# TRANSCRIPT OF DEVELOPMENT DRUMS [EPISODE 21 – RANDOMIZED EVALUATION]

**Host: Owen Barder. Guest: Rachel Glennerster**

**Listen to the podcast: http://developmentdrums.org/350**

**Owen Barder**
Welcome to a new series of Development Drums, a podcast which takes an in-depth look at key issues in international development. This is Owen Barder from Addis Ababa in Ethiopia. In this 21st edition of Development Drums we'll be taking a close look at rigorous evaluation of development programs and in particular the role of randomized controlled trials.

My guest today is Rachel Glennerster, the Executive Director of The Abdul Latif Jameel Poverty Action Lab, which is a center in the Economics Department of MIT. Rachel has worked on debt relief under the reform of the International Monetary System at the IMF. She's worked on financial regulation at the Harvard Institute for International Development, and before that she worked at the U.K. Treasury.

Rachel, thanks for coming on to Development Drums.

**Rachel Glennerster**
Thank you for inviting me.

**Owen Barder**
Let's start by explaining to people the work of the Jameel Poverty Action Lab J-PAL. It was founded, what, six years ago with a mission of ensuring that poverty policy is based on scientific evidence. What do you mean by scientific evidence and what kind of evidence does J-PAL produce?

**Rachel Glennerster**
So J-PAL is focused on a very particular methodology of using randomized trials to measure the effectiveness of programs. So we take the concept of a randomized trial for medicine but then apply it to poverty programs. And, we use this methodology to answer all sorts of different kinds of questions from how to get more kids in school to how to improve – to examine what's the best way to empower adolescent girls in Bangladesh, reducing corruption, post conflict mitigation, also really all the different kinds of work that's covered by development.

So, we do – we run these randomized trials with partners on the ground. But we also train other people in how to undertake this kind of work and try and disseminate the results of the work so that it feeds into policy. So, our idea is that we could improve the effectiveness of poverty programs by giving people better evidence and more evidence about what works. So we both work on producing the evidence but also trying to get it into policy.

**Owen Barder**
So we'll talk in a second about the benefits and limitations of this approach of using randomized trials as the cornerstone of your work. Let's start by looking at the kind of findings that J-PAL has produced, and you mentioned there a series of policy interventions in the social sphere. Things like, how do you get kids into school or how do you empower girls. Which of these finding so far, and it's early days yet, but which do you think are the most important for poverty policy, and in other words which would make the most difference if they were better understood and taken up by policy makers?

**Rachel Glennerster**
I think the area where we know most, and therefore we have most to say about to policy makers, is in education, because we've tried different things in different countries and are finding some general lessons emerging from really a deep wealth of experiments. That and probably the other area I'd mention is the impact of user fees.

But on education, we've looked at a range of different ways to encourage children to get in school, and there I think quite a surprising result was that improving their health particularly through providing mass school-based de-worming was extremely cheap and effective way of increasing children's attendance at school.

**Owen Barder**
So let's just pause on this de-worming finding. So this is a scheme where kids when they show up at school are given de-worming tablets?

**Rachel Glennerster**
Yes, obviously this is only effective in areas that have high worm prevalence, but that is actually a remarkably large percentage of the developing world's population live in areas where their children are liable to get intestinal worms; about 400 million children in the world have these intestinal worms. And they make children anemic and that makes them less likely to go to school. So, you can mass-treat in areas where there is high worm prevalence and at the cost of about $0.50 a year you can treat children for these worms, make that – once treated they are healthier and more able to go to school.

**Owen Barder**
So the theory isn't that they come to school to get a worm treatment, as it might be say for a school feeding program. The theory is that something about having worms in the health conditions that causes – is stopping kids from going to school, and if you could only treat the worms then that will help them go to school.

**Rachel Glennerster**
Exactly, exactly. As they say they are likely to get anemic, and also the worms eat the nutrition that the child is meant to be receiving. So they get tired and lethargic and that makes them less likely to go to school. And I think health people had known about worms and known about the importance of treating them, but the fact that this is was an extremely effective education intervention was quite new and, as I say when compared with a lot of other ways of getting children in school turned out to be one of the most cost effective.

