



From the SelectedWorks of Martin Ravallion

January 2012

Fighting Poverty one Experiment at a Time: A
Review Essay on Abhijit Banerjee and Esther Duflo,
Poor Economics

Contact
Author

Start Your Own
SelectedWorks

Notify Me
of New Work

Available at: http://works.bepress.com/martin_ravallion/18

Ref: Journal of Economic Literature, forthcoming

Fighting Poverty one Experiment at a Time

A Review of Abhijit Banerjee and Esther Duflo, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*

MARTIN RAVALLION¹

Banerjee and Duflo offer a coherent vision for an economics of poverty and anti-poverty policy. Their economics is grounded in an effort to understand the economic and psychological complexities in the lives of poor people, informed by social experiments and field observations. Their preferred policies entail small reforms at the margin, also informed by experiments—specifically randomized control trials. While the book provides some interesting insights, I question how far its approach will get us in fighting global poverty.

1. **Introduction**

Abhijit Banerjee and Esther Duflo are two of the most influential scholars working on development today, and their beautifully written and accessible book, *Poor Economics*, is bound to attract wide interest. The book advocates actions by poor countries to reduce poverty—actions that are informed by data and careful thinking about incentives and behavior. That message will not be news to most development economists, but it is nonetheless an important message, which this book will bring to a broad audience.

There are two newer themes, just under the surface. The first concerns the type of evidence used in policy making and the second is about the type of question that evidence is used to address. On the first, the book puts a large weight on evidence from randomized control trials (RCTs). On the second, it emphasizes small policy reforms at the margin within existing environments—a “quiet revolution” of small but sure improvements. And these are the types of

¹ Development Research Group, World Bank. These are the views of the author and should not be attributed to the World Bank or any affiliated organization. For comments, the author is grateful to Pedro Carneiro, Michael Carter, Jishnu Das, Quy-Toan Do, Jed Friedman, Emanuela Galasso, Markus Goldstein, James Heckman, Aart Kraay, Peter Lanjouw, David McKenzie, Alice Mesnard, Berk Ozler, ErikThorbecke, Adam Wagstaff, Dominique van de Walle, Nicolas van de Walle and Michael Woolcock.

policies that are well suited to RCTs. Thus the book provides a coherent vision for how to do development economics.

This paper critically assesses that vision, and whether it is up to the task of generating the knowledge needed to effectively fight global poverty.

2. *A new way of doing economics?*

While Banerjee and Duflo have made many contributions to economics, they are best known as the co-founders in 2003 (along with Sendhil Mullainathan) of what is now called the [Abdul Latif Jameel Poverty Action Lab](#) (J-PAL). To quote the authors' bios page: "J-PAL's mission is to reduce poverty by ensuring that policy is based on scientific evidence," where "scientific evidence" is code for RCTs. J-PAL's use of RCTs is seen to constitute a "new way of doing economics" (*Poor Economics*, p.14) by giving the subject greater scientific status; as one J-PAL associate explained at a conference on development effectiveness I attended, "We are the guys in the lab coats." Economics has long had its methodological camps and now we have a new one—the "randomistas camp."

While *Poor Economics* draws on evidence from various (sometimes surprising) sources, RCTs are weighted heavily. Randomized experiments are not new to economic analysis and policy evaluation. What is new is (first) the degree to which experiments are seen to be the only credible approach and (second) their extensive application in developing countries. How one assesses this book's contribution depends in no small measure on how one assesses this "new way of doing economics." So that is where this essay will start.

The simplicity of an experiment is clearly part of its appeal to J-PAL and followers; for example, when I asked an economics graduate student why she was so keen to do an RCT she said: "I want to do something I can explain to my parents." But there is also a deeper critique of non-experimental economics. It is argued that the observational data, models and econometric methods traditionally favored by economists require too many assumptions for reliably inferring impacts. The assumption that seems to be held in deepest suspicion is the "exclusion restriction" required by the popular instrumental variables estimator (IVE), whereby certain instrumental variables are assumed to only affect outcomes via treatment and so can isolate a degree of exogenous variation in treatment. Failure of this assumption to hold biases the IVE. The

randomistas argue that only in a randomized design is the exclusion restriction beyond question. It is acknowledged that when experimenting on human subjects, selective compliance with the randomized assignment is to be expected. But it is claimed that the randomized assignment can be excluded, to be used as the instrumental variable. RCTs promise to cleanly identify the causal effect with few assumptions.

