â??How to Publish Statistically Insignificant Results in Economicsâ?•

Description

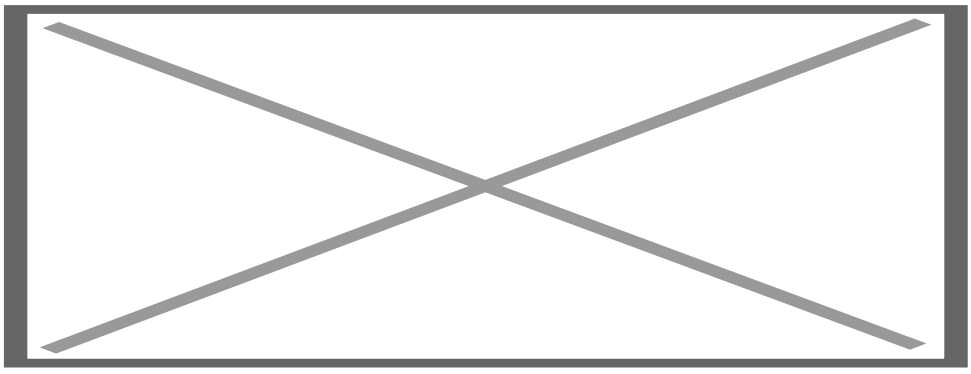
Tweet

â?¦being the title of a blog posting submitted by DAVID EVANS , 03/28/2018 on the World Bankâ??s Development Impact website here

â?¦and reproduced in full hereâ?¦

Sometimes, finding nothing at all can unlock the secrets of the universe. Consider this story from astronomy, recounted by Lily Zhao: â??In 1823, Heinrich Wilhelm Olbers gazed up and wondered not about the stars, but about the darkness between them, asking why the sky is dark at night. If we assume a universe that is infinite, uniform and unchanging, then our line of sight should land on a star no matter where we look. For instance, imagine you are in a forest that stretches around you with no end. Then, in every direction you turn, you will eventually see a tree. Like trees in a never-ending forest, we should similarly be able to see stars in every direction, lighting up the night sky as bright as if were day. The fact that we donâ??t indicates that the universe either is not infinite, is not uniform, or is somehow changing.â?•

What can â??finding nothingâ?• â?? statistically insignificant results â?? tell us in economics? In his breezy personal essay, MIT economist Alberto Abadie makes the case that statistically insignificant results are at least as interesting as significant ones. You can see excerpts of his piece below.



In case itâ??s not obvious from the above, one of Abadieâ??s key points (in a deeply reductive nutshell) is that results are interesting if they change what we believe (or â??update our priorsâ?•). With most public policy interventions, there is no reason that the expected impact would be zero. So there is no reason that the only finding that should change our beliefs is a non-zero finding.

Indeed, a quick review of popular papers (crowdsourced from Twitter) with key results that are statistically insignificantly different from zero showed that the vast majority showed an insignificant result in a context where many readers would expect a positive result.
For exampleâ?¦

- You think wealth improves health? Not so fast! (Cesarini et al., QJE, 2016)

- Okay, if wealth doesnâ??t affect health, maybe you think that education reduces mortality? Nuh-uh! (Meghir, Palme, & Simeonova, AEJ: Applied, 2018)
- You think going to an elite school improves your test scores? Not! (Abdulkadiroglu, Angrist, & Pathak, Econometrica, 2014)
- Do you still think going to an elite school improves your test scores, but only in Kenya? No way! (Lucas & Mbiti, AEJ: Applied, 2014)
- You think increasing teacher salaries will increase student learning? Nice try! (de Ree et al., QJE, 2017)
- You believe all the hype about microcredit and poverty? Think again! (Banerjee et al., AEJ: Applied, 2015)

and even

- You think people born on Friday the 13th are unlucky? Think again! (Cesarini et al., Kyklos, 2015)

It also doesnâ??t hurt if peopleâ??s expectations are fomented by active political debate.

- Do you believe that cutting taxes on individual dividends will increase corporate investment? Better luck next time! (Yagan, AER, 2015)
- Do you believe that Mexican migrant agricultural laborers drive down wages for U.S. workers? We think not! (Clemens, Lewis, & Postel, AER, forthcoming)
- Okay, maybe not the Mexicans. But what about Cuban immigrants? Nope! (Card, Industrial and Labor Relations Review, 1980)

In cases where you wouldnâ??t expect readers to have a strong prior, papers sometimes play up a methodological angle.

