

In praise of not knowing

Eric Blair

18 September 2005

The big ol' business report

I'd already commented on the process of producing one of those megasurveys that multinational institutions like to put out. The summary: you need to have some definite hypotheses before you start writing, and if your analysis gets a null result, you're screwed. This actually happened elsewhere in the report I've been dealing with: they spent about a month trying to run a regression that would find that remittance rates are positively correlated to growth. They had about ten years of data from a handful of countries, the data wasn't great, and there are causality problems—if your country is about to suffer a downturn, going abroad and sending money home is a great strategy, so a surge in remittances may predict an economic downturn.

In short, we don't know. We'll see what turns up in the actual report. But there's a clear and evident incentive to make claims toward what the report is supposed to say—and no matter what happens with the numbers, the report will *not* say 'we don't know the relationship between remittances and growth.'

The Realtor(R)

I have a pal who is interested in buying a house in Baltimore, so we went on a tour of Baltimore yesterday morning with a real estate agent (a Realtor(R) even). Saw lots of houses, asked lots of questions.

Me: Hey, so these appliances with stainless steel front panels; what do you think that adds to the sale price of the house?

Her: \$5,000.

We pass through my neighborhood, and there's a big ol' structure going up across from the university.

Another person: Hey, what are those gonna be?

Realtor(R): They'll be condos.

Me: Um, since I'm involved with the university, I can tell you they're gonna be student housing, with a Borders & Noble on the first floor.

My pal was impressed by the Realtor(R), who never emitted so much as an 'Um' before answering every question. Me, I was entirely turned off. For our Realtor(R), the correct answer to question (1) is: 'I don't know.', and the correct answer to (2) is: 'I don't know.' If you don't like these examples, it was

a two hour tour, during which she had a precise answer for every question, so other examples are available on request.

Our b-schoolers learn early on that confidence is essential, and derive a direct corollary that you should therefore never say you don't know. Let this type of comportment be 'business school confidence'. The reader will note that this sort of faux confidence is often observed in our President, who is the first President of the U.S.A. with a b-school degree.

“If we knew what we were doing, it wouldn't be called research, would it?”

It's a cliché that academics are more precise and methodical than their business-world counterparts, but there are enough businessmen who are precise and methodical. I believe that what actually separates the two worlds is that academics are allowed to say that they don't know.

In fact, even when academics say something, they're really just lowering the level with which they don't know. 'Before, I knew nothing, but now I know nothing in only about 5% of the states of the world.' The way we've been writing aforementioned report for multinational institution is that I've been sending in reports with the usual caveats and confidence intervals, and then they delete all of that and leave the results. This would be grounds for disciplinary action at any research institution in the world, but is evidently just another day's work for our business-oriented institution.

The good academic exudes confidence as well, but does so in a manner very different from business-school confidence. The good academic heads the seminar with a list of items which he or she knows with confidence. Then, when a participant raises his hand and asks about something else, the academic simply says 'I'm sorry, that's not on the list of things I know.' The audience walks out not thinking that the speaker is some sort of messiah, but with much more confidence in the short list of things that the speaker claims to know. The form of academic confidence explains a great deal of the academic world; e.g., why paper topics are so hopelessly narrow: it's the author's way of confessing ignorance about the whole world, save for a tiny sliver about which he or she can speak authoritatively.

In the borderline world between academia and the b-world, different people go in different directions. For example, the World Bank's World Development Report 2006 doesn't include a single confidence interval that I could find (given a decent skim—I'll say 95% confidence), even though it includes a number of regressions run in MSFT Excel by the authors. And this addresses my prior commentary on the Bank: the whole report is about inequality. Compare with this Urban Institute report which I chose because it happened to be on their home page right now. The executive summary makes no mention of statistical significance, and if you're a policymaker who slept through stats, you will have no problem reading the report. But if you look at the half of the report labeled "Notes on methods and terminology", those pesky little stars start to turn up. It's still not academic publication-level detail, but these guys took seriously the

problem of describing where the numbers came from, instead of presenting a façade that the numbers and regression results are indisputable and certain.

By giving b-school confidence and academic confidence parallel names, I'm maybe implying that they're both sort of OK in their context, but I don't believe this is so. B-school confidence is disingenuous, and is a form of misleading without technically lying. As you can see, b-school confidence has lately grated upon me.

Why not just throw in that extra column of numbers indicating that we could be wrong? People have great alibis for it, typically based on a Lake Wobegone story: oh, *you* understand the concept of statistical uncertainty, and *I* understand it, but our readers don't, or don't have time for it. True, many people *don't* get the minutæ of how one would apply a central limit theorem to produce a confidence interval, but they *do* understand what it means to be not-100%-certain.

This may also be a social norm issue: if b-school confidence is the norm, academic confidence looks out of place, and vice versa. But if it's a social norm issue, then we can do our part to change the norm: If you've been hanging out with your realtor(R) for two hours and she has a confident answer to all your impossible questions, don't entrust her with your business. If you do cite a report that doesn't confess to ignorance on any topic, point out that you looked for confidence intervals, but found none. Just as you would normally let the reader decide what level of confidence is acceptable(*), you should let the reader decide how to interpret a complete lack of such information.

(*)By which I mean, don't just report that this number has two stars attached and this other one has one star; give a significance level like 91.3%, and let the reader decide whether he or she is comfortable with that. Maybe print stars to suggest or guide, but not as the sole piece of info. But you knew this.