**Owen Barder**
So what some people might say is that it's a rather reductionist policy view. Because if you did treat people with de- worming tablets or – is it tablet or a powder? A tablet, okay. So if you did treat people with de-worming tablets then you'd have lots more kids wanting to come to school and then there wouldn't be enough schools and there wouldn't be enough teachers and there wouldn't be enough textbooks. So, you can say that it's the most effective thing to do but surely it wouldn't work without all the other things that were already doing a need to go on doing.

**Rachel Glennerster**
Well, that's why I said it's the most effective way of getting kids into school. We have a – we are now learning a lot more about how to improve the effectiveness of education in the classroom which is terribly important because there's a lot of evidence that children are going to school regularly and not learning very much in many countries, many poor countries in the world. A lot of children in India even when they go to school regularly in standard grades three and four, can't master very basic skills of reading, writing, even recognizing letters a substantial proportion find hard. So it's very important to, as you say, improve the quality of education in the classroom. And the interesting thing here is we're getting very similar policy conclusions from a lot of different work. And I think in different countries about how what are the best ways to improve learning in the schools. And I summarize those as teacher accountability, very important, very high levels of teacher absence across the developing world. So, if children are coming to school but the teacher isn't you really haven't moved that far forward. So, ways of making teachers accountable are very important in improving learning.

**Owen Barder**
What's an example of that? I think you were about to tell us, what's an example of it?

**Rachel Glennerster**
Yes. So, in one program, an NGO who was running their own schools, decided to introduce a monitoring system where the teachers had to take a picture of themselves with their children at the beginning and end of each day and their pay was tied to whether or not they could prove that they're actually at the school during the day and that increased teacher attendance also had a big effect on test scores. Obviously that could be implemented by this NGO; very unlikely to be accepted in by civil service teachers. So that experiment was in India. And in another experiment in Kenya, teachers were hired, local teachers were hired by communities and their contract was with the community, not with the central government. These teachers, extra teachers, were hired because of the big increase in enrollment in Kenya.

Lots of other African countries have seen big enrollments as more kids are going to school and school fees are reduced, that's great but as you say leads to lots more kids in the school, huge class sizes. So the communities hired the teachers and could fire them too. And again, those teachers showed up more often – they were less experienced than the government teacher, but they actually got much better test results, not least because they showed up more.

**Owen Barder**
Now, you personally are working I know in Sierra Leone on a community driven development scheme. Tell us about that and what it is that you are studying there.

**Rachel Glennerster**
So, across the world there are lots of community driven development programs, which vary quite a bit, but this is reasonably typical, it's where a community is given a grant to, and a lot of facilitation, a lot of help in thinking what are the needs of the community, facilitation to help them work together, make more participatory decisions; Sierra Leone is an example of a country with a very hierarchical system of decision making. So village chiefs have a lot of power, village elders have a lot of power.

Youths and women are often excluded from decision-making, and that is seen by many as a very important reason for the civil war which can be seen as an inter-generational war rather than interethnic war in Sierra Leone, so a rebellion of youth against this hierarchical system. So, this program was an attempt to provide communities with more decision-making power over development decisions in that community and a grant. But also to try and change the way in which decisions are made in communities and teaching people to be more participatory.

So, here interestingly the outcomes measures that we want to test are not classic economic outcome measures, but have decisions, are decisions made in a more participatory way, is there more trust between different community members? And there really the most interesting part of the work is trying to come up with measures of those difficult to measure concepts.

We've worked a lot with local groups; we spent a lot of time in the field figuring out how to ask these questions, and how to get people to reveal whether in fact decisions are being made in a participatory way. So we had – we actually setup in a sense an experiment in the field where people reveal how they act, we ask communities to make a decision about which gift they would like in response, you know, in exchange for being part of our survey, and observe how they make that decision and use that as one of the outcome measures. Do they call a meeting, does the chief just decide how many women come to the meeting. We're in the process of analyzing these results.

But, as I say, this is a big question in developmental of – is it possible as outsiders to come in and try and change these very deep seated ways in which communities have been working or are they so deep seated that this kind of program is really not very effective.

**Owen Barder**
And what do you think the answer to that question is going to be. I know you haven't published your

results yet.

**Rachel Glennerster**
We are just getting preliminary data back. But we still have one more round of surveys to do. I really don't know what the answer would be.