At the root of the new enthusiasm for experiments is a long-standing obsession amongst economists with selection bias based on unobserved variables. There are essentially two ways to address endogeneity concerns about participation in the intervention being evaluated (whereby participation is correlated with unobserved factors relevant to outcomes). The first way is to collect more data on those things that jointly influence outcomes and participation. The other way is to find a better instrumental variable, such as by doing an RCT.² It has not been established, and it is not obvious on *a priori* grounds, which approach dominates, taking account of costs as well as benefits. If data were chronically scarce and exogenous to research one would be chronically concerned about selection on unobservables. But with increasingly sophisticated data sets (including integrated, multi-purpose, surveys linked to geographic data), and much greater potential for tailoring data collection to the problem at hand, one would have expected selection on unobservables to have become less of a concern. Yet economists seem to worry more than ever about unobserved variables. Their worries have enhanced the influence of the randomistas to the point where many economics doctoral students and young academics now wander around looking for something to randomize. And they turn down evaluation opportunities when randomization is not feasible.

Not everyone has greeted the development randomistas with enthusiasm.³ *Poor Economics* pays little attention to the past critiques of RCTs,⁴ although the concerns that have

² The difference is clear if we note that the probability limit of the bias in the IVE (estimated β less true β) is given by $[\rho(z, \varepsilon) / \rho(z, x)][\sigma(\varepsilon) / \sigma(x)]$ (where ρ is the correlation coefficient, z is the instrumental variable, ε is the error term, x is the regressor and σ is the standard deviation). We can reduce the bias by either reducing $\rho(z, \varepsilon) / \rho(z, x)$ (such as by using an RCT) or reducing $\sigma(\varepsilon) / \sigma(x)$ (by collecting more data).

³ On the advantages and limitations of experiments in social policy making see James Heckman (1992), Heckman and Jeffrey Smith (1995) and Robert Moffitt (2004). Discussions focusing more on development applications, and J-PAL's work in particular, include Dani Rodrik (2009), Martin Ravallion (2009), Chris Barrett and Michael Carter (2010) and Angus Deaton (2010).

⁴ It can be granted that some of these issues are technical, though the authors do a fine job in explaining some other, equally technical, issues in a broadly accessible way.

been raised in the literature are important to assessing the book's vision for a new economics of poverty and policy. The claim that randomized assignment can help identification has not been at issue. Rather, the critics have pointed out that experiments are rarely so clean in practice, such that a number of assumptions are needed to draw valid inferences about the experimental population, let alone valid policy inferences—including assumptions that are not required by non-experimental studies.⁵ Biases can arise when the experiment influences the behavior of either the treatment or control groups, or staff in the field. Given the behavioral responses, it should not be presumed that RCTs necessarily dominate observational studies in practice.⁶

Spillover effects can also cloud inferences in both types of studies. For example, if one village gets the intervention and another within the same local jurisdiction does not, and this is known, then the local government can rationally re-allocate its own efforts accordingly, thus biasing the impact estimate, even with randomized assignment.⁷ More generally, the treatment and comparison groups are typically part of the same economy—trading in the same markets as well as being linked through shared political and social institutions—in which case spillover effects must be expected. General equilibrium effects are routinely ignored by the development randomistas.

Even the much vaunted claims about removing all selection bias with a simple RCT start to sound hollow once one recognizes that compliance with an experimental design will often depend on factors determining the expected impact of the treatment—factors that are known to those considering whether to take up the offer of treatment but unobserved by the researcher. Naturally, people make rational choices about whether to participate in an experiment, and they base their choices on things they know but we don't. Then the IVE no longer provides a consistent estimate of the mean causal impact even when the assignment to treatment is random.⁸

⁵ Heckman (1992) and Michael Keane (2010) discuss the (often implicit) assumptions made by RCTs.

⁶ For example, compliance has been found to be a bigger problem for experimental than non-experimental drug trials (Michael Kramer and Stanley Shapiro (1984).

⁷ Chen, Ren Mu and Ravallion (2009) demonstrate and quantify this mechanism for a poor-area development program in rural China.

