

## Big Picture and Big Questions in Development:

### Randomization, Method or Madness?

While working for the World Bank in India on improving the performance of local government (the Panchayat Raj Institutions) I was discussing with the Bank team designing a large scale rural water supply project how to engage the local government. After about an hour of fruitless back and forth, in which I argued for the engagement of elected local governments and the project team argued for project specific structures for implementation, I decided we needed to step back. I said, “Before we go on, perhaps it would help if you clarified your vision of success. What would success in the water sector in India look like to you?” They thought for only a minute and said: “Success is that every household in India lives within a half a kilometer of an improved water source capable of providing 40 liters a day of safe water.” I said, “I see why we cannot agree on a project design. My vision of success is that every household in India can turn on a tap, inside their house, 24/7, and take a hot shower.”

Many discussions of randomized controlled trials (RCTs) have, as motivation, something like the Bill Gates/Chris Blattman debate. Bill Gates had been promoting chicken ownership to address poverty in Africa. In an open letter, Chris Blattman pointed out that cash may be more cost effective. Professor Blattman: *“It would be straightforward to run a study with a few thousand people in six countries, and eight or 12 variations, to understand which combination works best, where, and with whom. **To me that answer is the best investment we could make to fight world poverty.** The scholars at *Innovations for Poverty Action* who ran the livestock trial in Science agree with me. In fact, we’ve been trying, together, to get just such a comparative study started.”*<sup>1</sup> [emphasis added]

I am not going to engage on either whether cash versus chickens is the best way to fight world poverty or whether RCTs or other empirical methods are the “gold standard” of evidence for that question<sup>2</sup>. Once one goes down the rabbit-hole one is already in Wonderland. I triply deny the premises of these discussions.

First, many discussions of RCTs start (implicitly) from the premise that low-bar goals of poverty (or water, or energy, or schooling, or health)—that I call “kinky development” (Pritchett) are the only, or even the major, goals of development. *I deny the premise* that arbitrary, low-bar, definitions of poverty (or other low-bar goals in water or [energy](#) or schooling) imposed by the world’s elite on the complex lives, wishes, aspirations, goals, and ambitions of the world’s men and women constitute development, or even a very major goal, except as a way station on track to larger goals. I argue these kinky goals lack moral or political legitimacy. The goals of

---

<sup>1</sup> <https://www.cgdev.org/blog/getting-kinky-chickens>

<sup>2</sup> But cannot resist the aside that the “gold standard” was abandoned because it was a bad way to run monetary policy so it is not clear why one would want to be the “gold standard” of evidence—is it really to make the analogy that standard is too stringent for practical application, like the gold standard itself?

First draft. For comment only.

Please to not circulate or cite without permission.

development were, are, and should be: broad based sustained increases in multiple dimensions of human well-being (Kenny and Pritchett 2013)—the same kinds of goals that you, dear reader, have for yourself.

Second, *I deny the premise* that national development is not a sufficient cause of improved human well-being—and even of achieving the low-bar goals. National development is a four-fold transformation in the ways in which human beings interact and cooperate with each other to achieve (1) productive economies, (2) capable organizations (including in the public sector), (3) responsive politics, and (4) social inclusiveness. I show that national development brings the human development people do care about. Countries that make progress in these four functionally defined domains of development achieve higher levels of human well-being: strongly, nearly always, and in nearly every dimension. This is empirically so, in spades, and is common sense as these four aspects of national development constitute a mechanism for nominating and solving obstacles to improved human well-being. As a consequence, as independent (rather than endogenous to national development) causal determinants specific, targeted, programs or “interventions” and their design are weakly important, are hardly at all important in achieving human well-being—and even debatably important in kinky goals.

Third, *I deny the premise* that the kind of evidence one can, even in principle, generate evidence from RCTs is especially useful, perhaps even useful at all, in raising human well-being (outside of their application in the natural sciences as in drug trials or agricultural field trials). I argue that the kind of knowledge that RCTs can reliably produce—abstract, codifiable, (dare I say) “scientific” knowledge about “interventions”—is neither (a) a key constraint to progress in national development nor even (b) a key constraint on the realization of effective programs, even conditional on national development. Moreover, such knowledge lacks sufficient external and construct validity to make it particularly or especially useful.

Thirty years ago Robert Lucas (1988) wrote:

*Is there some action a government of India could take that would lead the Indian economy to grow like Indonesia's or Egypt's? If so, what exactly? If not, what is it about the "nature of India" that makes it so? The consequences for human welfare involved in questions like these are simply staggering: once one starts to think about them, it is hard to think about anything else.*

If one generalizes Lucas to all development questions, not just those of growth (which is just one of the four independently important national development transformations) but also to “actions” that would lead to capable state organizations like central banks and post offices and police forces and environmental regulation, and “actions” that would lead to stable politics that respond to the aspirations of their citizens and respect human rights, and “actions” that would lead to

First draft. For comment only.

Please to not circulate or cite without permission.

more equal treatment of people, independent of their condition of birth (e.g. sex, race, ethnicity), then I fully agree it is hard to think about anything else but national development<sup>3</sup>.

What is strange is that some fraction of development economists have committed themselves to a method that does not allow them to productively contribute to any of these big questions about development. They are devoted to think about anything *but* these big questions. How is that Chris Blattman, an incredibly clever person, could possibly think, looking at a world in which China and Vietnam have eliminated extreme poverty in our lifetimes, looking at a world in which Yemen and South Sudan have spiraled into humanitarian disasters, looking at a world in which Venezuela has changed in our lifetimes from a country that attracted migrants from Spain into a source of crisis refugees, in which India has gone from poor house to powerhouse, in which the fifty years from 1950 to 2000 have seen more progress in nearly all indicators of human well-being than the previous 5000, that RCTs on the “cash versus chickens” are the “best investment” to “fight world poverty”?

In Wolfgang Pauli’s memorable phrase, RCTs are not even wrong, they are just irrelevant to the pressing questions of national development and broad based human well-being<sup>4</sup>.

#### 1) *Human well-being as the goal*

Nearly everyone who works in development would agree the goal is to promote higher levels of overall human well-being<sup>5</sup>. There, however, the agreement stops. What constitutes the relevant elements of “human well-being” and whether these are concrete or subjective assessments and how whatever concept one has of human well-being could be aggregated into an normative ordering of human well-being of a group of individuals/households are questions that have been debated for centuries. I pretend to no contribution to those debates. My modest objective is to create a classification that encompasses nearly all the widely used measures in order to one simple distinction between “broad based” and “low bar” measures.

Imagine there were a Rawls-like pre-birth “veil of ignorance” (Rawls 1970) state in which you had some notion of personal identity and could choose in which country you would be born, only knowing your probabilities of birth conditions (e.g. sex, race/ethnicity, social status, material well-being of your parents) would randomly represent the country demographics. One way of thinking about the existing variety of measures of human well-being is, which *components* of well-being and what about the *distribution* of these components would you like to know in order to choose your country of birth? Table 1 (adapted from Pritchett 2010) proposes

---

<sup>3</sup> With the obvious exception of some time to think about personal mobility across countries as movers acquire the national development of the place (see Clemens, Montenegro and Pritchett, forthcoming, for wages and Pritchett 2018 more generally) with little or no cost to existing residences (particularly in a Pareto sense) as “national” development is (mostly, within some limits), like “A”--in the air (Clemens and Pritchett, forthcoming).

<sup>4</sup> Or more precisely, they are wrong, but they weren’t wrong about anything interesting.

<sup>5</sup> This formulation incorporates the natural environment and animal life, both current and future, on human well-being and hence includes automatically includes sustainability.

First draft. For comment only.

Please to not circulate or cite without permission.

what might be the components (in columns) and the summary statistics of the distribution of those components (in rows) on which you might make a choice of country of birth. There are hundreds of issues about how to properly measure human well-being raised by this table and, for present purposes, I am taking no stance on any of those. No Sen's capabilities versus income debate here.

What I do want to highlight is that "low-bar" or "deprivation" indicators have the very special feature that they summarize only that bit of the overall distribution for any component (e.g. income, water, schooling, energy) below some certain threshold. Just as with the rural water supply anecdote, none of the improvements in household water availability (distance, volume, reliability, temperature, quality) above a low threshold matter *at all* to these low-bar measures. That is, while in a normative ordering one might want to put *more* weight on the tail of the distribution, this is possible in a variety of ways through "inequality" or "tail event" averse measures but which do not simply ignore all gains above a threshold.

Table 1: Measures of human development distinguishing by elements of human capabilities and empirical characteristics of those capabilities					
Empirical measure of the <i>distribution</i> of the component of well-being across individuals/HHs to create a normatively ordered aggregate	Possible components or measures of human well being				
	Material (Income, Consumption, Wealth)	Specific dimensions of material well-being (e.g. food, water and sanitation, shelter, electricity, etc.)	Life Capabilities (e.g. education, health and health living)	Political and Social (personal security and safety, treated with dignity and respect, participate in political and social life, etc.)	Subjective well-being (e.g. Happiness, affect, life satisfaction, not wanting to migrate, etc.)
Typical Level or Central Tendency (e.g. mean, median)	GDP per capita, average income	Average caloric in-take, electricity consumption per capita	Average years of education, healthy life expectancy	Satisfaction with government, political participation	Average level life satisfaction, fraction wanting to migrate
Inequality-adjusted typical level	Atkinson indices, inequality averse SWF measures	Measures with declining weights on higher levels of access	DALY or QALY measures	Equality of access, treatment in political/social life	SWF like aggregates with declining marginal weights
Inequality, Fairness, Equal Opportunity	Dispersion of income, Gini index, gender disparity, ethnic differences	Regional differences in access	Differences in access to education or health status by SES of parents, ethnicity	Identity based discrimination, limitations of political access, ethnic conflict	Differences in life satisfaction by identity, condition of birth
Volatility (measure of the variation over time for given individuals, insecurity, vulnerability)	Consumption volatility	Likelihood of periods of water shortage, food insufficiency	Risks of health shocks, epidemic/pandemics,	Frequency of political conflict, episodes of social violence	Risk perception measures, fear, uncertainty
<i>Low-bar deprivation (below a certain level)</i>	<i>FGT(<math>\alpha</math>) measures of poverty with a low bar</i>	<i>Percent with "access to safe water", percent without sanitation</i>	<i>Completion of primary school</i>	<i>Set of minimal negative rights protected, safety from violence</i>	<i>Fraction with negative affect, fraction below a given level of life satisfaction</i>

Source: Adapted from Pritchett (2010) Table 2.

First draft. For comment only.

Please to not circulate or cite without permission.

I.B) *Kinky Development goals are not legitimate development goals*

I refer to “low-bar” goals as “kinky” goals for two technical reasons. First, if one took the distribution of a dimension of human well-being and were able somehow to just eliminate that part of that distribution below the line and hence eliminate the “deprivation” completely, the resulting empirical distribution would have a mass just at that line and hence a “kink” in the distribution at that point. Second, if one examines the derivative of progress with respect to improvements in the indicator then at the line the derivative falls to, and stays at, *exactly* zero. This creates in most low bar goal measures a discontinuity in the derivative (e.g. completing grade 6 counts towards progress in “primary completion” but completing grade 7 counts for zero) or, at best, in sophisticated deprivation measures like FGT(1) poverty measures, a kink in the derivative (a discontinuity in the second derivative).

My argument is that these kinky measures are fundamentally illegitimate as stand-alone goals or measures of human well-being. Kinky goals can be way-posts towards a destination (“on our way to prosperity we wish to eliminate low-bar poverty” or “on our way to an appropriately educated society we wish all children to complete primary school” or (similarly) *parts* of goals (“as part of our efforts in improving the quality of household water supply we want every household to have minimal access”)) but they are fundamentally illegitimate as descriptions of, or “goals” of human development. I argue kinky goals are fundamentally illegitimate in two ways: normatively and politically.

*Normative inadequacy of low-bar goals.* A widespread (if not universal) principle of “moral realism” is something like the “golden rule” (do unto others) or the Kantian categorical imperative (“Act only according to that maxim whereby you can, at the same time, will that it should become a universal law” (Kant 1785 (1998))<sup>6</sup>). I argue that low-bar goals are normatively illegitimate because no person, in their private behavior, treats the low-bar goals as a complete normative ordering of their own well-being (even restricted to a given component of well-being, like education or water or income or energy) and hence it violates fundamental moral principles to project those as goals onto others.

Poverty measures based on a “low-bar” line in per capita consumption like “dollar a day” or “two dollars a day” (these have obviously changed with inflation) of the Foster Greer Thorbecke (1984) type have been widely computed and promoted, often as a goal, if not “the” goal of development. When Blattman says “world poverty” he probably has in mind something like at low bar FGT measure. These  $FGT(\alpha)$  measures are (weighted) integrals of the distribution of consumption or income up to the poverty line and hence have, as part of their definition, that the measure is complete invariant to the part of the distribution above the line. This implies that

---

<sup>6</sup> (Parfit 2011) recently argued that three common approaches to moral questions the Kantian deontological, consequentialism, and contractualism ultimately converge to the same answers and that these are “correct” answers. First draft. For comment only. Please to not circulate or cite without permission.

gains to income/consumption above the poverty line *exactly zero* gain in the normative measure of poverty.

(Essentially) *no one* acts as if, or sincerely believes, that their own well-being has *zero* gain from income/consumption gains above a low-bar poverty line. Most economists grant there is declining marginal utility of income. Empirically measures of “life satisfaction” well-being are roughly linear in the natural log of income (this implies declining marginal utility) to as high as incomes as one can reliably measure and measures of “emotional well-being” increase up to very high levels. Stevenson and Wolfers (2013) show smoothed regressions of the Gallup life-ladder satisfaction question and household income for 25 countries which shows life satisfaction not only is not *exactly zero* above a low-bar poverty threshold, it is roughly *linear* in natural log of household income up to levels *twenty times* higher than even the “higher low-bar lines like “two dollars a day” (64,000 in HH income for a family of 4 is \$43 dollars a day)<sup>7</sup>.

Kahneman and Deaton (2010) show that in the USA “ladder of life” (Cantril scale) measures of life satisfaction improve, roughly linear in log income up to very high levels, while there is a decline to near zero of “emotional well-being” measures like positive affect with a plateau starting around \$75,000 in household income, with a HH size of 4, is \$51 dollars a day.

