Svenja Flechtner

Winning the Fight Against Global Poverty One Experiment at a Time?

The Economics Nobel 2019

Abstract: The Economics Nobel 2019 for Abhijit Banerjee, Esther Duflo and Michael Kremer for their experimental approach to alleviating global poverty has been received by economists with a wide array of reactions. This article introduces the experimental revolution in development economics and the work of Banerjee, Duflo and Kremer and then discusses a number of critiques that have been raised about it: methodological concerns relating to the internal and external validity of experiments in economics; a supposed focus of RCTs on the microlevel and a distraction of development economics away from structural issues; the scope of experimental research for poverty reduction; and normative concerns arising when experiments are used in development economics.

Keywords: Development Economics, Experimental Economics, Randomized Controlled Trials (RCTs), Nobel Prize

Zusammenfassung: Der Nobelpreis für Wirtschaftswissenschaften 2019, der Abhiit Banerjee, Esther Duflo und Michael Kremer für ihren experimentellen Ansatz zur Armutsbekämpfung auszeichnet, hat unter Ökonom*innen weltweit eine große Bandbreite an Reaktionen und Kommentaren hervorgerufen. Dieser Artikel porträtiert die experimentelle Revolution in der Entwicklungsökonomik sowie die Arbeit von Banerjee, Duflo und Kremer und diskutiert einige der geläufigen Kritikpunkte: Bedenken bezüglich der internen und externen Validität von Experimenten in der Ökonomik; ein möglicher Fokus von RCTs auf die Mikroebene und damit die Ablenkung von strukturellen Faktoren in der Entwicklungsökonomik; die Reichweite von Experimenten für globale Armutsreduktion; und normative Bedenken der Anwendung von Experimenten in entwicklungsökonomischen Kontexten.

Schlagwörter: Entwicklungsökonomik, Experimentalökonomik, randomisierte kontrollierte Studien (RCTs), Nobelpreis

https://doi.org/10.1515/ordo-2020-0022

Contents

1 Introduction

Before starting to write this comment on this year's economics "Nobel" for Abhijit Banerjee, Esther Duflo and Michael Kremer, I had already been confronted with a large number of praises and critiques, most of them either very critical or very enthusiastic. It seems that this year economists are particularly divided in their appreciation of the three laureates, who were rewarded "for their experimental approach to alleviating global poverty"[1], and upon the meaning and implications of the committee's choice. Many agree with the Nobel committee in that Kremer, Banerjee and Duflo have "introduced a new approach to obtaining reliable answers about the best ways to fight global poverty", while others see no particular merit or even warn of detrimental side effects of the now widespread use of experiments in development economics. Without doubt, this prize invites some reflection about the laureates' contributions to both academic economics and the global fight against poverty. Because the experimental approach is so widely used, these reflections do not only concern the laureates' own work but also the reception and application of experimental methods in the development economics community.

After briefly summarizing the experimental revolution in development economics and the work of Banerjee, Duflo and Kremer (section 2), I will attempt to place these contributions in a broader context and discuss them in light of a few critical perspectives (section 3). These include methodological concerns about experiments and, in particular, randomized controlled trials (RCTs) in economics; the supposed focus of RCTs on the microlevel and a distraction of development economics away from structural issues; the scope of experimental research for poverty reduction; and normative concerns arising when RCTs are used in development economics.

1 The press release from 14th October 2019 is available at https://www.nobelprize.org/uploads/2019/10/press-economicsciences2019–2.pdf.

2 Experimental economics in the field of development

The current experimental revolution began no more than 20 years ago (see e.g. Angrist and Pischke 2010). Experiments had been used by economists before, for instance to evaluate public policies in the US, but in their current shape – as random assignments – they became prominent not before the 1990s. In a situation of dissatisfaction with empirical research, often due to research design issues, experimental studies seemed particularly promising to "take the con out of econometrics" (Leamer 1983). The basic principle is that the random assignment into "treatment" and control" groups should allow researchers to identify the effects of a treatment on outcomes, overcoming selection problems. By now, this new identification strategy has been placed by some econometricians at the forefront of the "credibility revolution" (Angrist and Pischke 2010)[2], and it is held to "occupy a special place in the hierarchy of evidence, namely at the very top" (Imbens 2010, p. 407). Provided they are feasible to address a certain question, Imbens (2010, p. 401) thinks that "randomized experiments are superior to all other designs in terms of credibility".