- Do you believe that funding community projects in Sierra Leone will improve community institutions? No strong feelings? It didnâ??t. But we had a pre-analysis plan which proves we arenâ??t cherry picking among a thousand outcomes, like some other paper on this topic might do. (Casey, Glennerster, & Miguel, QJE, 2012)
- Do you think that putting flipcharts in schools in Kenya improves student learning? What, you donâ??t really have an opinion about that? Well, they donâ??t. And we provide a nice demonstration that a prospective randomized-controlled trial can totally flip the results of a retrospective analysis. (Glewwe et al., JDE, 2004)

Sometimes, when reporting a statistically insignificant result, authors take special care to highlight what they can rule out.

- â??We find no evidence that wealth impacts mortality or health care utilizationâ?¦ Our estimates allow us to rule out effects on 10-year mortality *one sixth as large*as the cross-sectional wealth-mortality gradient.â?• In other words, we can rule out even a pretty small effect. â??The effects on most other child outcomes, including drug consumption, scholastic performance, and skills, can usually be bounded to *a tight interval around zero*.â?• (Cesarini et al., QJE, 2016)
- â??We estimate insignificant effects of the [Swedish education] reform [that increased years of compulsory schooling] on mortality in the affected cohort. From the confidence intervals, we can *rule out effects larger* than 1â??1.4 months of increased life expectancy.â?• (Meghir, Palme, & Simeonova, AEJ: Applied, 2018)

- â??We can rule out even modest positive impacts on test scores.â?• (de Ree et al., QJE, 2017)

Of course, not all insignificant results are created equal. In the design of a research project, data that illuminates what kind of statistically insignificant result you have can help. Consider five (non-exhaustive) potential reasons for an insignificant result proposed by Glewwe and Muralidharan (and summarized in my blog post on their paper, which I adapt below).

1. The intervention doesnâ??t work. (This is the easiest conclusion, but itâ??s often the wrong one.)
2. The intervention was implemented poorly. Textbooks in Sierra Leone made it to schools but never got distributed to students (Sabarwal et al. 2014).
3. The intervention led to substitution away from program inputs by other actors. School grants in India lost their impact in the second year when households lowered their education spending to compensate (Das et al. 2013).
4. The intervention works for some participants, but it doesnâ??t alleviate a binding constraint for the average participant. English language textbooks in rural Kenya only benefitted the top students, who were the only ones who could read them (Glewwe et al. 2009).
5. The intervention will only work with complementary interventions. School grants in Tanzania only worked when complemented with teacher performance pay (Mbiti et al. 2014).

Here are two papers that â?? just in the abstract â?? demonstrate detective work to understand whatâ??s going on behind their insignificant results.

For example #1, in Atkin et al. (QJE, 2017), few soccer ball producing firms in Pakistan take up a technology that reduces waste. Why?

â??We hypothesize that an important reason for the lack of adoption is a misalignment of incentives within firms: the key employees (cutters and printers) are typically paid piece rates, with no incentive to reduce waste, and the new technology slows them down, at least initially. Fearing reductions in their effective wage, employees resist adoption in various ways, including by misinforming owners about the value of the technology.â?•

And then, they implemented a second experiment to test the hypothesis.

â??To investigate this hypothesis, we implemented a second experiment among the firms that originally received the technology: we offered one cutter and one printer per firm a lump-sum payment, approximately a monthâ??s earnings, conditional on demonstrating competence in using the technology in the presence of the owner. This incentive payment, small from the point of view of the firm, had a significant positive effect on adoption.â?•

Wow! You thought we had a null result, but by the end of the abstract, we produced a statistically significant result!

For example #2, Michalopoulos and Papaioannou (QJE, 2014) canâ??t run a follow-up experiment because theyâ??re looking at the partition of African ethnic groups by political boundaries imposed half a century ago. â??We show that differences in countrywide institutional structures across the national border do not explain within-ethnicity differences in economic performance.â?• What? Do institutions not matter? Need we rethink everything we learned from Why Nations Fail? Oh ho, the â??average noneffectâ?¦masks considerable heterogeneity.â?• This is a version of Reason 4 from Glewwe and

Muralidharan above.

These papers remind us that economists need to be detectives as well as plumbers, especially in the context of insignificant results.

Towards the end of the paper that began this post, Abadie writes that â??we advocate a visible reporting and discussion of non-significant results in empirical practice.â?• I agree. Non-significant results can change our minds. They can teach us. But authors have to do the work to show readers what they should learn. And editors and reviewers need to be open to it.

What else can you read about this topic?

- David McKenzie, on the use of novel methods to complement papers with unsurprising findings.
- Gonzalo HernÃ¡ndez Licona, on how null (or disappointing) results have affected policy in Mexico.
- There are more examples of published null results in the Twitter thread that delivered many of the examples above.
- The Economist as Detective is an article by Claudia Goldin. The Economist as Plumber is an article written by Esther Duflo.
- Lots of other stuff, surely.

Category

1. Blog
2. Website

Date
17/10/2026
Date Created
28/03/2018
Author
admin