

for the AEA session at Chicago, 2:30 Saturday, January 3

Two Vices: Proof and Significance

Deirdre McCloskey

University of Iowa

and

Erasmus University of Rotterdam

My propositions are two.

(1.) Statistical significance is bankrupt. All the “findings” of the Age of Statistical Significance since the 1940s, and especially the 1970s, are erroneous and need to be redone. That is, if the agreed estimates of this or that elasticity are correct, it is only by God’s grace. Economists have thrown away gigabytes of scientific time better spent on finding out How Big is Big--observing and estimating and arguing rather than pseudo-testing by Student’s t . It is scientifically irrelevant, for example, whether or not there is a statistically significant effect of the minimum wage on employment. Statistical significance is only about sample sizes. The scientific question is whether there is an effect large

enough to be interesting, and the sample size has almost nothing to do with the scientific question.

(2.) Proof without magnitude, the existence theorem, the non-constructive proof, is nearly useless as science. Ninety-eight percent of the “work” since the 1940s proving that assumptions A lead to conclusions C has been a waste of time, since a year or two later it was proven that alternative assumptions A-prime, A-double-prime, and so forth, lead to C-prime, C-double-prime, and so forth. And C-n-tuple-prime can be any conclusion whatever. Think of the theorems on factor price equalization in the 1940s and 1950s or the theorems on the efficacy of monetary policy under rational expectations in the 1970s and 1980s. It is scientifically irrelevant, for example, that under assumptions A-treble-prime there exists a competitive equilibrium. It may be a philosophical result worth having, if one realizes that it is about ideas and not about the world (as numerous enthusiasts for the so-called Coase Theorem, for example, do not). But the scientific question is whether the economy is within delta of equilibrium, where delta is a scientifically interesting degree of deviation from price equals marginal cost and the rest. The mathematics can provide functional forms, in the style of engineering. But the Math-Department existence theorems have almost nothing to do with the scientific question. Physicists have been using the Schrödinger

equation since the 1920s without knowing whether it has solutions in general. They do not even know if the three body problem has solutions in general. The mathematical embarrassment has not stopped physics for a nanosecond. We need to drop proof and move to simulation. We need to get away from the Math Department and get back to Physics, Biology, History as models of science.

I realize these propositions are hard to take. For one thing, who is she to say such rude things? I'm like the little girl who remarked loudly about the emperor on parade, "My goodness: he has no clothes!" (Believe me, you remember the Andersen story wrongly--it was a little girl!)

I'm making some surprising claims here, and I hope that how I make them is amusing to you. But don't get me wrong. I'm not just trying to be funny. My most surprising claim, in a way, is that I am correct and much of the economics profession is mistaken. As scientists you are right to be suspicious of someone who talks this way. It feels like the monetary cranks who send long, handwritten letters to professors of economics. But I am speaking from within the establishment in economics and unfortunately I am correct. I find this unspeakably sad, because it means that much of what we claim to have accomplished in economics since the 1940s, and especially since the 1970s, needs to be done over again. The work was wrong, because it made these two massive

rhetorical mistakes: mistaking statistical significance for scientific significance and mistaking existence theorems for quantitative science.

I want you to try to forget the characteristics of the messenger that might distract you. Try to forget that that I am merely an economic historian, that I live in the Midwest, that I am not at Princeton, that I am a Chicago-School libertarian feminist, that I am transsexual, that I stutter. Try, for the good of our wonderful science, to put my argument through your frontal lobe. As Oliver Cromwell said in a letter to the General Assembly of the Church of Scotland, "I beseech you, in the bowels of Christ, think it possible you may be mistaken." Think it possible that statistical significance and existence proofs have been worse than useless these fifty years, and that modern economics needs to be reformed as engineering mathematics and descriptive statistics and observation, observation, observation. Consider that for fifty years we have engaged with increasing excitement and diminishing scientific reward in a boys' game in a sandbox.