⁸ This stems from the fact that the error term in the standard regression of outcomes on treatment contains the interaction effect between treatment and the deviations from mean impact. Furthermore, this interaction effect can be expected to be positive on average for those who choose to take up the treatment. Thus the error term has a non-zero mean conditional on the randomized assignment, violating the exclusion restriction. See Heckman, Serio Urzua and Edward Vytlačil (2006).

The randomistas' assumption that randomized assignment satisfies the exclusion restriction ceases to hold. Estimating mean impact then becomes a more difficult econometric problem.⁹

This illustrates a more general point: the task of learning from experiments, given likely behavioral responses, inevitably leads us back into the rival camp of the “regressionistas,” who favor more structural models estimated using econometric methods. That graduate student may well end up having just as hard a time explaining what she is doing to her parents!

Policy makers have also come to doubt the benefits of relying so much on RCTs, which can be frustratingly uninformative about the questions they face in making better policies for fighting poverty. As Heckman (1992, p.218) puts it, in social settings RCTs “...may produce clear answers to the wrong question.” There can be unusually large individual losses following an intervention, affecting specific groups of the populations, and policy makers are often keen to know about these big losses. Yet, even under ideal conditions, an RCT only delivers an estimate of the mean impact on those treated in the experimental population, and we learn little or nothing about the distribution of impacts from a standard RCT. We cannot even determine the median impact. Also, the question of why the intervention did or did not have impact in that population remains most often open. Nor is it clear whether the intervention would have similar impacts in some other population (an issue I return to in section 4). To answer all these questions stronger assumptions and modeling will be needed. Again we are led back to the rival camp.

There are also concerns about the ethics of this new way of doing economics. The subjects of these experiments are seen merely as means to some end. If we don't know who needs the “treatment” and what it will do—as long as it does no harm—then deliberately withholding that treatment for the purpose of an experiment seems ethically harmless. But the randomistas surely exaggerate our ignorance about the efficacy of the things we do in the name of fighting poverty. We know that de-worming tablets (say) work almost always. Then the RCTs discussed in this book are ethically worrying. Furthermore, the principle of “informed consent” does not seem to get much respect from the new developmental randomistas; clearly many people did not know that they were being experimented on (notable when it is groups, such as villages, that are randomized, rather than individuals). And there are even risks of doing real

⁹ For further discussion see Heckman, Urzua and Vytlačil (2006).

harm. For example, Banerjee and Duflo discuss one (non-JPAL) experiment on corruption in India that knowingly augmented the large number of drivers on Delhi's roads who clearly do not know much about how to drive properly. These are complex issues, and ethical concerns need to be properly weighed against the gains from knowledge. The ethics of experiments has long been taken seriously in medicine, and has been a major concern about social experiments in the US.¹⁰ But the topic has received scant attention from the new development randomistas.

In summary, while RCTs have a place in the toolkit for policy research, they do not stand alone, above all other types of evidence. It is not clear that a research strategy that relies heavily on RCTs constitutes a “new way of doing economics,” especially when one notes that the main inferential problems of the old way are still there, on top of some new concerns.

3. *Lab coats in the field*

I expect many readers of this book will share my surprise at how much the authors have been influenced by a very different type of evidence, namely casual observations from their field work—stories about specific individuals or families that the authors met. This is the kind of informal qualitative work that micro-development economists often do, and it can sometimes be an important source of ideas and insights—complementary to more rigorous empirics. However, while there is a role for direct observation in the field, it is a type of data that has to be handled carefully. More than once, I came back from my field trips to villages confident of some position that I subsequently overturned in light of much better data for a representative sample. It is understandable that the authors wanted to give a “face to poor people” in a book such as this—to help (mostly well-heeled) readers realize that poor people are just like them, except poorer. But this is clearly not the “hard evidence” that J-PAL has been advocating. Nor would it appear to meet the standards of rigorous qualitative and “mixed-method” work in the social sciences.¹¹ At times Banerjee and Duflo appear to raise their casual interviews and observations from their field work to a worryingly high inferential level.

¹⁰ Ethical concerns were identified as the main reason why local employment centers in the US refused to participate in voluntary RCTs for assessing new training programs (Heckman, 1992).

¹¹ For a recent example of rigorous qualitative work on local development issues see Patrick Barron, Rachael Diprose and Michael Woolcock (2011).

