Forthcoming 2019
in David Colander, ed.,
symposium in teaching econometrics
in the *Journal of Economic Education*.

What Quantitative Methods Should We Teach to Graduate Students? A Comment on Swann's "Is Precise Econometrics an Illusion?"

Deirdre Nansen McCloskey
deirdre2@uic.edu
deirdremccloskey.org
and
Stephen T. Ziliak
sziliak@roosevelt.edu
https://blogs.roosevelt.edu/sziliak/

Our answer to Professor Swann's question is "yes," with a big, big amendment. It is not Swann's conventional notion of "precision" that is the illusion, not the biggest one, but the notion embodied in the present practice of quantitative methods in economics that researchers should grind away at theoretical econometrics, despite its evident and sharply diminishing returns. Econometrics, understood as regression analysis with null hypothesis significance testing in the absence of a substantive loss function, has yielded no major economic finding since its invention in the 1940s. By contrast, other quantitative methods, such as crude or not so crude simulations (such as Harberger triangles), historical inquiries (such as *The Monetary History of the United States*), massive experiments (such as episodes of hyperinflation in Israel or Argentina), or the scatter plots Professor Swann uses (such as the Phillips Curve when first articulated) have changed scientific opinions repeatedly and, in another sense, significantly.

Swann writes, "What is to be done? Some econometricians have told me they recognize [a problem, in Swann's case the assumption of independence in error terms] . . . , but believe it can and will be solved by further incremental innovations in

econometric method and data collection. I hope they are right, but there is a very long way to go." You're telling us.

Swann is to be congratulated for being one of the few econometric sophisticates to have taken on the message of Morgenstern (1963) that economic "data" contains massive substantive errors. But then Swann then uses an expression which irks us Latinists, "given data." The Latin means "things given," so that "given data," to the amusement of the very learned Department of Economics at the London School of Economics in the 1930s, means "given things given" (Hayek 1945). We believe Swann would agree that one of the main problems in quantitative economics is that the economists do in fact receive the one-in-a-thousand sample from the Department of Labor as "given." Scientists on the contrary should be going out and getting the facts, striving for *capta,* things seized, and therefore knowing how the sausage is made, because the economists made it. They don't nowadays know, except in economic history and in experimental economics, and young economists are not taught the quantitative skills that would encourage them to do so.

The novelist and retired professor of English, David Lodge, begins his semi-autobiographical novel, *Deaf Sentence* (2008), with a hilarious six-page riff in which the elderly if randy British academic protagonist with failing hearing is standing at a cocktail party speaking close up to a young woman in a red silk blouse. He is "nodding sagely and emitting a phatic murmur from time to time. . . . The room is full of noise, a conversational hubbub . . . causing [the crowd] to shout even louder to make themselves heard." Economists will recognize the hubbub as a result of the spillovers of noise, giving each person an incentive to speak louder, which yields overfishing of the quiet of the room—rather as econometricians overfish "significant" results from the common pool of quarterly national income "data" since the War. Linguists call it (Lodge explains) "the Lombard Reflex, named after Etienne Lombard . . . [namely,] that speakers increase their vocal effort in the presence of noise." That is, econometrically speaking, the signal coming from the usual party-goer is correlated with the noise. Uh oh. Errors in variables, which is Professor Swann's way to worry.

The economic researcher, like Lodge's protagonist, is commonly "deaf enough to make communication imperfect." Or as Professor Swann outs it, "noisy data becomes a problem when it dominates the signal we want to observe." Well, not quite. It becomes a *serious* problem only when there is reason to believe that the noise *biases* the signal. If you are confident in your economic model, which in the Lodge case means having high priors about what the young woman is saying, the "precision of the estimate" measured by R^2 or t tests or whatever is irrelevant. You are trying to find out the coefficient, the oomph of one variable on another, and without bias or inconsistency

you have a clear path to discovering it. Noise be damned.

That is, the noise can be large, as in Lodge's cocktail party, but if your hearing is normal or your priors strong you can usually interpret what the person next to you is saying. Not, alas, the elderly professor in the novel, though equipped with hearing aids, because "the woman seems to be an exception to the rule of the Lombard Reflex. Instead of raising the pitch and volume of her voice . . . she maintains a level of utterance suitable for conversation in a quiet drawing room." Hopeless.

