


PODCASTS

READING TIME 36 MINS

## EDWARD MIGUEL ON THE "REPLICATION CRISIS" IN ECONOMICS AND HOW TO FIX IT: TWO THINK MINIMUM

September 28, 2021  [Robert Hahn](#) and [Edward Miguel](#)

EVIDENCE-BASED POLICY



Two Think Minimum  
Edward Miguel on the "Replication Crisis" in Ec...

SOUNDCLOUD

Share

3:24 48:41

Cookie policy 249

[Two Think Minimum](#) · [Edward Miguel on the "Replication Crisis" in Economics and How to Fix It](#)

Dr. Bob Hahn:

Hello, and welcome to the Technology Policy Institute's podcast, *Two Think Minimum*. I'm your host, Bob Hahn. Today is September 13th, 2021, and I will be speaking with Professor Edward Miguel, who is the Oxfam Professor of Environmental and Resource Economics and Faculty Director of the Center for Effective Global Action at the University of California, Berkeley. We will be talking about his book, *Transparent and Reproducible Social Science Research*, written with Garrett Christenson and Jeremy Freese. This podcast is part of our series on evidence-based policy. Ted, welcome to *Two Think Minimum*.

Dr. Edward Miguel:

Thanks, Bob, for having me.

Dr. Bob Hahn:

So, let me start out with a simple question. Why did you and your coauthors take on the subject and write this book? Because I know it's a big undertaking.

Dr. Edward Miguel:

It definitely takes time to write a book and a lot of thought. It's something that we talked about for a couple of years before doing. Garrett, Jeremy, and I over the last decade, and even more than a decade now, have been pushing our own disciplines. Jeremy's in sociology. Garrett and I are economists. We've been pushing our disciplines towards open science, towards research transparency in various ways, and we felt the need, we really felt compelled to put a lot of what we were doing and learning in one place to make it accessible so that students and policymakers and seasoned researchers had a go-to resource for all the progress that was being made in this area. So, one of the kind of cool aspects of this is since we do come from different disciplines, it forced us to use examples and language that we hope travel broadly across the social sciences. Everywhere from, I think advanced undergrads up to seasoned researchers will hopefully find the book useful.

Dr. Bob Hahn:

So, in the first part of the book, you define the problem, and in the second part of the book, you neatly solve the problem. So, let's hold off on how you neatly solve the problem until you help our listeners learn about what the problem is exactly. So, do you want to take a swipe at that?

Dr. Edward Miguel:

Definitely. The starting point for a lot of what we're doing in the book, and really for the open science movement as a whole, is the growing body of evidence and the number of high-profile cases where scientific research fails. It fails to produce credible work. There've been high profile instances of research fraud. There is more and more evidence of publication bias, meaning certain types of research tends to get published, maybe because

statistically significant findings are more likely to be published, or findings that conform to other researchers' expectations are more likely to be published.

So, there's this growing body of evidence that these issues are really major concerns in economics and sociology and in other fields. And again, in the first part of the book, we document some of those problems. And even since, in the last year, year or two, since we wrote the book, there's just more and more evidence accumulating about how pervasive things like publication bias really are.

So, that was a starting point. There definitely have been some high-profile instances of fraud that have gotten a lot of attention. Just a couple of weeks ago, and I'm not sure, Bob, how much you've heard about it, but there was a very high-profile case of a famous psychologist, who's also well-known to economists because he works in the behavioral economics space, some of whose work has been found to possibly be fraudulent. Data might have been made up. He hasn't admitted to anything, but it really looks like there's growing evidence that data was probably made up. There've been other instances of the same kind in recent years in political science and psychology. So, when you see those sorts of cases of very high-profile research being published in leading journals, that had data that was made up, that's really like an alarm bell that should be going on for all of us doing social science research that we need to do better, and we need to detect these problems earlier.

Dr. Bob Hahn:

Let's drill down on that a little bit if I might. You mentioned in the book, I think, I can't remember exactly, probably reflecting my age, but you mentioned, I think this professor in the Netherlands who had either fabricated data or something, and that was high profile a few years back, I guess when you were writing the book. How big a deal... is fraud what it's about here, in your opinion, and how big of a problem is fraud, and what do you mean by publication bias?

Dr. Edward Miguel:

Yeah. Fraud is one of the problems. My own sense, as an active researcher and in being involved in a lot of research and reviewing and whatnot, is I think fraud is an issue for a very small share of research. So, like people making up their data, that's not the norm, and I certainly wouldn't want people to latch onto what we've written and say, "Oh my God, these guys are saying most data's made up." Definitely not, definitely not. But there are such instances. We see those instances of fraud as sort of like the tip of the iceberg or an indication that our practices are not where they need to be. So, we should be able to detect fraud. We should be able to discover those things, and very often it appears that we don't, or we don't for many years. So yes, there are high-profile instances of fraud.

Dr. Edward Miguel:

We don't think that's the main problem, but when those cases do occur, they're deeply corrosive to science. They get a lot of attention. They, I think jeopardized a lot of other research. They lead the public to question the credibility of our work. So, I think they're just very, very destructive.

Publication bias is far more pervasive, and I think it's more, a lot of the tools that we developed to try to deal with these problems are really focused on more of this run of the mill publication bias. So, for research studies to attain a high profile, very often, they need to be published in leading journals. They need to go through the peer review process where other scientists, other economists, other leading scholars, get to examine that work, ask questions to the author, and then eventually an editor makes the decision that it is high quality work that belongs in a leading journal.