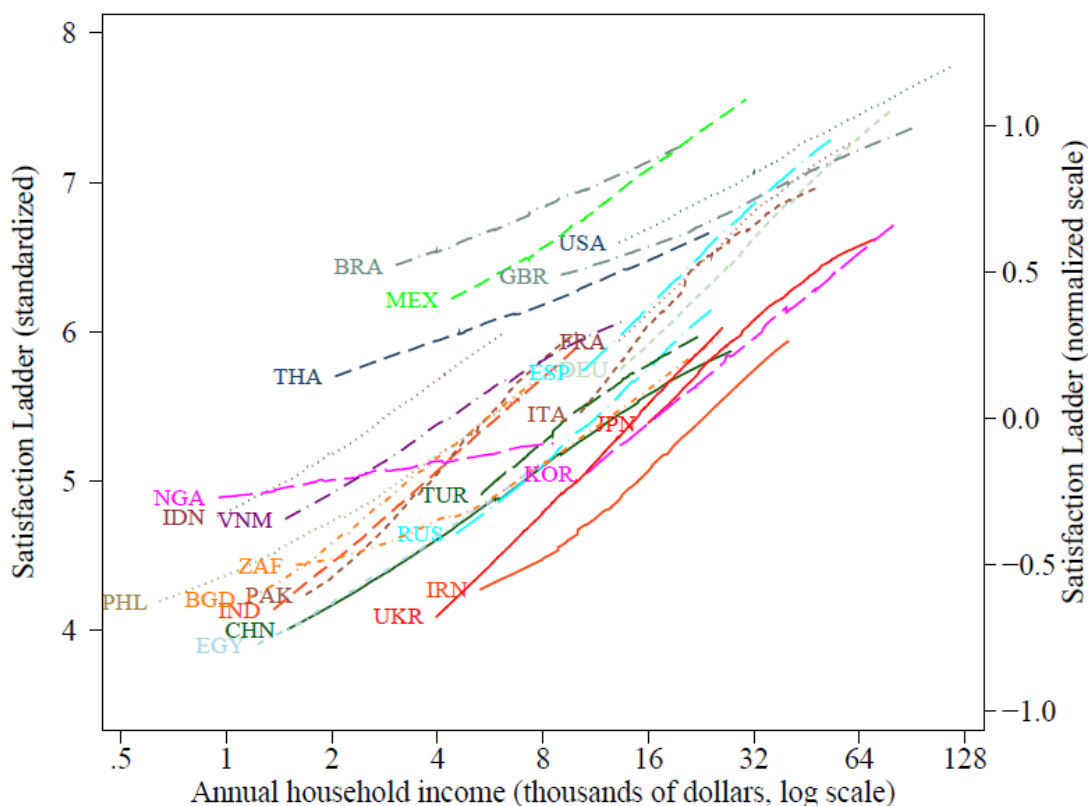
---

<sup>7</sup> This is consistent with sources, such as the Gallup surveys from countries around the world which also show that life satisfaction is roughly linear in natural log income all the way up (Diener et al 2010).

First draft. For comment only.

Please to not circulate or cite without permission.

**Figure 1: There is no line: the relationship between life satisfaction and household income doesn't go to zero as poverty measures imply**



*Source: Stevenson and Wolfers 2013.*

Beyond such direct evidence it is worth pointing the obvious absurdity of a low-bar poverty line as descriptive of individual's normative assessments. Consider work. To reach the income for a household of 4 of twice the current (inflation adjusted, PPP) poverty line of \$1.90 one household member would have to work 40 hours a week, 52 weeks a year at a wage rate of \$2.66/hour (40 hours/week\*52 weeks/year\*2.66 \$/hour=3.80 \$/day\*365 days/year\*4 people/HH=5,548\$/HH). If there were really zero marginal utility to income we would expect labor supply to diminish sharply above that level and people would be indifferent to wage gains above that level. Both of those are obviously false, not just near any low-bar poverty line but far, far, up the income distribution.

It is worth pointing out that nearly all people do not act as if there are discrete, low, levels in per capita consumption, or access to water, or education, or health services or housing or (nearly) any component their own human well-being and that gains below those levels count and gains above those levels count for zero (or even that there is any discrete and sharp fall). This would imply zero willingness to pay for improvements above those levels, which is massively counter-First draft. For comment only.

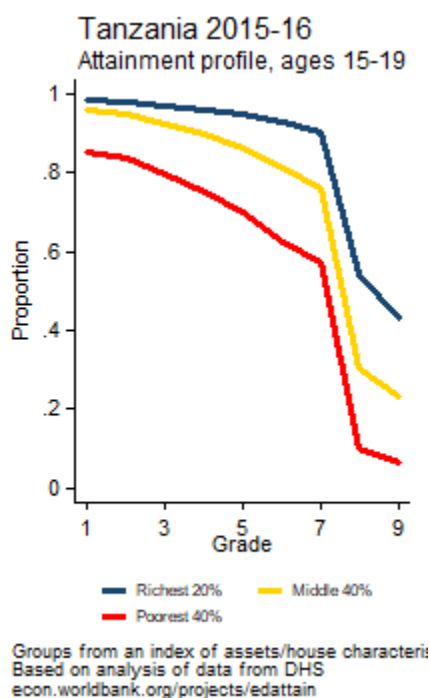
Please to not circulate or cite without permission.



factual, as evidenced by the existence of thick and vibrant markets in all of those services for quality/quantity levels far, far above the low-bar level. Studies in water access found household's willingness to pay for private household connections was often twice as high as for the "low-bar" goal of a public standpipe (World Bank Water Demand Research Team, 1993).

Tanzania under Nyere, for instance, took the "universal primary completion" goal quite literally and seriously (for reasons of his own) and hence created a system in which the government did not provide secondary schools nor did they allow private schools. This led to a grade attainment profile like Figure 2, in which roughly 44 percent of all children terminate their education after grade 7 (the end of primary) and the grade completion of children from the poorest 40 percent of households falls from 57.5 percent completing grade 7 to only 10 percent completing grade 8. On the "low bar" goal of "primary completion" this fall in attainment is a matter of indifference to students and parents as it counts for zero in development goals. I strongly suspect, that is not their view.

**Figure 2: In Tanzania over 40 percent of children terminate their education at the end of primary school (grade 7): Is that a matter of indifference to parents because the MDG was met?**



*Political legitimacy.* The second way that low-bar goals are illegitimate as development definitions or human development goals (except as way-stations or sub-goals) is political. While I was working with the World Bank in India I had a memorable conversation with a very experienced and very senior Indian policy maker. His message was: "The World Bank wants to

First draft. For comment only.

Please to not circulate or cite without permission.

talk exclusively about “the poor” but, in so doing, you are making yourselves irrelevant to the Government of India. First, India is a democracy and in a democracy we have to get more than 50 percent of the votes. Headcount poverty is 26 percent. We cannot have a political agenda of ‘poverty reduction’ as an exclusive or even organizing principle and win elections. Second, once we are elected the government of India we must represent and encompass the interests of all citizens. We cannot tell people that just because they are now above some level of well-being their interests are no longer of concern to their government.”

The donors and development organizations have, in recent decades, been pushing two very contradictory messages: (a) all countries should be “democracies” that respect human rights, rule of law, and hold free and fair elections, (b) their major (if not exclusive) goal is “poverty reduction” perhaps extended to few other low-bar goals (the MDGs). This is a development agenda with very few developing countries. Table 2 shows that in all but four of the 20 largest population developing countries even the 40<sup>th</sup> percentile is above the “extreme poverty” threshold. Hence the “poverty agenda” cannot be the organizing principle or foundation for a democratic government in developing countries with 4 billion people.

This a fundamental question about what “development assistance” means. One definition is that it is a mechanism whereby multi-lateral and bi-lateral organizations help (legitimately constituted and selected) developing country governments and their citizens achieve their own locally nominated legitimately constituted goals. Another definition is that “development” organizations, multi-lateral and bi-lateral (and philanthropic), choose what *they* define to be the goals *they* (through what process?) regard as legitimate development goals and then extend assistance and cooperation to developing country governments and citizens only to the extent it advances the development agenda as defined by the goals *of the development organizations*. The Millennium Development Goals and the promotion of the “poverty” agenda (defined by an unnecessarily low-bar standard) were a decisive shift towards the latter approach.

I argue the low-bar approach is politically illegitimate in two senses. First, it cannot be the foundation of the political agenda of a democratic (or even legitimately responsive state) as these to not, as argued above, reflect the actual aspirations of their citizens (except in the most extremely poverty ridden states). Second, one cannot pretend there is a development “partnership” or that developing country governments “own” the goals if the agenda is set by global elites and expected to then be reflected back from the “recipient” government.

First draft. For comment only.

Please to not circulate or cite without permission.

Table 2: An “extreme poverty” goal is extremely exclusionary—gains to the middle income groups in nearly every country in the world are not included as “progress” by a goal of eradicating extreme poverty						
Country Name  ( <i>All</i> middle income groups are excluded by an extreme poverty goal in the countries in bold)	Pop'l	Gains to citizens at the specified percentile of the distribution included in a “poverty” goal?				
		80th	60 <sup>th</sup>	40th	60 <sup>th</sup>	60 <sup>th</sup>
		Extreme Poverty (\$1.25 per day)			USA poverty line \$15 per day	Half of US poverty line
<b>China</b>	<b>1357.4</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>India</b>	<b>1252.1</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Indonesia</b>	<b>249.9</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Brazil</b>	<b>200.4</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	No
<b>Pakistan</b>	<b>182.1</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
Nigeria	173.6	No	No	Yes	Yes	Yes
Bangladesh	156.6	No	No	Yes	Yes	Yes
<b>Russian</b>	<b>143.5</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	No
<b>Mexico</b>	<b>122.3</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	No
<b>Philippines</b>	<b>98.4</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Ethiopia</b>	<b>94.1</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Vietnam</b>	<b>89.7</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Egypt</b>	<b>82.1</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Iran</b>	<b>77.4</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>Turkey</b>	<b>74.9</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	No
Democratic Republic of Congo	67.5	Yes	Yes	Yes	Yes	Yes
<b>Thailand</b>	<b>67.0</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
<b>South Africa</b>	<b>53.0</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
Tanzania	49.3	No	Yes	Yes	Yes	Yes
<b>Colombia</b>	<b>48.3</b>	<b>No</b>	<b>No</b>	<b>No</b>	Yes	Yes
Total population of countries in which gains to the "middle class" (by each definition—80 <sup>th</sup> , 60 <sup>th</sup> , 40th) is excluded by an "extreme poverty" goal	4640	4572	4523	4193	0	541
Percent of the population of world's 20 largest developing country whose middle class (by various definitions) is excluded by an "extreme poverty" goal		98.54%	97.48%	90.37%	0%	11.66%
Source: <a href="https://www.cgdev.org/blog/development-agenda-without-developing-countries-politics-penurious-poverty-lines-part-i">https://www.cgdev.org/blog/development-agenda-without-developing-countries-politics-penurious-poverty-lines-part-i</a>						

First draft. For comment only.

Please to not circulate or cite without permission.

*I.C) The connection between kinky goals, like the MDGs, and “solutions” and “interventions” and RCTs*

During a visit to Tanzania in 2013 President Obama took time out of one hopes is a busy schedule to promote the Soccket which is a soccer ball with the feature that kicking it around charges a battery that allows the powering of a small light<sup>8</sup>. This was the brain child of Harvard students<sup>9</sup>. The marketing of this type of “solution” demonstrates the illegitimacy of “kinky” goals. According to IEA data from 2010, energy consumption in Tanzania was 78 kWh/year per person. A refrigerator in the USA uses 459 kWh/year. Put another way, US holiday decorative lights use more electricity than Tanzanians use for all purposes. No American would ever define their “problem” such that the Soccket was a “solution”—only by the world super-elites imposing “goals” for others they would never accept for themselves can they imagine themselves the heroes by bringing “solutions.” Mitigating the consequences of the lack of a solution is not a solution.

The adoption of the MDGs in 2000 had an important distorting effect on development discussions. Since the beginning of “development” there had been a healthy debate about exactly which measures of human well-being to use and a health debate about whether the instrumentally best way to pursue those goals for human well-being was “systemic” (national development (with the attendant debates about how to pursue those objectives) or “programmatic.” The kinky development approach of the MDGs created the intellectual environment in which the *randomista* movement was even a possibility.

Imagine that, as I have argued it should have (Pritchett 2006), the World Bank adopted a definition of “poverty” that reflected the ownership share weighted poverty lines of its member countries (about \$13/day to \$30/day) rather than the penurious poverty line it actually adopted based on the *lowest* poverty lines of any of its members of “dollar a day”<sup>10</sup>? With a high poverty line the idea of searching for “rigorous evidence” about the causal impact of “interventions” or “programs” on poverty in order to allocate available resources to best reduce poverty is just obviously ridiculous. After all, in many countries average per capita consumption is well below any plausible upper-bar poverty line so “redistributive” or “targeted” interventions can obviously make no dent in headcount poverty at those poverty lines<sup>11</sup>.

---

<sup>8</sup> This paragraph draws on this recent article: <https://nonprofitchronicles.com/2018/12/09/todd-moss-when-it-comes-to-energy-big-is-beautiful/>

<sup>9</sup> Appendix 1 reproduces an email from the late John Briscoe to Drew Faust pointing out that Harvard was deliberately encouraging this view of Harvard students contributing to condescending “solutions.”

<sup>10</sup> Debates about who is “middle class”—which one could define as “those not poor”—often focus on lower lines for the middle class of around 10 dollars/day (e.g. Birdsall 2010, Pew 2015.).

<sup>11</sup> Of course with an FGT( $\alpha$ ) poverty measure with  $\alpha > 0$  those further below the poverty line count more for the poverty measure than those near it (and increasing so as  $\alpha$  is larger)—but there is at least a trade-off as reduction in incomes of all those below the poverty line count for something, as opposed to low-bar lines where reductions in the consumption of all those above the line count for exactly zero.

First draft. For comment only.

Please to not circulate or cite without permission.

A good example of the connection of the kinky development/MDG and RCT connection is the literature on conditional cash transfers (CCTs). There is a myth that the spread of CCTs is a “success” case of RCTs as there was an RCT that shows “CCTs work” and therefore based on this “rigorous” evidence other countries adopted the CCT approach<sup>12</sup>. However, every element of this story is dicey.

First, any trained economist’s instinct is that adding a condition to a cash transfer is either (i) irrelevant as the condition is infra-marginal (e.g. “food stamps” less than total optimal budget for food the transfer inclusive income) or (ii) welfare *reducing* to the recipient (e.g. if it has an impact because the recipient was at a “corner solution”)<sup>13</sup>. So any empirical evidence that the condition *per se* casually increased, say, child school attendance, is consistent with the interpretation the condition was *welfare worsening* for the recipients (including perhaps the student) over the unconditional transfer. It is at least possible the household and student was making the optimal consumption/schooling decision and the condition in the CCT pushed them off their welfare optimizing allocation to lower self-assessed well-being. It is only if one adopts a “poverty” plus “years of schooling” as objectives (which again, are not anyone’s actual welfare function) that a targeted cash transfer with conditions on the recipient doing X (e.g. forcing child into school) that empirically “works” to raise schooling relative to the no condition counterfactual can be *automatically* associated with “CCT (adding the C (condition) to the CT is better than plain CT.” Since the evidence from Mexico was that years of schooling went up (and not really by very much) but there was no evidence learning increased (perhaps because children less likely to learn because they were in terrible schools had, optimally, given their unfortunate choice set, dropped out?) it was hard to claim this was terrific evidence the condition in the CCT was welfare enhancing to the recipients.

Second, the PROGRESA program in Mexico was an attempt to consolidate *existing* expenditures in weakly targeted, poorly executed, “in-kind”, politicized programs into a better designed, better targeted, better implemented scheme that gave cash. Given that one political advantage of in-kind programs—like providing food—is the impression (false, of course, with full fungibility) that the in-kind transfer cannot be ill-spent by the household (e.g. men drinking it up). Hence, the addition of the “conditional” to the consolidating cash transfer was essentially a political posture to both (a) deflect the concern that, relative to in-kind, money would be ill-used and (b) create conditions that were about “building human capital” that has positive resonance with tax-paying non-recipients<sup>14</sup>. In this respect the “C” in CCT was a welfare

---

<sup>12</sup> As one example: “Due in large part to the high-quality evaluation of Mexico’s PROGRESA, CCT programs have become common in Latin America and are beginning to spread to other parts of the world.” (Baird, McIntosh, Ozler 2010).

<sup>13</sup> [Hammer \(2017\)](#) makes this point about normative evaluation.

