



Knowledge or its Adoption?

Lant Pritchett

Senior Fellow, Center for Global Development

CGD Note

August 2018

When I criticize the fad of RCTs in development (which I do occasionally), I often get the sarcastic response, “What should we do, go back to doing *growth regressions*?” The under-40 generation defend their fad, whatever its flaws, as at least better than the previous fad. It is conventional wisdom that development economics did far too many growth regressions and that theory was simplistic, empirical work was sloppy, and therefore *nothing* was learned. Moreover, it was argued that problems of pathways of causality both amongst the many potential covariates (the covariate “robustness” problem) and between covariates and growth (the “adequate identifying instruments” problem) would and even in principle *could* never be adequately resolved. In many ways, the rise in development of the RCT agenda of carefully controlled experiments to measure causal impacts of specific identifiable programs or “treatments” was a direct, allergic-type, reaction to the real and perceived negative excesses of growth research generally, and growth regressions in particular.

I argue, though, that we did learn two very important things from growth research, and these were *learned* from research in the strong sense that they changed people’s views from a previous view that was incorrect.

A doesn’t converge

One thing we did *learn* from growth research is that convergence in total factor productivity (TFP) was not common. This was *learned* from empirical growth research in the strong sense that most people doing development in the 1950s and 1960s thought there would be convergence in TFP.

It would not be a caricature of the Solow (1956) model and its aftermath in the 1950s and 1960s to conclude that American economic growth could usefully be decomposed into “factor accumulation” and “TFP.” Moreover, while formally people recognized TFP was residually measured and hence, strictly speaking, “a measure of our ignorance,” it was not uncommon to think of TFP growth as “technical progress.” I was taught in graduate school at MIT that the production function A represented “sets of blueprints” of what was technologically possible (subsuming organization into technology), and this set of the possible with science and technology (and organizational) advanced to account for a significant fraction of output-per-hour growth. Reading Robert J. Gordon’s magisterial 2016 book *The Rise and Fall of American Growth*, which still has as its centerpiece decompositions of growth, I am convinced this is still a productive way to think about American economic growth. It is demonstrably the case that during the twentieth century, science and technology created vast new possibilities (e.g., electricity, internal combustion engines, telephones, jet travel, air conditioning, improved medicine). Alfred Chandler and the business history school emphasize that new forms and practices of organizations (the rise of managerial capitalism) and professions led to the “scale and scope” that brought these potentials into everyday life.

If one understood “ A ” in the aggregate production function as codifiable technical knowledge—how medicines affect disease, how fertilizers affect plant growth, how to produce steel, how telephones transmit sound, etc.—then it was easy to think of A as a “public good” that was non-rival and non-excludable. In a post-colonialist world in which political sovereigns were interested in progress in their country (and maybe even the well-being of their people), it was easy to imagine that governments would have every incentive to bring this available knowledge to bear in promoting growth in their country. This was an obvious, and widely accepted, narrative of the two pre-World War II development successes: Russia and Japan. The idea of the “advantages of backwardness” was premised on the perfectly plausible notion that it must be easier to transplant, adopt, and adapt existing knowledge, already within the frontiers of technology and organizational practice, than to push the frontier.

In this intellectual context, everything about the “first generation” development research and practice is pretty clear. If A converges rapidly—because, after all, the knowledge of how penicillin and nitrogenous fertilizers and electricity work is “in the air,” like Jefferson’s metaphor of the light of a candle that all can benefit from—then the key constraint on convergence in incomes is the speed with which resources can be mobilized, from domestic and foreign savings, to invest in physical and human capital (and it is a complete myth that human capital was ever underacknowledged). The convergence of A with low K/L meant returns on K would be high and growth dynamics—the speed of convergence—would be determined by savings. Hence the famous 1954 Arthur Lewis quote:

The central problem in the theory of economic development is to understand the process by which a community which was previously saving, and investing, 4 or 5 per cent of its national income or less converts itself into an economy where voluntary saving is running income or less converts itself into an economy where voluntary saving is running at about 12 to 15 per cent of national income or more. This is the central problem because the central fact of economic development is rapid capital accumulation.