Even aside from the dubious character of the messenger, the message itself sounds unbelievable. We've been told over and over since the 1940s that these two techniques--statistical significance and proof by existence--are the very essence of Scientific economics. What is confusing is that in superficials the

techniques look like science, and impress outsiders. But in their serious rhetoric they have nothing to do with the chief scientific question, How Big is Big?

In sciences that are doing real scientific work the scientists spend their days asking questions like, How big a meteor could or did cause a nuclear winter 60 million years ago? The question is always “how big?” Theory and observation are both involved, and sometimes mind-boggling statistical theory or deep mathematics. Look at Brownian motion. It’s not the use of statistics or mathematics that’s at stake. It’s the dramatic misuse of them in significance and proof, which despite their rhetoric of number are disconnected from real science. No question of How Big has ever been solved by consulting a significance test or an existence proof. Or more exactly, it’s been pseudo-solved: “I have a statistically significant coefficient on January in my regression, and therefore the market is importantly imperfect”; or, I have an existence theorem about free trade, and therefore GATT is importantly good.” By contrast, every line the geological or paleontological or historical scientists write is directed, and obviously directed, at discovering How Big is Big.

We all agree that there are two things we need to do as scientists if we are going to find out the world. Watching and thinking. Observing and theorizing. Feeling and imagining. Listening and speaking. Using history and using models.

Devising stories and devising metaphors. Attending to yin and attending to yang. Being female and being male. Accepting with love the facts in the world and projecting with courage our ideas onto the world.

We obviously need both, preferably in the same person, a sort of scientific androgyny. Economists are fond of defending the split of watching and thinking (with less pay for watchers: women, you see) by appeal to specialization. Sure. But if you don't then exchange the results the economics is not being correctly appealed to, is it? The theoretical physicists read the physical equivalent of the Journal of Economic History with a passion. Unless the thinking and the watching are brought together in a scientific argument, such as Wilson's sociobiology or Stephen Gould's and Richard Lewontin's punctuated equilibria or the Alvarez's meteor story of mass extinctions or Simon Kuznets' account of modern economic growth, nothing scientific happens. You get what professional historians sneer at as antiquarianism, mere piling up of facts; or what professional physicists sneer at as math-mongering, mere piling up of proofs. This is not controversial. Two centuries ago Immanuel Kant said that facts without concepts are blind, and concepts without facts not much use, either.

But you've got to do the real thing. You have to be really thinking--about the world. And you have to be really watching--the world. Modern economics

had become largely what the physicist Richard Feynman used to call “cargo-cult science” (Feynman 1985, p. 340). In New Guinea after the Second World War the natives built “airports” with coconut shells for “lights” and “planes” out of cardboard, because they wanted the cargo planes and their prosperity to come back. But the natives weren’t doing real cargo work. The coconuts looked a little bit like the real thing, as the contents of the American Economic Review looks a little bit like the Physical Review, but on closer inspection nothing real is happening. Barbara Bergmann once apologized ironically for her vulgar attachment to asking How Big: “Such simulations seem sleazy to those economists for whom respectable microeconomics is synonymous with solving optimization problems” (1990, p. 99). The sleaze, of course, is on the other foot. Nothing of a scientific, How-Big-Is-Big character is done by existence theorems or statistical significance. It’s cargo-cult science.

I can’t make here all the many arguments that support such a position, or deal with the few replies. If you’re serious about remaking economics as science ^{read} the books and articles by the scores of economists, statisticians, and even mathematicians since the 1940s who have tried to stop the cargo-cult in the sandbox, such as Ken Arrow’s blast against statistical significance in 1959, or Bruno Frey and Reiner Eichenberger’s attack on empty and ephemeral formalism

in 1997, or my books of 1994, 1997, and 1998 on these themes. Or you can read in other sciences, and see how they actually work. If you read Michael Summerfield's textbook, Global Geomorphology (1991) you will wonder why economics cannot be this way.

To those of you who are not familiar with the arguments I and others have been making about economics, let me sketch them. They are simple, so simple that only an idiot like me could see them clearly. You'll imagine you have idiotically simple replies to them. Believe me, you don't. The case against statistical significance and blackboard economics is crushing. In the long run you are going to look like a fool if you merely sneer at me and continue in the sandbox. Please: stand up, brush the sand off, and get down to scientific work.