For example, during a field trip to some villages in West Bengal in early 2009 Banerjee and Duflo asked people about impacts of the global financial crisis. Their expectation was that reduced demand for construction work in India's cities due to the crisis will have led to a fall in remittances from migrants to their families in these villages. However, Banerjee and Duflo heard nothing in the villages to suggest that there was such an effect, and they concluded that the crisis had little impact in these villages. But what is the counterfactual here? The impact of the crisis in the booming economies of India's biggest cities must presumably be judged against the counterfactual trajectory in the absence of the crisis. This may well be hidden from view to the families of rural migrants, who continue to report higher earnings in the cities than the villages. Field work alone cannot tell us much about such things.

The authors' observations from field work lead them to pose some interesting questions, which they then explore further, invoking seemingly plausible conjectures about behavior, drawing creatively on both economics and psychology. Why do mothers not inoculate their kids? Why do farmers not save the (seemingly) small amounts needed to finance the next growing season's fertilizer purchase, even when sufficient yield gains can be expected? Why don't undernourished people spend more on nourishing food? Why are so many kids absent from schools when they should know the benefits of education? All interesting questions, though it was not clear to me why these questions were chosen and not others.

For example, the chapter on education tells the story of a 40-year old widow, Shantarama, in Karnataka, India. It seems her youngest (school-age) children are not in school by their choice, despite the fact that her older children are relatively well schooled. Banerjee and Duflo also quote statistics indicating that primary enrolment rates are now reasonably high in developing countries, but so too are absentee rates. All this leads them to pose the question (p.72):

“So if the failure of schools in developing countries to attract children can't be explained by problems of access, or lack of demand for educated labor, or parental resistance to educating their children, then where is the snag?”

They go on to discuss other factors inhibiting attendance and learning. However, one wonders if the authors might have been too hasty. It is clearly a huge step from the story of Shantarama and some aggregate education statistics to the question posed above. Are we to think that Shantarama's story is somehow typical of all developing countries? And while there is now near-

universal primary school enrollment in many developing countries, the drop off at the secondary level is often large, and then issues of access and returns remain relevant.

It is not always clear that the evidence presented supports the book's conclusions. Apparently the reason Shantarama's younger kids were not in school was not because it cost too much. And the authors point to an experiment in Malawi that found that there was no extra impact on schooling of adding a schooling conditionality to transfers targeted to households with adolescent girls. The conditionality is a pure price effect: essentially it lowers the net cost of schooling, on top of the direct income gain. So neither Shantarama's story nor the experiment tells us that schooling is responsive to its own price. Yet, Banerjee and Duflo (p.81) conclude that:

“Unless we can fully erase differences in income, public supply-side interventions that make education cheaper would be necessary to get close to the socially efficient outcome: making sure that every child gets a chance.”

Whether one agrees or not with this conclusion, it is not implied by either the field observations or the RCTs reported in this book.

There is an interesting epilogue to this example, carrying a further lesson. A recently revised version of the paper on Malawi that Banerjee and Duflo had relied on has exploited better data on schooling attainments than the self-reported measures used in the early version of the paper. The new results indicate a sizeable gain from the schooling conditions (Sarah Baird, Craig McIntosh and Berk Ozler, 2010). Better data dramatically changed the experimental result. So Banerjee and Duflo may still be right that incentives really do matter to schooling. But one is also led to wonder how many of these experimental findings might be overturned by better data. Doing an RCT does not diminish the importance of reliable data.

4. *Small may be beautiful, but is it big enough?*

A book that aims to understand how best to fight global poverty might be expected to say something about country experience in reducing poverty. There has been a huge expansion in the amount of data on poor economies over the last 30 years—both macro data and micro data.¹²

¹² The following observations can be verified from the World Bank's [PovcalNet](#) website for poverty monitoring. The following claims are robust to the choice of the poverty line over the range found amongst low-income countries.

Based on that data, China is clearly at the top of the list of performers in reducing poverty. South Asia is seeing sustained progress against poverty. Since the mid-1990s, so too is Sub-Saharan Africa as a whole, though there the progress has been at a slower pace, uneven within the region, and more fragile. These diverse country experiences offer hope for understanding why some countries do better against poverty than others.