**Owen Barder**
Interesting. And what about your work on microcredit in India?

**Rachel Glennerster**
So, this is a study which has been getting quite a bit of attention recently. It's very hard to do an impact evaluation of microcredit. I mean this is an attempt, the first attempt to do a randomized evaluation of the impact of microfinance on a whole community. So, what we did is work with the microcredit organization Spandana, who has a pretty typical microcredit product of group, lending in groups to women. They had worked in other parts of India but we're moving into Hyderabad. And they picked a larger number of communities within the city than they could work in, but all of them were ones where they would potentially like to work eventually. And we randomized which communities to move into.

We then compared the whole community that now had access to microfinance with a whole community that didn't have as much access to microfinance to look at the communitywide effects. And we found some pretty interesting results, I'd call them mixed.

In some ways microfinance delivered on its potential, it increased the number of businesses that were started. Interestingly, businesses that already existed were not hurt and indeed benefited from the arrival of microfinance. This was interesting because you could think that microfinance by getting more businesses started would create more competition for the existing businesses and that they would actually do worse. But in fact they did better, they were able to borrow and invest more in that business and they saw profits rise.

But we saw very different results for different kinds of people in the community. As I say, those who already had businesses benefited. Those who were likely, the kinds of people who are likely to start new businesses started more businesses. They invested more in durable goods; they actually cut back on other kinds of consumption. So they weren't financing their business just from the microfinance loan but also from saving their own money and investing that.

Now, a third group of people who were not likely to start businesses, we predict using sort of the initial data, those people increased consumption in general and not investment. So, as I say, you have these three very different responses to microfinance. And, overall, we see no increase in consumption from microfinance.

Some people have seen that as a bad sign for microfinance. But I don't see that as a bad sign at all, because you've got some people who are saving and investing, some people who are using the loan to go out and just consume. We did – we were quite surprised that we saw no impact on empowerment in the – of women in the community, no change in the way decisions were made despite the fact that all these loans were going to women.

We also found no impact on increasing enrolment at school, which is something that a lot of microfinance organizations have said that microfinance could help. But again I'm not surprised with that. Actually in this area a lot of kids go to school already so it would be surprising if you increase that any further.

**Owen Barder**
But overall, this is a more positive story than some of the more recent stories that have said that microfinance has very little impact on investment and growth. I mean you are seeing some people starting new businesses as result of this?

**Rachel Glennerster**

Yes, actually a lot of the newspaper reports about the study are saying there is no impact. You know, there is very little impact. This is, some of the headlines on our study have said, small change, suggesting that this isn't very much impact. I would say one thing to stress is this was a follow-up after 18 months.

So it's quite normal that you would expect people to be investing and you wouldn't see higher income at this point, maybe there will be higher income in the future, we don't know. The potential danger sign is those people who took the loan and just consumed it, are they going to be further in debt in the future? Again, 18 months is probably too short to know for certain whether they are getting themselves into trouble.

**Owen Barder**
All right, now your work, the work of J-PAL generally is mainly about what kinds of policies work, you know, how you hold teachers accountable or what kinds of microfinance institutions have what kind of effect. Rather than a question of whether aid works and whether foreign aid in particular works, although things like your work in Sierra Leone is looking at whether it's possible for outsiders to change hierarchical decision making and structures. But more generally, does the gamut of work that you're doing lead you to a view about whether aid works, whether it could work better, what the overall impact of aid is?

**Rachel Glennerster**
I don't think we can say, in fact I think it's a rather misleading question to say, you know, does aid work? There are so many different things aid money is spent on. We wouldn't expect to have a single answer to that. I do think that it's – our work suggests that we could be much cleverer about using aid. Even also a lot of areas where that we spend aid money on that we don't really have much evidence about their effectiveness.

So, I think there is no question that we can improve the effectiveness of aid by doing more rigorous evaluation, by using aid money on the things that we find are most effective. So, and also, I think an important role for aid is to invest in finding out what works because that's a very useful aid to government who in most parts of the world are the biggest spenders on anti-property to help them make their own policies more effective. But as you say, our work is not just about aid. We see finding out what works as important for aid and donors. But one of our biggest consumers of our work are developing country governments themselves.

**Owen Barder**
You're listening to Development Drums with Owen Barder, and my guest is Rachel Glennerster the Executive Director of the Abdul Latif Jameel Poverty Action Lab. You can subscribe to Development Drums on iTunes, or you can download all the episodes from the Development Drums website at developmentdrums.org. At the website, you can also leave comments about the episodes and you can download transcripts of the shows.