Statistical significance is about the sample size. But that's seldom the scientific question. End of argument. If you need a purchasing power parity equation for speculation on currency exchanges you'll need a tighter fit than would be all right for macroeconomic policy. The use to which the finding is to be put should figure in the statistical inference, as Neyman and Pearson declared in 1933. Yet it does not. Significance levels are plucked from the air. The decision theoretical procedures that Neyman and Pearson advocated have not been followed. R. A. Fisher won. There is therefore practically no overlap between a

finding of statistical significance as it is understood in modern economics and a finding of substantive significance.

As for existence proof, it is about. . . existence. But it is about existence within epsilon where epsilon may without loss of generality be so small as to be unimportant. Existence is not the same thing as “exists and by God is important.” You prove under assumptions A that the North American Free Trade Agreement was good for the U.S., conclusion C. Next month your opponent can modify one of the assumptions to get A-prime--by now there are whole schools of thought devoted to such scholasticism, so he won't have trouble finding an assumption to alter--and prove the contrary, C-prime. What has been gained? Nothing.

Well, something: we have gained a proof by repeated example that blackboard economics doesn't amount to anything scientifically, that you can't prove anything about the world on a blackboard unaided by magnitudes. Look at the battle on the blackboard among monetarists and Keynesian and rational exectors down to the present. When someone asked Edison if he had been depressed by the failure of the 1500 filaments he had tried before carbonized cotton for the lightbulb, he replied, “No, of course not: I learned a great deal scientifically, namely, that those 1500 were not the right ones!” Since the 1940s we have learned that all those theorems are not the right ones.

The right theorem we have instead discovered is the A-prime, C-prime theorem:

Metatheorem on Theorems:

The A-Prime, C-Prime Theorem

For each and every set of assumption A implying a conclusion C, and for each alternative conclusion C' arbitrarily far from C (for example, disjoint with C), there exists an alternative set of assumptions A' arbitrarily close to the original assumption A, such that A' implies C'.

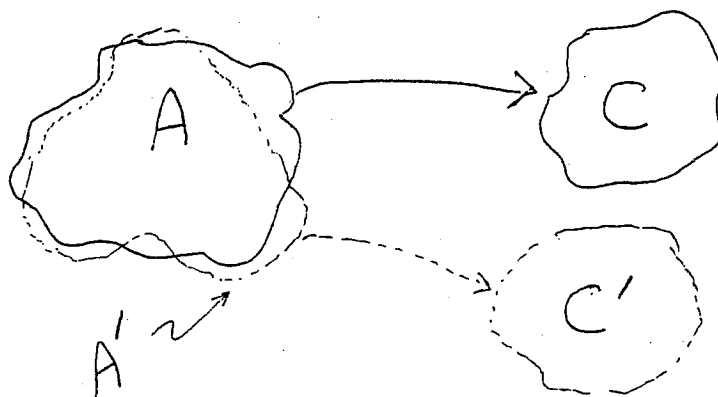


Figure 1.

Small Changes in Assumptions, If Carefully Chosen,
Can Make Big Differences in Conclusions

See if you don't think the Theorem fits the intellectual history of most fields of economics these days, from international trade to industrial organization.

It's easy to hear my arguments wrongly, and therefore not feel impelled to put them through your frontal lobes. I know this from experience. I have been saying these things for some years, and in response to them the unsophisticated people merely get angry at me. They think I am undermining economics for no good reason (such as its mistakenness). They can't grasp the arguments, and are indignant that I would make them, whatever they are.

Usually I get agreement from the sophisticated people. But they try to say, We know that. ("We" don't. As Stephen Ziliak and I showed in 1996, for example, nearly all the best economists grossly misuse statistical significance.) And then, having agreed I am correct . . . they go on teaching the same nonsense to their graduate students. First-year programs in economics have long been mere terror, the imposition of indefensible intellectual values by examination. It's like some sort of strange compulsion, the repeated washing of hands or pounding ones head on the wall. The sophisticates agree statistical significance and blackboard economics are indefensible. Then they go back to washing and pounding.