<sup>14</sup> That the conditions were, at least in part, political signaling is evidenced by the very, very weak enforcement in the Brazilian CCT. The implementers of the program wanted to do the CT (cash transfer)—they did not want to actually enforce the C conditions in ways that eligible HHs didn’t receive the CT—and really only needed the C as political rhetoric.

First draft. For comment only.

Please to not circulate or cite without permission.

*reducing* aspect of program design that was undertaken due to *political* concerns for support for CCT, not technical considerations that CCT was better than CT for recipients.

Similarly, the “independent” and rigorous evaluation of the program was partly an effort to learn program impact but also a tool with two political objectives. One, by doing an evaluation one has “independent” academics actually seeing program implementation so in a case where a central authority was strict targeting they fear will be undermined or politicized at the local level the existence of data collection and data access by independent evaluators is a useful device. Two, the evaluation is there to convince everyone the implementation. Second, showing that the C in the CT “worked” to increase “human capital” was a *political* defense mechanism, not an interesting research finding.

Third, when PROGRESA was adopted Mexico was a relatively high income country with relatively high income inequality with a relatively strong tax base. These are conditions in which a transfer scheme could potentially be an optimal. However, it is not at all evident that cash transfer programs (conditional or not) are welfare improving when overall incomes are low, inequality is low and the direct and indirect costs of taxation are high and hence public funds have high opportunity cost.

The very strong role of World Bank funding in diffusing CCTs may suggest that, rather than persuading domestic citizens and policy makers on the basis of evidence, the adoption was driven by the preferences and needs of organizations like the World Bank who had, for their own political reasons, adopted “low-bar poverty reduction” as its agenda. In this case the “evidence” from Mexico was (a) never compelling about any broader measure of well-being and impact evaluations of CCTs have focused on specific, low-bar, indicators and (b) never “rigorous” about being an attractive policy in countries in which it was adopted through international isomorphism.

*II) National Development Does Bring Human Development, Strongly and Kinky Development Goals, Completely*

*National development* is a transformation of an *intrinsically social* grouping (country or region or society) to higher levels of capabilities in multiple dimensions: economic, political, administrative, and social. Crudely put, national development is about countries like Haiti or India or Bolivia or Indonesia “getting to Denmark”—achieving the high levels of economic, political, administrative, and social *functional* capabilities that Denmark possesses<sup>15</sup>.

For many (not all) people, national development is a means to the end of improved levels of human well-being, not an end, in and of itself. The strength of the relationship between national

---

<sup>15</sup> Pritchett and Woolcock (2003) introduced the metaphor of “getting to Denmark” as a measure of “development” with the immediate proviso: “By ‘Denmark’ we do not, of course, mean Denmark. Rather, we mean the common core of the structure of the workings of the public sector in countries usually called ‘developed’ (including new arrivals like Singapore).”

First draft. For comment only.

Please to not circulate or cite without permission.

development and measures of human well-being is an empirical question which we show is extraordinarily high for omnibus measures of typical well-being. Moreover, national development eliminates “low-bar” deprivation. We show these empirical findings have simple theoretical grounding. I also show that observed poverty reduction has been driven almost exclusively by national development, both in specific instances and across countries, with “programs” or “interventions” playing little or no role—and this, strangely, happened in precisely the period in which a small group of economists focused on the least important determinants of poverty reduction.

The idea of “development” and its self-conscious promotion as an objective by national and international agencies and organizations grew out of the post-World War II global system and the end of colonialization. As former colonies became new sovereign states in the aftermath of the War: Philippine independence was agreed to July 4, 1946, Indonesia declared independence from the Netherlands August 17, 1945 (recognized by the Dutch in 1949), India and Pakistan became independent states August 15, 1947. These were just the beginning and by the early 1960s (formal, official) colonialism was largely over. While obviously there were “developing” countries who were never colonized (e.g. Ethiopia, Thailand), many countries were independent since the 1800s (nearly all of Latin America), or become independent states earlier (e.g. Iraq in 1932) the post WWII geostrategic order and the proliferation of sovereigns of de-colonialization were the context in which ideas and organizations doing “development” emerged.

The very word “development” implies a change over time in which something becomes a better, more mature, more advanced version of what it has the ontological nature to become. A human develops from zygote to mature adult, a frog from zygote to tadpole to frog. Rocks neither “develop” to become frogs nor, through erosion, to become sand as the first is impossible and the latter not directional. What is it that “develops” with “development”?

#### II.A) *National Development as a four-fold transformation of countries*

By one definition of development what “develops” is intrinsically a socially constructed aggregate—usually defined by political boundaries, typically a country<sup>16</sup> (though one can ask if states/provinces/regions within countries are more or less developed, e.g. Tamil Nadu versus Bihar, Italy’s South versus North). I will use exclusively country examples from here on, but this is for convenience, not as definition.

A country has (at least) four important dimensions that have been of concern since the beginning of development.

---

<sup>16</sup> One needs to be careful here as while “nation” or “nation-state” are often used casually as synonyms for “country” this language brings in massive ideological baggage about what a “nation” is and its relationship to sovereign states recognized as “countries.”

First draft. For comment only.

Please to not circulate or cite without permission.

*Economic development.* This is usually understood as the productive capability of a place. A country's labor productivity, as measured by GDP per worker, is one possible indicator of economic development, though there can be many others (e.g. like Hausmann et al measures of economic complexity), and GDP can be adjusted in many ways (e.g. green accounting). These measures are *not not not* measures of human well-being (though they might be related to such causally), they are the measure of the productivity of a politically defined geographic space.

*Administrative development.* Measured in one way or another as the capability of (mostly state) organizations to accomplish public purposes<sup>17</sup>. Countries have organizations: armies, central banks, post offices, police forces, courts, land registries: an entire range of organizations--ministries/agencies/bureaus--with purposes. One can, at least in principle, order the capability of these organizations for accomplishing their public purpose from "weak" to "strong"—for instance, Chong et al ( ) rank post offices. While there is of course variation within countries in the capability of organizations (Kaufmann et al ), one measure of development is some aggregation of the administrative capability of the state—what is the range of things it could do, if it were so minded? The Fragile States Index, as one example, ranks countries from 0 (best) to 10 (worst, most fragile) on their broad based provision of public services and Denmark is .9, Indonesia 5.6 and Haiti 9.4.

*Political development.* This is obviously hugely value laden and, like anything said about politics, political, but *descriptively* when people described the "development" of states they usually had in mind some notion that those in political power and exercising sovereign power in a country: (a) were responsive to the needs, wishes, wants, desires of the citizens of the country and that political processes allowed those to be expressed by citizens and aggregated in fair and legitimate ways and (b) respected at least some set of "negative" rights that preserved liberty and security of the person (and perhaps in addition some "positive" rights). I am avoiding the word "democracy" although this obviously was used frequently as a short-hand for (a) but one view was that something like "democracy" was itself a "right" of the type (b) and another view was that one had in mind the goals of (a) and made a means-to-end argument that politically processes that could, broadly speaking, be taken to instantiate "democracy" achieved (a). I am not claiming people were correct in their beliefs, but I do claim that most people working in and around believed that states with more of (a) and (b) were more "developed" states than those with less. The Fragile States Index, for instance, has two distinct measures, one for "state legitimacy" (not "democracy") and one for "human rights and rule of law" and for State Legitimacy Haiti is 8.7, Indonesia 4.8 and Denmark .9 while for Human Rights and Rule of Law Haiti is 7.4, Indonesia is 7.3 and Denmark is 1.2. The Polity2 measure on a 10 is democracy, -10

---

<sup>17</sup> In our work *Building State Capability* we distinguish between the *capability* of organizations, which is a feature of an organization, and *capacity* as a feature of individuals and point out that capability or an organization is not the aggregation of the capacity of the individuals. This is to emphasize there are two distinct concepts, but we acknowledge one could just as well use the words exchanged (e.g. capacity as a feature of organizations) and, as long as one were consistent about distinguishing the two concepts, achieve the same goal.

First draft. For comment only.

Please to not circulate or cite without permission.



is complete autocracy of the POLITY IV project has been 10 for Denmark since 1915 (with the interregnum of WWII), in Indonesia was -5 in 1998 (last year of Suharto), jumped to 6 in 1999 and was 9 by 2017, in Haiti this was 0 2010 to 2015 and 5 in 2016 and 2017.

*Social Development.* Even more value laden and hence, if anything, more political than political development is the notion of how citizens/members of a common society treat each other changes as an intrinsic part of development. While these ideas were flawed in many ways (and in many ways reprehensible projections of social constructs of colonialists and colonialism) there was an important notion that “social equality”--in the sense that people were treated by other people equally independent of their social identities (kin, hereditary class, clan, tribe, ethnicity, race, sex, religion)--was, in and of itself, part of development. These are obviously historically constructed values of the Western experience and do not have universal validity, but I would argue were often bundled into notions of “modernization” and “development” for good or ill. Today of course this is most obvious in the views that development needs to be gendered and that societies that do not treat the sexes equally are less “developed” at least in one important sense, than those that do.

Each of these four dimensions of a country’s development are *intrinsically* and *ontologically* social—there are important elements of each of these aspects of development that cannot be meaningfully individuated. Two examples.

There is a huge literature showing that country level “A” (the labor productivity residual) plays a large role in explaining cross-national differences GDP per capita. Who owns “A”? A is a spatial “public good” and hence is a property of a place. When (low skill) people move to places with high A their income converges to that of similarly skilled natives quickly and nearly completely (Clemens, Montenegro, and Pritchett, forthcoming, Clemens and Pritchett, forthcoming).

Similarly, there is a very large literature about the role of “institutions” and their impact on economic productivity and other outcomes. But no one person “has” an institution, institutions, by any definition, such as “roles of the game” or “formal and informal norms” are fundamentally a social process, not individuated.

Our tendency to anthropomorphize is powerful but “the market” does not have a hand, invisible or otherwise<sup>18</sup>, governments or organizations or institutions or systems are not people (or even unitary actors) nor simple aggregations of the people in them, they are their own phenomena.

## *II.B) National Development and Measures of “Social Progress”<sup>19</sup>*

---

<sup>18</sup> Jokes like that of James Carville that we wanted to be reincarnated as the bond market because then he could intimidate anyone, shows the ubiquity of anthropomorphizing of complex social phenomena.

<sup>19</sup> This section draws on the blog post: “Turns out national development really does bring development.”

First draft. For comment only.

Please to not circulate or cite without permission.

The [Social Progress Index](#) is an effort of the Social Progress Imperative to create a new and better way to compare the human and social development performance of countries. An explicit part of their agenda is to *not* use GDP per capita or other measures of national development, but rather focus on direct measures of human well-being. As this index is similar to many others, I will use it to explore the strength of the connection between “national development” and measures of “human development.”

The Social Progress Index (SPI) has three components called: 1) basic human needs, 2) foundations of wellbeing, and 3) opportunity. Each of these three components are built of four sub-indicators which are each themselves built up from raw data. For instance, 1) basic human needs has four sub-components: 1.1 nutrition and basic medical care, 1.2 water and sanitation, 1.3 shelter, 1.4 personal safety. In turn, each of this is based on specific indicators, for instance, sub-component 1.2 water and sanitation is based on data on: 1.2.a access to piped water, 1.2.b rural access to improved water source, and 1.2.c access to improved sanitation.

The SPI is a thoughtful and careful attempt to provide a measure of social progress based on 53 distinct indicators—which include economic, education and health indicators but also religious tolerance, freedom from crime, political rights.

Empirically, the Social Progress Index (SPI) is almost perfectly correlated with national development. If one regresses the Social Progress Index on (ln) GDP per capita (productive economy), the POLITY2 measure of autocracy/democracy (responsive polity) and World Governance Indicator of Government Effectiveness (capable administration) the correlation across 128 countries of the actual values of SPI with the predicted values is .961 (an R-Squared of .924). This is an amazingly strong association. Different cross national measures of the *same thing* from different sources or methods—like “years of schooling of the adult population” or “child mortality”—don’t have correlations as high as .96, just due to pure measurement error<sup>20</sup>.

Figure 1 shows the scatter-plot of the actual values of the Social Progress Index and the predicted values. The boxes divide the two SPI and National Development Index (predicted SPI) into terciles and emphasize the white space as white space in this graph means something—it is the combination of country experiences in social progress-national development that just don’t happen. The upper left is white space: no country in the bottom third of national development is anywhere near the top third of social progress. The lower right is white space because there is no

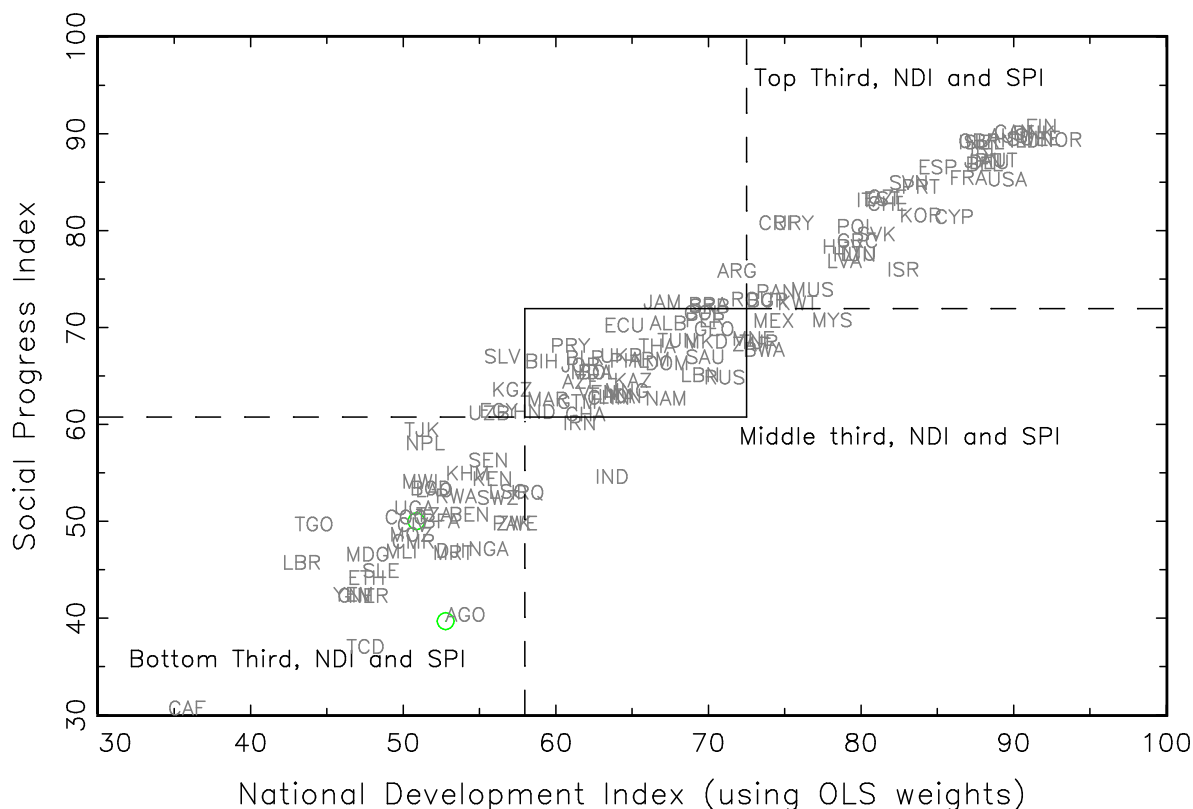
---

<sup>20</sup> Two points about the goodness of fit. First, the fit is amazingly good given how much measurement error people think there is in measuring concepts like “social progress” or “national development.” Suppose the true relationship was that national development and social progress were *perfectly* correlated and hence national development predicted social progress *exactly*, but that each was a concept measured relative to the true concept with a small amount of measurement error—a noise to noise plus signal of only 4 percent. Then the *observed* R-Square would be .924 just from measurement error alone. Second, I did all this empirical work in a few hours and hence have had no time to data mine or functional form mine or goodness of fit mine in any way. With time and effort I could crank the R-Squared (a bit) higher but that this is just the “slap in the standard data available on my computer into a standard regression” goodness of fit makes .924 truly amazing.

