American Economic Association

Teaching Undergraduate Econometrics: A Suggestion for Fundamental Change
Author(s): Peter E. Kennedy
Source: *The American Economic Review*, Vol. 88, No. 2, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association (May, 1998), pp. 487-492
Published by: American Economic Association
Stable URL: http://www.jstor.org/stable/116972
Accessed: 19/08/2010 20:26

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=aea.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*.

http://www.jstor.org

Teaching Undergraduate Econometrics: A Suggestion for Fundamental Change

By PETER E. KENNEDY *

Contrary to the belief of most econometrics instructors, upon completion of introductory statistics courses the vast majority of students do not understand the basic logic of classical statistics as captured in the sampling-distribution concept. They have learned to do mechanical things such as compute a sample variance, run a regression, and test an hypothesis, and they know they can pass the course by remembering how these techniques work. They view statistics as a branch of mathematics because it uses mathematical formulas, so they look at statistics through a mathematical lens.

What they are missing is the statistical lens through which to view this world, allowing this world to make sense. The sampling-distribution concept is this statistical lens. My own experience discovering this lens was a revelation, akin to the experience I had when I put on my first pair of eyeglasses: suddenly everything was sharp and clear. In this paper I discuss this issue and suggest a means whereby instructors of undergraduate econometrics courses (i.e., following introductory statistics) can provide the missing statistical lens.

I. The Importance of the Sampling-Distribution Concept

"Constructivism," a recent theory of learning, has been widely accepted in education communities and is the guiding theory for much research and reform in mathematics and science education. According to this theory, students bring to the classroom their own ideas, and rather than passively adding to these ideas as material is presented in class, they actively restructure the new information to fit it into their own cognitive frameworks. In this way they are "constructing" their own knowledge, rather than copying knowledge delivered to them through some teaching mechanism. Joan Garfield (1995 p. 30) provides this explanation:

> Regardless of how clearly a teacher or book tells them something, students will understand the material only after they have constructed their own meaning for what they are learning. Moreover, ignoring, dismissing, or merely "disproving" the students' current ideas will leave them intact—and they will outlast the thin veneer of course content.
>
> Students do not come to class as "blank slates" or "empty vessels" waiting to be filled, but instead approach learning activities with significant prior knowledge. In learning something new, they interpret the new information in terms of the knowledge they already have, constructing their own meanings by connecting the new information to what they already believe. Students tend to accept new ideas only when their old ideas do not work, or are shown to be inefficient for purposes they think are important.

An important consequence of this theory is that students whose personal cognitive framework does not match that of the instructor/textbook find it difficult to learn what the instructor wants them to learn, because what is taught in the classroom appears to be a set of unrelated ideas that must be memorized. The result is frustration, poor understanding of course material, and a dislike of the subject. I believe that this problem characterizes econometrics, and that the main culprit is that students come to econometrics, at both the undergraduate and graduate level, not understanding adequately the sampling-distribution concept.

As illustrated below, this concept provides a unifying logic that permits estimation,

* Department of Economics, Simon Fraser University, Burnaby, BC, Canada V5A 1S6.

hypothesis testing, and econometric's algebraic manipulations to be seen in context:

(i) Using a formula β^* to produce an estimate of β can be conceptualized as the econometrician shutting his or her eyes and obtaining an estimate of β by reaching blindly into the sampling distribution of β^* to obtain a single number.
(ii) Because of (i) above, choosing between β^* and and a competing formula β^{**} comes down to the following: Would you prefer to produce your estimate of β by reaching blindly into the sampling distribution of β^* or by reaching blindly into the sampling distribution of β^{**}?
(iii) Because of (ii) above, desirable properties of an estimator β^* are defined in terms of its sampling distribution. For example, β^* is unbiased if the mean of its sampling distribution equals the number β being estimated.
(iv) Because of (iii) above, econometricians spend a lot of algebraic energy figuring out sampling-distribution properties, such as mean and variance.
(v) The properties of the sampling distribution of an estimator β^* depend on the process generating the data, so an estimator can be a good one in one context but a bad one in another; β^*'s sampling-distribution properties need to be recalculated for every different data-generating process.
(vi) All statistics, not just parameter estimates, have sampling distributions. For example, under the null hypothesis, an F statistic will have a sampling distribution described by the F table found in statistics textbooks.