And that has been a pillar of the scientific process for centuries. I mean, you know, the journal nature goes back to the time of Isaac Newton, where researchers would send in letters and reports to editors, and then they would decide to publish some of them. So, this is a kind of longstanding process, incredibly important for determining scholars' advancement in the field, promotions, getting jobs, et cetera. So, a tremendous amount of effort is focused on publication. Those sort of concerns around publication bias revolve around findings that, or evidence that, certain types of scientific findings are much more likely to be published than others, even if the underlying research design and data quality

are similar. And so, the kind of most famous manifestation of this problem is the issue that empirical work, so say statistical work that attains a high level of statistical significance, meaning we can reject that there's no effect in the data. Those studies are much more likely to be published than studies that show null effects.

So, null effects are, "Well in the data, we can't reject that there was no relationship between Variable A and Variable B." Now, overall, we should care about research design, the question being asked, the quality of the data, and if a really well-designed experiment, for instance, yields a null result, that's really valuable scientifically, and those results should see the light of day and get published.

But there's a lot of evidence that null results have a hard time getting published, that editors and journal referees are really drawn to highly statistically significant results. And the problem with that is imagine there's a research area where there's a big question. "Is there a relationship between Variable A and Variable B?" And a lot of people are trying to study it.

And if all the good studies that were conducted got published, we'd see clearly, "Look, the bulk of the evidence shows no relationship between A and B." But if there is publication bias, it means those null results are less likely to be published, but a potentially spurious kind of, you know, random result, that's just obtained through chance, that shows a relationship between A and B, gets published. Now, all of a sudden, in the scientific record, there's this evidence saying A and B are highly correlated, or A causes B, even though that's not true. And that matters for public policy because then policymakers are looking at the research literature and saying, "Oh, you know, this program," let's say variable A is a public policy program, "this program improves child health, or this program reduces child hunger."

If that's a spurious finding and it's getting published, policymakers could make the wrong decision. They could make bad choices. They could put money into programs that don't work. So, the stakes are very high. And, you know, as I mentioned, just in the last few years, there's growing evidence that publication bias is even more of a problem than we thought it was before.

Bob Hahn:

So, in the book you talk about this, I assume he's a professor, Robert Rosenthal, who coined the term, "the file drawer problem." Is that related to this problem with publication bias that you're talking about?

Dr. Edward Miguel:

Exactly. So, Bob Rosenthal's work from really decades ago started to shine a spotlight on this issue. There was even some earlier work back in the fifties by Sterling that showed the same thing and said, "Wait a minute! Why are all the articles being published in psychology journals showing statistically significant results? In my own lab, like most of the time when we run stuff, we don't find relationships. Why is everything that's published, like literally 90 something percent of articles, significant?" People have pointed out the same thing for economics articles, for political science, sociology.

The traditional cutoff for determining statistical significance is 95% confidence. So, if you're at least 95% confident that you're rejecting the null hypothesis of no effect, then that means it's significant. Those were the results that were more likely to be published. And so, Bob Rosenthal, as you mentioned, coined this term, "the file drawer problem," that researchers or people in labs were running all these experiments, but if they didn't get statistical significance, those results just sat in a file drawer where it never got published.

I mean, related to this publication bias issue, there's some very interesting and damning research that plots out the P-values, the significance levels of published findings across different disciplines and shows for instance, that there are like huge spikes in the density of findings that have statistical significance just below 5%, and there's really no naturally occurring distribution of P-values that would look that way. So, that's just very strong evidence that there's widespread publication bias.

There have also been some other studies that have been done that have looked at all research that was funded from a particular grant. So, the National Science Foundation in the US funds tons of research, social science research, and in particular, they funded, you

know, starting around a decade ago, these experiments called "time use experiments in the social sciences," and they funded scores of studies.

Researchers went back and tracked what had happened to all of these research teams that had applied for money and gotten funding. These were high quality studies funded by the NSF, just to see what had happened to the research, including like what types of statistical findings they had found, and so they were able to classify the statistical findings of these studies into statistically significant findings, stronger findings, and then also null findings. And when they looked several years later, the majority of the null findings had never been published. Actually, the majority had never even been written up in working paper format. They'd never been circulated to colleagues.

But the majority of statistically significant findings had been published, and that's really like the essence of publication bias.

Dr. Bob Hahn:

Is that "the file drawer problem?" You don't publish the stuff that's not significant?

Dr. Edward Miguel:

Exactly, that's "the file drawer problem." "The file drawer problem" could take a few different forms. So, the most common one is, you know, you don't publish stuff that's not significant. Sometimes people get significant results, but they're unexpected. They go in the wrong direction, theoretically, and I think all of us in economics know about this. Economic theory makes some predictions. My colleague here, Dave Card, did famous work with Alan Krueger in the nineties on the minimum wage and was finding unexpected results that raising the minimum wage did not hurt employment, et cetera. And there was initially a lot of pushback against that work from economists being like, "Wait, this can't be right. This data can't be right." So, that's another type of file drawer problem. If you produce evidence, maybe it's significant or not, but you produce evidence that goes against the views or beliefs or theories of your colleagues, it may be hard to publish that work. Even if you have a really good research design and high quality data.

Dr. Bob Hahn:

I want to talk with you a little bit later about that, because you had a lot of great examples related to the minimum wage, and I have some fairly strong views about that. Not that I've done any path breaking research in the area, but I have my opinions.

Dr. Edward Miguel

That's great.

Dr. Bob Hahn:

So, you penned an article, I think in Bloomberg, a few years back saying economics is having a replication crisis, and I seem to recall an article that I think you cited in your book, by Kamer and some other people, that talked about the fact that they were able to replicate a non-trivial number of laboratory experiments or something like that. So, you know, did Bloomberg invent that title for you, or is that what you believe, and where's the evidence?