First draft. For comment only.  
Please to not circulate or cite without permission.

country in the top third of national development anywhere near the bottom third of social progress. Pretty much all the countries in the middle third by national development are at or near the middle third on social progress (with India (IND) as a striking low performer for its national development). The historically “developed” countries (the pre-expansion OECD) are very tightly clustered in the upper right

Figure 1: The tight relationship of Social Progress Index and National Development



Not just the overall Social Progress Index, but each of the three components (Basic Needs, Foundations of Well-Being, and Opportunity) and, with one exception<sup>21</sup>, all of the 12 sub-components are strongly associated with national development. Since both the indicators and the three national development indicators are all scaled to 0 to 100 the coefficients are comparable across indicators and across national development indicators. Note I am not making an “economic fundamentalist” or “neo-liberal” argument that this association is all economic growth. For some sub-components, like Personal Safety, the level of GDPPC isn’t that important but Administrative Capability is. For other sub-components, like Personal Rights or Environmental Quality, GDPPC isn’t important but Administrative Capability and Political

<sup>21</sup> It seems strange that “Health and Wellness” is the sub-component for which national development isn’t important, but this indicator includes “obesity” which is negatively correlated with GDPPC and “suicide rate” which has a very low correlation.

First draft. For comment only.

Please to not circulate or cite without permission.

Responsiveness are both important. On the other hand, for some sub-components, like Water and Sanitation, Shelter, or Access to Basic Knowledge essentially only GDPPC matters (and GDPPC alone (in a simple linear functional form) has very high explanatory power).

Table 1: The Social Progress Index—and all of its components and sub-components—are strongly associated with three indicators of national development							
Social Progress Indicator, its three components (Basic Human Needs, Foundations of Well Being, Opportunity) and the four sub-components of each component	Economic Productivity ((ln) GDP per capita, PWT8.0, rescaled 0 to 100)		Administrative Capability (World Governance Indicators, Government Effectiveness, rescaled 0 to 100)		Political Responsiveness (Polity IV Project, Polity 2, rescaled 0 to 100)		R-Squared of regression on national development indicators
	OLS coeff.	t-stat.	OLS coeff.	t-stat	OLS coeff.	t-stat	
<b>Social Progress Index</b>	0.50	12.26	0.39	7.96	0.13	5.41	<b>0.924</b>
<b>Basic Human Needs</b>	0.69	11.06	0.26	3.48	0.00	0.05	<b>0.817</b>
Nutrition and Basic Medical Care	0.70	8.92	0.00	0.00	0.05	1.17	0.646
Water and Sanitation	1.09	10.45	0.13	1.27	0.08	1.70	0.756
Shelter	0.96	11.02	0.16	1.86	0.02	0.42	0.780
Personal Safety	0.06	0.62	0.69	7.28	-0.08	-1.73	0.587
<b>Foundations of Well-Being</b>	0.56	8.46	0.38	5.81	0.18	5.90	<b>0.856</b>
Access to Basic Knowledge	0.94	9.36	-0.03	-0.31	0.08	1.80	0.672
Access to Info and Comm.	0.56	8.54	0.24	3.80	0.19	6.39	0.811
Health and Wellness	-0.20	-1.77	0.49	4.56	0.02	0.45	0.247
Environmental Quality	0.13	1.69	0.51	6.63	0.17	4.62	0.701
<b>Opportunity</b>	0.29	4.68	0.53	8.91	0.25	8.67	<b>0.870</b>
Personal Rights	-0.08	-0.86	0.53	6.02	0.54	12.98	0.774
Personal Freedom and Choice	0.12	1.52	0.70	9.43	0.00	0.07	0.750
Tolerance and Inclusion	0.21	1.98	0.39	3.81	0.15	2.91	0.516
Access to Advanced Ed.	0.88	10.17	0.23	2.76	0.09	2.30	0.802

Source: Author's calculations.

The Fund for Peace presents a Fragile States Index that has multiple components. Two of those, “Public Services”<sup>22</sup> and “State Legitimacy”<sup>23</sup> are potential alternative empirical proxies for

<sup>22</sup> The description of this indicator is: *The Public Services Indicator refers to the presence of basic state functions that serve the people. On the one hand, this may include the provision of essential services, such as health, education, water and sanitation, transport infrastructure, electricity and power, and internet and connectivity. On the other hand, it may include the state's ability to protect its citizens, such as from terrorism and violence, through perceived effective policing.*

First draft. For comment only.

Please to not circulate or cite without permission.

the concepts of “administrative capability” and “political responsiveness.” A regression the overall Social Progress Index on GDP per capita, FSI: Public Services and FSI: State Legitimacy (all scaled to 100) the R-Squared is .947 with all three indicators having powerful roles<sup>24</sup>.

Similar strong associations of measures of human development and national development are produced using a variety of other indicators of either human development or national development. For instance, the Human Development Index (HDI) was developed by the UNDP, based loosely on Sen’s ideas of “capabilities,” including measures of basic education and health. This was intended as a counter-weight to the overemphasis on GDP per capita as a development indicator. However, by 2010 due to convergence in the education and health indicators (and the tight relationship of those to GDPPC) the HDI was almost perfectly correlated with GDP per capita alone.

One could move to subjective indicators of overall well-being. If one uses the Cantril “ladder of life” measure of average subjective well-being the R-squared of the three national development indicators is .66 (with the main explanatory factors being GDPPC and Administrative Capability). While this is lower than for the Social Progress Index, it is worth noting that the association of life satisfaction and the SPI is almost exactly the same: .67—so the two omnibus measures of “human development” have the same cross-national correlation as the “life ladder” measure of subjective well-being and national development.

The World Happiness Report has developed another index of human well-being based on the empirical relationship of seven factors (like “perceptions of corruption”, “Healthy life expectancy”, “social support” and measures of affect) to the “ladder of life” measure of subjective well-being. One of those seven elements (as above) is GDP per capita and so we obviously cannot correlate this overall index to national development. But when I construct an equally weighted index of those other six elements of their happiness index and regress that the three indicators of national development in their data (which is GDP per capita, and two indicators of “delivery quality” (administrative capability) and “Democratic quality” (political responsiveness) which are based on the indicators in the World Governance Indicators) the R2 across 120 countries is .788. Again, the correlation between this six element index and the directly observed ladder of life measure is .81. The three indicators of national development is about as tightly correlated with omnibus measures of human well-being as alternative measures of human well-being are correlated amongst themselves.

---

<sup>23</sup> The first sentence of the description is: *The State Legitimacy Indicator considers the representativeness and openness of government and its relationship with its citizenry.*

<sup>24</sup> As footnote 5 suggests, tinkering about with other variables would likely produce some even higher R2 results.

My conjecture is that country's average level of pretty much any human well-being is going to be related to measures of national development and will be more strongly related the more comprehensive the measure (and the more accurately measured). As we see, comprehensive measures, like the Social Progress Index, are incredibly tightly associated with national development—and the various human well-being measures are as tightly associated with national development as they are amongst themselves. The motivating idea of development—that national development would lead to higher levels of human development—in many, many, dimensions of human well-being—was not wrong. National development really does produce human development.

### *II.C) National Development and Kinky Goals, Low-bar poverty, among others*

There is asymmetry in the potential for improving human well-being between national development and targeted interventions aimed at kinky development (Kenny and Pritchett 2013). National development does everything that the targeted interventions are intended to do, and more. That is, national development is *sufficient* for achieving low-bar goals *and* national development promotes improves broad based human well-being (as we showed above). In contrast, the targeted interventions *at best* provide *some* progress on the kinky goals and *by design* provide no help at all on broad based human well-being.

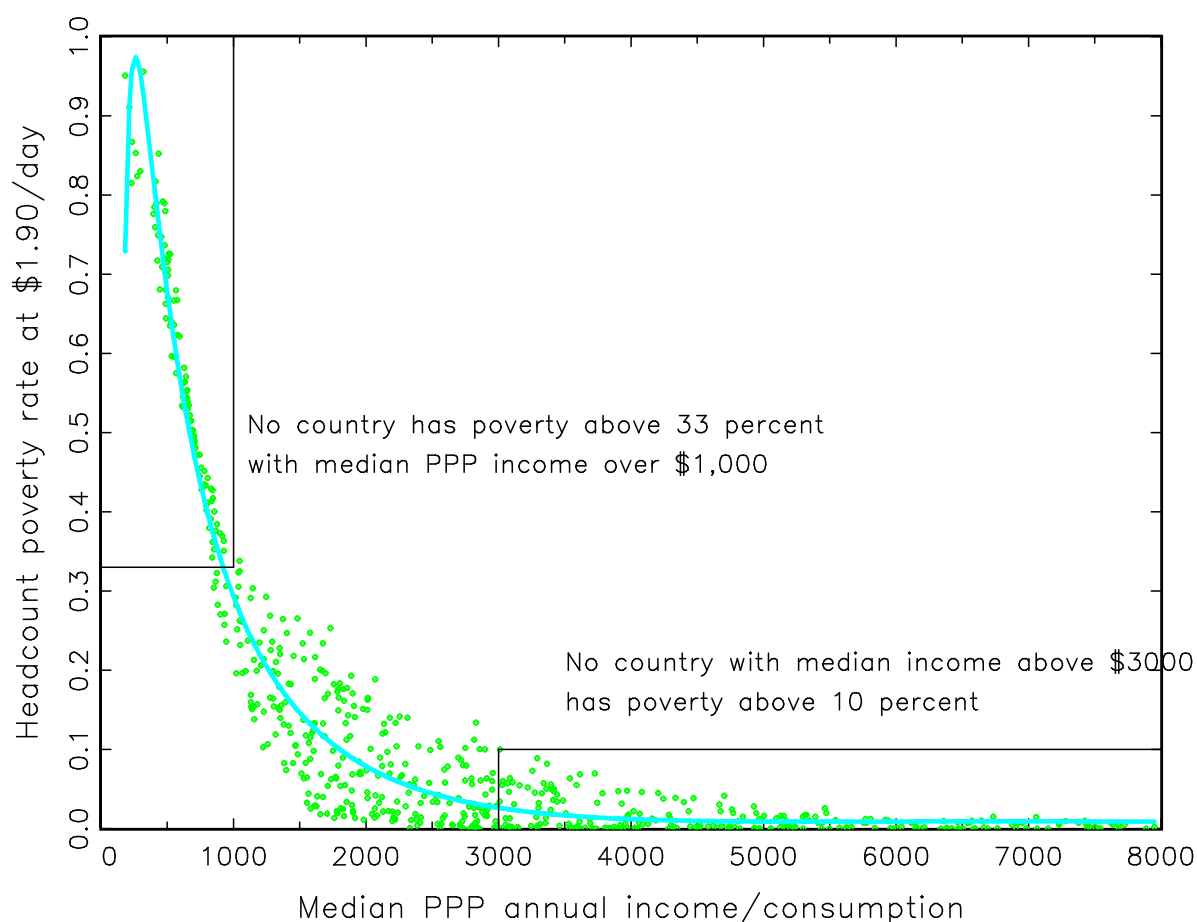
*Income/consumption poverty.* I start with the standard World Bank data on X “dollar a day” measures of poverty and show that inclusive economic growth (increases in median income/consumption) is necessary and sufficient to eliminate poverty. Moreover, as one moves to higher, and hence more reasonable, poverty lines the role of growth becomes *everything*—across the 810 country/year measures of poverty at \$5.20 a day the correlation with median country/year income is .98. And, as Aart Kraay (2006) has demonstrated this cross-sectional relationship holds across (long) country spells (to a very large extent this section is just an update of one part of Kraay's work).

National development, and in this case, just economic growth, essentially eliminates “low bar” or “dollar a day” (now, with inflation, P\$1.90 a day), even at quite modest levels of income. Using data from the World Bank and all country/year pairs with actual survey data one has over 800 country/year observations on measured poverty rates and on median income/consumption in the data source used to calculate that poverty rate. No country with median annual income above P\$3,000 (about the level of Peru or Mongolia around 2010) has low-bar poverty more than 10 percent. By P\$5,000 (about the level of Costa Rica) essentially no country has low-bar poverty above 2 percent. Also, no country with median income above 1,000 (about the level of Bangladesh in 2010) has low-bar poverty more than a third of their population. Increases in the income of the typical (median) person is *sufficient* to drive low-bar poverty.

First draft. For comment only.

Please to not circulate or cite without permission.

**Figure 3: The level of income/consumption is sufficient to eliminate high levels of “dollar a day” poverty and attain low levels of “dollar a day” poverty**



Source: Author’s calculations with data from [PovcalNet: the on-line tool for poverty measurement developed by the Development Research Group of the World Bank](#)

Moreover, *just* the level of income/consumption (more precisely, the median, a measure of the central tendency of the distribution of income) explains essentially all of the cross-national variation in headcount poverty rates. The R-Squared of a flexible functional form in median income/consumption is .964 for “dollar a day” poverty<sup>25</sup>. At higher poverty lines the fit gets even better and for poverty at P\$5.50 (which is still far below the poverty lines of OECD countries) the R-Squared is .986. If you want to know a country’s headcount rate of poverty *all* you need to make a *very* precise prediction is median income/consumption.

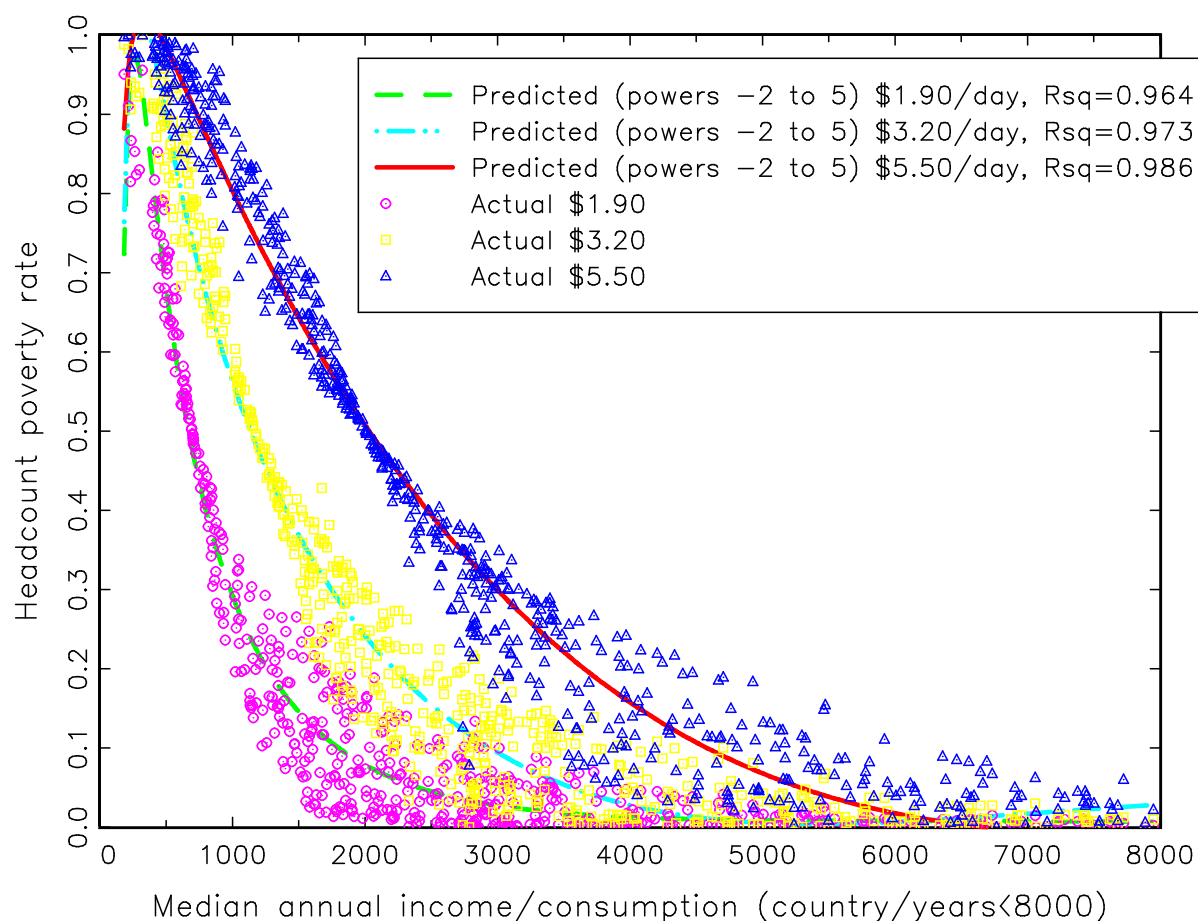
<sup>25</sup> I use the powers from -2 to 5 to allow maximum flexibility in the relationship. With 810 degrees of freedom the loss in “degrees of freedom” from this is inconsequential. Using a flexible functional form is not “data mining” as we have no reason to impose any particular functional form and the very nature of poverty measures as partial integrals generates non-linearity as low-bar poverty, by construction, has to fall to (essentially) zero as the distribution shifts right-ward.