Development economics was one of the fields – perhaps *the* field – where the experimental turn took off in practice. Development economist Michael Kremer was among the first to run a number of experiments in Kenya, starting with a randomized evaluation of the impacts of an NGO health sponsorship programs in Kenyan schools. He quickly involved a number of other researchers and students in these projects and, in 1996, started to work with Abhijit Banerjee on randomized impact evaluation in India.

All three Nobel laureates have contributed highly cited RCT studies on topics such as the impacts of microfinance (Banerjee et al. 2015a), the impact of deworming on schooling in Kenya (Miguel and Kremer 2004), the reaction of Kenyan farmers to a nudge aimed at increasing their use of fertilizers (Duflo et al. 2011b), a remedial education program in India (Banerjee et al. 2007), the impact of tracking of the academic performance of Kenyan pupils (Duflo et al. 2011a). Further, the laureates have published well-known evaluations of natural experiments, for example concerning the impacts of gender quotation of Indian village council heads (Chattopadhyay and Duflo 2004; Beaman et al. 2012), or of a large school construction program in Indonesia (Duflo 2001).[3] They have also contributed to spreading and developing the

2 This strategy includes random assignment in general, not only randomized controlled trials. Noting that randomized controlled experiments are not time-consuming, costly, and sometimes infeasible, Angrist and Pischke explicitly emphasize natural or quasi experiments as another type of research designs with (as good as) random variation.

3 Let me recommend David Evans's blog, where the author has published a helpful overview of Kremer's (https://www.cgdev.org/blog/quick-guide-100-research-articles-economics-nobel-winner-

method in economics, publishing toolkit papers and handbook entries (e.g. Duflo et al. 2007, Duflo and Banerjee 2009).

3 RCTs: a revolution for critical thinking about development or rather a critically thin basis?

3.1 Methodological concerns

Because random assignment is understood as a superior identification strategy, many economists share the view that experimental methods can provide a clean estimate of the effect of a treatment on the outcome variable (see e.g. Bandiera 2019 for a recent statement with reference to the Nobel award). But the position of RCTs as the supposed "gold-standard" on top of the hierarchy of empirical methods has also been questioned on technical grounds. Among the main concerns are questions of both external and internal validity.

3.1.1 Internal validity

With respect to internal validity, prominent critiques have been advanced by Deaton and Cartwright (2018). A single randomized experiment provides an estimate of the average treatment effect (ATE), provided it is well-conducted and unbiased. In a world where treatment and control groups were *fully* identical, we would be able to estimate the ATE exactly. But this is not the case in economics even when researchers control for specific covariates: some potential other causes will be distributed unevenly across treatment and control groups. Provided that unobserved covariates are distributed evenly over the population from which these groups are randomly drawn, repeating the same experiment could help: repeated experiments would identify the correct result *on average*. With one single ATE estimated, however, it is not possible to know which point of a larger distribution of estimates it represents. In other words, it is possible that one estimation comes up with a result that is rather an outlier than representative of the true average effect.

Further, Young (2018) recently criticised that many randomized experiments are analysed with econometric methods that overestimate the statistical significance of treatment effects. Specifically, he applied different randomization tests to 53 pub-

michael-kremer), Duflo's (https://www.cgdev.org/blog/quick-guide-100-publications-by-economics-nobel-winner-esther-duflo) and Banerjee's works (https://www.cgdev.org/blog/quick-guide-100-publications-economics-nobel-winner-abhijit-banerjee).

lished experiments and reduced the number of significant results (at the 1 % level) by 22 %. In joint tests of all treatment effects, the share was even reduced by 49 %. Young argues that these differences in results produced by different tests are due to different ways in which these tests handle extreme values in covariates (leverage). His exercise certainly raises doubts about the internal validity of a considerable share of published RCT studies.