Dr. Edward Miguel:

Yeah. So, you mentioned the Kamer article, and there's others sort of since then. There's work by Andrews and Casey, and then there's some related work by Brian Nosek and colleagues and psychologists. There's been this kind of cottage industry around trying to understand replication, and I see it as a glass half-full, half-empty situation. So, in some of the psychology studies, they find that "only a third of studies replicate," meaning they get something even close to the original estimate and a lab experiment. The Kamer et al. is more like two thirds of economic studies broadly replicate and a third don't. So, I don't know. Is that a crisis or not? I think it depends on the perspective of the viewer, like two-thirds of the economics experiments replicating is better than a third of psychology, we could be like, "Well, you know, maybe on average, social psychology has had even more of a crisis."



And some of these really high-profile examples of fraud and whatnot are also coming out of social psychology. So, that feels like a field where there's been a lot of hair pulling and angst about replication, and maybe economics experiments are doing a bit better. And in the area that I work in, in development economics, we do large scale field experiments. There's even some more recent evidence of even less publication bias and fewer problems with replication. Some of Abel Brodeur's recent work shows that in RCTs, these large-scale field experiments, we don't even have that spike at the P-values right around 0.05, that other types of empirical work have.

So, some fields are doing better than others. There's a variety of patterns. So, I see it as a glass half-full, half-empty situation, but there are serious problems. Whether that's in a third of studies, or 20% of studies, or 50% of the studies, we could debate depending on the sub-field, but there is a problem. A lot of research is falling short. By temperament, I'm actually more of an optimist, Bob, and kind of forward-looking. I tend to focus on the positive. So, I see that we're making progress. The tools we talk about in the latter half of the book are starting to really make a difference. So, I'm not someone to focus too much on the negative, but we can't ignore the problems that are there.

Dr. Bob Hahn:

I'm going to press you on that a little bit as an optimist and talk about an article that you talk about in the book, by John Ioannidis, which is titled, "Why Most Published Research Findings are False." Are you persuaded by that argument, or are you more persuaded by the fact that we have a problem here that needs to be addressed?

Dr. Edward Miguel:

Yeah. So, whether it's most or not, again, I think it is going to depend on the subfield and the time period of studies. I think he uses some basically logic. I mean, it's a really interesting piece. Ioannidis' piece is one of the most influential, most cited scientific papers of the last 20 years. And he says, "Well, under these assumptions that seem pretty plausible, the majority of published findings actually would be false when we're thinking about, you know, significance versus null findings and things like that." You know, certainly

in realms where null findings can't get published or like almost never published, his claim would be plausible. I mean, if you never saw null findings, you only saw like spurious false positives, then he'd probably be right. In my own field of development economics, and I think this is true for lots of empirical labor and other fields, I'm not convinced that the majority of findings are false at all. And you know, for those of us that do large scale field experiments that take years to do, where maybe millions of dollars are spent on the intervention and the research, like we're going to write up those results and get them published if that's been like our main activity over several years. And I think that helps deal with some of the concerns over publication bias. I'm more concerned with smaller scale experiments and whatnot, where let's say a psych lab or experimental economics lab could do dozens of these a year, and so if one experiment with 200 people "fails" and it never gets published, it disappears into the file drawer. So, I do think it's field specific, and again, in my own field of development, people tend to write up the work they spent years working on.

Dr. Bob Hahn:

Let's talk about that a little bit, and maybe you can unpack this a little bit for our listeners. You mentioned the acronym RCTs. I just want you to explain that for 30 seconds or a minute, and you mentioned the words field experiment. And then once you explain that, maybe you can give an example of work you're doing or have recently done. I have a question about replication in this area.

Dr. Edward Miguel:

For sure. So, you know, RCTs or randomized controlled trials, the methodology of carrying out RCTs has a long pedigree going back to agricultural experiments over a century ago, and then health experiments. I think people are very familiar with medical trials, maybe more than ever in this era of COVID. We're all following the latest evidence from the latest Pfizer trial or whatever. So, you know, in a randomized controlled trial, there's a population that could be, say people in our case, in development economics, maybe there's households that are eligible or could receive a cash transfer to try to help them get out of poverty.

And a random subset of those households are chosen to get the cash transfer, and then there's a control group that doesn't. In a medical trial half would get a vaccine, half wouldn't. And then what we do is we follow up with that population over time to try to estimate how the paths of these two groups, or there could be multiple treatment groups, how their paths diverge over time. Are there differences in their outcomes? And then we use statistical analysis to see if those differences are meaningful between the groups.

We, starting in the 1990s and development economics, and I was very lucky to be involved first as a research assistant with Michael Kramer, and then co-author, and then carrying out my own RCTs. I was very lucky to be involved in the early wave of field experiments and development starting in the mid-nineties. I was mainly working in Kenya, and then [inaudible] and others in India, and Paul Gertler and others in Mexico, were carrying out these early field experiments that apply this methodology to economics interventions, interventions relevant to economics. It could be cash transfers. Some of the early work we did was in school. So, projects on the economics of education, and we carry out this methodology. There could be a population in need. There are limited resources with which to assist this population. So, the beneficiaries or the recipients of a certain intervention are chosen randomly and compared to a control group.