First draft. For comment only.

Please to not circulate or cite without permission.

I am perfectly away that some of this tight fit is “baked in” as poverty measures and medians are just different summary statistics of the *same* distribution (I know how poverty is calculated, see Pradhan et al 2001). That is precisely my point, what the data are telling us is that unless the whole distribution shifts right (what one might call “inclusive growth”) poverty is unlikely to change as the other elements of the distribution that affect poverty (the distribution of the left tail) apparently do not differ by much across countries relative to the differences in the median. These “left tail” differences, of the type that targeted transfer programs would induce, play only a tiny role in variations actual poverty rates relative to the huge effects of differences in median income.

**Figure 4: The median of the distribution of income/consumption of a country predicts the level of poverty *exactly* for high poverty lines and near *exactly* even for low poverty lines**



Source: Author's calculations with data from [PovcalNet: the on-line tool for poverty measurement developed by the Development Research Group of the World Bank](#)

This super-tight correlations of measured poverty and median income/consumption also hold in changes over time within countries. Figure 5 compares the change in “dollar a day”

First draft. For comment only.

Please to not circulate or cite without permission.





countries over 60 on a 0 to 100 additive index of national development (about the level of Poland) this primary completion rate is over 95 percent.

Similarly with under-5 child mortality. 28 of the 30 countries with a national development index over 60 (the level of Costa Rica and Poland) had under-5 child mortality under 10 per 1000 (the two were Trinidad and Tobago (an oil rich country) and Mauritius). In contrast, only 2 of the 92 countries with that index less than 60 have child mortality below 10 (Malaysia and Cuba). Hence, national development isn't quite "necessary and sufficient" for achieving low child mortality—but pretty close.

Back to the opening anecdote. The percent of "people with at least basic water services" is higher than 97.5 percent for all countries with an NDI index above 60 (except, again TTO). National development is sufficient for reaching this low-bar water access goal. All of the 40 countries where less than 75 percent of the population had basic water services had national development indices less than 40. (Though in these case there are low national development countries with high access to basic water, so it is not necessary).

#### *II.D) "Theory"*

What is odd is that anyone would ever doubt a very tight empirical relationship between national development as the four-fold transformation and human well-being—and especially for low-bar goals around necessities. Four-fold national development is a well-being machine.

Obviously the most basic of economics shows that for private goods when the budget set of a household expands they have more choices and (even without strict assumptions on "optimizing") are more likely to achieve higher levels of well-being.

But what about "public goods" and "market failures"? What I mean conceptually by a "responsive polity" is a government to which and through which problems are nominated and what I mean by "capable administration" is the ability to effectively deploy state directed resources to solve problems (this could be through contracting with the private sector, regulating the private sector, or PPPs). A rich "laissez faire" society might have terrible environmental quality. But national development includes a responsive polity and a capable administration and hence with higher national development one would expect the government to act in ways that improve environmental quality. Indeed, just as an indicator (not "proof") one sees that the Social Progress Index of "environmental quality" was strongly correlated with national development—through capability and polity, not GDP per capita.

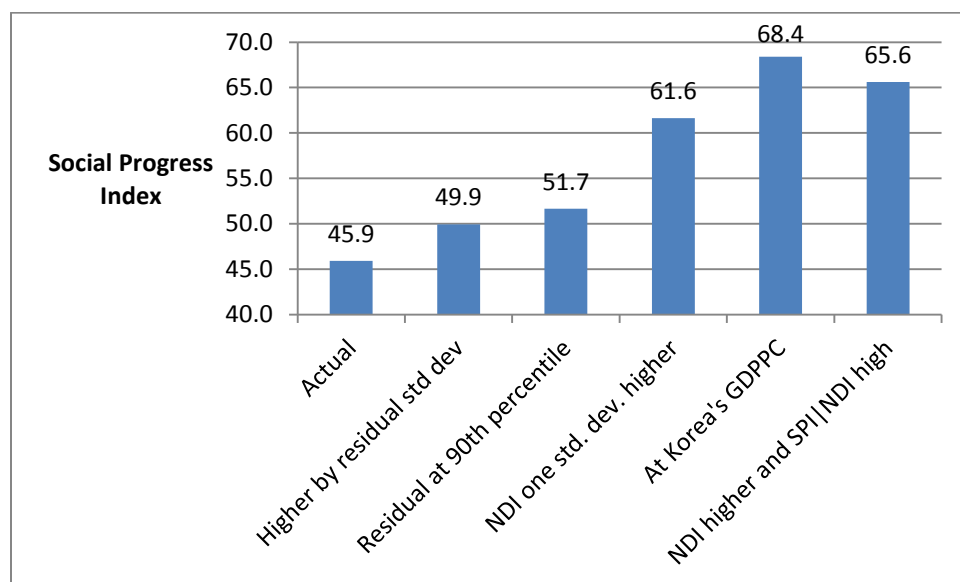
For similar reasons, one would expect the relationship with national development to be even stronger for “necessities” than for general well-being. An economist’s definition of “necessity” is something for which marginal utility gets very high as consumption of it falls and, related, something for which the price elasticity (especially at low levels) is expected to be very low. Hence, one would expect that the budget expansion path of a necessity would be stronger at lower levels as one’s efforts are focused on necessities and then with falling budget shares. Indeed, an Engel curve—that food share declines with aggregate income/consumption (roughly linear of the share against the log)—is the best documented fact in economics.

The strong and tight relationship implies that the range of the overall SPI around a given NDI is quite small relative to the range of SPI. A typical country in the bottom third of SPI, like Mozambique, has an actual and predicted SPI (conditional on NDI) of about 46. The regression residuals have a standard error of about 4 so that even if Madagascar were an over-performer and had SPI performance one standard error better than predicted for its NDI its SPI would be 50. That would move it ahead of Pakistan and Zimbabwe (at 49) to about the level of Tanzania or Uganda. Alternatively, the 90<sup>th</sup> percentile of the residuals is 5.75. So if Madagascar were among the 10 percent best SPI performers for its level of NDI it would be at 51.75, reaching Rwanda. But even as an “outlier” for its NDI Mozambique would still reach the SPI of Laos or Bangladesh.

In contrast, the predicted gains from national development progress are both powerful and unlimited on the upside. If Madagascar improved its GDP per capita, government effectiveness and polity scores by one standard deviation each, its predicted Social Progress would rise to 61.6—ahead of Egypt (60.74) and approaching Morocco (61.9) or China (62.1). Or, suppose Madagascar had experienced the growth performance of Korea and had Korea’s GDP per capita—but same administrative capability and political responsiveness, that (counter-factual) growth alone would have raised its predicted SPI to 68.4—ahead of Thailand or Turkey.

There is no reason why a country could not have both national development progress and improve its SPI relative to NDI at the same time (though this is a residual so this simple exercise alone would not tell us how to do that). Moving ahead in NDI by one standard deviation and having SPI higher for NDI, Madagascar reaches 65.6.

**Figure 6: Illustrating various scenarios for Madagascar (a lower third SPI country)**



Much of the “development” versus “targeted interventions” debate has been a barrel full of red herrings. There was a caricature that development practitioners (or even economists) believed that “development” meant exclusively GDP per capita, which is not, and has never been, true. By “human development” economists have always had in mind some measure of overall human well-being and by “national development” economists have always recognized the need for a strong state to address public goods (like law and order, the environment) and the need for a responsive polity and a social transformation to more equal treatment are embedded in nearly every vision of development.

### *III) The Strange (timing of the) Rise of the Randomistas*

One instructive question to examine is the timing of the rise of use of RCTs in the field of development. Figure 6 shows the rise of impact evaluation using randomization per year, rising from a baseline of next to nothing and then exploding after around 2000. Why?

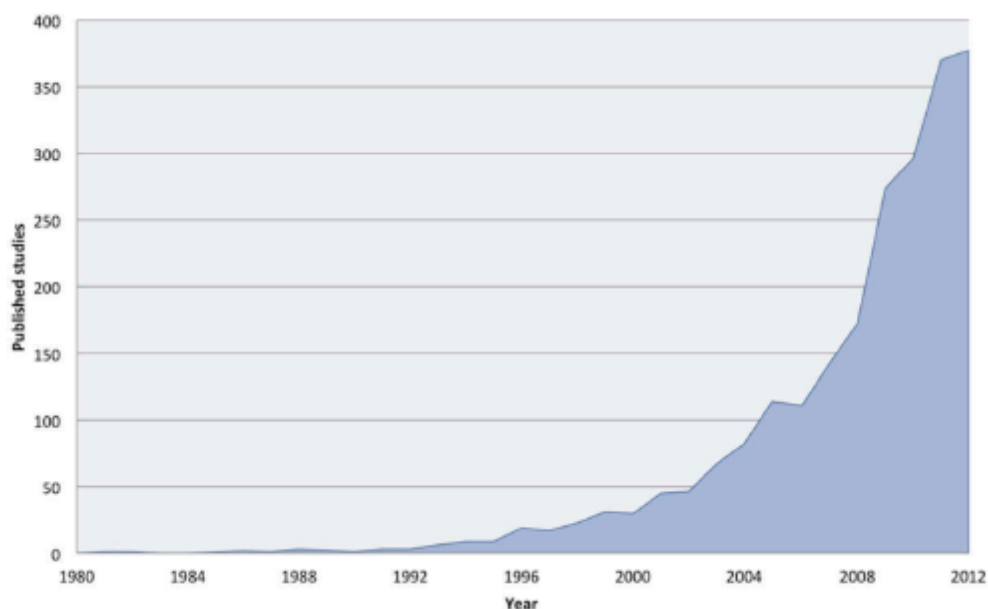
The timing of the rise cannot be explained as the result of a technological or scientific advance that made this possible—like say, the expansion of cell-phones or the uptake of a new vaccine. All of the techniques used in RCTs are very old and the practice of using RCTs to evaluate social policy is also old. Mathematica Policy Research was founded to do large scale experiments in social policy, and in 2018 is celebrating its 50<sup>th</sup> anniversary (older than many development randomistas). The US saw several large, sophisticated, RCT evaluations of social programs by the 1970s in a variety of fields: the income maintenance experiments (four experiments in seven different sites started between 1968 and 1971), the RAND health insurance experiment began in 1971, the Kansas City Preventive Patrol experiment in 1972. By the 1980s there were at least four large firms/organizations in the USA carrying out large, sophisticated

First draft. For comment only.

Please to not circulate or cite without permission.

RCTs on social policy issues (Mathematica Policy Research, RAND, Abt Associates, MDRC). And RCTs in developing country settings are not new. The Matlab experiment in Bangladesh started in 1977. Authors from the World Bank (Dean Jamison and Stephen Heynemann) published results of an experiment on radio instruction and textbooks from Nicaragua from the 1970s (Jamison et al 1981).

**Figure 7: The rise of the RCT based impact evaluation only really got underway post-2000**



*Source: Cameron, Mishra and Brown 2016.*

The rise of RCTs in the developing world is also not the spread of success from the USA (or other leading practitioners) to developing countries, if anything the puzzle is the opposite. That is, none of the *randomistas* had experience in the USA and were attempting to transmit that experience, nor did any of the large scale organizations actually doing RCTs in the USA play a key role. For a movement promoting the idea that using more rigorous evidence from RCTs will improve development outcomes, the deafening silence on lessons from the experience of RCTs in the USA, which was over 30 years old in 1998, is striking. This is likely so, at least in part, because it is not widely interpreted that this overall experience was a “success.” On the big picture level, no one I know of points to the USA and says: “look at the wonderful social policies in the USA compared to other countries (or compared to the USA counter-factual): that success is because of their commitment to using rigorous evidence from independent RCTs to inform policy design.” Even by 1986 there was a high profile conference to ask: “why aren’t the income maintenance RCTs [by far the large experiments in economics to date] influencing policy?” In the conference overview, Alicia Munnell (a prominent US policy economist) notes that the policy that the experiments were designed to evaluate, the negative income tax, was largely out

First draft. For comment only.

Please to not circulate or cite without permission.

of favor and that “Although the last of the four experiments ended in 1982, the major lessons of the experiments are neither widely known nor well understood.” (Munnell 1986). Heckman and Smith argued the use of the findings of the RCT that were used to cut the budget of one component of the Job Training and Partnership Act by 80 percent were likely misinterpreted and that the findings were very sensitive to a range of factors and that the caution that “experimental impact estimates, like those from nonexperimental analyses, require careful interpretation if they are to provide a reliable guide to policymakers.”