You can also sign up for an email which tells you whenever a new episode is available. Alternatively, you can sign up for our Facebook group, where you'll find information about future episodes; and you can suggest questions you'd like me to ask.

Let's turn to some of the methodological questions; and in particular, the use of randomized evaluations. That's absolutely at the heart of what the Poverty Action Lab focuses on. Tell us why this is important.

**Rachel Glennerster**
Yes, so let me first say that while we focus on randomized evaluations, we don't think that that's the only approach that is useful to evaluating programs. It's what we do, and we think there's more room to spend a lot more time doing and evaluating more programs using this methodology.

Why do we think it's a useful methodology and there's scope for increasing its use? Because it's – as I said in the example of microfinance, it's actually very difficult often to evaluate the impact of a program. If you think about microfinance, a lot of the ways that people have evaluated its impact in the past has been to compare outcomes for women in a community, who take up microfinance; and compare those outcomes

with women in the same community who haven't taken up microfinance. And then the assumption is that any difference in outcomes from those who take up microfinance, is – to those haven't – is due to microfinance.

But if you think about it, women who sign up for microfinance when it arrives in their community are different in very important ways from those who don't take out the program. They tend to be more entrepreneurial, they may be more motivated, they may be bigger risk takers. And all of those factors that made them sign up for microfinance will also affect their outcomes.

So these women may start more businesses, but they may have started more businesses even without microfinance, because it's exactly the ones who want to start a business who sign up for microfinance. So it's virtually impossible, it is impossible, to disentangle those two effects, the effect of microfinance versus the effect of the attributes that makes them want to take up microfinance. It's virtually impossible to distinguish those two.

**Owen Barder**
Let me just put this in my terms and you can tell me whether I've got this right. So we usually can describe reasonably well what has happened to – for example, somebody who takes up microfinance loans. But the difficulty is describing with any degree of certainty what would have happened without some policy interventions.

So we can see that something's happened with the policy intervention, but we can't exactly describe whether the policy has worked, because we can't observe what would happen to those same people if the policy intervention had never existed. Is that…?

**Rachel Glennerster**
Exactly, and randomization gives you a comparison group who are as close as you can get to seeing the group that did get the program, and what would have happened to them if they hadn't got the program.

**Owen Barder**
So explain the importance of randomization in this process. Why does randomization give you a better counterfactual?

**Rachel Glennerster**
Because – take the example of our slums. Instead of comparing women who took up microfinance with those who didn't, we take a group of slums and we randomize which ones Spandana will go into. And we can be sure, if we have large enough sample size, that the slums that they went into were the same in every other aspect to the ones that they did go into, with the one exception that one got the program and one didn't. So you're isolating the single factor that's different is, is the program.

So it's nothing – you're flipping a coin to determine who gets the program. So it can't be determined by anything else, by motivation and closeness to the road on all these other factors, people will be identical except whether they get the program.

**Owen Barder**
So one big problem that often happens in development interventions is you do a pilot study, but you often start off with somewhere where you think a program is likely to work or where you already know somebody in the community and you pick them for that reason. Of course, that is a normal – that – it may turn out that the reason why something works is because of – it's the reason that you knew them, you knew that person in the village that suggests they're more entrepreneurial and outward going. So this takes away all the reasons why you picked one community over another community by picking them at random.

**Rachel Glennerster**
Right, exactly. And when people roll out, they often roll out first to either communities where they know someone, or that are near the road or near the hotel with a hot shower. And those communities may be richer, more educated, and therefore the program may seem to be working. But actually it's just the fact

that people who are richer and more educated are closer to the road and that's why those communities are better off.

**Owen Barder**
Now you were careful at the beginning of this to say that you're not saying that this is the only possible form of rigorous evaluation. But randomized evaluation has at times been called a gold standard. Which to me it conveys the idea that it's a technique that you should use if you can, and you should use a different kind of method only if you have considered randomized evaluation and you can't figure out how to make it work.

**Rachel Glennerster**
Well, I think it's worth making two distinctions here. One is, if you want to do a rigorous impact evaluation, I would say if you can do a randomized evaluation that's going to be better. Virtually every other method of doing a rigorous quantitative impact evaluation are all trying to mimic a randomized evaluation. All the other methodologies that are used by economists are really trying to mimic a randomized evaluation but where you can't.