Finally, some concerns are of a more practical nature: participants could find ways of becoming a member of the treatment group or disappear over the course of follow-up surveys; or there could be spillover effects, which would blur the lines between treatment and control groups (e.g. Fletcher and Marksteiner 2017). Recent approaches have used dual-level stratified randomized designs to identify spillover effects within and across units of randomization, such as schools or villages (Angelucci and De Giorgi 2009; Haushofer and Shapiro 2018). In the face of these and other problems, it has even been argued that there are situations – for example, when there are interactions between different treatment activities as participants select into one activity as a consequence of having been selected for a previous one – in which other regression techniques could be more suitable for impact evaluation (Elbers and Gunning 2014).

3.1.2 External validity

Experiments are run in specific places and environments. Internal validity provided, it still remains unclear to what extent the results of a policy intervention would replicate in a different setting. Concerns about external validity have accumulated especially since a number of RCT studies were not successfully replicated. For instance, the positive impact of class size on students' outcomes found in the US (Krueger 1991) has not been found in Kenya (Duflo, Dupas and Kremer 2012) or India (Banerjee et al. 2007; see Pritchett and Sandefur 2014 for a longer discussion). In many other cases, we simply do not know whether a result would hold if it were to be tested elsewhere: replication studies (Sukhtankar 2017) or parallel tests in different settings as in Banerjee et al. (2015b) are the exception. But in some fields, there have been enough studies of the same topic or very similar policy interventions for economists so start thinking about meta-analysis. For example, Meager (2019) estimates the average effect as well as the heterogeneity in effects across seven RCT evaluations of microcredit, using Bayesian hierarchical models. In this case, she found heterogeneity of effects to be reasonable and largely due to sampling variation.

Still, when results differ substantially, it is difficult to assess why. In developing countries, external validity seems too often to be forgotten when researchers write about behavioural patterns of "the poor" as though the behaviour of people living in poverty was globally homogeneous. In this context, it has also been criticised that randomistas do not make sufficient effort to explain why they found a certain result

under certain circumstances, thereby making it harder to understand whether it can be generalised or not (Vivalt 2015). Unfortunately, RCTs by themselves do not tell us why they worked, i.e. which contextual factors were decisive in producing a certain outcome. Considering all these limitations to external validity, Pritchett and Sandefur (2014) recommend that policy-makers base their decision on OLS estimates from the same context rather than on experimental evidence from different contexts.

3.2 Do RCTs overemphasize the microlevel?

RCTs did not only appeal to economists as an identification strategy that promised to overcome selection problems: another reason was a deep dissatisfaction with the effects of public policies (Reddy 2019). The Washington Consensus or persistent poverty after decades of development aid come to mind. This was fertile ground for economists to become increasingly interested in "what works" in terms of concrete policy interventions.

Because of this focus on small and tractable problems rather than on the big questions, randomistas have faced a lot of criticism: they focus on the microlevel only, and regard individual people – often the addressees of policy interventions – and their behaviour as the main hindrance to overcoming poverty (e.g. Berndt 2015).

Not all experimental interventions focus on the micro level. Furthermore, some policies that start out local might be scalable, for instance when education policies or deworming treatments can be implemented at the country level, based on a local impact evaluation (see e.g. Banerjee et al. 2016 for an attempt at scaling up and Banerjee et al. 2017 for a discussion). Further, structural questions are not fully ignored even when the behaviour or decision-making of specific people is at centre stage. Specifically, recent contributions have analysed how a life of poverty impacts decision-making (e.g. Mani et al. 2013).

What is true, though, is that structural causes of poverty enter mostly as a background condition impacting behaviour or policy interventions. Structural conditions are not fully ignored – but the focus is on how they impact individual behaviour. While this is important and interesting, the global fight against poverty cannot be fought, much less won, without also directly addressing structural issues like trade policies, fiscal policies, climate change, monetary policies, or global taxation. Not all relevant issues and policies can be studied and addressed through an RCT. Furthermore, even if global poverty could be ended one experiment and one intervention at a time, this would not be the end of the story: economic development brings along distributional changes and conflicts. For example, while conditional cash transfers in countries such as Brazil have certainly contributed to improving the prospects of many Brazilian children, it seems evident that redistributional policies behind them have also fuelled distributional conflicts. This is to say that processes of economic development rarely happen conflict-free. Hence questions of politics and power must

also receive attention, alongside more local and technical questions pertaining to specific policies.