You mentioned some stuff I've been working on, Bob. I've worked on a bunch of field experiments through the years. A recent one where some work just got published in the last year, focuses on electrification in Kenya, rural electrification, huge policy issue. How does making this big infrastructure investment in electrification effect households' economic outcomes and other outcomes? I mean, it could affect kids' schooling. It can affect health. It could affect, you know, lots of things. We studied this in collaboration with the utility in Kenya, Kenya Power, as well as the Rural Electrification Authority. So, we were very lucky and put some blood, sweat, and tears into forging this collaboration with government entities, and they were willing to randomize subsidies that allowed households to get connected to the national electricity grid.

We did this across 150 villages. So, it was like a pretty large-scale experiment. Some villages got electrification, others didn't, and then we followed up the economic and social and educational outcomes of thousands of households over four to five years to say,

"What happened when households got connected to electricity?" So, that's kind of a pretty typical example of what a modern field experiment might look like in development economics today.

We're focusing on a policy issue of importance. We're trying to understand costs and benefits, and also trying to understand behavioral choices in ways that are relevant to economic theory.

Dr. Bob Hahn

So, is there a big takeaway or you guys are still analyzing the data?

Dr. Edward Miguel

Yeah, there was, and we published a paper last year. We published two papers last year, one in the *Journal of Political Economy* and one in the *Journal of Economic Perspectives* on it. Our own prior view going into this experiment was electrification is going to be transformative. It's gotta be, right? It's this big, you know, it opens up all these possibilities. It's very expensive. It over a thousand dollars per household to connect a household just because of the equipment and the staffing and the planning, designing, et cetera. All those phases are pretty expensive, and we went in and measured outcomes.

And the big takeaway after four to five years is really minimal economic impacts on rural Kenyan households of electrification. So, the effects were kind of much more muted than we had expected. On most outcomes, we don't have statistically significant effects. We have impact estimates close to zero. Of course, the question that went through our heads and everybody else is like, "Wait, what went wrong? You give households electricity. Why don't they start a business and just do all this new stuff? Why aren't their kids learning more in school?" We also don't find test score impacts on kids' outcomes.

And what we found in our survey data was these households in rural Kenya are really poor, and even after getting electrification, they typically don't buy more appliances. Like they can't afford a refrigerator. So, even if we give them a thousand dollar subsidy to get connected to the grid, if they can't buy the equipment they would need for a business, or a

refrigerator, or other appliances, they don't do much. Basically, what they do is they buy bulbs, electrical bulbs. So, they have light. That's a positive, and they can charge their cell phones. So, that's a positive, but that's pretty much all that the median household uses it for in these poor rural Kenyan regions. There's a big contrast with cities in Kenya, where people are a lot richer. There, they use lots of power and there are big benefits, you know, industry commercial benefits, et cetera.

So, we are very focused on this issue and the paper of rural electrification, we want to make the contrast with electrification in richer populations where almost certainly there would be larger impacts, but at least in the short run, it doesn't feel like electrifying poor, rural, agrarian households is going to do very much for them unless you also subsidized appliance purchases and subsidized consumption of electricity, et cetera. So, that was the takeaway, and it's a kind of surprising finding. Part of the reason I mentioned this example is it gets back to this whole question about publication. We managed to get our paper published despite the fact that we had lots of null results, and a lot of scholars would be very concerned to write the paper that we wrote and say, "Hey, your prior beliefs about the benefit of this intervention are wrong. We have a null result."

A lot of people would think that's sort of the kiss of death for publication. Now, we managed to get the paper published. And maybe we got a little bit lucky. One of the things that we did in this project, and it relates to the second half of the book and the tools for solving these problems, is we wrote something called a pre-analysis plan before we analyze the data. So, that's one of the tools that we advocate for in the book that we think could be very valuable for the research community. The pre-analysis plan laid out the kinds of hypotheses we had, and the statistical analysis we would use to test them. So, when we write up our paper, we say, "Hey, everybody kind of agrees these are important outcomes, and we designed our study to have adequate statistical power, and here are results, and they're null." And so it brings a bit of scientific rigor into the process. It assures the readers of the paper that we're not just data mining or doing things to get a certain result. So, that probably did help us get published, I think.

Dr. Bob Hahn:

So let's talk about, well, I want to talk about two things. One is you mentioned this idea of field experiment or going out into the field and doing an experiment, say on electrification in Kenya. It strikes me, and just listening to you, that's not going to be so easy to replicate because one, it's expensive. Two, you got to have the agreement of utilities, and who knows where to do it? So, is that a big issue in and of itself, or is that kind of separate from what you're talking about in the book?

Dr. Edward Miguel:

I think it's very related. I totally agree. And even beyond the very important issues you raised. Over time, conditions changed. So, it's not... a lab experiment is a bit different. You put people in the lab with the exact same protocol. It feels like the same choice problem. But Kenya today is actually quite different than Kenya when we started our project six or seven years ago. Because in the interim, the government has rolled out mass, rural electrification, even though there wasn't much evidence of benefit to it. And so the whole choice environment and people's exposure to electrification and whether their neighbors have it or don't has changed. And so, that's actually another rationale for having pre-analysis plans. If you're doing a big field experiment that's really expensive, where it's hard to forge partnerships, where there's a dynamic environment, it's pretty important to get it right the first time. Because there may not be an equivalent setting somewhere else.

And you're going to want to understand how things work right here right now, and that's a strong reason to pre-specify what you're doing and try to make that analysis as credible as possible. But yeah, you're right. Like replicating that in other settings could be impossible.

You know, the word replication is a bit tricky because different people use it for different things and actually different fields. It has different meanings in psychology versus econ, et cetera. There's the sort of most basic form of replication, which would just be to like grab our data. And then literally reproduce the analysis of the published paper. People these days are calling that, just to be precise, they're calling it computational reproducibility. So, that kind of lays that out. Sometimes people call that verification, but then there's this broader issue of replicability that ties into external validity. Like here's a finding, does it

replicate? Meaning, does it travel somewhere else? Is there external validity? And that could be really difficult to achieve for the reasons that you were saying.