*III.A) 1950 to 2000 was the greatest 50 year period for human progress, ever, by a lot*

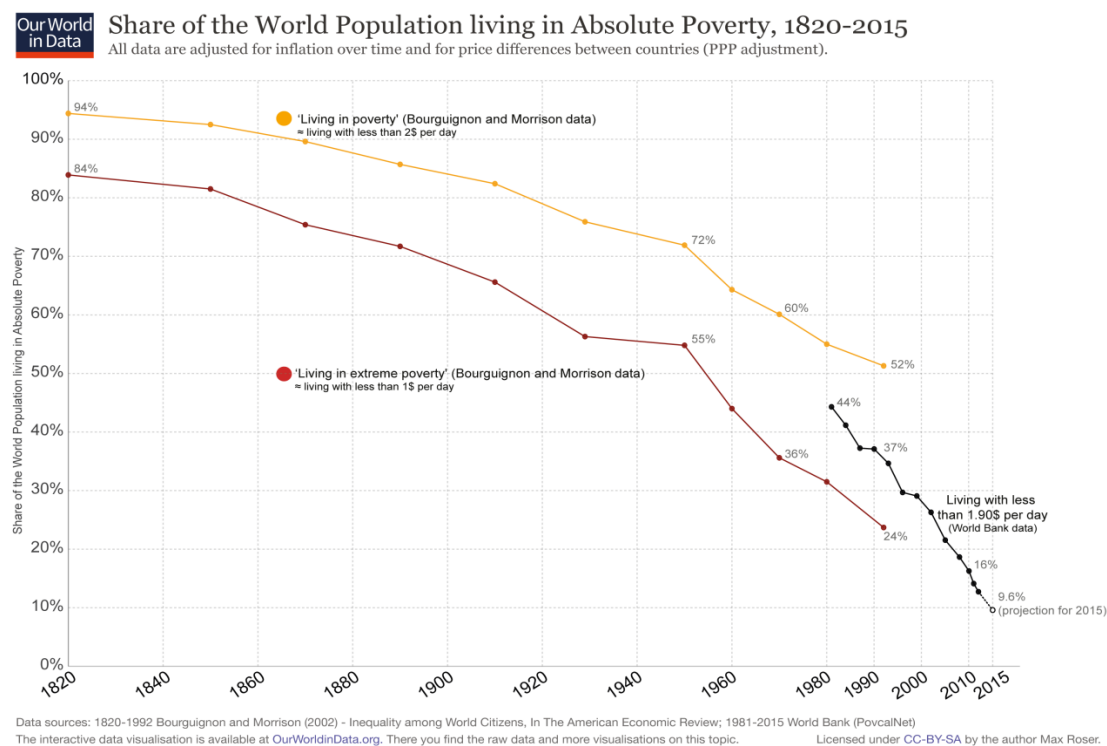
The Rosling’s *Factfulness* (2018) is a reminder that even before focusing on counterfactualness, knowing the facts can be helpful. In this section I will show several graphs taken from the *Our World in Data* as they have done the great service of compiling data from a variety of sources into wonderful graphics. The fact is that the *randomista* movement arose 50 years into a “development era” that had been the greatest 50 year period for improvement of (nearly all) measures of human well-being in the history of man—by a large margin. So the timing of the rise of RCTs cannot easily be explained as a response to “failure” in the development enterprise, for which “lack of rigorous evidence” could be an explanation.

*Poverty.* Figure 8 shows the evolution of “dollar a day” poverty over time. Since the highest this could ever be is 100 percent then the progress in all of human history up to 1820 was 16 percent of humanity escaping poverty. If we call this roughly 5000 years of the rise of large scale complex states (e.g. Egypt, Harappa, China) then this is poverty progress of .0032 percentage points per year. Between 1820 and 1950 absolute poverty fell from 84% to 55%, .22 percentage points per year. Between 1950 and 1992 it fell from 55% to 24% or .74 percentage points per year (three times faster than the pre-development era rate of decline). According to the World Bank data from 1981 to 1999 global extreme poverty in OWD it fell from 42.8 to 28.02 percent, a fall of .82 percentage points per year and then fell further to 10.67% percent in 2013.

Factually, the *randomista* movement arose in the period which the world was experiencing the most rapid reduction in extreme poverty rates in the history of man, by a factor of 3. The *Science* article referenced by Chris Blattman in the introduction, which in many ways could be claimed to be the first “rigorous” evidence for a successful anti-poverty program with any cross-national external validity was published in 2015, by which time global extreme poverty rate had fallen below 10 percent. Before there was any rigorous RCT evidence on how to reduce global poverty, global poverty had been massively reduced.

First draft. For comment only.  
Please to not circulate or cite without permission.

**Figure 8: Since 1950 global poverty rates had been in steep decline**



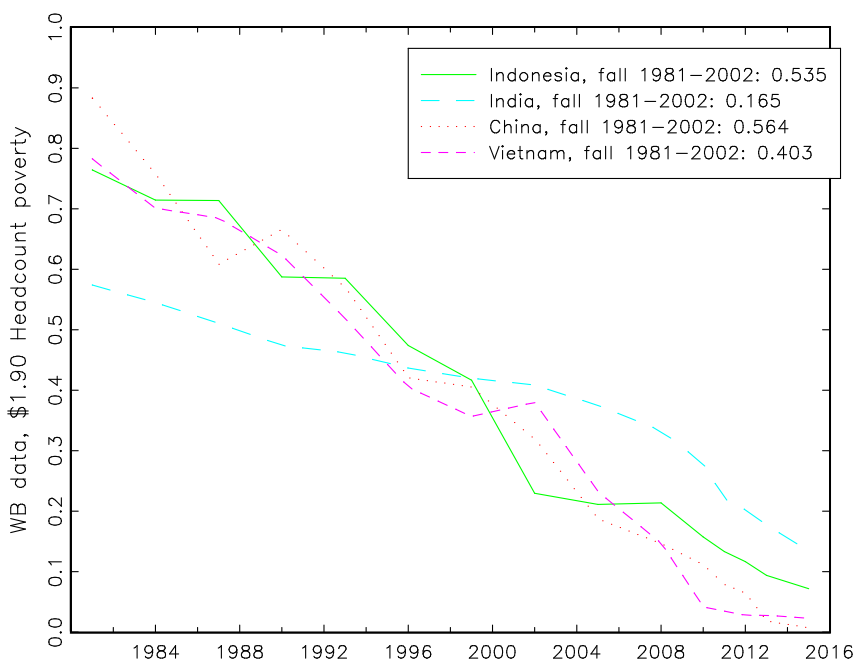
Around the year 2000 it was already obvious that the world was experiencing some of the most rapid reductions in rates of extreme poverty in the history of mankind. According to the World Bank's estimates between 1981 and 2002 (prior to RCTs) headcount poverty fell in Indonesia by 53 percentage points, in China by 56 percentage points, in Vietnam by 40 percentage points. By 2015 extreme poverty was virtually eliminated in Vietnam and China and down to 7.2 percent in Indonesia. It is widely accepted that little to none of this progress was due to direct, targeted programs (or improvements in the distribution of income)<sup>27</sup>.

<sup>27</sup> I try not to be ad hominem, but there is a kind of self-contained surreality to publishing a book in 2011 with the sub-title "How a new economics is helping to solve global poverty"—ignoring the fact that the "old economics"—a focus on promoting economic growth—had already solved the problem whenever it was practiced and economic growth was achieved and sustained. Extreme poverty fell, on average by 4 million people a month in East Asia and South Asia between 1990 and 2011. I would love to see the estimates for the total poverty reduction, ever, on any serious basis, from the application of the "new economics" to supposedly "solving" global poverty. For instance, the Graduation program costs \$4500 per household and let's say half of participants "leave poverty" and suppose there are 5 people per household then this is about \$1800 per person out of poverty. To produce a reduction of 4 million (one month's progress) at those costs it would take \$7.3 billion dollars.

First draft. For comment only.

Please to not circulate or cite without permission.

**Figure 9: In several countries the most rapid reductions in extreme poverty in history in several countries had been underway for 20 years by 2000**



Source: Author's calculations with data from [PovcalNet: the on-line tool for poverty measurement developed by the Development Research Group of the World Bank](#)

**Schooling.** Lee and Lee (2016) have a new data set with schooling information for 111 countries from 1870 to 2010. Figure 8 shows the evolution of the average years of schooling of the young cohort, aged 15 to 24, as this captures progress more clearly, for the population weighted average of 86 developing countries (without China)<sup>28</sup>. Schooling years are only 2.39 in 1950 and progress was .027 years of additional schooling per year. Then from 1950 to 2010 (a cohort who was 5 in 2000 and hence almost certainly unaffected by any learning from RCTs) average schooling rose to 8.32, increasing .099 years per year. Progress in the development era was 3.7 times faster than before (and orders of magnitude higher than human history up to that point). The fraction of a 15-24 cohort with no schooling fell from 54.2 percent to 10.3 percent. Before there was any rigorous RCT evidence on how to expand schooling (or evidence that CCTs “worked”), 9 out of 10 children had enrolled in school.

There now a “learning crisis” as many of the schooling systems that have effectively expanded are not doing a good job of actually conveying skills (Pritchett 2013, RISE 2015, World Bank 2018) and there is an ongoing shift to focus on that, but it is worth noting that, at the target that was set of universalizing schooling, there was amazing, impressive, progress.

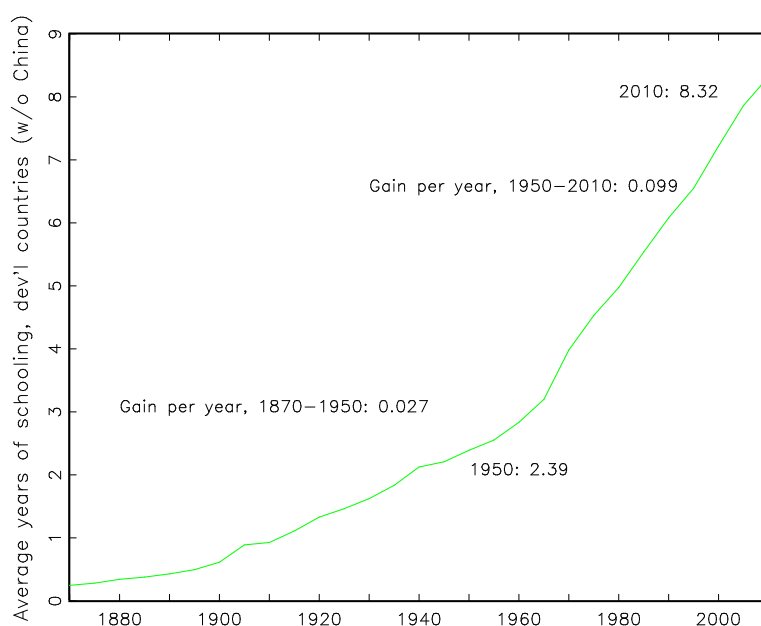
<sup>28</sup> This excludes those Lee and Lee (2016) list as “Advanced Economies. China is excluded because the reported years of school jump from .63 in 1935 to 3.88 in 1940 which seems implausible on the face of it (given that the Japanese invaded in 1937 and it was a pretty chaotic time). Given China’s population affects the whole aggregate and nothing really important rides on including China in this calculation.

First draft. For comment only.

Please to not circulate or cite without permission.



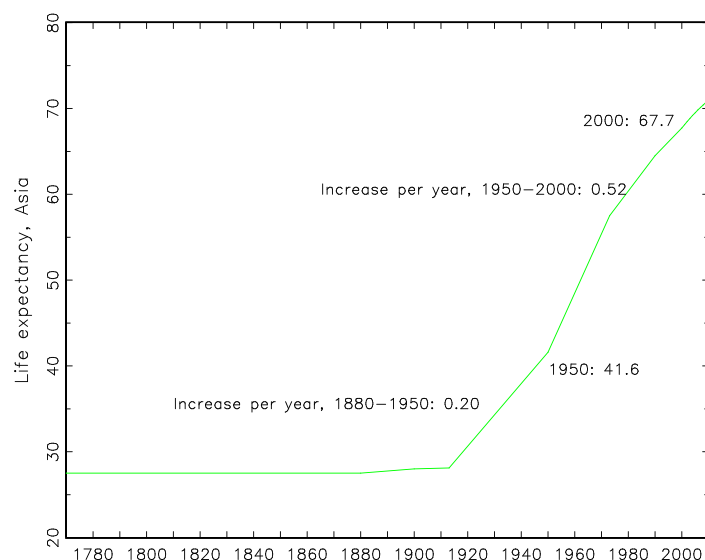
**Figure 10: Schooling expanded massively during the development era 1950-2010.**



*Source: Author's calculations with data from Lee and Lee 2016.*

*Life expectancy.* Health is, of course, an area in which the sciences, including the use of RCTs, have been deployed to improve medical treatments and health conditions more generally. Figure 11 also shows that life expectancy in Asia was improving very rapidly in the 1950 to 2000 period, far faster of course than in the period prior to 1870, but even faster than from 1880 to 1950. Again, life expectancy was increasing both globally, and a very rapid pace in Asia and every developing region (with a set back in Africa due to HIV/AIDS) from 1950 to 2000 before the onset of the *randomista* revolution.

**Figure 11: Rapid improvements in life expectancy in Asia in the development era**



*Source: Author's calculations with Our World in Data, from Riley (2005).*

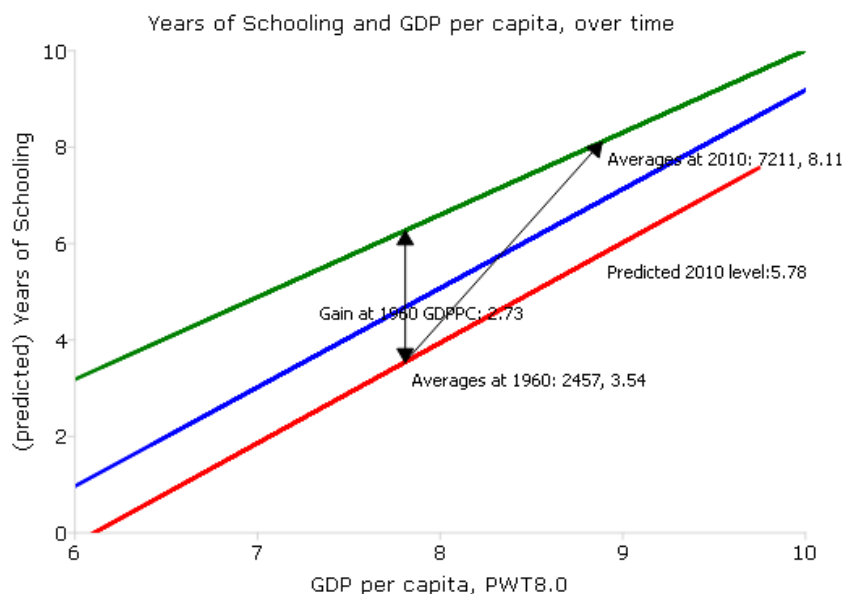
For both schooling and life expectancy it is worth noting that these indicators were rising *faster* than growth in GDP per capita would have predicted. That is, there is a rhetoric that the “development enterprise” put too much emphasis on growth and not enough on other aspects of well-being, like health and education (and other basic needs and the environment). But the fact is that life expectancy and years of schooling rose substantially *even for countries at the same level of GDP per capita*—the entire cross sectional relationship shifted *upwards* over time. This is well known for life expectancy and the upward shift in the Preston curve is well known and well-documented<sup>29</sup> and Figure 12 shows this is also true for years of schooling. A country who saw no progress in GDP per capita from 1960 to 2010 still had an average increase of 2.73 years of schooling. If the increase in income alone were responsible for the expansion moving along the 1960 gradient, schooling would have only increased to 5.78 versus the actual of 8.11. The world was even more successful at getting kids into school than progress in income alone can explain.

<sup>29</sup> Our World In Data has a nice [visual](#).

First draft. For comment only.

Please to not circulate or cite without permission.

**Figure 12: Schooling rose even faster than expected from progress in incomes**



Source: Pritchett, forthcoming

One of truly astounding things about *Factfulness* (2018) is just how completely and totally wrong about these trends people in the rich world on, as the Roslings put it, much worse than chimpanzees.

When asked as multiple choice, “In the last 20 years the proportion of the world population living in extreme poverty has... (a) almost doubled, (b) remained more or less the same, (c) almost halved”—random guessing would produce about 1/3 getting the right answer. Instead in France 4 percent got it right, in the USA 5 percent, and outside the Nordic countries (where there might already be a “Roslings” effect) the highest was 10 percent.

When asked “In all low income countries across the world today how many girls will finish primary school... (a) 20 percent, (b) 40 percent, (c) 60 percent.” Again, on average only 7 percent got this right (it is 60 percent) and the majority picked 20 percent, which are conditions that exist in countries that cover at most a few percent of the world’s population.