Readers will not find any discussion of this topic in *Poor Economics*. And it is not because of space limitations. Rather, Banerjee and Duflo do not think it useful to address such questions. They believe that development economics has gone astray by being “fixated on the big questions” (p.3). Better to focus on narrow questions and small interventions, rather than to run large and ambitious non-experimental evaluations, or try to use non-experimental micro data to understand social and economic behaviors relevant to policy impacts, or use general equilibrium analysis, or (worse still!) run cross-country regressions, which try to answer questions such as why some countries have succeeded against poverty while others have not. This preference for small questions comes hand-in-hand with their preference for RCTs, which are only well suited to such questions. So this “new way of doing economics” comes with severe restrictions on the types of questions economics can address.

A key issue in assessing this book’s strategy for fighting poverty is whether the type of learning process and the knowledge generated can reliably guide public action. There are a number of reasons for doubting that it can.

One reason lies in the potential for poverty traps. Focusing on small but sure gains for poor people may not get us very far if there exists the type of dynamic poverty trap that Jeffrey Sachs (2005) and others have assumed in arguing for a large expansion in development aid. By this view, poor people are stuck in a low-level equilibrium, and a large boost will be needed to get them out of it. (And similarly, sufficiently large negative shocks will create destitution.) So the question of whether such poverty traps exist is crucial to assessing the approach to policy making proposed by this book. Banerjee and Duflo do discuss poverty traps, with a very nice

non-technical explanation early on in the book. However, we do not come to closure on the issue of whether poverty traps exist and the authors ignore much of the literature on the subject.¹³

At one point, the authors claim the existence of a poverty trap based on a graph (p. 201) of the relationship between individual wealth at one date and that for an earlier date, using data for Thailand. This is not, however, a vindication of the Sachs view, since a poor person can get out of the kind of poverty trap they illustrate in the Thai data with only a small nudge at the right time. However, it is not clear what this kind of data and analysis can really tell us about poverty traps. Suppose that there is a unique equilibrium at each date for each individual, but the equilibrium is shifting over time for most individuals, being influenced by other variables in the process of economic development. Then one can obtain a great many possible cross-sectional graphs between individual wealth at one date and that at an earlier date, whether or not there are poverty traps in the underlying dynamics. This type of data has little power for the purpose that Banerjee and Duflo use it for. A much better test for a poverty trap is to estimate a suitable nonlinear dynamic model with controls, estimated on micro panel-data with sufficient observations over time.¹⁴ Of course, this requires identifying assumptions, though they would seem no less plausible than those Banerjee and Duflo are implicitly making in their interpretation of these Thai data.

Even if there are no poverty traps that frustrate small reforms, there are doubts stemming from the restrictions on the type of questions addressed by this “new way of doing economics.” The interventions for which RCTs are feasible constitute a non-random subset of the things that are done by governments in the name of “fighting poverty.” Actual development programs are not (of course) randomized; that would make little sense given their objectives. It is unlikely that we will be able to randomize road building to any reasonable scale, or dam construction, or poor-area development programs, or public-sector reforms, or trade and industrial policies—all things that developing countries do a lot of, and we do need to know whether they work or not. Indeed, it is not clear that RCTs are particularly well suited to programs involving the types of commodities for which market failures are a concern, such as those with large spillover effects in

¹³ Relevant empirical work using non-experimental data includes Michael Lokshin and Ravallion (2004), Carter et al. (2006), David McKenzie and Christopher Woodruff (2006), Alice Mesnard and Ravallion (2006), Francisca Antman and McKenzie (2007) and Kraay and Raddatz (2007).

¹⁴ This is the approach followed by Lokshin and Ravallion (2004). Also see Antman and McKenzie (2007).

production or consumption. Our knowledge needs to improve where there are both significant knowledge gaps and an *a priori* case for government intervention.

Another likely bias in the learning process is that J-PAL's researchers have evidently worked far more with non-governmental organizations (NGOs) than governments. And they have chosen to work with some of the best NGOs, notably in India, where there is a variance in NGO quality. A small program run by the committed staff of a good NGO may well work very differently to an ostensibly similar program applied at scale by a government or other NGO for which staff have different preferences and face new and different incentives.