This implies that the goal of a development organization, say a bank, say a World Bank, should be to mobilize investible resources to augment domestic savings and perhaps transmit those savings via investment projects that would also transmit the knowledge of the technical frontier.

These ideas were so powerful in part because they were grounded in common sense and practical observation. Who could deny there had been technical progress? Before there weren’t cars, now there are cars. Before people died of diseases that are now easily treated. Who could deny that scientific knowledge was a public good (of course, the whole premise of protection of intellectual property like patents was that it was otherwise a public good)? Who could deny it was hard to mobilize savings when consumption levels were very low?

Good thing the fad of “growth research” came along and documented the facts. Bosworth and Collins (2003) (among many others) decomposed growth in output per person across lots of countries from 1960 to 2000 into TFP growth and factor accumulation. What was striking was that for most developing country regions (Latin America, Africa, Middle East), TFP had grown *more slowly* than in industrial countries: (measured) A was *diverging*. Even in relatively high-growth regions (East Asia excluding China, South Asia), the more rapid rates of growth were not due to convergence of A (it grew at 1 percent in these regions, exactly the industrial country rate) but faster factor accumulation.

Table 1. Decomposition of growth into the growth of factors and the growth of A (TFP) shows most developing country regions were *diverging* in A from 1960 to 2000

Region (number of countries)	Growth in output per worker	Contribution by component (percent per year)		
		Physical capital per worker	Education per workers	Total factor productivity
Industrial countries (22)	2.2	.9	.3	1.0
Africa (19)	.6	.5	.3	-.1
Latin America (22)	1.1	.6	.4	.2
Middle East (9)	2.1	1.1	.4	.5
South Asia (4)	2.3	1.0	.3	1.0
East Asia (7) (except China)	3.9	2.3	.5	1.0
China (by period)				
1960-1970	.9	0.0	.3	.5
1970-1980	2.8	1.6	.4	.7
1980-1990	6.8	2.1	.4	4.2
1990-2000	8.8	3.2	.3	5.1
1960-2000	3.9	2.3	.5	1.0
USA (by historical periods), Gordon 2016				
		Capital deepening	Education	TFP
1890-1920	1.50	0.65	0.35	0.50
1920-1970	2.82	0.59	0.33	1.90
1970-2014	1.62	0.78	0.22	0.62

Source: Bosworth and Collins 2003, Table 1. Gordon 2016, Figure 1.2.

By the early to mid-2000s, many of the major academics in the field of growth/development had written papers arguing against the “A converges/factor accumulation” view of growth dynamics (e.g., [Hall and Jones \[1999\]](#); [Easterly and Levine \[2002\]](#); [Rodrik, Subramanian, and Trebbi \[2004\]](#); [Acemoglu, Johnson, and Robinson \[2001\]](#)) and positing that something deep, like “institutions” (rather than “endowments” or “factors” or “policies”), explains the levels and dynamics of growth. [Caselli \(2005\)](#) showed the standard growth accounting suggested that in 1996 data, only about 35 percent of the 90th-10th percentile gap in levels of per capita income was explained by differences in physical and human capital. [Grier and Grier \(2007\)](#) wrote “Only Income Diverges: A Neoclassical Anomaly,” showing that the cross-national data showed convergence in many of the standard growth determinants/correlates but continued divergence in incomes.

[Comin and Mestieri \(2018\)](#) continue and contribute to this literature by showing that technology diffusion across countries can be well modeled as a combination of *technological adoption*—how long it takes for a given technology to arrive in a country—and *intensity of use*—how widespread the technology becomes. They conclude that the Great Divergence (or, as some call it, Divergence, Big Time) is driven primarily by differences in technology diffusion; 75 percent of the increase in the income gap between the Western and non-Western countries during the period 1820 to 2000 is driven by diverging aggregate TFP driven by differences in the process of technological diffusion of adoption and intensity of use of available technologies.