Now that's very different from saying that those other methodologies aren't useful. Obviously, sometimes you can't do a randomized evaluation. So you can use those. It's also very different from saying that's the only useful method of evaluation, which we would never say.

**Owen Barder**
But you are saying it is the best.

**Rachel Glennerster**
It's the best of the rigorous impact evaluations.

**Owen Barder**
Okay.

**Rachel Glennerster**
But there are a lot of other kinds of evaluation that needs to be done. It's too expensive to do rigorous impact evaluations of every single program out there. It's just not possible. It's much better to focus your resources on doing a few really good impact evaluations. And in your other programs, do other kinds of evaluation, process evaluation. Make sure you know that you did what you were planning to do. The money didn't get stolen.

All of that is evaluation. It's desperately important; it has to be done in every single program. Rigorous impact evaluation should only be done in a few – in enough situations to answer key questions. But they don't need to be done in every single program.

**Owen Barder**
Now there are quite a lot of – I don't know if skeptic is the right word – but there are worries from people about whether randomized evaluation is really appropriate in a wide range of cases. And let's pause on what some of those questions, what some of those challenges are. And one is an ethical concern. You've used a couple of times in this conversation the word 'experiment', and I think a lot of people are uncomfortable with the idea of experimenting on humans. The idea that we pick some group of people and provide largesse to them, and we pick – more importantly, we pick some other people and deny them those benefits. So what do you say to those people who think that randomized evaluations are playing God with people's lives?

**Rachel Glennerster**
We would never ask or work on a program that led to fewer people getting a service than they would if we haven't been working on the evaluation. That would be unethical. We wouldn't do it. We won't be allowed to do it by universities. Where we work is with programs that have limited budgets, that can only reach a certain number of communities, or schools, or individuals. And then in a sense everyone who has worked

in development knows this. You are choosing people to get programs even though many other people need the help. So I'll give you an example. I'm working with Save the Children in Bangladesh to empower adolescent girls. They were planning to work in three sub-districts in Bangladesh even though girls in the neighboring sub-districts needed this program just as much as the girls in those three sub-districts. When we came along and helped them do an evaluation, we suggested that they expand. They are now working in five sub-districts. But instead of getting every single girl in those three sub-districts, they're getting some of the girls in five sub-districts, the total number of girls, absolutely the same; the total number of communities, the same. It's just sort of spreading it out. It's just using a different rule to decide how to allocate.

So, unfortunately, people in development are playing god all the time. This is just a slightly different way of choosing the people that you're going to work with.

**Owen Barder**
I have heard some people in fact say that this is a better way of playing God, because at least it's a throw of the dice and not some un-transparent invisible decision making process.

**Rachel Glennerster**
It's interesting when you – a lot of people are concerned about the ethics of this in developed countries. I've never had a problem in a developing country. When we go to the communities and explain that they are all going to be entered into a lottery to decide who gets the program, people are really pleased. They say this is fair. They say, 'you mean we will actually get a fair chance?' 'We won't have to bribe someone to get this government program?' 'We will have as much chance as our neighbor who is maybe more politically connected?'

So yes, it's seen in most communities as a very fair way of making these decisions. Indeed, we have done evaluations of programs that were randomized by governments. Even though they never thought of doing an evaluation, the randomization was implemented by the government itself because it was seen as the fairest way to make allocation decisions.

**Owen Barder**
So the second class of challenge to this approach is that these kinds of interventions and particularly the social policies, you can't really generalize from one experiment to reach any generalizable conclusion. So and the obvious example of the de-worming program, it clearly makes sense in communities that suffer from that kind of worm. But you clearly couldn't generalize to communities where there aren't worms. And particularly where you are looking, for example, at how communities are empowered by external interventions, it seems quite possible that that only teaches you lessons about that particular community at that particular place at that particular time. To what extent are these kinds of experiments really giving us broader insights? Are they not just a kind of expensive way of doing aid to a particular community with no real generalizable lessons?