For example, *Poor Economics* describes the authors' experiment aiming to incentivize the immunization of children in India's roving immunization camps. A seemingly small incentive in an RCT done with an NGO was found to bring statistically significant gains in the immunization rate, though still a long way short of achieving the benefits of full immunity. The subsidy appears to be needed to compensate for the small current cost (in time or money) incurred by some in getting the vaccinations—a cost that may weigh heavily on today's choices by poor people, even when that cost matters very little in the longer-term, relative to the (considerable) gains from immunization. However, as Jishnu Das (2010) points out, the scope for scaling up this RCT and attaining similar outcomes may be limited by new factors that come into play at scale. For example, the problems of getting public health workers to turn up for work that Banerjee and Duflo discuss elsewhere in the book will surely re-surface in the scaled-up immunization program. Das argues instead for a more fundamental re-think of India's entire model of immunization camps—precisely the type of big change that *Poor Economics* shuns.

Heterogeneity across jurisdictions in the political acceptability of RCTs can also be a source of bias in the learning process advocated in this book. Even for those programs for which randomization is an option, the ethical and political concerns raised by social experiments will have greater salience in some settings than others. The menu of feasible experiments will then look very different in different places, and the menu will probably change over time.

In short, this vision for a new development economics applies selectively to policies and settings and so it can be expected to generate selective knowledge about what works and what does not. Indeed, I can see no obvious reason why RCTs compensate for the distortions

generating existing, policy-relevant, knowledge gaps, stemming from the combination of decentralized decision making about project evaluation with the inevitable externalities in knowledge generation.

Learning from RCTs raises a further concern. Banerjee and Duflo tell us very little about the institutional-implementation factors that might make a given program a success in one place, or at one scale, but not another. The same obsession with unobserved variables—that the randomistas contend can only be convincingly dealt with by randomization—also points to concerns about extrapolation beyond the experimental environment, where unobservables must surely be expected.

The external validity of RCTs does not get adequate attention in this book, though this neglect is not uncommon; as Nancy Cartwright (2010, p. 69) puts it: “Of course all advocates of RCTs recognize that internal validity is not external validity. But the gap is far bigger than most let on.” There are many ways that a program, even if it aims for only incremental improvements, can work differently at scale to what is found for a pilot. An estimate of the impact on schooling of a tuition subsidy based on an RCT can be deceptive about the national program, which may well alter the structure of returns to schooling.¹⁵ A pilot run by an NGO can easily fall below the radar screen, while a scaled-up version of the same intervention is handicapped by governance or political economy constraints. More than once I felt Banerjee and Duflo were going too far in drawing broad lessons about the best “way to fight global poverty” from their field observations and experimental trials in specific settings. The risk is not only to current policy making, but also to the direction taken by future policy-oriented research.

For example, how confident can we really be that poor people all over the world will radically change their health-seeking behaviors with a modest subsidy, based on an experiment in one town in Rajasthan, which establishes that lower prices for vaccination result in higher demand? As Das, Shantayana Devarajan and Jeffrey Hammer (2009) note with reference to the same experiment, “...the mere existence of an inverse relation between prices and consumption can’t possibly be sufficient to justify any particular subsidy.” Similarly, an experiment planting fake crime reports in police stations (also somewhere in Rajasthan) is used to test police

¹⁵ For a full discussion of general equilibrium treatment effects see Jaap Abbring and Heckman (2007, section 4).

responsiveness and show that police who are rewarded by their conviction rates are loath to register crimes. The rate of police registration of reported crimes did improve over time after the experiment. The inference drawn is that even moribund institutions can be improved with small reforms (like hiring people to report fake crimes). One or two experiments such as this might convince us that it is possible to find effective incremental reforms to bad institutions, but governments and citizens should be wary of concluding that they do not need to worry so much about these more structural problems. That is not established.

Drawing lessons from research for policy requires great care. It is hardly surprising that *Poor Economics* draws heavily on the authors' own work and that of their students. Nonetheless, given that Banerjee and Duflo are so committed to eliminating all selection bias from our inferences about policy, it is surprising that they do not do more to alleviate readers' potential concerns that they do not provide a comprehensive review of the evidence on each issue. The book makes use of a "scorecard" on some of the development policy debates reviewed, with "supply wallahs" and "demand wallahs" each chalking up points based on the various empirical studies cited. This is a nice device for keeping the reader's interest. But we are told nothing about how the studies were chosen to make the tallies of scores. The sample is clearly not exhaustive, so there is some sort of latent selection process. Any reasonably skeptical reader will no doubt wonder whether the scores can be trusted.