Comin and Mestieri (2018) measure the “intensity of use” of technologies across countries, conditional on adoption (think of an S-curve of technology penetration in the use-time space and “adoption” is the horizontal shifter [years of lag from discovery to adoption in a country] and “intensity of use” is the vertical shifter of the penetration). They find that over time that although *adoption* has been speeding up (intuitively the spread of the PC to first adoption was much faster than the use of the ship), the *intensity of use* has been *diverging*. They find that, even for simple technologies invented over 100 years ago (e.g., railway freight, mail, electricity, tractors), the poorer countries (10th percentile) and median country are still far behind in the intensity of use compared to the Western country average.

Table 2. Estimates of the log-intensity of use parameter relative to Western countries

(selected) Technology	Invention Year (sorted)	N	Mean	Standard Deviation	P10	P50
Railway Freight	1825	43	-.33	.49	-.95	-.33
Mail	1840	45	-.31	.35	-.79	-.30
Electricity	1882	75	-.74	.59	-1.49	-.65
Tractor	1885	87	-1.20	.89	-2.43	-1.19
Fertilizer	1910	92	-.97	.78	-1.93	-.91
Harvester	1912	70	-1.44	1.13	-3.17	-1.36
Synthetic Fiber	1931	45	-.76		-1.93	-.69
All technologies (25 in original)		1189	-.76	.85	-1.94	-.59

Source: Adapted from Comin and Mestieri (2018).

The argument that differences in outcomes are the result of differential country adoption of widely known technology (that has been embedded effectively in a variety of organizational forms and practices) accords well with studies of particular functions.

In “Letter Grading Government Efficiency,” Chong et al. (2014) show that on the simple task of returning misaddressed foreign mail, a function for which all countries have an identical official policy (as signatories to an international convention that commits them to return misaddressed foreign mail to the sending country), country performance ranged from zero (none of 10 letters returned, ever) to 100 percent (all letters returned). Obviously, zero of this difference in efficacy on this task can be attributed to the availability of mail A. How is it there are countries where the mail is not reliably delivered?

Das et al. (2012) use trained “standardized patients” presenting with symptoms of three common conditions (unstable angina, asthma, and dysentery) to assess medical care in practice in rural Madhya Pradesh, India. They find that existing medical care from first-contact practitioners does not reach “do no harm.” For dysentery, the correct treatment was recommended only 12.7 percent of the time versus 7.9 percent for an unnecessary or harmful treatment. For unstable angina, the correct treatment was recommended 31.2 percent of the time versus 55.5 percent of the time for which a harmful or unnecessary treatment was recommended. For asthma, less than half got the correct treatment but 62.7 percent got a harmful or unnecessary treatment. Clearly none of this observed outcome has anything do to with medical A—knowledge of the correct diagnosis and treatment of dysentery is well known. Moreover, Das et al.’s paper shows that the trained practitioners in public sector clinics have at their disposal checklists of what should be done in response to each of the conditions the standardized patients presented with—but the practitioners just don’t do even a small fraction of those checklists. Medical A has nothing to do with this.

The cumulative impact of this evidence from growth research was like the Rutherford experiment (actually, I learned from Wikipedia, researchers in his lab, Geiger and Marsden) firing alpha particles at gold foil and having them bounce straight back. A lack of convergence in income itself might not have been surprising as perhaps A would converge but industrial countries could, with higher incomes and savings rates, be able to maintain more rapid factor accumulation. But the opposite happened: most of the lack of convergence was because A (the residual) did not converge and it appears to be because the use of known technologies—based on established and completely accepted and widely known and practiced science—did not diffuse and were not adopted. This meant the mechanics of financial flows, premised on the idea that returns were high because A was high relative to K/L or HK/L, was less important.

