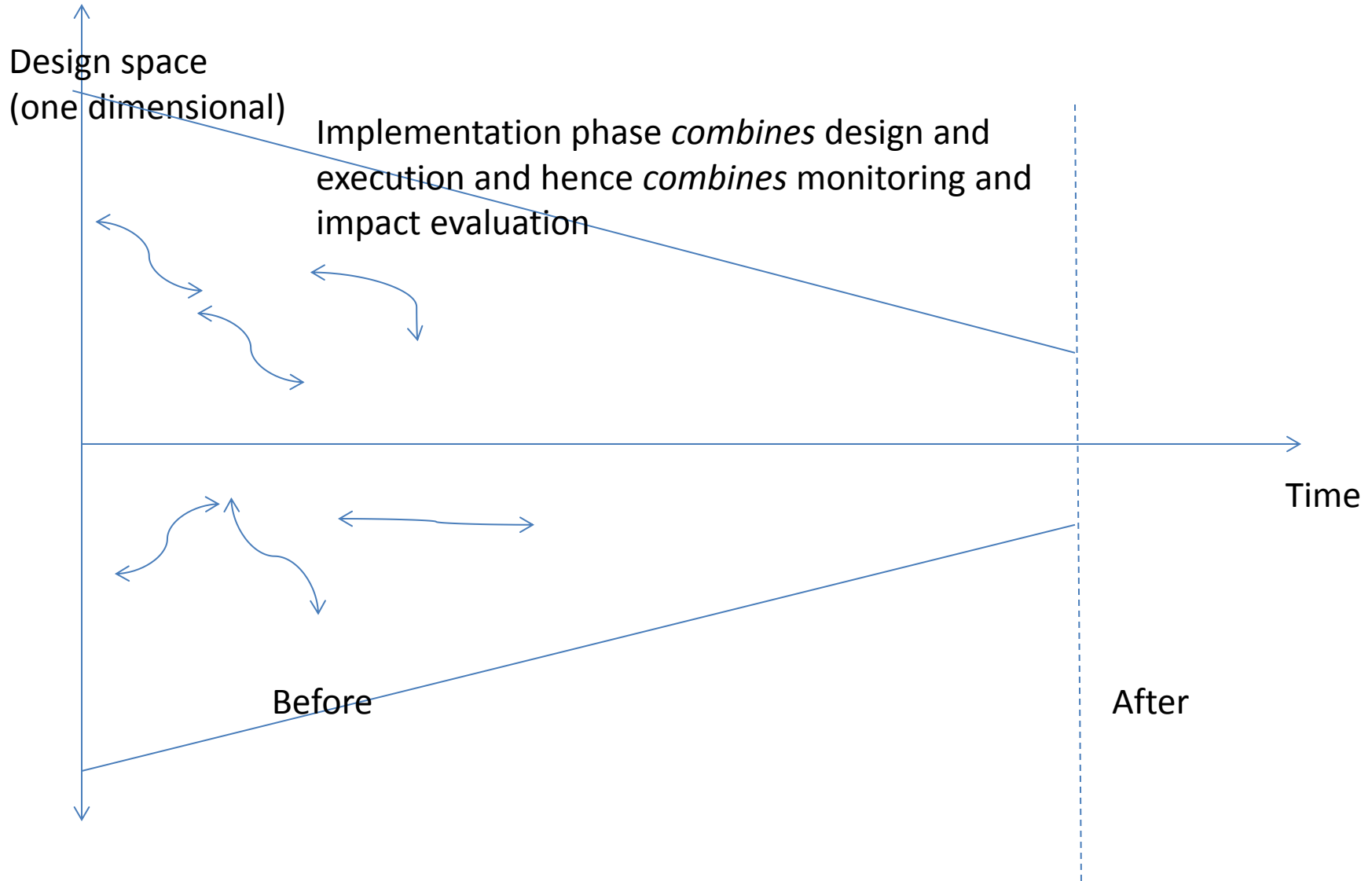
2018: Debate over. <i>Every point to non-RCT advocates.</i>	
Topics important for development	National Development leads to better well being. National development is ontologically a social process (markets, politics, organizations, institutions). RCTs have focused on topics that account for roughly 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 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 an RCT will change is complete wack 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 methods dominant RCTs in finding good policy designs.
External validity	External validity (a) <a href="#">logically incoherent when existing evidence has variance</a> , (b) <a href="#">RCTs worse predictors of impact than OLS</a> , (c) <a href="#">reviews show massive variance</a> . If experiments were the hallmark of science alchemists would win Nobel prizes.