"Small but sure" would be a more credible research strategy if we could attain J-PAL's original vision of doing enough high quality RCTs to span all the relevant dimensions of variation in impact by scale and context, as proposed by Banerjee (2007). This would get around the worries about those troublesome unobserved variables in the non-experimental population, even if internal validity is assured (though the aforementioned concerns remain). We would end up with a complete map of what works and what does not in every circumstance. To some observers (me included) that vision seems absurdly ambitious. And it is clear from this book that J-PAL still has a long way to go.

However, there are also encouraging signs that J-PAL may be changing track, such that their experiments are being used more often to understand social and economic behavior. For example, J-PAL's experiments have taught us more about demand functions, such as by randomly assigning different prices for bed nets to help prevent malaria. And they have taught us

something very interesting about some of the ill-informed, and even life threatening, beliefs held by poor people about health care. Hopefully we will also see experiments that will help understand the behavior of policy makers, though this will be harder.

The use of experiments—invariably in combination with non-experimental data and methods—to estimate structural parameters in economic and political behavior may ultimately prove far more important to sound policy making than their direct contribution in evaluating specific policy interventions. That way we can better understand why an intervention does or does not have an impact, and simulate alternative designs that might enhance its impact. The experimental purists may well be uncomfortable with this approach, since it requires essentially the same sort of theory-based empirics that they found so unacceptable for identification purposes. But it would be progress.

5. Conclusions

Poor Economics never treats poor people in a simplistic way. They are understood to have essentially the same desires, psychological foibles and time inconsistencies as all of us; everything is just a lot harder for them than for the well-off. This is a perspective that is fundamentally sympathetic with the needs and desires of poor people, rather than being alienated from them or judgmental of them. Banerjee and Duflo see things that poor people know full well, such as the fact that being poor does not mean that one spends every extra cent on some tasteless starchy staple; other needs—the tastiness of the diet, entertainment and social inclusion needs—are important for poor people, like anyone else.

While this is not a book full of specific policy recommendations about how best to fight poverty, some do emerge. The importance of access to reliable public information comes out often (though some of the contingent factors in the efficacy of information programs might have got more attention). The authors also advocate free or subsidized distribution of certain goods and services with large payoffs (though they say little about how this is financed or how markets would respond). They point to the high social returns to improving early childhood nutrition through micronutrient supplementation—for those who do not exercise choice over things that matter long into their future. Their overarching message for policy makers is to fight poverty one step at a time, based on evidence about what works and what does not.

This may remind readers of Deng Xiaoping's famous characterization of economic reform as a process of "crossing the river by feeling the stones under the water." Along with many developing countries, China's leaders came to realize (in the late 1970s) that their reform agenda had to be based on evidence, which they then collected. (For example, the government's early "experiments" in observing how farmers responded to new institutional arrangements for contracting out the collective's farmland were crucial to the decision to introduce the "household responsibility system.") But there are two important differences. First, while China's reformers were selective and cautious, there was nothing "small" about their reforms. Second, the non-ideological pragmatism of China's leaders since the late 1970s came with an equally pragmatic approach to evidence. Both are necessary; strong methodological priors about what type of data one can learn from have no place in sound and comprehensive policy making for fighting poverty. Success will continue to depend on our ability to make the most of unclean empirical analyses drawn from a wide range of noisy data sources.

It is not uncommon to find that the researcher's preferred methodology has dictated what questions are researched. Here the randomistas are no different. Sometimes the search for something to randomize leads to an important question and Banerjee and Duflo are good at finding such questions. But one has to wonder if we could do better as a community of researchers in aligning our research topics more closely with the questions (yes, often "big questions") faced by governments and civil society groups in attempting to fight poverty. Important knowledge gaps persist.

One can question whether *Poor Economics* adds up to a "radical rethinking of the way to fight global poverty." But it usefully re-affirms an important message: poverty can be alleviated through well-informed, well-thought out and relevant public actions. Poor countries are not doomed to stay quite so poor, the cycle of self-fulfilling expectations of poor service delivery to poor people can be broken, better public programs and policy reforms can be devised and shown to work, even while many deeper problems of market and governmental failures remain. But please let us not neglect those deeper problems.