### *II.B) Interpreting the lack of demand for RCTs*

The upsurge in RCTs cannot, in the first instance, be interpreted as an uptick in demand from development organizations like the World Bank or other development organizations or, even less, from developing country governments. Interestingly, a large part of the rhetoric of the *randomistas* was devoted to criticizing development agencies and developing country governments precisely because of the lack of interest in funding RCTs.

First draft. For comment only.

Please to not circulate or cite without permission.

There are two ways of interpreting the lack of demand, both of which have strong elements of validity.

One interpretation is that, for the most part, those actually engaged in the development enterprise did not believe that RCTs brought value for money to the questions they thought were important bottlenecks to progress *and that they were right*. That is, those in development who believed that (a) improvement in broad based measures of human well-being were important and that (b) national development was the most important key constraint to that progress and that (c) sector wide reforms of organizations, institutions, and incentives were more important than the specific details of programmatic design didn't see any hope that RCTs would contribute significantly to improving performance. They saw that RCTs were methodologically (only) well-suited to a narrow range of development problems and those, by and large, were not the most important issues. Moreover, by 2000 experienced practitioners of development (both national and global) knew (or were coming to appreciate) that where the "best practice" model of top down adoption of what worked elsewhere was going to work, it had worked, and that most of the problems moving forward were the "wicked" kind of problems for which the standard approach, even if fortified by some "rigorous" evidence from somewhere else, was not going to work (e.g. Pritchett and Woolcock 2004).

In this view the strong demand from academic and think tank based researchers for RCTs was the result, not of an assessment of what practitioners actually needed, but the outcome of methodological struggles within a more or less autonomous discourse of research. The arguments for more RCTs looked exactly like the self-interested pleading and faith-based advocacy for more resources for a favored "silver-bullet" from lots of groups.

The other view was among those that thought that (a) development enterprise should be more focused on kinky goals, (b) targeted programmatic interventions could make important progress on those kinky goals, even conditional on levels of national development (and hence, say, holding constant resources available to governments, political structures, capabilities of state organizations, economy-wide productivity, "institutions", etc.) and that (c) "knowledge" of the kind that RCTs could generate would be instrumental in getting wide-scale adoption of effective programs (as opposed to either wide scale adoption of ineffective programs or the lack of scale for effective programs).

In this view, the lack of demand by governments and development agencies for RCTs was invidious: people either were not interested in effectiveness, they were defensive about their programs, they didn't understand the importance of counter-factuals for causal inference, etc. An alternative explanation was the fact that "knowledge" was a public good and individual actors or governments did not want to internalize the externality.

### *III.C) RCTs for national development questions don't pass a smell test*

First draft. For comment only.

Please to not circulate or cite without permission.

We know how to produce high levels of human well-being, high levels of national development. But this only gets us back to Lucas's (1988) about "actions"—"if so, what exactly?" It may well be that actions being taken to promote "capable administration" are not effective. It may well be that, in certain conditions, elections alone will not produce a responsive polity. It may well be that "liberalization" won't automatically promote growth. No development practitioner of my generation has any remaining hubris about knowing "what, exactly" will promote national development.

Part of the rise of the *randomistas* was a reaction to previous, confidently given (even imposed by conditionality) advice how to accelerate growth, based supposedly on evidence and research, that didn't work, reliably at least. A legitimate counter-reaction was to stop giving "policy advice" except when it could be based on "rigorous" evidence. Unfortunately, that stance, applied strictly, precludes advice on important questions for which there will be decisions and actions taken, and, unfortunately, fools will rush in where academically pure angels fear to tread.

One could take the stance that it was impossible *in principle* do research that would have impact on actions to promote national development. This could be either because it was (a) impossible to come up with good research that would inform recommendations in a way that increased reliably probability of success or (b) because it was impossible (unlikely) that research could, even if correct, given the large and complex institutional, political, and social dynamics, have an impact on actions and hence outcomes. Again, both of those points are well taken and development economists certainly can, and have been, taken to task for both their excess certitude and their naiveté (though of course one suggests they were too influential and wrong and one that they were not influential).

On the other hand, it is not the case economic growth is unchanged and invariant. Pritchett, Sen, Kar and Raihan 2016) document that there are *shifts* in growth rates that create episodes of positive accelerations and negative decelerations relative to counter-factual growth rates that produce gains and losses of enormous magnitude, trillions of dollars in the case of large countries like China and India (gains) and Brazil (losses from deceleration). While it is hard to explain completely the timing and magnitude of those episodes (Hausmann, Pritchett, Rodrik 2005) there are obvious cases in which the adoption of a new policy stance is associated with large and sustained episodes of poverty reducing growth (the four countries in Figure 9, for instance). Even if research on the determinants of growth is only of small probabilistic help the gains are so massive the returns to such research can be astronomical (Pritchett 2018).

But perhaps RCTs can be applied to questions important for national development? There are several points worth making.

First, there is no evidence for that view. Some countries have high levels of national development today and others low levels. None of that existing variation is explained by some

First draft. For comment only.

Please to not circulate or cite without permission.

countries relying more than others on rigorous RCT evidence. Moreover, some countries have made very rapid progress in national development (e.g. Korea, Chile) and others very slow progress or reversal (e.g. Venezuela, Liberia). No one makes the claim that any of the variation in growth of national development to date has been driven by differential adoption of RCT evidence. That RCTs are not a *necessary* condition for national development and rapid improvements in human well-being is obvious. That they even might come to be a contributing factor of any importance is, at this stage, would be entirely an article of faith.

Second, no serious scholar that has ever studied the complex processes whereby countries come to have productive economies, whereby countries come to have a responsive polity, whereby countries come to have capable administrations, whereby societies come to have more equal treatment—and there have been hundreds and hundreds of such scholars—have ever proposed that these kinds of marco-historical, social, contested, contingent, processes are of the type for which RCTs estimates could, even in principle, be important.

Third, I have argued previously that reviews of the RCTs to date do not suggest that they are located in domains or about interventions of the type that could contribute to, say, generating and sustaining an episode of rapid economic growth sufficient to produce very different levels of income. This is based on the pretty simplistic, but rough and ready useful, notion that if X is an important cause of Y then certain things should be true of the empirical association between X and Y, like that higher Y over time should be associated with higher X over time (and of course all of this is “all else equal” and one can create complex dynamics in the relationships such that some of these correlations are not true even if X is a powerful determinant of Y).

Table 2: RCTs to date are not focused on interventions that even plausibly pass a four part “smell test” of being important determinants of economic growth (filling it out is an exercise for the reader)				
The 20 “interventions” on which there have been sufficient rigorous impact evaluations to make comparisons about generalizability (Vivalt 2014)	Four criteria for whether “X” is an important determinant of economic growth			
	Done more in richer than poorer economies? (e.g. more in Denmark than Haiti)	Done more today than historically in now rich economies? (e.g. more today in Denmark than in 1870)	Done more in rapidly progressing countries than stagnant countries? (e.g. done more in Korea than Ghana)	Country’s progress accelerates/ decelerates when a country does more/less of it? (e.g. done 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				
Performance pay				
Rural electrification				
Women’s livelihood programs				

Source: <https://www.cgdev.org/blog/your-impact-evaluation-asking-questions-matter-four-part-smell-test>

I am reasonably confident this extends to other components of national development. Nothing about why some countries have effective post offices and others don’t or why some countries have responsive polities and others do not is associated with improved cook stoves or micro-nutrients or deworming.

*Analyzing the distribution of private goods.* A deep and fundamental problem is that RCTs work well for “interventions” than can be individualized, where one person gets a “treatment” and one person is a “control”: one person gets a drug, another gets a placebo, one

First draft. For comment only.

Please to not circulate or cite without permission.

set of women gets (the possibility of) micro-credit, another has no (incremental gain to) access to micro-credit. This is because to get statistical power to differentiate impact one needs an adequate number of observations in each category, “treatment” and “control.” This is why there have been many studies of cash transfers (both unconditional and conditional), because it is easy to give it to this household and not that household and hence estimate the “treatment on the treated.” But, as Hammer (2017) points out these are, perhaps by definition, the least interesting or policy relevant questions to ask or answer as they often involve private goods that are rival and excludable and so, other than a gain from redistribution (which may itself be substantial of course), there is typically no “rationale for public sector action” in the standard public economics. Often studies show that demand curves slope down (if it is cheaper people will do more of it) which cannot either be news nor can that, in of itself, have any policy implication for private goods. As Hammer (2017) suggests there is something of an unholy alliance between people (politicians, philanthropist) who want to give away something in a way they get direct credit for, which means distributing private goods, and people willing to do RCTs on the impacts of this, that cannot, in and of itself, have implications for action.

*III.D) It is not obvious RCTs work, even for the narrow questions on which they might have worked*

Now, on page 39, we are to back to where many discussions about RCTs start: “Suppose we (development agency, NGO, philanthropist) want to design an intervention (project/program/policy) that promotes a specific narrow objective (e.g. cook-stoves, water treatment, water storage)<sup>30</sup> how useful are impact evaluations based on randomization going to be?” And we are back to this point having spent 38 pages to triply deny the premise this is a particularly interesting or important question as: (1) the “kinky development” this framing embodies has little normative or political legitimacy in developing countries and broader measures of well-being should be the goal and not some narrowly circumscribed (global super-elite driven) definition of goals or who constitutes the “world’s poor”, (2) national development will accomplish all of the goals that these programs are setting out to accomplish, (3) RCTs are not much use in promoting national development.

But, even on this narrow question, the original vision that “independent impact evaluation” using RCTs was going to produce “rigorous” evidence on “what works” and that was going to raise program effectiveness has, rightly, been almost completely abandoned even by its original proponents.

This is an entirely separate paper, but there are two fundamental issues.

First, RCTs produce (at best) a consistent (in the statistical sense) estimate of the average causal impact of a *specific program* implemented in a *specific place and time* and one might

---

<sup>30</sup> Or, say, “the ‘Ways the World’s Poor Borrow, Save, Farm, Learn, and Stay Healthy’  
First draft. For comment only.  
Please to not circulate or cite without permission.



think of this as a “rigorous” estimate. But, as we knew would be the case (“[we knew fire was hot](#)”), these estimates have neither external validity (the impact is not the same across contexts) or construct validity (variations in project design often produce big changes in project outcomes). Hence, rigorous evidence isn’t. First, evidence from “there” has no special status “here”—as however rigorous it was “there” to bring it “here” requires all kinds of unverified assumptions that are not rigorous. Second, it is easy to show examples in which relying on RCT evidence from “there” actually *increases* the prediction error of impact versus using the simplest possible OLS evidence from “here” (see Pritchett and Sandefur 2015, Pritchett 2018).

Third, construct validity implies that discussing the evidence of a “class” of project types, like “teacher training” or “microcredit” is like discussing the evidence about “motor vehicles”—certainly almost nothing can be true and interesting of a category that includes semi-trucks and motorcycles. Evidence from education interventions can show *order of magnitude* differences in learning outcomes from minor variants of the same project in the same context (Kerwin and Thornton 2018).

There are ideas about addressing the “generalization puzzle” (Bates and Glennerster 2017) that results from a lack of external and construct validity, and these seem to be eminently sensible. But when the executive director of JPAL is saying things like: “The challenge is to pair local information with global evidence and use each piece of evidence to help understand, interpret, and complement the other” it means the first generation insistence that RCTs would provide “rigorous” evidence about “what works” and that this evidence could be followed to improve effectiveness is now completely and totally over. Rigorous evidence that is “paired” with “local information” certainly no longer has any claim to be “rigorous.”

Second, the idea that knowledge of the type RCTs can produce is a key constraint to better programs, implemented at scale, has not been demonstrated and seems dubious, for three reasons.

One, the constraint might be that “interventions” that can be carried out in a collaboration between NGOs and researchers (or at small scale with governments as an experiment) are not political viable. This certainly appears to be the case for interventions like “cameras in classrooms” (Duflo, Hanna, and Ryan 2012) or “contract teachers” (Duflo, Dupas and Kremer 2012) or even “teaching at the right level” (Banerjee et al 2007). Variants on “teaching at the right level” have been shown to work again and again, but the politics of scaling these insights in India (where the earliest and best evidence has been generated) have yet to fall into place (Banerji 2015).

Two, another constraint might be that the organizations available for implementation may not implement with fidelity the policy as designed. The transfer of “best practice” has been fraught with failures. A community development program widely thought to have had

First draft. For comment only.  
Please to not circulate or cite without permission.

reasonably positive outcomes in Indonesia (KDP) was “replicated” in Kenya and resulted in massive corruption that vitiated any positive impact. Bold et al (2013) show the exact program shown to “work” in one part of Kenya when implemented by an NGO (Duflo Dupas and Kremer 2012) also worked at scale with an NGO but did not when the Ministry itself implemented.

Three, More broadly, it seems that the effective *adoption* of techniques and practices *known* to be effective is the more serious constraint that a lack of knowledge of what is effective. In a world where TFP has *diverged* during most of the development era it seems hard to believe that the availability of knowledge in a technical, codifiable, transmittable in journal articles can be regarded as the key constraint.

### *III.E) In which I say three nice things about the potential uses of RCTs*

It would be madness to oppose method *qua* method. While the first generation fad of “RCT as independent impact evaluation of project to provide rigorous estimates of causal impact and hence reliable evidence of what works to be used by all” is over, there are three positive ways in which RCTs are increasingly (?) being used.

*Partnership with organizations as part of a learning strategy.* There were two limitations of the idea that “independent impact” evaluation was using RCTs was important (even in the limited space). One was that one could identify, evaluate, and scale interventions with knowledge generated “independently” of the organization responsible for implementation. The other was that one always needed “impact” evaluation—whereas in practice many interventions fail, not because they successfully produce “output” but the “outputs” don’t improve “outcomes” for intended beneficiaries (though many fail for this reason, see the above about private goods), but because the implementing organization is unable to produce the intended activities or outputs. RCTs can be used *in partnership* with organizations in order to explore ways the organization can be more effective at delivering *outputs*. Organizations like [IDinsight](#) are emphasizing the use of DFE (decision focused evaluations) versus KFE (knowledge focused evaluations) that bring the help from evaluators inside the learning cycle of the organization (Shah, Wang, Fraker and Gastfriend 2015). In a partnership approach with organizations one can integrate the learning into implementation and combine rigorous methods for evaluation with qualitative data and case studies to address the truly complex nature of development interventions (Bamberger, Rao, Woolcock 2010 (2016)), Woolcock 2013) as in the Social Observatory approach.

Strategies for building the organizational capabilities of state organizations, like PDIA (Andrews et al 2017) and other “adaptive” approaches, emphasize the need for organizational feedback loops, “experiential learning” as part of an integrated “MeE” (Monitoring, experiential learning and impact Evaluation) approach to “crawl the design space” (Pritchett, Samji and Hammer 2012). Randomization (like that used in A/B testing of alternatives) can be an integral First draft. For comment only.  
Please to not circulate or cite without permission.

part of that. It is easy to show that with complex and rugged response surfaces the speed of the feedback loops used to explore the response surface can be much more important than a slower more “rigorous” approach that holds smaller numbers of treatment arms fixed over longer periods to do impact evaluation (Nadel and Pritchett 2016).

My argument is that I am not anti-RCT, I am pro-RCT, I want the use of randomization across alternatives to be orders of magnitude more widely used as a learning device by organizations. But that is not going to happen with “independent impact evaluations” done by academics. To expand the use of RCT techniques they will have to be more organizationally relevant and helpful, faster to get to findings and providing useful midcourse corrections, and cheaper (should not be a surprise that if one wants quantity demanded to be higher, price should be lower).