Dr. Bob Hahn:

So, you cite one of my favorite papers in your book, Ed Leamer's... just because of the title. It's a great paper, but also the title is fantastic, "Let's Take the Con Out of Econometrics," and when I think about this stuff, and we can talk about the solutions a little more in a bit, I don't think it's a [inaudible]. I think it's a good idea to go in and say, "This is what I plan to do with the data."

But the fact of the matter is, when I do research, and maybe I'm an old fogey, is I don't think of everything I might have before I actually look at the data. So, in some sense, what I want to do is be honest that I went in with these priors. I tested these things, but I also thought about this other thing, or I saw this other correlation and it looks interesting and either worth pursuing or whatever, without necessarily overclaiming for it. But noting that I found something interesting, even if it wasn't in my pre-analysis plan. What do you think about that? And how would you address that?

Dr. Edward Miguel:

For sure. Everything you said, Bob, makes sense. And no one can sort of foresee everything that will occur in an experiment in data collection, in their analysis, and it's impossible to do. So, we'd also be leaving a lot of really good results on the table if we ignored everything we discovered by just serendipity while analyzing the data. So, the approach that we push for in the book, and that I think is really becoming the norm in economics around pre-analysis plans, is using the pre-analysis plan as a tool to specify what you had thought of in advance and which hypotheses you can run and really understand what that P-value is. Like, this is my one test, my one shot. I pre-specified it. This is an important outcome for me. Great. You can highlight that with the pre analysis plan, but then may clear other work that's more exploratory. Other work where you followed a more circuitous path to get there.

All that exploratory work's incredibly valuable. A lot of my own papers have been exploratory papers. I find them some of my most interesting work in many cases. It's just very important that the reader know that those hypotheses were not pre-specified. And unfortunately, in our field of economics and other social science fields, there is a tradition of doing all the exploratory work. And this is what Ed Leamer talks about, right? Doing all the exploratory work, but then writing up the paper as if this one regression you ran was like the thing you thought of in advance and not acknowledging that you really ran 50 regressions. So, as long as we're clear on what we did and what's exploratory, I think that's great. And hopefully the pre-analysis plan can help draw that line.

Another thing is the first pre-analysis plan I wrote was for a project using data from Sierra Leone on a local governance reform project, and it was really experimental for us. We didn't know what we were doing. We were kind of coming up... I think we coined the term pre-analysis plan when writing up that paper, which came out in 2012. This is with Kate Casey and Rachel Glenister. We wrote up our pre-analysis plan. We did the analysis, and then we realized in writing up the paper, we forgot to pre-specify like the most basic hypothesis, like was the program actually implemented in this village? So, we write up in the paper like, "Guys, we're sorry, we forgot the most basic thing."

And that happens, and it happens to all of us. And, you know, I don't know if I've ever seen a paper in economics based on a pre-analysis plan where the author doesn't say, "Here's some additional stuff we didn't think of." Like it's always going to happen. As long as people are clear about it, then I think we maintain our sort of scientific rigor and clarity. I hope.

Dr. Bob Hahn:

I think that's a great way of dealing with part of the problem, but part of the problem is how we shape and incentivize human beings who entered the world of econs or economics. So, one of the things I was delighted to see in your book, I think in one of the early chapters, was a discussion of ethical research. Because I actually learned nothing when I was in graduate school about ethical research. What I learned about ethical research was in my 10th grade chemistry class, where they talked about how you write up



a lab report, and you have to be honest about what you do at each step in the process. But I guess my question to you is do you think it would be valuable to teach graduate students in all disciplines, something about ethical research, and what you might do in order to pursue that objective? And what would that look like?

Dr. Edward Miguel:

I totally support it. I feel like there's actually a lot of blind spots and gaps in graduate training. I think in many ways, economics training today in some ways reflects where the field was a while back. I started grad school in 1996 and, you know, over time, economics has become a more data driven field, and we're using big data sets, and we have better econometric methods and more computing power, and just a larger and larger share of the field is empirical research, more and more of it's experimental. And so in some ways we're becoming more like the chemistries or physics or other hard sciences, but I don't think our training reflects that. So, you know, for those of us in development economics, I think ethical research would be incredibly important to teach. It's rarely taught. When we're working in international field work settings with poor populations, there's a range of ethical issues one encounters all the time.

The way grad students learn about those issues now, is they work as a research assistant or they work on a research team and they basically have an apprenticeship with someone like me or my colleagues in development economics, where they learn the folk wisdom and they learn the standard practices in the field, but they're typically not taught it in a class. But it goes beyond fieldwork, there's issues of honesty and the whole scientific ethos that aren't taught but are really central to how we really should do our work. So, I think there'd be a lot of value there. I did teach a graduate course here a few years ago and it became really, in some ways the seed of the book, Bob. I taught this course; it had these different weeks, and like the different weeks kind of became the outlines for the chapters of the book.

I found it extremely exciting to do because we just don't have that kind of material in economics or other social sciences, typically. I guess the other social sciences are better than we are. They may have classes on some of these issues, but in economics, we

typically don't. I do hope the book can be used by folks at the advanced undergrad level, master's level, all the way up to the PhD level, to start bringing some of these issues into graduate curricula. And either the whole book as a semester long course, or pieces of it in particular courses, I think there'll be a lot of value to that.