II. Do Students Understand the Sampling Distribution Concept?

There is much research, summarized for example by Garfield and Andrew Ahlgren (1988), showing that a large proportion of university students in introductory statistics courses do not understand many of the concepts they are studying. But I cannot refer to this research to defend my claim that students do not understand the sampling-distribution concept, because this research relates mainly to the psychology of probability in which students can be shown to use misleading heuristics to compute probabilities. Instead I defend my claim as follows:

(i) I have sampled scores of students at the beginning graduate and upper-level undergraduate level, finding that very few are able to explain what is a sampling distribution. Most think it is a distribution pictured by drawing a histogram of the sample data.
(ii) I have delivered to new graduate students an expository lecture on the sampling-distribution concept and the role it plays in statistics/econometrics, and through anonymous surveys after this lecture have discovered that almost all students admit that they did not realize the important conceptual role of the sampling distribution in econometrics.
(iii) I frequently encounter students with A grades in their introductory statistics courses who clearly have no understanding of statistics beyond a mechanical ability to apply standard statistical procedures.
(iv) I often see graduate students with an impressive ability to derive statistical formulas but a remarkable inability to explain how to evaluate those formulas via a Monte Carlo study, even after Monte Carlo procedures have been explained.
(v) Others have expressed a similar sentiment. Ruth Hubbard (1997 p. 11) complains that "When graduate students from a variety of disciplines approach me for help with analysing or interpreting their data, I always begin by asking if they understand the statistical term 'standard deviation.' Their standard response is, 'It is something that you calculate from a formula.' In most cases no amount of probing succeeds in clarifying either the formula or what it measures." Vijaya Duggal (1987 p. 26) contends "that more than three-fourths of students finishing a course in quantitative methods who know the mechanics of hypothesis testing have no conceptual

comprehension of the sampling distribution of the mean.''

How has this state of affairs come about? It is not because the introductory statistics books ignore sampling distributions; they all have plenty of good material on this concept and give it appropriate emphasis. And it is not because instructors ignore this dimension of introductory textbooks; all instructors swear that they teach this concept thoroughly. I suggest three reasons why students fail properly to understand this concept. First, it is difficult. Students can visualize a distribution of sample observations, but a sampling distribution is at a higher level of abstraction, where sample observations yield a single value of a statistic, not an entire distribution. Second, texts typically introduce the sampling-distribution concept in the context of the sample mean statistic but do not follow it up adequately when discussing regression, the central focus of econometrics. Third, and I believe most important, as Hubbard (1997 p. 1) quotes Lauren Resnick ''We get what we assess, and if we don't assess it, we won't get it.'' Students are utility-maximizers. Typically their exams require an ability to interpret regressions, calculate t statistics, and perform hypothesis tests, but not an ability to explain the concept of a sampling distribution. They cannot imagine, and instructors and textbooks seldom provide, examples of exam questions that in any meaningful way probe student understanding of this concept.

Good instructors realize that students have mostly forgotten their introductory statistics material, or never learned it properly in the first place, and proceed on the basis that this introductory material needs to be reviewed. Unfortunately, this review suffers from the same problem noted earlier: there is little motivation for students to learn properly the concept of a sampling distribution and the role it plays in econometrics because instructors' expositions of this concept, however clear, are seldom accompanied by appropriate example exam questions. Such questions provide motivation and, more importantly, force students to work out answers. Although lecturers like to think otherwise, brilliant expositions seldom cause students fully to understand; such understanding comes through working out problems based on the concept to be learned. An old Chinese proverb bears repeating: I hear, I forget; I see, I remember; I do, I understand.