From a theoretical or methodological point of view, it is clear that many phenomena at the macroeconomic or structural level cannot be randomized, and that natural experiments are not easily available. Defenders of the experimental approach readily recognize how random assignment is often time-consuming, costly and sometimes infeasible (Angrist and Pischke 2009), and Imbens (2009, p. 401) even expresses his "main concern" that the rise of randomized experiments may lead researchers to avoid questions that cannot be tackled through experimental methods. Therefore, experimental methods can never be the only or main tool; fighting poverty cannot be successful just on the basis of the experimental approach. As Reddy (2012) has put it, one cannot have the cake and eat it: there is a trade-off between focusing on narrow, tractable and identifiable problems on the one hand and addressing the complex problem of global poverty on the other.

The Nobel laureates Banerjee, Duflo and Kremer have all three published papers that address more structural issues: Kremer on the O-ring theory of economic development (Kremer 1993), Banerjee of the legacy of colonial institutions in India (Banerjee and Iyer 2005), and Duflo on health services (Banerjee et al. 2004), for example. Hence it is all the more lamentable that there is a lot of rhetoric about how the experimental method has transformed the fight against poverty and seems more relevant and promising than work that addresses other questions. When asked about it, of course, randomistas readily acknowledge that the bigger questions are also important. But at the same time, the small questions are often given much more priority and importance. For example, in *Poor Economics*, Banerjee and Duflo (2010) write that rather than asking "What is the cause of poverty?" or "Is foreign aid good for economic development?", economists should study particular factors that make the lives of poor people difficult, and interventions that can improve their situation. It is not the latter recommendation that is problematic, but the suggestion of an "either-or" approach. Perhaps this comes from the conviction that the big questions are harder to answer. The Nobel announcement captures this well: the "smaller, more precise, questions are often best answered via carefully designed experiments". However, the big questions of how to fight global poverty and achieve economic development may involve, but cannot entirely be divided "into smaller, more manageable, questions". And if it is true that the bigger questions are more difficult to answer than the small ones, then the time and energy of the smartest economists worldwide should go into developing methods and approaches that can address them nevertheless.

3.3 A sense of proportion

A related concern about experimental methods in development is not so much about what they can – in theory and practice – deliver, but how they have been received and

understood by the development economics community. Experimental methods in development have not only been praised for their academic merits, but also with regards to their – achieved and future – potential to decrease global poverty. Here, too, it seems good advice to keep a sense of proportion.

As the Nobel Popular Science Background report states, the laureates "have shown how the problem of global poverty can be tackled by breaking it down into a number of smaller – but more precise – questions at individual or group levels. They then answer each of these using a specially designed field experiment". In reality, and as discussed above, not each of the relevant problems in the context of global poverty can be tackled in this way. There are highly relevant issues that escape this treatment.

This is no devastating critique: any method can only achieve so much. What is problematic, however, is the suggestion that experimental methods could achieve much more, and address any kind of problem related with persisting poverty. Without addressing the other big questions that RCTs cannot address, the eradication of global poverty is hardly imaginable.

All methodological qualifications aside, the statement that it is worthwhile to know something substantial about the effects of policy interventions, rather than putting time, effort and resources in interventions that do not bring the desired effects, is certainly convincing. Yet the interpretation of the laureates' contribution by the Nobel committee – sometimes fuelled by randomistas' own statements – likes to make a much bolder case of a "field that studies the causes of global poverty and how best to combat it" (Royal Swedish Academy of Sciences 2019).

Revolutionizing impact evaluation is not equivalent to revolutionizing development economics, development policy, global capitalism or the fight against poverty. "Poverty reduction" is not equivalent to "economic development" more broadly conceived, nor with "development aid" or with "aid policy". Economic development necessarily involves people and societies as agents, not only as randomized study participants, and has a whole lot with politics, power, interests and the functioning of the global economy. Making small steps to improve are welcome and desirable, but the claim that this could be a sufficient ingredient to achieve the large transformations needed to end global poverty, and eventually a good life for all, seems exaggerated.