This was also a super hard question to work on because the standard neoclassical growth setup produced “Solow invariance”—if production functions are constant returns to scale in factors, and markets are competitive, then factor payments exhaust product and hence there is nothing left over to pay for improving A. The “exogenous” growth models were deeply exogenous, and it was not at all clear the Romer-esque endogenous growth models, with their focus mainly on the long-run (steady state) growth of A, were a useful approach. Early versions had “scale effects” that might have been a feature or a bug, but Charles Jones showed very early on that the prediction of early endogenous growth models—that the level of knowledge increased the rate of growth of A—seemed a lot like a very big bug as it was massively counterfactual for the rich countries’ growth.

All of this is just so that the younger generation can understand the puzzlement of the previous generation of development academics and practitioners with the new fad of RCTs and the founding of IPA (in 2002) and JPAL (in 2003) at exactly the time this new consensus was emerging from growth research. Research revealed that it wasn’t typically A (particularly as interpreted as technical knowledge) or even “policies” that constrained developing country outcomes for the most part but the adoption and diffusion of known A across organizations (both public and private) in developing country settings.

Of course, as a (pre-RCT fad trained) economist, I could understand the private interests of the actors. As a younger or junior faculty, it was great to have a “new” method to deploy that allowed you to write and publish papers; the superiority of causal claims based on clean identification via experimental assignment met that criteria. (Although the claim to “new” was pretty limited as in 2002 there were already at least four organizations in the US alone with a long-standing expertise in doing randomization in social experiments: Mathematica, which had begun a social policy experiment in 1968; Rand Corporation, which began fieldwork on the Health Insurance Experiment in 1976; MDRC, which was launched in 1974; and Abt Associates, founded in 1965. So “new” meant “new” to the sub-field of development economics as the use of randomized experiments to assess social policy was at least 30 years old in 2002/2003.) And as with any researcher or faculty, it was great to attract funding to do what you wanted to do and, given the known bias of many funders for the “new” and “innovative,” and with little fear of the fad, I can understand the interest in selling funders on RCTs.

But what I have never understood is the lack of any realistic, empirically formed, theoretically grounded “theory of change” of how this new research fad would have impact on the course of events in the developing world. There has always seemed to be to me a pretty obvious dilemma with some pretty sharp horns.

On one horn, one could claim that RCTs would produce “gold standard” evidence about causal impacts of the type of “knowledge” that was “technical” and “codifiable” and “scientific” and appeal to analogies like the use of RCTs as the standard for drug trials. That is, the claim might be that RCTs would produce A of the “technical progress” type and that this better A would improve development outcomes. This horn of the dilemma had its

obvious academic attractions but seemed at odds with obvious facts. How can one work in a country that doesn't deliver the mail reliably, where medical practice doesn't reach "do no harm," and that generally has low and non-converging TFP—so is obviously not using A (as codifiable technical knowledge) that has been available for decades, if not centuries—and think that "more A" of the type like the old understanding of A has any major part to play in accelerating development progress?

On the other horn, one could claim that RCTs would produce evidence about how to get organizations to be more effective at using the A they had and hence, perhaps, how to make development organization projects more effective. But then this is very unlike drug trials or agricultural field experiments as it is not clear there is "scientific" knowledge in the usual sense of knowledge that has "external" validity and "construct" validity and hence can be applied with confidence, even if, at one time and in one space and with one set of implementers, one could "rigorously" demonstrate impact. "Here is knowledge about how to get your post office to work better by applying widely known A" hardly seems like the kind of *general or widely applicable* knowledge that one could imagine even an RCT could generate usefully.

As Deaton (2010) puts it:

Finding out how people in low-income countries can and do escape from poverty is unlikely to come from the empirical evaluation of actual projects or programs, whether through randomized trials or econometric methods that are designed to extract defensible causal inferences, unless such analysis tries to discover *why* projects work rather than *whether* they work.

But *why* projects work or not, includes, among other factors, organizational and institutional features about specific contexts which are demonstrably not the type of "knowledge" that can, even in principle, be regarded as codifiable.

Fifteen years into the RCT fad, my take is that the fad has, by ignoring what was learned from previous research about the development process in the attempt to create a "blank slate" on which a "new" methods could write results, has been even less useful to policies and practices in development than the fad of growth research (with all its faults) that it replaced.