Let's get this much clear. I am not arguing Against Statistics, Against Theory, or Against Mathematics. I am a quantitative economic historian, in love with statistics, theory, and mathematics. I view existence theorems such as "there should not exist any math in economics" as silly. How would you prove such a

Methodological proposition? If we are going to think about and observe the economic world we need statistics, theory, and mathematics. But that it is silly to oppose the existence of statistics, theory, or mathematics in economics does not determine their optimal amounts or their forms. It's the difference between existence and How Big is Big. That formal methods should exist in economics is obvious. But it is not obvious that the methods should be mistaken for substantive science.

“But she’s Going Too Far.” This is the last refuge of people who do not want to put the argument through their frontal lobe. They say, “Well, yes, McCloskey has something there.” (They often cannot reproduce my two points accurately if asked, so they do not know what they are agreeing to.) “But she goes too far. Surely, significance and proof are of some use. After all, many intelligent people have built careers around them.” My reply is economic, and scientific. How much significance and proof should we have? Their existence, I repeat, is not at stake. It's their centrality, their gigantic role in modern economics, that is the problem. Right now the arguments that economists claim are decisive are perhaps 40 percent statistical significance and 50 percent proof, leaving 10 percent for observation, common experience, empirically disciplined simulation, and

fittings of the size of coefficients on a scale of Bigness that makes scientific sense. That's 9 to 1 against science.

How far is it from optimal? In perhaps 2 percent of economic problems the sampling error is the central issue. And perhaps 2 percent of the economic scientists should be assigned to rummaging through A, A-prime, A-double prime for ideas we may have missed otherwise. So the 9 to 1 should be reversed. Economics would look more like geology or physics than what it looks like now: philosophy and mathematics.

We need to Watch, we need to Think. What has happened in modern economics is that Statistical Significance has been confused with Watching and the Existence Theorem with Thinking. The received methods of modern economics elevate almost-irrelevant intellectual values--large sample size and strict logical consistency--into The Values of the science. It's crowded out real science. Please, boys, let's get out of the sandbox.

Works Cited

- Arrow, Kenneth. 1959. "Decision Theory and the Choice of a Level of Significance for the t-Test." In Ingram Olkin and others, eds. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford: Stanford University Press.
- Bergmann, Barbara R. 1990. "Micro-to-Macro Simulations: A Primer with a Labor Market Example." *Journal of Economic Perspectives* 4 (Winter): 99-116.
- Feynman, Richard, as told to R. Leighton, edited by E. Hutchings. 1985. *"Surely You're Joking, Mr. Feynman!" Adventures of a Curious Character*. New York: W. W. Norton.
- Frey, Bruno, and Reiner Eichenberger. 1997. "Economists: First Semester, High Flyers and UFOs." Pp. 15-48 in P. Van Bergeijk, A. Bovenberg, E. Van Damme, and J. Van Sinderen, eds., *Economic Science and Practice*. Cheltenham, England: Edward Elgar.
- Leamer, Edward. 1983. "Let's Take the Con Out of Econometrics." *American Economic Review* 73 (March): 31-43.
- McCloskey, Deirdre, and Stephen Ziliak. 1996. "The Standard Error of Regression." *Journal of Economic Literature* 34 (March): 97-114.
- McCloskey, Deirdre. 1994. *Knowledge and Persuasion in Economics*. Cambridge: Cambridge University Press.
- McCloskey, Deirdre. 1996. *The Vices of Economists; The Virtues of the Bourgeoisie*. Ann Arbor and Amsterdam: Michigan and Amsterdam university presses.
- McCloskey, Deirdre. 1997. *The Rhetoric of Economics*. Revised edition. Madison: University of Wisconsin Press.
- Summerfield, Michael A. 1991. *Global Geomorphology*. Harlow, England and New York: Longman and Wiley.