And it may be hopeless in the econometric case, too, which is Swann's point. "We cannot expect precise results in such an environment," he writes. We would gently reply to him that "precise," meaning low variance of the estimate, modulo sampling, should not be the focus. "Sensitivity analysis shows that the apparent precision of reported econometric results is generally an illusion, because it is highly dependent on error term independence assumptions." True. "As we know nothing about this variable, or collection of variables, it may seem plausible enough to assume that x and u are independent. But is this the only plausible assumption?" He's making the point that Hank Houthakker made to graduate students at Harvard in 1964 when teaching econometrics. Like as not, Hank would say, there is life on Mars. Therefore, $P = 50\%$. Whoops.

But such a mistake—though it well illustrates the angels-dancing-on-the-head-of-a-pin character of theoretical econometrics—is not the crux. "Why can't we make a realistic assessment of the accuracy and reliability of econometric studies?" The main reason is that economists and others have, disastrously, since R. A. Fisher and then the Cowles Commission, mixed up significance with significance. Swann's "accuracy and reliability" is another way of falling back into the error of taking sampling error to be the main problem one faces in economic science. It isn't. The main problem is substantive bias and irrelevancy. The lamppost of sampling theory casts a bright light, to be sure. But the economists and the others are mistaken to believe that they should therefore drag every scientific problem under it, considering that they have lost their scientific keys out in the dark.

Swann refers briefly to our writings 1983 to the present trying to make the point, but does not seem to have entirely grasped it. No wonder: very few statistical practitioners have grasped "our" point over the century that it has been reiterated by Gosset, Egon Pearson, Jeffreys, Borel, Neyman, Wald, Wolfowitz, Yule, Deming, Yates, Savage, de Finetti, Good, Lindley, Feynman, Lehmann, DeGroot, Chernoff, Raiffa, Arrow, Blackwell, Friedman, Mosteller, Kruskal, Mandelbrot, Wallis, Roberts, Granger, Press, Berger, and Zellner. A little technical sophistication combined with the ease of

Fisherian two-standard-deviation rule overrides common sense (McCloskey and Ziliak 2007, 2012, Ziliak and McCloskey 2013). "While McCloskey and Ziliak were right to warn us that it is dangerous to use *t*-statistics as a measure of economic significance," Swann writes, "these *t*-values can be used to measure precision—so long as the assumptions of econometric theory are satisfied." He is going in the wrong direction. "Precision" is not what one seeks, usually, in science. "Accuracy and [sampling] precision" are not the same thing, if one takes "accuracy" to mean "relevant numbers large enough to matter."

We can only admire, on the other hand, Swann's call to admit candidly in a serious scientific study "the weakest link in the analysis." "In reporting the results, we should make that 'Achilles heel' clear to readers." Right. But then: "and not suggest greater precision than this weakest link." Uh oh: again, "precision," The phony confidence comes from *t* tests, dragging the problem under the lamppost of sampling error. "I believe that the independence assumption is indeed the 'Achilles heel' of econometrics." No, the heel is not the problem. It's more like not having a leg to stand on.

Modulo sampling, Swann makes a good case that "In a world of low signal-to-noise ratios … parameter estimates are very sensitive to the independence assumption. While the independence assumption may be plausible, in the sense that we have no evidence to reject it, we should only rely on it if we are sure that any form of dependence between x and u is quite implausible. But in most cases, we cannot possibly be sure of that." Yet he's still measuring effect-size oomph with sampling variance. That's the real problem with the techniques of econometrics since Cowles. It's the "modulo sampling."

Yet Swann provides a good statement of the dire situation in econometrics: one is normally supplied in econometric studies "with no context, no definition of the variables, no understanding of what the axis units mean, and no idea of what the observations relate to." The reporting conventions for econometrics, inherited from the distant age of main-frame computation, do not provide such qualitative interpretations for the quantities (see Ziliak 2018). When you know them you can understand the speaker in the cocktail party. Swann quotes Verbeek's fine sarcasm: "econometrics is much easier without data." Swann then complains about "the very peculiar position of having a so-called estimator for a parameter that depends on our assumptions, and not on our data. That should ring alarm bells!"

Yes, but the call to arms is to common sense, and to Bayes.