REFERENCES

- Abbring, Jaap H., and James J. Heckman. 2007. "Econometric Evaluation of Social Programs. Part iii: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation," *Handbook of Econometrics, Volume 6B* (Amsterdam: North Holland).
- Antman, Francisca and David McKenzie. 2007. "Poverty Traps and Nonlinear Dynamics with Measurement Error and Individual Heterogeneity," *Journal of Development Studies* 43(6): 1057-1083.
- Baird, Sarah, Craig McIntosh and Berk Ozler. 2010. "Cash or Condition? Evidence from a Cash Transfer Experiment," [Policy Research Working Paper 5259](#), World Bank.
- Banerjee, Abhijit. 2007. *Making Aid Work* (Cambridge, Mass.: MIT Press).
- Barrett, Christopher B., and Michael R. Carter. 2010. "The Power and Pitfalls of Experiments in Development Economics," *Applied Economic Perspectives and Policy* 32(4): 515-548.
- Barron, Patrick, Rachael Diprose and Michael Woolcock. 2011. *Contesting Development: Participatory Projects and Local Conflict Dynamics in Indonesia* (New Haven: Yale University Press)
- Carter, Michael R., Peter D. Little, Tewodaj Mogues and Workneh Negatu. 2006. "Poverty Traps and Natural Disasters in Ethiopia and Honduras," *World Development* 35(5): 835-856.
- Cartwright, Nancy. 2010. "What are Randomized Controlled Trials Good For?" *Philosophical Studies* 147: 59-70.
- Chen, Shaohua, Ren Mu and Martin Ravallion. 2009. "Are There Lasting Impacts of Aid to Poor Areas? Evidence from Rural China," *Journal of Public Economics* 93: 512-528.
- Das, Jishnu. 2010. "Improving Immunization Coverage in Rural India," [British Medical Journal](#) 340: c2553.
- Das, Jishnu, Shantayana Devarajan and Jeffrey S. Hammer. 2009. "Lost in Translation," [Boston Review](#) March/April.
- Deaton, Angus. 2010. "Instruments, Randomization and Learning about Development," *Journal of Economic Literature* 48: 424-455.

- Heckman, James J. 1992. "Randomization and Social Policy Evaluation." In *Evaluating Welfare and Training Programs*, ed. Charles F. Manski and Irwin Garfinkel, 201 - 30. Cambridge and London: Harvard University Press.
- Heckman James and Jeffrey Smith. 1995. "Assessing the Case for Social Experiments," *Journal of Economic Perspectives* 9(2): 85-110.
- Heckman, James, Serio Urzua and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics* 88(3): 389-432.
- Keane, Michael P. 2010. "Structural vs. Atheoretic Approaches to Econometrics," *Journal of Econometrics* 156(1): 3-20.
- Kraay, Aart and Claudio Raddatz. 2007. "Poverty Traps, Aid and Growth," *Journal of Development Economics* 82(2): 315-347.
- Kramer, Michael S., and Stanley Shapiro. 1984. "Scientific Challenges in the Application of Randomized Trials," *Journal of the American Medical Association* 252(19):2739-2745.
- Lokshin, Michael and Martin Ravallion. 2004. "Household Income Dynamics in Two Transition Economies," [*Studies in Nonlinear Dynamics and Econometrics*](#), 8(3): Article 4.
- McKenzie, David and Christopher Woodruff. 2006. "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises," *Economic Development and Cultural Change* 55(1): 3-42.
- Mesnard, Alice and Martin Ravallion. 2006. "The Wealth Effect on New Business Startups in a Developing Economy," *Economica*, 73: 367-392.
- Moffitt, Robert. 2004. "The Role of Randomized Field Trials in Social Science Research," *American Behavioral Scientist* 47(5): 506-540.
- Ravallion, Martin. 2009. "Evaluation in the Practice of Development," *World Bank Research Observer* 24(1): 29-54.
- Rodrik, Dani. 2009. "The New Development Economics: We Shall Experiment, but How Shall We Learn?" In *What Works in Development? Thinking Big and Thinking Small*, ed. Jessica Cohen and William Easterly (Washington, D.C.: Brookings Institution Press).
- Sachs, Jeffrey. 2005. *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals* (New York: Millennium Project, United Nations).