Dr. Bob Hahn:

So, how do we deal with the fact that researchers, who're very bright people, are often very passionate about their subject and have very strong priors? I'm going to tell you a story. I don't know if it's true, but it's based on my memory. I was sitting in on a physics course at Cal Tech when I was in graduate school, and we had a fantastic physics teacher by the name of David Goodstein, and he showed us the notes from Millikan's oil drop experiment on something that you may or may not know how to spell, called microfiche, where he actually showed Millikan's notes from God knows when, 1907 or whatever...

Dr. Edward Miguel:

I'm from the microfiche era, Bob. I remember spending hours as an undergrad at the library, looking up stuff...

Dr. Bob Hahn:

Here's the short version of the story, which may be apocryphal. I may not have gotten it right, but he shows us Millikan's notebook where he's doing the oil drop experiment, and in the margins you see at some page 17 or whatever, I'm making this up, "This result is exactly what I wanted, published!"

You know, I guess the point I'm trying to make, is all of us have strong biases about what we might expect, like your Kenya electrification or Millikan's oil drop experiment or whatever. Is a pre-analysis plan really enough to get at that bias, or is this just a part of human nature?

Dr. Edward Miguel:

There's no doubt it's part of human nature, and I love that example. Since you're talking about Cal Tech physicists, let me counter with a great story and quote from Richard Fineman, another famous Cal Tech physicist. He was still there when you were there.

So Fineman, you know, who's written some popular books and is kind of a legend. I was an MIT undergrad. So, Fineman was a legend among MIT undergrads. He gave a commencement address in, I think 1973 or 1974 at Cal Tech, his famous commencement address where he kind of deals head on with the point you just said, and he said... it's really, you could read the address as all about scientific integrity and scientific ethos and the problem you dealt with. And he says, "The easiest person to fool with your research is yourself. Because if you get what you want, you're going to be fooled and you're going to push it. So, you need to protect your research from yourself." It's just amazing to find. I think pre-analysis plans can help with that. Pre-analysis plans layout in a way that you can't have selective memory about what you plan to do.

It protects the research from yourself and your own bias that you want to get a finding like, "Hey, we really thought there going to be benefits of electrification. We laid out what we thought were the best test of that hypothesis. We get null results." Okay, we have to deal with that. Pre-analysis plans can also help you from partners.

So, in that Sierra Leone project I briefly alluded to. The project we were evaluating, this local governance reform project, was carried out with the government of Sierra Leone with lots of World Bank funding, and there were a lot of interested parties in that research. We showed our partners our pre-analysis plan, and they agree that what we were testing and the data we were collecting were good tests. So, when we came back with null results in that project too, we had already kind of gotten their sign-off that these were the key tests. It really, I think on almost like a political level can protect researchers from pressure from donors and politicians and other interested parties. So, I think it serves a number of purposes even beyond like the statistical benefits of understanding exactly what P-values mean because you know how many hypotheses have been run.

Dr. Bob Hahn:

Let's talk about something that most people, at least in America, are familiar with, the issue of the minimum wage. And I think economists have done a ton of studies on this. I think many of which have been found the minimum wage may have an adverse impact on employment, and some of which I think you talked about a little bit earlier, Card and Kruger and others have not. What's your take on that issue, but perhaps more importantly, what's your take on good and not so good academic ways of trying to adjudicate issues like that? You have several really cool ideas in your book, and I just wanted you to do some free flow on this.

Dr. Edward Miguel:

First, I'll just say I'm not a labor economist, and I haven't followed all the latest work in this area. I do know there's been a ton of research on it. I do know that some of the work looking at publication bias has found lots of publication bias in this area too. So, I think that, in many ways, it's like a perfect example to illustrate a lot of these issues. One way to get around this relates to pre-specifications. So, one of the examples we give in the book is this case from the late nineties, where Card and Kruger and David Neumark and others who were really active in this area, were kind of convened by another colleague of mine here at Berkeley, David Levine, who was an editor of a journal at the time and said, "Hey, why don't you all write down what you think of as like the right data to use, the right statistical models to use, to understand the impact of the minimum wage on employment, and why don't you do that in advance of the next federal minimum wage increase?"

So, there's no way to data mine this because the federal minimum wage increase at that time hadn't taken place. The data hadn't been released. So, it's very close to a pre-analysis plan even though it's not experimental research, it's research that's using the natural experiment as a policy variation. David Neumark actually participated in this. Card and Kruger didn't for various reasons. But David Neumark participated. He wrote a working paper where he laid out the data, he would use the test he would run, and the data came in, and he wrote a paper that got published in the journal. So, there really are ways forward.

The example that David Levine, and David Levine's my colleague here, so, I've talked to him about it. He was very much inspired by Danny Conoman, the Nobel prize winner in Economics. Danny Conoman in psychology, when he was working in a literature that had conflicting results, he would try to get together with folks who had different views or different findings in what he called "adversarial collaboration" and said, "Look, as scientists let's agree, this is the right test, and then let's just run this lab together and get the result. And then we'll all agree on... sort of we'll adjudicate between these different views."

And that's what David Levine was trying to do by bringing Card and Kruger and David Newmark and others together. So, we can be creative. There's a lot more we can do in this area where scholars with different views can get together and try to agree scientifically on what the right test is before seeing the data, before seeing what the results look like, to try to reach a kind of robust conclusion.

Dr. Bob Hahn:

So, let me push you a little bit on this, and it may be outside of your comfort zone, but let's talk about climate change. So, there was a recent book, *Unsettled*. I can't think of the author right now, where he talks about setting up something like an adversarial process. You know, a red team and a green team or whatever, to sort of kick the tires, you know. Do you think that works for bigger public policy issues like climate change, or is that really out of your comfort zone, and it's better to do in things like minimum wage and the Conoman type problems you were talking about a minute ago?