§

What to do, really? We can only agree with Swann's worry that "education in econometrics has become quite unbalanced." But the imbalance is not about technical matters of error terms. It is the imbalance coming from the now routine of three terms of econometrics, meaning Cowles-Commission regression, which goes absurdly deeply into theoretical techniques on the head of a pin without any other quantitative methods being taught, or for that matter mentioned. Economists should know that diminishing returns applies in teaching as much as in agriculture. Swann correctly argues that economists "should give up the idea that econometrics can be a 'universal solvent' for all empirical questions in econometrics, and accept instead that the economics profession needs to use a wider range of empirical techniques." "A common objection to this suggestion is that many of these other techniques are 'woolly' and 'imprecise'. But from what we have seen in the sensitivity analysis, the same could be said of many econometric estimates." We stood up and cheered.

"I am not suggesting," Swann writes, though, in drawing back from the precipice of common sense, "that econometricians should learn to do qualitative case studies." We do not know why not. He argues that "the two methods are completely different, and call for quite different skills and personal qualities." So what? Economists write English and manipulate matrices, which call for quite different skills and personal qualities. But economists use both if they are serious about doing economics. The solution, Swann wisely says, is to "make proper space in the economics discipline for those economists who have learned a lot about empirical economics by using other techniques, and not dismiss these people as heterodox or irrelevant." We stood up and cheered again.

On the other hand (more backing away from the precipice) Swann says that "econometrics would remain a core part of the curriculum, but students should also learn that good econometrics supplements formal econometrics with other empirical research methods." On the second clause of the sentence, we can only give more cheering. But not on the first clause. We do not know why the con of econometrics of the Cowles sort should remain the "core." The core of an apple, after all, is not the nutritious or even the digestible part. As Ed Leaner said long ago, let's get the "con" out of econometrics. But then let's get the "tric[k]s"out. Then the narcissism of the "me." Then all that's left is a cry of pain, "eo."

What then should constitute the quantitative education of an economist?

The three terms are fine. But only one term should be Cowles econometrics, that is, regression analysis. At present the graduate programs have three terms of regression, and young economists emerge thinking that it is the sole method of

confronting evidence quantitatively. A terminological confusion arises from the very naming of people who specialize in (only) regression analysis as "econometricians," that is, quantitative economists. The same thing has happened since the 1970s in the teaching of economic theory, which is—again because of the mere name—handed over in the first year of graduate programs to otherwise unemployable mathematical economists, who insist on existence-theorems using a "real" (ha, ha!) analysis useless for actual sciences such as engineering or physics or meteorology, though justly favored in the math department, with its fascination with universal, un-computed, Greek-style proofs by contradiction.

That one term of regression analysis should include serious instruction in substantive loss functions, to replace Fisher's vacuous ceremony of two standard deviations. The exposition should be decision theoretic, instructing the students in maximum likelihood, too. It should get deeply into the matter of computational errors arising from variant computer programs (Stokes 1997, 2005).

For the other two terms the students should be instructed in all the other quantitative methods, beyond regression analysis, the ones actually used by applied economists or by scientists generally. It would be a good idea to have such courses taught by actual applied economists, not by theoretical econometricians. When the economist Arjo Klamer, who was well trained in the Dutch econometrics of the 1970s, went around to Dutch "Departments of Econometrics" he discovered that none of their members could so much as name any actual economic scientists. They had devolved into departments of theoretical statistics, grinding away at existence theorems (Klamer DDDD).

Let us list the quantitative methods that a serious empirical economist should be instructed in, at least to the level at which she knows how to get started in taking them up and becoming truly expert later in her career. Anyone who has been reading economics for many decades can give vivid and scientifically persuasive examples of each. We repeat: Cowles-Commission econometrics has not achieved a single similarly persuasive finding.

Applied mathematics, especially error analysis and convergent approximation, such as the well-named Newton's Method. Stop teaching and "using" (if that is quite the word) real analysis.

Simulation, with sensitivity analysis, such as is commonly used in agricultural economics, taking parameters from agronomical experiments, say, and is the central method of engineering and, recently, architecture.

Computable general equilibrium methods, as a subset of simulation.

Historical economics, providing comparative perspectives, considering that comparison (e.g., medieval China and medieval Europe; laminar flow of fluids vs. chaotic flow; continental plates vs. margins of plates) is one of the strongest quantitative methods in science.