*Evaluating government programs in actual practice, at scale.* Muralidharan and Niehaus (2017) recently acknowledge the criticisms of the usefulness of small scale, boutique, academic and NGO led, use of RCTs as a guide to government programs. They make the case for experimentation at scale of government programs as they are implemented in practice. Certainly the authors have produced a number of examples of such evaluations and their potentially usefulness (e.g. Ree et al 2015, Muralidharan, Niehaus and Sukhtankar 2016). Of course these are exactly the arguments made 50 years ago (and subsequently) for doing RCTs of programs in the USA and it would be interesting to have the discussion of what the net contribution to social policy outcomes evaluations at scale have played in the USA (relative to a counter-factual) and whether one expects the contribution to be higher or lower in places like India, and why. There is no question that, particularly in a very, very, large country like India there are going to be many instances of super-cost effective evaluations at scale, but that is not to claim these are the most important, or even one of the top ten most important, elements of what will lead India to have better outcomes.

*Using RCTs to study impacts of system type interventions on endogenous outcomes.* The third interesting and helpful use of RCTs is to undertake the evaluation of perturbations to the system itself, rather than of interventions with specific individuals. Markets, for instance, as a system for producing outcomes through interactions of sellers and buyers might work better (or worse) depending on the quality of information consumers can use. Andrabi, Das and Khwaja (2017) demonstrate the impact of providing “report cards” with student and school specific information on how information provision changed prices charged, differential responses of supply, and hence influenced test scores and enrollments. Andrabi, et al, 2018 show the impacts of providing unconditional grants to private schools to examine how it changes their behavior and have different treatments arms in which either only one school or all private schools in a village received the grant, will allowed not just the investigation of potential financial market inefficiencies but also how market structure determines the impact of education subsidies. Both

First draft. For comment only.

Please to not circulate or cite without permission.

of these clearly have “treatments” but the impact of the treatment is mediated through market responses.

These instances in which the “intervention” is a perturbation to elements of the system that determines behaviors of suppliers and demanders and hence outcomes, is over and above the point that general equilibrium effects might be different from “treatment on the treated.” Filmer, et al (2018) shows that since the prices of perishable foods rise when beneficiary households receive cash transfers the impact on nutrition of non-beneficiaries is negative and hence the total impact cannot be inferred from observing “treatment” versus “control” households alone. Similarly Muralidharan and Sundararaman (2015) show the impact of vouchers not only on those receiving the vouchers but also on those in the village who did not receive the voucher to examine potential spillover effects (and in this case do not find them).

### *Conclusion*

*Speak to us only with the killer's tongue,  
The animal madness of the fierce and young...*

Conrad Aiken, *Sursum Corda*

Short Answer: Madness.

The pioneers of developing in the post WWII era, political, practical, and academic, and I would guess, especially those from the “developing” world, and particularly the newly independent states, has an aspirational vision of national development. They believed that if they managed to “modernize”: have a productive economy, capable organizations for administration, a politics that served their citizens and achieved greater sense of social equality (though this latter was undoubtedly more contested) the citizens of their countries would be freed from deprivation and have higher levels of well-being, certainly materially and potentially in other dimensions as well. To them national development would deliver human well-being, and eliminating the worst manifestations of deprivation would be way-stations on the path to loftier goals.

They could not have imagined how right they were. Today omnibus measures of human well-being, like the Social Policy Index, are unbelievably tightly linked to three simple measures of national development. With modest amounts of measurement error it is possible *all* of observed cross-national variation in the Social Progress Index is associated with national development. Similarly, the country level measures of income/consumption poverty are associated with the median of the distribution of income/consumption at levels of correlation of .96 to .98. Growth, increases in median income/consumption, reduces poverty and reaching high levels (essentially) eliminates low-bar poverty. Full. Complete. Total. Stop.

First draft. For comment only.

Please to not circulate or cite without permission.

The first 50 years of the “development era” from 1950 to 2000 saw more progress in (nearly) every individual/household measure of material well-being: health, education, income, poverty reduction, access to electricity, that in all of previous human history.

By 2000 we had seen low-bar poverty essentially eliminated in the four dragons (Korea, Taiwan, Hong Kong, Singapore) and brought to very low levels in all of the high performing East Asian economies—by national development, not poverty programs. We were witnessing right before our eyes the economic transformation of countries like China, India, and Vietnam where the onset of rapid growth was reducing poverty at rates and in magnitudes never before seen in human history.

And it wasn't just income. Health was improving, even faster than income growth alone would suggest. The Preston curve was shifting up over time—more life expectancy even at the same levels of income. Years of schooling was expanding even faster than income growth alone would suggest, again, more years of schooling at the same levels of income. The Human Development Index, meant to capture more than just income, became very highly correlated with income because their indicators of health and education converged so strongly.

Of course, there were in 2000 of course places in which human well-being was not high and not progressing: but those were precisely places where national development was failing.

Into this environment a small group of development economists proposed the notion that a (key) factor limiting progress in human well-being, particularly the kind of “low-bar” goals endorsed by the MDGs in 2000, was the lack of rigorous evidence about the impact of development projects because they relied on “before” and “after” comparisons because they lacked randomization of “beneficiaries” into “treatment” and “control” that would allow rigorous estimates of causal impact.

While the idea that causal identification required a defensible estimate of the counterfactual was well-known and widely accepted (and had been for decades) the suggestion this insight was particularly important for actual development outcomes was, on the face of it, preposterous based on known facts. Moreover, while persuasive arguments were presented on narrow technical issues, the larger “theory of change” of just how better causal identification of projects would lead to better outcomes was never even argued with any evidence, just asserted.

All of the momentum of RCTs in development came from discourses in the rich countries: debates inside the academic discipline of economics, the shift to a kinky approach to development among development agencies, the increasing influence of philanthropists, end of the Cold War and the shift in US politics towards developing countries. (Almost) none of it came from developing countries or development practitioners.

There are still important challenges in raising broad-based, omnibus, measures of human well-being (and reaching low-bar goals on the way to those) and more and more rapid national

First draft. For comment only.  
Please to not circulate or cite without permission.

development will be needed to achieve that. Perhaps some tiny role in that can be served by RCTs (though it is hard to impossible to “power up” system level interventions and evidence about key national development challenges isn’t “technical” knowledge that could ever, even in principle, achieve external validity) but let’s get back to helping countries with their challenges, using whatever tools, techniques, analysis, evidence, wisdom, we can gather. The idea that RCTs of cash versus chickens would be important is madness, masquerading as method.

## References

- Andrabi, Tahir; Jishnu Das; Asim I. Khwaja; Selcuk Ozyurt and Niharika Singh.** 2018. "Upping the Ante : The Equilibrium Effects of Unconditional Grants to Private Schools." *Policy Research Working Paper;No. 8563*.
- Andrabi, Tahir; Jishnu Das and Asim Ijaz Khwaja.** 2017. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." *American Economic Review*, 107(6), 1535-63.
- Andrews, Matt; Lant Pritchett and Michael Woolcock.** 2017. *Building State Capability: Evidence, Analysis, Action*. Oxford, UK: Oxford University Press.
- Bamberger, Michael; Vijayendra Rao and Michael Woolcock.** 2016. "Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development." *World Bank Policy Research Working Paper No. 5245*.
- Banerjee, Abhijit V; Esther Duflo; Shawn Cole and Leigh Linden.** 2007. "Remedying Education Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics*, 3(122).
- Bates, Mary Ann and Rachel Glennerster.** 2017. "The Generalizability Puzzle." *Stanford Social Innovation Review*, Summer.
- Bold, Tessa; Mwangi Kimenyi; Germano Mwabu; Alice Ng'ang'a and Justin Sandefur.** 2013. "Scaling up What Works: Experimental Evidence on External Validity in Kenyan Education." *Center for Global Development Working Paper 321*.
- Duflo, Esther; Pascaline Dupas and Michael Kremer.** 2012a. "School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools." *MIT Economics Working Paper 7660*.
- Duflo, Esther; Rema Hanna and Stephen; Ryan.** 2012b. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4 (June)), 1241-78.
- Filmer, Deon; Jed Arnold Friedman; Eeshani Kandpal and Junko Onishi.** 2018. "General Equilibrium Effects of Targeted Cash Transfers: Nutrition Impacts on Non-Beneficiary Children." *World Bank Policy Research Working Paper No. 8377*.
- Foster, James; Joel Greer and Erik Thorbecke.** 1984. "A Class of Decomposable Poverty Measures." *Econometrica*, 52(3), 761-66.
- Hausmann, Ricardo; Lant Pritchett and Dani Rodrik.** 2005. "Growth Accelerations." *Journal of Economic Growth*, 10(5), 303-29.
- Heckman, James and James Smith.** 2000. "The Sensitivity of Experimental Impact Estimates (Evidence from the National Jtpa Study)," D. Blanchflower and R. Freeman, *Youth Employment and Joblessness in Advanced Countries* Chicago: University of Chicago Press, 331-56.
- Kahneman, Daniel and Angus Deaton.** 2010. "High Income Improves Evaluation of Life but Not Emotional Well-Being." *PNAS*, 107(38), 16489-93.
- Kant, Immanuel.** 1998 (1785). *Groundwork of the Metaphysics of Morals*. Cambridge UK: Cambridge University Press.
- Kenny, Charles; Pritchett, Lant.** 2013. "Promoting Millennium Development Ideals: The Risks of Defining Development Down " *Center for Global Development, Working Paper*, (338).

First draft. For comment only.

Please to not circulate or cite without permission.

- Kerwin, Jason T. and Rebecca L. Thornton.** 2018. "Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures."
- Kraay, Aart.** 2006. "When Is Growth Pro-Poor?" *Journal of Development Economics*, 80(1), 198-227.
- Lee, Jung-Wha and Hanol Lee.** 2016. "Human Capital in the Long Run." *Journal of Development Economics*, 122, 147-69.
- Muralidharan, Karthik and Paul Niehaus.** 2017. "Experimentation at Scale." *Journal of Economic Perspectives*, 31(4), 103-24.
- Muralidharan, Karthik; Paul Niehaus and Sandip Sukhtankar.** 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review*, vol 106(10), pages 2895-2929., 106(10), 2895-929.
- Nadel, Sara and Lant Pritchett.** 2016. "Searching for the Devil in the Details: Learning About Development Program Design." *Center for Global Development Working Papers*, 434.
- Parfit, Derek.** 2011. *On What Matters*. Oxford UK: Oxford University Press.
- Pritchett, Lant.** 2010. "Birth Satisfaction Units (Bsu): Measuring Cross-National Differences in Human Well-Being." *UNDP Human Development Report Research Paper 2010/03*.
- \_\_\_\_\_. 2006. "Who Is Not Poor? Dreaming of a World Truly Free of Poverty." *The World Bank Research Observer*, 21(1), 1-23.
- Pritchett, Lant; Salimah Samji and Jeffrey Hammer.** 2012. "It's All About Mee: Using Structured Experiential Learning ("E") to Crawl the Design Space." *HKS Faculty Research Working Paper Series*, (forthcoming).
- Pritchett, Lant and Justin Sandefur.** 2015. "Learning from Experiments When Context Matters." *American Economic Review*, 105(5), 471-75.
- Pritchett, Lant; Kunal Sen; Sabyasachi Kar and Selim Raihan.** 2016. "Trillions Gained and Lost: Estimating the Magnitude of Growth Episodes." *Economic Modeling*, 33(June), 279-91.
- Pritchett, Lant; Woolcock, Michael; .** 2004. "Solutions When *the* Solution Is the Problem: Arraying the Disarray in Development." *World Development*, 32(2 (February)), 191-212.
- Rawls, John.** 1970. *A Theory of Justice*. Cambridge Mass.: Belknap Press of Harvard University Press.
- Ree, Joppe de; Karthik Muralidharan; Menno Pradhan and Halsey Rogers.** 2015. "Double for Nothing? The Effects of Unconditional Teacher Salary Increases on Performance."
- Rordriguez, Fransisco; Shelton, Cameron A.** 2013. "Cleaning up the Kitchen Sink: Specification Tests and Average Derivative Estimators for Growth Econometrics." *Journal of Macroeconomics*, 38(PB), 260-73.
- Rosling, Hans; Ola Rosling and Anna Rosling Ronnlund.** 2018. *Factfulness: Ten Reasons We're Wrong About the World--and Why Things Are Better Than You Think*. UK: Hodder and Staughton.
- Woolcock, Michael.** 2013. "Using Case Studies to Explore the External Validity of 'Complex' Development Interventions." *WIDER Working Paper, No. 2013/096*.

First draft. For comment only.

Please to not circulate or cite without permission.



## Appendix 1: Excerpt from an email from John Briscoe to Drew Faust

Drew,

I hope you are having a relaxing summer.

You have no reason to know, but I am South African, and have lived (and worked) in many countries around the world (including Mozambique, Bangladesh, India and Brazil).

In my six years at Harvard dozens of students have come to me because they knew of my experience with development, and wanted to engage with this challenge. Of course there have been many nuances, but this conversation is typical:

TWO UNDERGRADUATE STUDENTS: Hi Professor. We respect so much what you have done in development. We are concerned about the billions of people who don't have water. We want your help in starting a water filter factory in Uganda.

BRISCOE: Great that you want to do something for others! But....have you ever started a business?

STUDENTS: No

BRISCOE: Have you ever been to Uganda?

STUDENTS: No.

BRISCOE: I have seen hundreds of efforts like yours. Almost all of them produce lovely (but temporary) websites of smiling faces. But two years later you have moved on and everything goes back to where it was. On the other hand, there has been enormous progress in bringing water to the unserved -- over every day for the last 20 years 280,000 people have acquired improved water services, almost all from better and better organized local water utilities. By the way... where do you come from?

STUDENTS: Chicago

BRISCOE: Who provided your water in Chicago?

STUDENTS: We don't know...

BRISCOE: It is the water utility of Chicago. And what worked in Chicago (and China and Brazil and South Africa) is the only proven path.... Now I have a good friend, a Ugandan who manages the Uganda Water Authority. He has done an incredible job. Spending the summer working for him would be a great experience -- learning from an African leader and using your skills on things where he needs help. Would you like me to contact him?

STUDENTS (looking dismayed): Hmmm..... but ... no thanks professor, what we really want is to found our own NGO for this....

BRISCOE: With respect, I think you are treating Africans (have you ever met one?) as objects of charity and pity. And, with less respect, in cases like this I see the main "benefit" is for you to feel that you are saviors of poor, incapable people....

**First draft. For comment only.**

**Please to not circulate or cite without permission.**

And so, undeterred, because there is so much "what wonderful people you are!" from the system at Harvard, they enter one of the competitions you have set up and actually win it and get a big grant to go and "follow their dreams".

I am pleased that there is now an active debate taking place in the media, stimulated by the deceptions wrought by Nicolas Kristof (an icon, of course, at Harvard). One of the many discussions of this is here... <http://www.theatlantic.com/international/archive/2012/03/the-white-savior-industrial-complex/254843/>

My main point is that Harvard is, in my view, perpetuating a fine-sounding but fundamentally paternalistic and exploitative way of interacting with Africa. Students are rewarded for going down this white-savior track, at all levels, including from you. It seems to me that there needs to be a serious conversation about the university's role in perpetuating this vision of how concerned students can and should engage with the challenges of development.

First draft. For comment only.  
Please to not circulate or cite without permission.