3.4 The forgotten normative side

So far, we have implicitly assumed that RCTs are probably not the one-and-only solution to achieve economic development, but that they could at least make a reasonable contribution. However, some criticism also points towards potentially detrimental effects. There are several reasons why this may happen.

A considerable part of RCTs carried out in the Global South are not related with "what works" in terms of aid interventions or policies, but with understanding the behaviour of "the poor". Especially in behavioural development economics, one main

question is why the poor do not take up apparently worthwhile investment opportunities (in more nutritious food, preventive health, etc.) but rather buy television sets, spend on festivals and so forth (Banerjee and Duflo 2010; Kremer et al. 2019). As Banerjee and Duflo point out, these "'indulgences' are not the impulsive purchases of people who are not thinking about what they are doing"; but in a world full of insecurity and deprivation, the poor might rather "focus on the here and now, on living their lives as pleasantly as possible" (p. 38). Many (Western) researchers seem unsatisfied with this life choice and have designed RCTs aimed at changing this apparently hedonistic behaviour. For example, it has been tested whether better information about the low likelihood of winning in the lottery changed the demand for participation in Thailand (it did not) (Zenker et al. 2018). This individual example aside, this type of intervention would warrant some reflection on normative implications. Indeed, the emergence of nudging in the rich world has led to an intense debate about the welfare implications of libertarian paternalism. In turn, with the poor as the target of nudges (e.g. in Duflo et al. 2011), these considerations are largely absent. When academic researchers calculate that a more intensive use of fertilizer in Ethiopian agriculture is economically worthwhile, or that higher savings would be beneficial for poor households, this seems sufficient to legitimize an intervention that might interfere with "the basic human need for a pleasant life" (Banerjee and Duflo 2011, p. 37), while questions of autonomy and (democratic) legitimacy are not discussed. This is all the more problematic when researchers have only incomplete information about the motives of the behaviour that a policy aims to change.

Further, there are concerns related to people's and society's agency as opposed to the idea of the economist as a plumber (Duflo 2017). As argued above, economic development is a process of structural and societal change. It is essential that the members of a society play an active role in this process. In contrast, some policy interventions address people in poverty as mere recipients or objects of a randomized project, rather than agents of change. This seems even to imply that they must endure negative consequences when a randomized intervention suggests so: for example, the Nobel Committee understood from one experiment conducted in Kenya (Duflo et al. 2015) that "the employment of contract teachers is generally a cost-effective way to improve student learning" (Royal Swedish Academy of Sciences 2019, p. 3). It is not hard to imagine how such a policy recommendation would lead to disruptions and detrimental side effects for society, as precarious working conditions and instability in the school system can hardly be the goal of development policy. Indeed, Hoffman (2018) reports disruptive consequences of an experiment installed (and abandoned) in Kenya as a replication of the limited-contract teacher project. Now, even if the Nobel Committee's interpretation is not exactly the lesson of the reported study, it points to an important limitation of policy recommendations of experimental studies: these cannot generate policies in any automatic way (Drèze 2018). Normative implications of policies to be implemented must be weighted carefully; policies should not simply be installed on the basis of what, apparently, "works".

4 Concluding remarks

The economics "Nobel" is not awarded to people whose work – and the work of those who follow them – is perfect in all regards and from all perspectives. This is particularly relevant when the prize is awarded at a time when the achievements of the laureates are still very recent and their careers in full swing. I wish to conclude by emphasizing that the academic contributions of Abhijit Banerjee, Esther Duflo and Michael Kremer to the field of economics have been huge in the sense that the field of economics would look very different today, had they decided to become real plumbers or made a completely different vocational decision. All three have devoted enormous efforts to the development and application of a tool that was new to economists. They have achieved a significant increase in attention, in economics and beyond, for questions of economic development and poverty reduction. That said, the merits of their work could be appreciated even more if they were not often presented as the one and only approach to the study of the effectiveness of policy interventions. The Nobel announcement stated that the laureates' "experimental research methods now entirely dominates development economics". If this is true, is it certainly nothing to be happy about, as no single method should dominate the field. In this sense, the field of development economics will now hopefully go on to work towards ways in which the experimental approach to economic development can be brought in (more) active exchange with other methods, social sciences disciplines and people – not objects of randomization.