**Rachel Glennerster**
Well, the first thing to realize is that potential criticism and it's something we have to take very seriously is true of all rigorous impact evaluations of any programs. So it's not that randomization is less generalizable. It's just whenever you are studying a particular program, you have to worry about generalizability. The second thing is that's something we can test. We can test and we are increasingly doing that the extent to which we find similar results in different kinds of communities.

Now you would, as you say, expect that to be different. De-worming is more likely to generalize to other areas where kids have worms because our bodies all work the same way. And if we are anemic, we are tired whatever country we live in. So that's more likely to generalize and indeed that is an example where people have tested it and seen whether it generalized. And then you get remarkably similar results in urban India when you address anemia, rural Kenya, and the South of the United States when hook worm was abolished. So three different continents, three very similar results.

In terms of empowerment, I would agree with you, that that scenario where we would have to be very

careful and test different things, we haven't done a series of tests on empowerment in different context. In other places though, as I say in improving education, we found that very similar approaches have worked in India and Africa, for example. So we keep doing this. To a remarkable extent when we test something in a different context we've found very similar results. So, I don't know, up to – well, let's take empowerment we have done the same thing in two different states, very different states within India. Maybe you say, well, it's all within India, but one with very high, much higher than average female literacy, one with lower than average female literacy. The same program had very similar results. I think the other thing you have to be careful with is thinking about how the program works. Understanding in detail how it works helps you to understand how it might generalize. Also testing deep fundamental principles rather than just kind of a surface program also helps you to get at the underlying reasons why something may work and therefore help you to think about how it generalizes.

**Owen Barder**
That rather segues into a third criticism, which is that this approach has tended to focus too much on whether projects work and not enough on why they work. It's generating kind of evidence about the success rates and failure rates, but not enough on – there may be small variations in the way that particular projects are implemented and followed through that hugely affect whether they are effective or not and this approach doesn't really pick that up enough that is not – it's not fine grained enough to pick up those different nuances on how programs are implemented?

**Rachel Glennerster**
Right, it certainly is the case that testing a whole variety of different ways of implementing a particular program takes a lot of sample size and therefore is hard to do. But actually, that's not our only way of getting evidence. When we do this kind of project, we collect an awful lot of information about the mechanisms through which it works. And we can see – in the chain of causation we can see if a project fails where it fails. We've worked on empowering humanities to demand better education in India.

We can see that the program did lead to increased information about how to take action. They did know more about who to go to. But they then didn't take any steps. And we can look at a whole range of steps and see which ones they took and which ones they didn't take. And that can help us understand how and why a program may fail. So we get a wealth of information from these evaluations, which may not be covered in the headlines, but if you think about my discussion on microfinance, we're looking in quite a lot of detail about who responded and how they responded and what happened and that can help us understand some of these more deep lessons.

We are always testing a particular program and policy. So if it isn't implemented well, you don't maybe know whether it would have worked if it had been implemented better. But we do a range of things. Sometimes we try and test a best possible case scenario of this type of policy and other times you want to test a sort of standard run of the mill to see does this kind of program work in the way that it's normally implemented. And those will give you very different answers.

**Owen Barder**
So a fourth challenge is the idea that this approach applies reasonably well to small questions. How do you get more teachers in school or how do you get farmers to use new seeds. But that it doesn't apply to the big question how do you stimulate economic growth, how can you make government more accountable and in fact one of our listeners, CAFOD, have asked, would you agree that randomized controlled testing approach is difficult to apply to advocacy programs which seek to promote policy change. So that's a nuance on the same idea that this is good for measuring the impact of the kinds of programs where what you are doing is delivering widgets, so getting kids in school, but not for the big kind of social changes that probably underlie development process.

**Rachel Glennerster**
I wouldn't agree with that. There obviously are some things that we can't test in this approach. You know, fixed versus flexible exchange rates, it would be difficult to randomize because the unit would be a country. On the other hand, there are assumptions that go into deciding whether you should have a fixed versus floating exchange rate that you would be able to test.

But let's look at some of the things we've tested. Maybe you argue these are small, having user fees or not for education and health is something that we can very easily test. And how to improve health systems delivery, how to improve teacher accountability, now I don't think those, maybe some people think those are small questions, I don't think they are small questions. I think they are fundamental to how do you make effective government.