Dr. Edward Miguel:

I love the idea of scrutiny for any issue. And in fact, the bigger the problem, the more it should be scrutinized. So, I think the idea of having teams go after models and results is important. There's some kind of examples of this, again from psychology. So, kind of more limited lab work so far, but Brian Nosek and the team at the Center for Open Science have been doing these crowdsourced multi-investigator projects where folks are given the same dataset and just, they maybe get 20, 30 different teams analyzing data to try to see what conclusions are robust, and they've done this several times. So, I don't see why that

couldn't be done for larger scale problems. There's actually, you know, a variety of efforts underway in economics right now to try to promote more replication type activities.

Here at Berkeley, I'm heading an initiative called the Berkeley Initiative for Transparency in the Social Sciences (BITSS), which really helped to fund and sort of helped bring the textbook to fruition. The team here at BITSS has just set up a platform called the Social Science Reproduction Platform (SSRP), that just went live a few weeks ago, to be a kind of standard platform where people could replicate empirical results of any kind, and there could be dozens of replications of the same paper. We also allow people not just to replicate a whole paper, but a very specific claim. So, if there's like one table or one regression result that interests you, you can just focus on that one, get the publicly available data, or get access to the data and run the analysis. And then we allow for extensions, robustness checks, improvements, to just always recognize that research can always be improved, and more scrutiny is typically going to lead to better outcomes. So, that's one initiative. And then I know Abel Brodeur at University of Ottawa is also trying to organize a kind of more coherent, coordinated replication effort in economics as well.

So, I do feel like that would be valuable, Bob, and nothing is probably more important, right, for the survival of humanity than understanding the impacts of climate change and dealing with it and measuring facts and planning for it. So, yeah, I think scrutinizing the models and the data would be ideal scientifically.

Dr. Bob Hahn:

The author's name I was trying to think of was Steven Koonin, but let me try to play devil's advocate for a moment. You throw around words like transparency and reproducibility. You're an economist. I mean, you know, in principle, no one can be against transparency and reproducibility if they were free, but we know that there are costs and benefits. Have you thought carefully about what the costs might be?

So, let me throw out a couple of ideas. How does it affect, for example, the quality and quantity of publication that we're going to see if we undertake some of the ideas that you recommend, like a pre-analysis plan?

Dr. Edward Miguel:

These are great points. And you know, in the book we talk about some of these, but actually, I've elaborated recently, I wrote a piece that came out in the *Journal of Economic Perspectives*, just came out a couple months ago, where I try to elaborate on cost and benefits more explicitly, because again, as economists, we're trying to think through the trade-offs. Even if there are benefits to what we're proposing, if the costs are onerous, then it may not be worth adopting.

So, I think you're totally right to think about those issues. In terms of things like pre-analysis plans, some researchers have been collecting survey data to try to figure out how much time it takes to write a pre-analysis plan. Just simply, is this a day-long activity, a year-long activity? Like how much time does it take, and do the researchers who write pre-analysis plans think it's worthwhile to put this time in?

And so, there's a recent paper by George Ofofu and Dan Posner that does this. They actually surveyed scholars in economics and political science. They're political scientists, and they surveyed hundreds of folks who have been using pre-analysis plans. The median amount of time that it takes to write a plan was something like two to four weeks in their data. So, a kind of moderate amount of time, you know. It's not as long as writing a paper, but it's a non-trivial amount of time. And then they asked folks, was it worthwhile? Like, were there some benefits to this that you perceived? 70% of these folks who are familiar with pre-analysis plans said they thought it was worthwhile. So, you know, it seems like the bulk of the people who are doing this see benefits. I'll say for myself, and their survey data indicates something similar, there were like two really important benefits, personal benefits to me, like private benefits as a researcher, that I found from writing a pre-analysis plan.

One is when you put a lot of time and thought into writing down your statistical and econometric specifications in advance, it really does save you time on the backend. Like when the data comes in, sure, maybe it took me a couple of weeks to think through this plan and whatnot, but then it's faster to do the analysis. So, there's kind of like a

displacement of the work time. You have to do it earlier. So, that could be costly. You have to do it before getting the data, but it does save you time later. So, that's one aspect of it.

The other thing that I think a lot of us have found with pre-analysis plans is really carefully writing down what you're going to do with the data, what models you're going to run, what your outcome measures are going to be, actually can improve the research. So, if you do this early in the process, like when you're in the field still collecting data, you might modify your survey questions, because when you think it through you go, "Oh, wait. As I think this through, I really need to collect this variable because it really helped me establish a certain mechanism or whatnot." So, I do think that there's this sort of personal benefit to the researcher, and then also a social benefit, because if you're actually thinking through your research more, that's better research, but there are costs. It takes time. Some researchers find it very onerous. There's been a lot of debate in economics recently about how detailed do these plans need to be. So, the American Economic Association has a registry to post RCT trials and also to post pre-analysis plans. The basic data you need to put on the registry is pretty sparse.

I personally write pre analysis plans that are much more detailed. So, that's what I've opted to do. But to be honest, my own personal view is even the sparse information is really valuable. The sparse information lays out, here's my sample. Here's my sample size. Here's my design. Here are my main outcomes. Here's how the randomization is going to look. That alone is really valuable. If you want to get into like every single econometric model you can think of, the way I do. I find that valuable personally, because I like to think through the research and improve it in advance, but even sparse pre- analysis plans can be valuable. So again, I've taken this kind of optimistic view on things, Bob, which is like, we're moving in the right direction, and even a relatively light-touch move in the right direction is still a step forward for the field.