References

Angelucci, Manuela and Giacomo De Giorgi (2009), Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption? In: American Economic Review, 99 (1), pp. 486–508.

Angrist, Joshua D. and Jörn-Steffen Pischke (2010), The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. In: Journal of Economic Perspectives, 24(2), pp. 3–30.

Bandiera, Oriana (2019), Alleviating poverty with experimental research: The 2019 Nobel laureates, VOX CEPR Policy Portal, 21 October 2019. Available at https://voxeu.org/article/alleviating-poverty-experimental-research-2019-nobel-laureates, (28.10.2019).

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukherji, Marc Shotland and Michael Walton (2016), Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India. In: National Bureau of Economic Research Working Paper 22746.

Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukherji, Marc Shotland and Michael Walton (2017), From proof of concept to scalable policies: challenges and solutions, with an application. In: Journal of Economic Perspectives, 31(4), pp. 73-102.

Banerjee, Abhijit, Shawn Cole, Esther Duflo and Leigh Linden (2007), Remedying education: Evidence from two randomized experiments in India. In: Quarterly Journal of Economics, 122(3), pp. 1235–1264.

Banerjee, Abhijit, Angus Deaton and Esther Duflo (2004), Wealth, Health, and Health Services in Rural Rajasthan. In: American Economic Review, 94(2), pp. 326–330.
Banerjee, Abhijit and Esther Duflo (2009), The experimental approach to development economics. In: Annual Review of Economics, 1(1), pp. 151–178.
Banerjee, Abhijit and Esther Duflo (2011), Poor economics: A radical rethinking of the way to fight global poverty. Public Affairs.
Banerjee, Abhijit, Esther Duflo, Rachel Glennerster and Cynthia Kinnan (2015a), The miracle of microfinance? Evidence from a randomized evaluation. In: American Economic Journal: Applied Economics, 7(1), pp. 22–53.
Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert and Christopher Udry (2015b), Multifaceted program causes lasting progress for the very poor: Evidence from six countries. In: Science, 348(6236): 1260799, DOI: 10.1126/science.1260799.
Banerjee, Abhijit and Lakshmi Iyer (2005), History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. In: American Economic Review, 95(4), pp. 1190–1213.
Beaman, Lori, Esther Duflo, Rohini Pande and Petia Topalova (2012), Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. In: Science, 335(6068), pp. 582–586.
Berndt, Christian (2015), Behavioural economics, experimentalism and the marketization of development. In: Economy and Society, 44(4), pp. 567–591.
Chattopadhyay, Raghabendra and Esther Duflo (2004), Women as policy makers: Evidence from a randomized policy experiment in India. In: Econometrica, 72 (5), pp. 1409–1443.
Deaton, Angus and Nancy Cartwright (2018), Understanding and misunderstanding randomized controlled trials. In: Social Science & Medicine, 210, pp. 2–21.
Drèze, Jean (2018), Evidence, policy, and politics. In: Ideas for India, 3 August 2018. Available at https://www.ideasforindia.in/topics/miscellany/evidence-policy-and-politics.html (29 October 2019).
Duflo, Esther (2001), Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. In: American Economics Review, 91(4), pp. 795–813.
Duflo, Esther (2017), Richard T. Ely Lecture: The Economist as Plumber. In: American Economic Review, 107(5), pp. 1–26.
Duflo, Esther, Rachel Glennerster and Michael Kremer (2007), Using randomization in development economics research: A toolkit. In: Schultz, T. Paul and John A. Strauss (Eds.), Handbook of development economics, Vol. 4. North Holland: pp. 3895–3962.
Duflo, Esther, Pascaline Dupas and Michael Kremer (2011a), Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. In: American Economic Review, 101(5), pp. 1739–1174.
Duflo, Esther, Michael Kremer and Jonathan Robinson (2011b), Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. In: American Economic Review, 101(6), pp. 2350–2390.
Duflo, Esther, Pascaline Dupas and Michael Kremer (2015), School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools. In: Journal of Public Economics, 123, pp. 92-110.
Elbers, Chris and Jan Willem Gunning (2013), Evaluation of development programs: randomized controlled trials or regressions? In: The World Bank Economic Review, 28(3), pp. 432–445.
Fletcher, Jason and Ryne Marksteiner (2017), Causal Spousal Health Spillover Effects and Implications for Program Evaluation. In: American Economic Journal, Economic Policy, 9(4), pp. 144–166.