Again do you have closets for woman? How do you best combat corruption? Is it from the bottom up with community empowerment or from top down with more auditing? These are questions that policy makers are struggling with and there were big debates, again what's the best way to promote HIV prevention. In some ways, is there any bigger question than that in some parts of Africa. So it's true that we are not directly answering a question like how do you make a country grow. But I would argue that's too big a question for anyone to be able to answer very simply or clearly.

What we do, is we take that and to take that question and divide it up into chunks. You know, big part of growth is improving agricultural productivity, what you'd need to do to improve agricultural productivity? We can answer that question. So, I'd argue we've been looking for simple answers to too big questions. We need to take those big questions, divide them up into sections, test them rigorously, build them back together and then you will have very specific things you can do to get growth. You know, is it micro finance? Is that going to be part of the answer or not? So, yes they are smaller than maybe these big questions, but I think in the end, they build up to important policy questions.

**Owen Barder**
And your answer to CAFOD's question about…

**Rachel Glennerster**
Oh, advocacy.

**Owen Barder**
Testing advocacy, right.

**Rachel Glennerster**
Yes, that is a hard question to answer. But it's not impossible, I think, we've gone into areas that were seen a few years ago as impossible to use this methodology to answer. Advocacy is something where we are starting to sit down with the people who do it and brainstorm about how we can use this methodology to answer the question. Again, you may be only be able to get a bits of it, but we always encourage people who are, who we work with to take what they do and break it down into segments and test different bits together. And you may not get all of it, you may not be able to test all of the program, but there were definitely going to be bits of advocacy that you could test in this way.

**Owen Barder**
You've mentioned a couple of times, your research into user fees in health and education. What approach have you been taking on user fees and what are your findings so far?

**Rachel Glennerster**
So we've looked the different arguments both for and against user fees. The argument against is that user fees depress demand for critical health and education services a lot, especially by the poor. Those in favour have argued that user fees can help target products and services to those who need the most. It can make sure that those – if someone pays for something they're more likely to use it. So it makes more likely these products will be used. So we've tested those points. The way that's done is by either randomizing the price at which people are offered products individually or you know, one clinic will be will give out bed nets for free and other clinic will charge a small user fee for them. And then we can follow up and see how much do the user fee depress demand and get paying for it, make people more likely to use it.

And overall, the response is actually very consistent and very clear across countries and across products, small user fees have a very big impact on demand. So to give you one example, going from zero to $0.75 for insecticide-treated bed nets in maternity clinics in Kenya reduced demand by 75%. So $0.75 is a tiny

fraction of the cost of the bed net and yet 75% of mothers were not taking the bed net if they had to pay $0.75 for it.

Very similar effects in de-worming. You get a sharp drop in attendance at school if people have to pay for uniforms which were a cost of going to school, so similar to save charging a fee for school. Chlorine, so lots of different products in Zambia and Kenya, both very similar results.

**Owen Barder**
And the result is that at around zero, the impact of even a very, very small price is to deter an enormous number of people from taking up the service?

**Rachel Glennerster**
Yes, and it has none of positive effects that people had hopes they would have. So it doesn't help you target products on those who need the most. In most of the cases, there was no screening effect i.e. those who were not likely to use it not, you know, screening those out by charging. There was no evidence of the psychological effect, that if you pay for something you're more likely to use it.

**Owen Barder**
What about the impact on the suppliers, because it's often said that is people who have the commercial incentive they will be more active in getting people to use a bed net or to take it up, you'll create a chain of people with an interest in higher take up?

**Rachel Glennerster** Well it certainly doesn't promote take up. It's not helping the suppliers promote take up and it's not helping the suppliers promote use. But the one thing that could possibly happen is that the providers don't turn up to work if they can't pocket the user fees. And that's something that is worth looking in to. I would argue that given the huge negative impact on demand of user fees it's much better to try and motivate your providers to show up to work in other ways and we talked about some of those other ways earlier.

So, because if you use user fees as a way to get your suppliers to show up to work, you're having a terrible effect on the poor not getting the product. So, better to come up with other ways to motivate suppliers, and there is also issues of potential corruption of people announcing that they're getting rid of user fees but providers on the ground still requiring bribes to give them the service. Those are complicated difficult issues that we're working to look into and to try and find ways to address. So I am not saying that all the questions about user fees have been answered, but it's clearly very bad for take up.

