On weights and coding errors: odd coincidence or dress rehearsal?

Angus Deaton, October 2013

The recent criticism of Carmen Reinhart and Ken Rogoff 's work on debt and growth has resembled nothing so much as a public pillorying. According to their critics, a team at the University of Massachusetts, their results are vitiated by a coding error and by their choice of weights. According to Reinhart and Rogoff, the critics' results are identical to their own preferred results, but which were ignored by their critics. Yet none of this has deflected a firestorm of public criticism and the public dismissal of Reinhart and Rogoff's work.

I do not comment on the substance of this debate here, leaving it in Reinhart and Rogoff's capable hands. Instead, I note an earlier and less-known incident in which my own (joint) work came under criticism by economists at UMass—one of whom, Michael Ash was involved in both cases—and where there are a number of almost uncanny similarities to the more important current debate.

The substantive question

In 2003, Darren Lubotsky and I published a paper in the journal *Social Science and Medicine*. Our topic was whether income inequality is a health risk to people who live in particularly unequal cities or states. The idea is that income inequality is like a toxic pollution, harming everyone who lives with it. My own view, then and now, is that there is no such effect, though I also believe that the extreme income inequality that we see in the US today is a threat to public health, but through quite different (essentially political) mechanisms.

Our analysis showed that the correlation between higher mortality and income inequality arises from a failure to control for the racial composition of the population in each city or state. In cities (states) with a larger fraction of blacks, the difference between the average incomes of blacks and whites is larger, perhaps because employers do not regard blacks and whites as fully substitutable in production. This drives a link from fraction black to income inequality. But blacks also have worse health than whites—in part because of an *apartheid* healthcare system that treats blacks less well than whites—so that the fraction black is also linked to overall mortality. Those two links induce a strong positive correlation between income inequality and mortality. That this correlation is spurious is documented by its vanishing when we control for fraction black, by the fact that mortality rates of blacks and whites separately are uncorrelated with income inequality, and by the fact that income inequality and mortality are uncorrelated across space in other settings where race is not a salient factor.

The topic has acquired some political baggage. The inequality as pollution story is often favored on the left, particularly in Britain, and those who argue against it have sometimes been accused of doing so on political grounds. Of course, this is but a pale shadow of the political importance of the debt questions in Reinhart and Rogoff's work.

A challenge

In December 2005, I received a letter from Michael Ash and Dean Robinson at UMass asking questions about the data in our paper and saying that they could not replicate our results. All of the data that we had used were (and are) publicly available on government websites, but they require processing and organizing to be useful. Such replication queries from other scholars are routine, and while we believed that our procedures were clearly enough specified in the paper to permit replication, we were happy to help guide others who are less familiar with this kind of calculation. Darren Lubotsky, who had done the original data assembly, corresponded with Professor Ash over a period of time, and provided explanations, code, and data to allow him to replicate what we had done. Correspondence then stopped, and we assumed that the matter had been resolved.

We were not so fortunate. Almost three years later, Ash wrote to us to say that *Social Science and Medicine* was going to publish their paper criticizing our work, as was later confirmed by the editor. Different journals may have different policies, but most journals seek the opinion of the authors being criticized *prior* to making the publication decision. Of course, authors cannot be given a veto, but they will often be able to persuade an editor that the comment is worthless—as was the case here—and thus prevent unnecessary public controversy.

Ash and Robinson's note claims that Lubotsky and I had made a coding error in specifying the weights in our regressions, and that without the error, inequality retains its significant positive effect on mortality. There was no coding error, but we had chosen weights that downplayed the larger cities and states compared with the weights that Ash and Robinson had chosen. And it turns out that except for one case (in *one* specification in *one* data period and with only *one* of their alternative weighting schemes) *all* of our results are unaffected by the change in weights. In spite of this, Ash and Robinson make the extraordinarily misleading statement in the abstract that "correcting the error changes the basic results of the paper with respect to inequality and mortality in a relevant and substantive way," Ash and Robinson's preferred specification has exactly the same result as our paper, that income inequality is not a health hazard.

Our response, which was published together with Ash and Robinson's note, makes all of this clear. To our knowledge, there has been no subsequent controversy.

General implications

It is hard to imagine *any* applied paper that would be immune to this sort of attack. Weighting can always be argued over: in standard regression analysis, it is not supposed to matter, and when it does, it is usually unclear what is being estimated. So if you want to debunk a paper, working through it equation by equation, trying out a range of weights, you will eventually find something that changes. You can then cry "coding error" and hope that the rhetoric shifts the burden of proof back to the original authors.

In our case, as in Reinhart and Rogoff, neither the coding error (in our case there was none) nor the choice of weights has any effect on the main results. In our case, Ash and Robinson simply ignored the results that did not support their charges, and claimed that their results were different from ours in a "relevant and substantive way." With Reinhart and Rogoff, they referred only to an early paper, ignoring updated results. But the effect is the same, to magnify a tiny or non-existent problem and claim that it threatens the whole enterprise whereas, in fact, nothing of the sort is true.

There is also the question of publication. In our case, I believe that *Social Science and Medicine* should have shown us the paper prior to the publication decision, but they did allow us to publish a response alongside the critique. In the Reinhart and Rogoff case, Ash et al did not submit their paper to a journal where there might have been a chance of an appropriate professional response, but sent it directly to the world press, copying it to Reinhart and Rogoff on the same day.

Such smear methods appear to work, and provide a handy template for others on how to disguise political attacks as legitimate scientific commentary. While it is naïve to think that science can ever be insulated from politics, if these methods of attack are widely replicated, and if journals and newspapers are prepared to abet them, it will make it much more difficult for serious policy-relevant researchers to do their job. Scholars will also be much less willing to share data than is currently the case; doing so allows anyone who is unscrupulous enough to turn your cooperation into against you.