Haushofer, Johannes and Jeremy Shapiro (2018), The long-term impact of unconditional cash transfers: experimental evidence from Kenya. Poverty Action Lab Working Paper, January 2018. Available at https://www.povertyactionlab.org/sites/default/files/publications/The-long-term-impact-of-conditional-cash-tranfer_Kenya_Haushofer_Shapiro_January2018.pdf, (28.10.2019).

Hoffmann, Nimi (2018), When are experiments corrupt? In: Journal of Contemporary African Studies, 36(4), pp. 532–552.

Imbens, Guido W. (2010), Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). In: Journal of Economic Literature, 48(2), pp. 399–423.

Kremer, Michael (1993), The O-ring theory of economic development. In: Quarterly Journal of Economics, 108(3), pp. 551–575.

Kremer, Michael, Gautam Rao and Frank Schilbach (2019), Behavioral development economics. In: Bernheim, Douglas, David Laibson and Stefano DellaVigna (Eds.), Handbook of Behavioral Economics – Foundations and Applications. Vol. 2. Amsterdam and Oxford: North Holland, pp. 345–458.

Krueger, Alan B. (1999), Experimental estimates of education production functions. In: Quarterly Journal of Economics, 114(2), pp. 497–532.

Leamer, Edward E. (1983), Let's take the con out of econometrics. In: American Economic Review, 73 (1), pp. 31–43.

Mani, Anandi, Sendil Mullainathan, Eldar Shafir and Jiaying Zhao (2013), Poverty impedes cognitive function. In: Science, 341(6149), pp. 976–980.

Meager, Rachael (2019), Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments. In: American Economic Journal: Applied Economics, 11(1), pp. 57–91.

Miguel, Edward and Michael Kremer (2004), Worms: identifying impacts on education and health in the presence of treatment externalities. In: Econometrica, 72(1), pp. 159–217.

Pritchett, Lant and Justin Sandefur (2014), Context Matters for Size: Why External Validity Claims and Development Practice do not Mix. In: Journal of Globalization and Development, 4(2), pp. 161–197.

Reddy, Sanjay (2012), Randomise This! On Poor Economics. In: Review of Agrarian Studies, 2(2), pp. 60–73.

Reddy, Sanjay (2019), 'Economics' Biggest Success Story Is a Cautionary Tale. In: Foreign Policy, 22 October 2019. Available at https://foreignpolicy.com/2019/10/22/economics-development-rcts-esther-duflo-abhijit-banerjee-michael-kremer-nobel/, (28.10.2019).

Royal Swedish Academy of Sciences (2019), Understanding Development and Poverty Alleviation. Scientific Background Report on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019, 14 October 2019. Available at https://www.nobelprize.org/uploads/2019/10/advanced-economicsciencesprize2019.pdf (29.10.2019).

Sukhtankar, Sandip (2017), Replications in development economics. In: American Economic Review, 107(5), pp. 32–36.

Vivalt, Eva (2015), Heterogeneous Treatment Effects in Impact Evaluation. In: American Economic Review, 105(5), pp. 467–470.

Young, Alwyn (2018), Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. In: Quarterly Journal of Economics, 134(2), pp. 557–598.

Zenker, Juliane, Andreas Wagener and Sebastian Vollmer (2016), Better knowledge need not affect behavior: a randomized evaluation of the demand for lottery tickets in rural Thailand. In: The World Bank Economic Review, 32(3), pp. 570–583.

Copyright of ORDO: Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft is the property of De Gruyter and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.