**Owen Barder**
One of you colleagues Esther Duflo, he has recently been awarded one of the MacArthur grants, said a few years ago, I quote, "creating a culture in which rigorous randomized evaluations are promoted, encouraged and financed has the potential to revolutionize social policy during the 21st century, just as randomized trials, revolutionized medicine during 20th century."

Is that a way you see this going? Do you see there being the kind of change in development policy in this century that we saw in medicine last century?

**Rachel Glennerster**
I'm very encouraged by what I've seen over the last few years. There was a lot of concern that people in development wouldn't want to see rigorous randomized evaluations. NGOs would be nervous that they would be proved to not be having an effect. Instead we've seen NGOs and governments take up this idea with a lot of enthusiasm. We've seen an incredible burst of interest in this methodology. We've run training programs to train other people to do it and that are always massively oversubscribed. My phone is ringing off the hook with people wanting to do this kind of work, and also I think most encouragingly we are seeing governments come to us and say, tell us what you are finding, tell us what works, we want to improve the effectiveness of our programs and we want to learn from the best possible evidence.

So I am seeing a revolution, and a few years ago none of the big international NGOs were doing very much

of this kind of work and now most of them are doing this kind of work. Similarly, governments or donor governments around the world are showing a lot of interest. The World Bank is doing hundreds of these rigorous randomized trials. It started in the Africa region, within a couple of years they were doing a hundred across Africa. They are working with governments in Africa. Most recently at a workshop in Addis we had governments from across Africa and the enthusiasm from the representatives from governments to do this kind of work was just really inspiring.

**Owen Barder**
It seemed in the early days that a lot of this was driven by evaluators to whom this kind of work could have been outsourced and that very few organizations were thinking strategically about how randomized control trials could improve their strategy, could influence their operations. Do you think that's changing? Is this getting into the strategic life blood of decision making by both by governments and by aid agencies or is it still a bit of a nice to have add-on that some people who are interested are experimenting little bit on an increasing scale?

**Rachel Glennerster**
I think when an organization first takes on this kind of work they tend to pick a program where there happens to be interest and not think so strategically about what's my most important program to test. But then over time, as they get more comfortable with the approach you see them being more strategic. I think an example of this kind of strategy is our work that we are just launching with the Gates Foundation Agriculture group. To look systematically about what do we know and what do we not know, about how to promote the use of new technologies that could help small farmers in Africa. What do we know? What do we not know? Where are the gaps? Let's fund the research to fill those gaps.

We've done a similar project with the Nike Foundation. What do we know? What do we not know about empowering adolescent girls? What are the questions? Let's design some evaluations to answer those specific questions. So again the World Bank is thinking much more strategically, a lot of people want to do this work and they are choosing the projects that answer their most important questions. So I think it's unsurprising that when people do their first one, they kind of experiment around the edges, but as they get more comfortable, they think more systematically about this.

**Owen Barder**
And one last question which comes from another listener Jimmy Greer asked on Twitter. What, if possible, do you think is the single biggest objection or obstacle if there is one to scaling up randomized trials?

**Rachel Glennerster**
I mean I would say, as I said before you don't want to do these everywhere. So they are expensive to do, not that randomized trials are more expensive than other rigorous impact evaluations, but rigorous impact evaluations are expensive. So I don't want, I think they can be scaled up from where they are now, but I don't want every projects in every place to be doing them.

In terms of – maybe that's a hard one for me to answer, of course. I don't see many of the criticisms as fundamental to saying these shouldn't be done. Obviously they need to be done in the right place at the right time to answer sensible questions, because they are expensive and that's how I see the future of these as part of an overall evaluation strategy, answering your biggest question that you most want answered in an accurate way.

**Owen Barder**
Rachel Glennerster thanks for joining me on Development Drums.

**Rachel Glennerster**
Thank you very much.

**Owen Barder**
You have been listening to Development Drums and my guest today has been Rachel Glennerster, the Executive Director of the Abdul Latif Jameel Poverty Action Lab. My guests next time will be Roger

Thurow and Scott Kilman authors of ‘Enough’, a book about why the world's poorest starve in an age of plenty. And let me say welcome and thanks to Anna Scott, the new producer of Development Drums and thanks to Development Initiatives.

If you want to suggest topics for this series of Development Drums please go to our website at developmentdrums.org or go to our Facebook group.

From me, Owen Barder in Addis Ababa, thanks for listening.

*****