Dr. Bob Hahn:

Okay. So, I can easily accept the fact that there may be social benefits that exceed the social costs. I find it a little bit harder as an economist to suggest that economists are



behaving sub-optimally with respect to their own private research agenda. Not saying it's impossible, we're all human beings. But anyway, that's my take on it.

Dr. Edward Miguel:

And they're writing plans now. I mean, pre-analysis plans have taken off. Like it just was a technology that like no one had thought of, and now it's like it's out there. And at first people were a little confused, but if you look at the AEA registry, it's like not only are the total, the cumulative numbers increasing every month of registered studies. The month-by-month numbers are increasing, it's like exponential. So, I think in a lot of empirical fields, really experimental fields, pre-analysis plans have become the norm. So, I think it's actually a pretty quick adoption curve that basically in a few years, we've reached that point. So, I hope it is privately beneficial and socially beneficial.

Dr. Bob Hahn:

Yeah. Me too. Let me ask you, we got to close in a couple of minutes, but let me ask you one other question that I assume you've spent some time thinking about, because you mentioned the words public policy. There are a lot of studies that are actually not just funded by the government, but done by the government or done by consultants who work for the government or whatever. Would you apply the same type of standards that you're talking about in academia to those studies? And if so, why? And if not, why not?

Dr. Edward Miguel:

I think it would be very useful to do so. And here at BITSS and this Berkeley initiative, we have a program that we call Open Policy Analysis, OPA. A colleague of mine, Fernando Hoces, who's a PhD in public policy, has a lot of experience in public policy-making data analysis, and he actually worked on the minimum wage. I mean, this is a topic we've come back to. He's been advancing this agenda with us in some ways, Bob.

The value of transparency is always there. It's always going to be useful for government research, for research that directly informs public policy, there's a very strong rationale for making it really easy to look under the hood and allow the population, citizens, other

researchers, those, you know, researchers with different views, to really understand like where is the data coming from, making it available to them, making the code available, you know, making it very clear how people reach their conclusions, what their priors were. I think it can increase the credibility of research and be important for the legitimacy of the policy that results from that analysis, hopefully.

I mean, that's aspirational. So, I do feel like a lot of these tools could be applied, but most public policy-oriented research and most research carried out by government institutions are not necessarily complying with these principles yet. So, I think there's a way to go, but in an open democratic society, it feels like it would be a very natural step to promote open policy analysis.

Dr. Bob Hahn:

I would simply add to that, that it may be a big step from where we are now based on my experience in the regulatory arena.

Dr. Edward Miguel:

I'm sure you're right.

Dr. Bob Hahn:

We've been talking with Professor Edward Miguel of UC Berkeley. Ted, thanks for joining *Two Think Minimum*.

Dr. Edward Miguel

Thanks so much for having me, Bob.

[DOWNLOAD FULL PUBLICATION](#) 

## SHARE THIS ARTICLE



## GET THE LATEST ON TECH POLICY IN YOUR INBOX

ENTER YOUR EMAIL ADDRESS

**SUBMIT**

## VIEW MORE PUBLICATIONS BY ROBERT HAHN AND EDWARD MIGUEL

### RECOMMENDED READS



JUNE 23, 2021

### ADAM GAMORAN ON EVIDENCE-BASED POLICY (TWO THINK MINIMUM)

PODCASTS

MAY 22, 2021

### EFFICIENCY AND EQUITY IMPACTS OF ENERGY SUBSIDIES

RESEARCH PAPERS

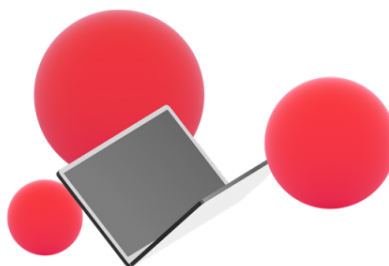
APRIL 20, 2020

## EQUITY IN COST-BENEFIT ANALYSIS

RESEARCH PAPERS

---

### EXPLORE MORE TOPICS



ANTITRUST AND COMPETITION

BIG DATA

BLOCKCHAIN

BROADBAND

CONTENT MODERATION

ECONOMICS AND METHODS

ECONOMICS OF DIGITIZATION

EVIDENCE-BASED POLICY

INTELLECTUAL PROPERTY

MISCELLANEOUS

PRIVACY AND SECURITY

### RELATED ARTICLES

JUNE 23, 2021

**ADAM GAMORAN ON EVIDENCE-BASED POLICY (TWO THINK MINIMUM)**

---

MAY 22, 2021

**EFFICIENCY AND EQUITY IMPACTS OF ENERGY SUBSIDIES**

---

MAY 11, 2021

**TPI DISTINGUISHED SENIOR FELLOW BOB HAHN ARGUES FOR BRINGING MORE SCIENCE TO POLICYMAKING**

---

APRIL 12, 2021

**JON BARON ON EVIDENCE BASED POLICY AT ARNOLD VENTURES (TWO THINK MINIMUM)**

---

JANUARY 28, 2021

**NICK HART ON FOUNDATIONS FOR EVIDENCE-BASED POLICYMAKING ACT**

---

DECEMBER 8, 2020

**TEN TECH PRINCIPLES FOR THE BIDEN ADMINISTRATION**

---

JUNE 29, 2020

**"ROBERT SHEA ON EVIDENCED BASED POLICY'S IMPACT AND POTENTIAL" (TWO THINK MINIMUM)**

---

APRIL 20, 2020

**EQUITY IN COST-BENEFIT ANALYSIS**

---