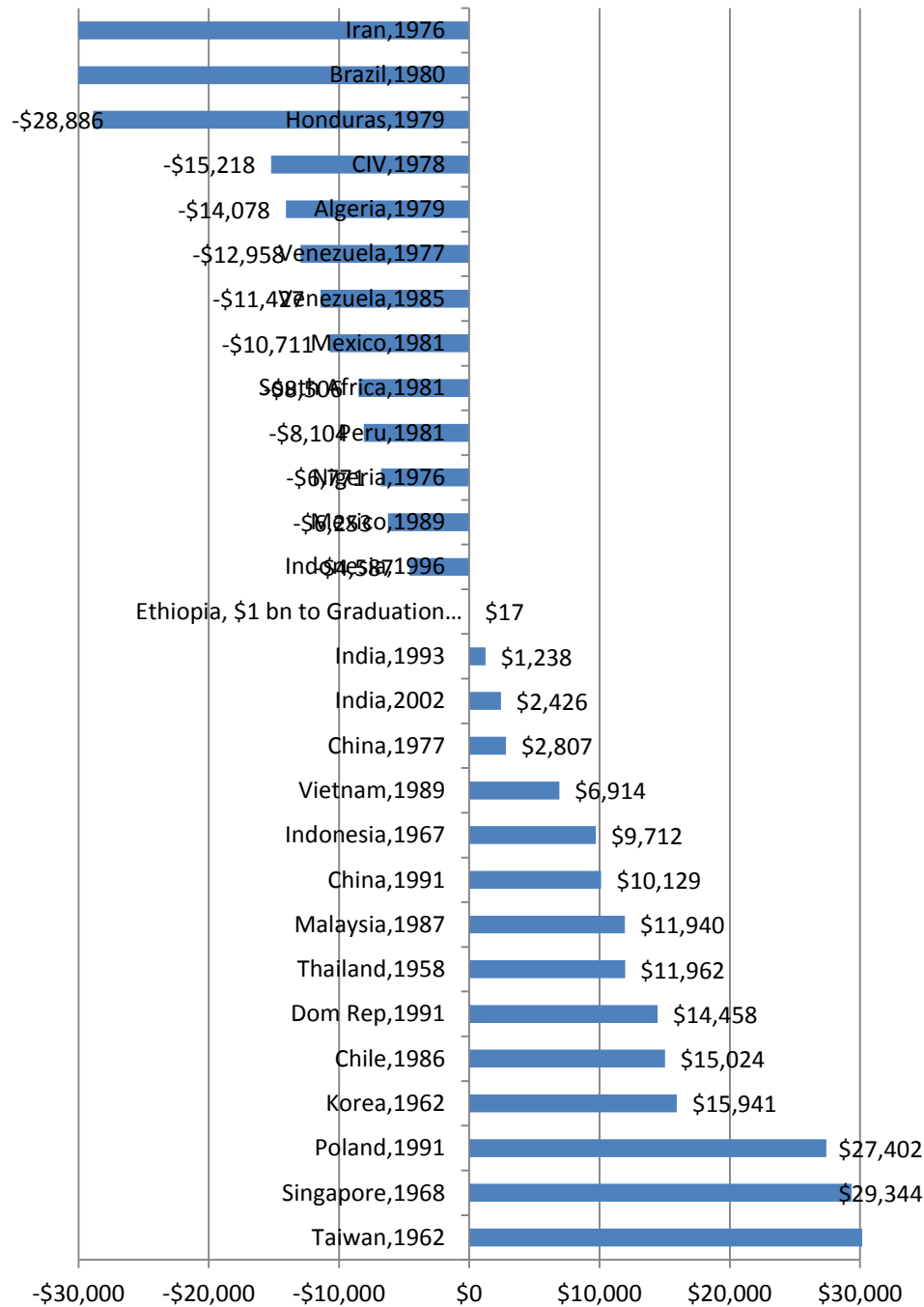
To this model of learning in the project cycle the RCT movement brought was mono-maniacal about one (known) thing: before and after isn't with and without



# What is radically different in “adaptive” approaches (like PDIA) versus the emphasis on “rigorous evidence”



	<b>RCT 2.0 –“learned from experience” and conceded on all key points and hence changed the practice of doing RCTs from “independent impact evaluation” to more MeE (Monitoring experiential learning, and impact Evaluation) approaches</b>
Topics important for development	Still stuck on this point.
Organizational capability and learning	<p>“Crawl the design space”—worth with local partners in the <i>design</i> phase and build implementation feedback loops to build towards effective interventions and capability simultaneously.</p> <p>This gives up on the notion of “independent” evaluation as now the “intervenors” and “evaluators” are the same people.</p> <p>This gives up on the priority of “impact evaluation” (from outputs to outcomes involving causal claims about impact on beneficiaries) to “efficacy”—helping organizations get from inputs to activities to outputs.</p>
Political economy	Working with governments on the generation and use of “evidence” as a broader issue than just doing an RCT. Conceded on the “special” role of RCTs.
Construct validity	Completely conceded. The specifics of program design have to be worked out instance by instance in an iterative way.
External validity	Completely conceded. Evaluation costs have to be amortized over the specific project as there cannot be claims of generalizable knowledge.



The gains (and losses) in NPV per person in the economy from big growth accelerations (or decelerations) are *orders of magnitude* larger than the per person NPV of the best available development programs.

The recently reported NPV per person gains from the Graduation approach done by BRAC in five countries *for the targeted households only* are on the order of \$1700 per person (off spending of \$1000 per person).

Say Ethiopia spends a US\$ 1 billion to benefit 1 million people for benefits of 1.7 billion. In a country of 100 million people this is U\$17 per person in NPV. This is two *orders of magnitude* (100 times) less than India's 2002 growth acceleration, three orders of magnitude less than Brazil's 1980 slow down.

*Source: Estimates of gains/losses adapted Pritchett et al 2016.*

The 20 interventions on which there have been sufficient rigorous impact evaluations to make comparisons about generalizability (Vivalt 2014)	Done more in more developed countries than less developed economies? (e.g. Denmark more than Malawi)	Done more today than historically in developed economies? (e.g. Denmark today more than in 1870)	Done more in rapidly progressing countries than stagnant countries? (e.g. More in Korea than Ghana)	Country's progress accelerates/ decelerates when a country does more/less of it? (e.g. More in China post 1978 than pre 1978)
Conditional cash transfers				
Deworming				
Improved stoves				
Treated bed nets				
Microfinance				
Safe water storage				
Scholarships				
School meals				
Unconditional cash transfers				
Water treatment				
Contract teachers				
Financial literacy training				
HIV education				
Irrigation				
Micro health insurance				
Micro nutrient supplementation				
Mobile phone based reminders				
Def				

What the RCT agenda has mostly been working on (by availability to do a review) doesn't pass a simple four part "smell test" for being important to development