National income accounting, no longer taught in economics, which allows students to see that assumptions about depreciation run the show, or that calculations of the balance of payments are deeply mischievous.

Graphing, the computer-assisted techniques, which have arrived at a high level (see Tufte's books, such as *The Quantitative Display of Quantitative Information*, DDDD). Why do some economists still think there is a Phillips Curve? Because when he wrote the article it leapt off the chart. Why do some no longer think so? Because the chart changed, statistical significance be damned.

Experiments, especially the experiments on groups of subjects, an actual *social* science such as Vernon Smith and Bart Wilson perform, as against the clumsy reinventions of psychology, without a finding of aggregate importance, typical of behavioral economics.

Questionnaires, which economists think they do not use—but of course the unemployment rate is a survey.

Introspection of a serious sort—not "how I feel about the law of demand this afternoon" but "what really would I do if the price of gasoline doubled?" The seriousness would entail philosophy on a level beyond the simpleton's version economists rely on.

Interviewing, of the sort that anthropologist use, evoking answers that matter scientifically. You don't have to believe everything that comes out of an informant's mouth ("Witchcraft made me ill"; "Marginal cost doesn't matter to our business") to learn from well-crafted interviews.

Intelligent listening, of the sort that guided Ronald Coase's career.

Field work inside actual economic entities, such as business firms, nonprofits, and governmental offices, or indeed households such as we all live in.

Walkabout economics, as the Irish economist and TV personality David McWilliams calls it, the acute observation of economic behavior in ordinary life, of the sort Armen Alchian did (McWilliams 2018, p. 95).

Archival work, that is, the historian's techniques of achieving *capta*, taught by having students actually do some of it.

Narrative techniques, such as evolutionary biology routinely uses.

All of this is what should actually be taught, if economics is to advance as a science.

Works Cited

Hakek, Friedrich. 1945. "The Use of Knowledge in Society." *American Economic Review* **35** (4): 519–530.

Klamer, Arjo.

Lodge, David. 2008. *Deaf Sentence: A Nevel.* London: Penguin.

McCloskey, Deirdre Nansen, and Stephen T. Ziliak. 2008. "Signifying Nothing: A Reply to Hoover and Siegler," March, 2008, *Journal of Economic Methodology*.

McCloskey, Deirdre Nansen, and Stephen T.. Ziliak. 1996. "The Standard Error of Regression." *Journal of Economic Literature* 34 (1): 97–114.

McCloskey, Deirdre Nansen, and Stephen T. Ziliak. 2012. "Statistical Significance in New Tom and Old: A Reply to Thomas Mayer on Statistical Significance," *Econ Journal Watch* 9(3): 298–308.

McWilliams, David. 2018. *Renaissance Nation: How the Pope's Children Rewrote the Rules for Ireland*. Dublin: Gill.

Morgenstern, O. 1963. *On the Accuracy of Economic Observations*. 2nd ed. Princeton: Princeton University Press.

Stokes, Houston J. 1997. *Specifying and Diagnostically Testing Econometric Models*, Quorum Books / Greenwood Press, Westport CT, Second Edition.

Stokes, Houston J. 2005. "The Sensitivity of Econometric Results to Alternative Implementations of Least Squares." *Journal of Economic and Social Measurement,* 30: 9-38.

Tufte, Edward R. 1983. The *Visual Display* of *Quantitative Information*. Cheshire, CT: Graphics Press.

Verbeek, M. 2000. *A Guide to Modern Econometrics.* New York: Wiley.

Ziliak, Stephen T. 2018. "How Large Are Your G-values? Try Gosset's Guinnessometrics When a Little "*p*" is NotEenough." Forthcoming 2018, *The American Statistician* (Special Issue: Statistical Inference in the 21st Century).

Ziliak, Stephen T., and Deirdre Nansen McCloskey. 2004. "Size Matters: The Standard Error of Regressions in the *American Economic Review*." *Journal of Socio-Economics* 33 (5): 527–546.

Ziliak, Stephen T., and Deirdre Nansen McCloskey. 2008. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives*. Ann Arbor: University of Michigan Press.

Ziliak, Stephen T., and Deirdre Nansen McCloskey. 2013. "We Agree That Statistical Significance Proves Essentially Nothing: A Rejoinder to Thomas Mayer." *Econ Journal Watch* 10(1): 97–107.