

The Dismal Science Remains Dismal, Say Scientists

 [wired.com/story/econ-statbias-study/](https://www.wired.com/story/econ-statbias-study/)

Adam Rogers, *Wired*, 11.14.17



Getty Images/WIRED

When Hristos Doucouliagos was a young economist in the mid-1990s, he got interested in all the ways economics [was wrong](#) about itself—bias, underpowered research, statistical shenanigans. Nobody wanted to hear it. “I’d go to seminars and people would say, ‘You’ll never get this published,’” [Doucouliagos](#), now at Deakin University in Australia, says. “They’d say, ‘this is bordering on libel.’”

Now, though? “The norms have changed,” Doucouliagos says. “People are interested in this, and interested in the science.” He should know—he’s one of the reasons why. In the October issue of the prestigious *Economic Journal*, a [paper](#) he co-authored is the centerpiece among a half-dozen papers on the topic of economics’ own private replication crisis, a variation of the one hitting disciplines from psychology to chemistry to neuroscience.

The paper inhales more than 6,700 individual pieces of research, all meta-analyses that themselves encompass 64,076 estimates of economic outcomes. That’s right: It’s a meta-meta-analysis. And in this case, Doucouliagos never meta-analyzed something he didn’t dislike. Of the fields covered in this corpus, half were statistically underpowered—the studies couldn’t show the effect they said they did. And most of the ones that were powerful enough overestimated the size of the effect they purported to show. Economics has a profound effect on policymaking and understanding human behavior. For a science, this is, frankly, dismal.

One of the authors of the paper is John Ioannidis, head of the Meta Research Innovation Center at Stanford. As the author of a 2005 [paper](#) with the shocking title “Why Most Published Research Findings Are False,” Ioannidis is arguably the replication crisis’ chief inquisitor. Sure, economics has had its outspoken critics. But now the sheriff has come to town.

For a field coming somewhat late to the replication crisis party, it’s ironic that economics identified its own credibility issues early. In 1983 Edward Leamer, an economist at UCLA, [published](#) a lecture he called “Let’s Take

the Con Out of Econometrics.” Leamer took his colleagues to task for the then-new practice of collecting data through observation and then fitting it to a model. In practice, Leamer said, econometricians fit their data against thousands of statistical models, found the one that worked the best, and then pretended that they were using that model all along. It’s a recipe for letting bias creep in.

At about the same time as Leamer wrote his paper, [Colin Camerer](#)—today an economist at Caltech—was getting pushback for his interest in reproducibility. “One of my first papers, in the 1980s, has all of the data and the instructions printed in the journal article. Nowadays it would all be online,” Camerer says. “I was able to kind of bully the editor and say, ‘This is how science works.’” Observe, hypothesize, experiment, collect data, *repeat*.

Over time, things improved. By 2010, the field was undergoing a “[credibility revolution](#),” says [Esther Duflo](#), an economist at MIT and editor of the *American Economic Review*. A few top journals began to sniff out shenanigans like p-hacking, massaging data for favorable outcomes. They asked for complete datasets to be posted online, and for pre-registered research plans (so investigators can’t change their hypotheses after the fact). To publish in these journals, economists now have to submit the actual code they used to carry out their analysis, and unlike the old days it has to work on someone else’s computer.

Yes, open data, available code, and pre-registration don’t always guarantee reproducibility. “If I pick up Chrissy Teigen’s cookbook, it might not taste the same as it does at her house,” says Camerer, “even though she’s only 10 miles away and was shopping at the same store.” In 2015, economists at the Federal Reserve and Department of the Treasury [tried to replicate](#) 67 papers using data and code from the original authors; they were able to do it without calling the authors for help for just 22. It was a little grim.

One thing that did help economics: an increasing reliance on experimental data over empirical or observational research. Randomized controlled trials in the lab and in the field are getting more common. In another big-deal [paper](#), this one for the prestigious journal *Science*, Camerer’s team attempted to replicate 18 articles from two top journals. And the results were—well, let’s say the glass was half-full. All were statistically powerful enough to see the effect they purported to, and 11 out of 18 had “a significant effect in the same direction as the original study.”

Maybe more importantly, though, everyone was on board with the concept. “When somebody says ‘I want to replicate your study,’ usually it’s like when the IRS calls and says they want to check your math,” Camerer says. “But when we sent out letters to 18 groups saying, ‘We’re going to replicate your study,’ every one of them was quite cooperative.”

The problem is that only a few journals and subfields in economics have been willing to take up the new standards of controlled trials, openness, and reproducibility that other social sciences—behavioral psychology, most notably—have largely embraced. “Adoption of improved practices is idiosyncratic and governed by local norms,” Camerer says.

That leaves an awful lot of economics—and after failures like the inability to predict the housing crisis and ongoing political disagreements about things as fundamental as taxes and income levels, economics seems a little hard to trust. That’s where big meta-studies of meta-analyses come in, like the one Doucouliagos did with Ioannidis and Tom Stanley. This is the kind of work Ioannidis now specializes in—evaluating not just individual studies, like Camerer’s reproducibility paper, but entire bodies of literature, capturing all the data and stats embedded in many meta-analyses at once. In this case, that wasn’t randomized controlled trials. “The vast majority of available data are observational data, and this is pretty much what was included in these meta-analyses,” Ioannidis says.

The sort-of good news? According to his team, economics isn’t that bad. Sure, the statistical power was way too low and the bias was toward exaggerating effect sizes. “We have seen that pattern in many other fields,” Ioannidis says. “Economics and neuroscience have the same problem.” (So, OK, not great news for fans of brain scan studies.) But that also shows that Ioannidis isn’t just trying to nuke economics out of pique. “Not being an economist, hopefully I avoided the bias of having a strong opinion about any of these topics,” he says. “I just couldn’t care less about what was proposed to have been found.”

That paper should at least red-flag, then, the fact that while at the most elite level and in some fields, economics is working out its issues, elsewhere the familiar problems remain. The grungy spadework of reproducing other research still isn't rewarded by journal editors and tenure committees. Scientists still want to land papers in top-shelf journals, and journals still want to publish "good" results—which is to say, statistically significant, positive findings. "People are likely to publish their most significant or most positive results," Ioannidis says. It's called data-dredging.

Science is supposed to have mechanisms for self-correction, and work to bridge the credibility gap across different fields shows self-correction in action. Still, though, you'd like to see economics farther along, maybe, instead of getting its lapels grabbed by Ioannidis. "We're not very good at understanding how the brain works. We're not that great on models of human nature and connections to anthropology," Camerer says. "But economists are really good at understanding incentives and how we create systems to produce an outcome."

And yet credibility-increasing incentives don't yet exist within economics itself.

Journals and funding agencies have been slow, cautious even. Universities and institutions aren't paying people or tenuring them for the work. "Fields like statistics or psychology are sending strong signals that they care about people working on research transparency," says [Fernando Hoces de la Guardia](#), a fellow at the Berkeley Initiative for Transparency in the Social Sciences. "You don't see any of these folks placing in top economics departments." When he sent me a relevant paper by a colleague, Hoces de la Guardia pointed out that it wasn't his colleague's "job market paper," the piece of research a PhD student would use to find a job.

"One of the problems in raising these sorts of issues is finding the journal space for it," Doucouliagos says. "You're going to have bright scholars who would like to address these issues, but they're worried about being seen as cassetras." But maybe unlike Cassandra, if enough researchers and standard-setters see value in critiquing their own fields, they'll be better equipped to survive the future.