Econometrics textbooks do not help much. They are in too big a hurry to produce the theorems, proofs, and formulas that define theoretical econometrics. Some review introductory statistics but do so without much emphasis on sampling distributions and seldom provide meaningful questions that force student understanding of the sampling-distribution concept. I am amazed that econometrics textbooks pay so little attention to the sampling-distribution concept, but it is consistent with my observation that most econometrics instructors believe that their students already have a good understanding of this concept. This impression is reinforced by reading the very limited literature on teaching econometrics. Eric Sowey (1983) and his commentators, for example, ignore this issue. This is indeed a major part of the problem. If instructors are not aware of this deficiency in student understanding, they will not try to remedy it.

III. A Suggestion for Change

Two changes are needed. First, instructors need to produce for students a good exposition of what is a sampling distribution, couched in the context of regression coefficient estimation, followed by a discussion of the overall role of the sampling-distribution concept in econometrics. I have provided my own such exposition elsewhere (Kennedy, 1988a, b appendix A). An important ingredient of this exposition is an explanation of how, for a given data-generating context, econometricians learn the properties of statistics' sampling distributions. This involves discussion of the roles of mathematical derivations and asymptotic algebra and, most importantly, an exposition of Monte Carlo studies: how computer simulation can be used to investigate the properties of a sampling distribution.

Second, a way must be found to hammer home this lesson so that the perspective it provides for later study is not forgotten. I suggest the use of ''explain how to do a Monte Carlo

study'' problems. By forcing students to outline step-by-step a Monte Carlo study to examine an econometric issue, they are forced to spell out very clearly their understanding of this issue. For example, I ask students to explain how to conduct a Monte Carlo study to examine the implications of omitting a relevant explanatory variable. Instructors may wish to ease students into this kind of question by providing the Monte Carlo instructions and asking students to anticipate the results, as in the following example:

(a) Draw 50 x values from a uniform distribution between 3 and 12.
(b) Draw 50 z values from a standard normal distribution.
(c) Compute 50 w values as $4 - 3x + 9z$.
(d) Draw 50 e values from a standard normal distribution.
(e) Compute 50 y values as $2 + 3x + 4w + 5e$.
(f) Regress y on x and save the x slope coefficient estimate b1.
(g) Regress y on x and w and save the x slope coefficient estimate bb1.
(h) Repeat this procedure from (d) to get 1,000 b and bb values.
(i) Compute the averages of these sets of 1,000 values to get B and BB, respectively.
(j) Compute the variances of these sets of 1,000 values to get VB and VBB, respectively.

Should B or BB be closer to 3? Should VB or VBB be closer to 0?

Kennedy (1998b appendix D) contains additional examples. Based on many years of experience using this approach I suggest several dos and don'ts (discussed at greater length in Kennedy [1998a]) for ensuring its success:

(i) Only ask students actually to do a Monte Carlo study if you are willing to pay the high opportunity cost of having them learn how to program.
(ii) Be prepared for the question: Won't our choice of parameter values affect the results of the Monte Carlo study?
(iii) Be very careful evaluating students' Monte Carlo descriptions because small mistakes, and especially vagueness, can reflect major misunderstandings.
(iv) Begin with very easy questions such as finding the variance of the sample mean statistic.
(v) Do not be surprised that, once the mechanical steps of a Monte Carlo study are in hand, students' biggest problem is how to simulate the data-generating process. If they cannot explain how to generate raw data for a heteroscedastic-error model, for a two-equation simultaneous-equation system, or for a qualitative-choice model, can you be confident that they understand what any of your algebra is about?
(vi) In-class development of a step-by-step set of Monte Carlo instructions can be an effective means of generating classroom participation. Call on a student to suggest the first step, then go around the room asking each student in turn to supply the next step, or correct an incorrect step suggested by an earlier student.

As should be evident from the discussion above, this proposal requires that instructors spend considerable time on sampling distributions and Monte Carlo studies, something most do not now do. What should be given up to accommodate this new feature? I believe that the benefit of this innovation is so large that no instructor could possibly claim that the least-valued subset of their course has more benefit. Indeed, this innovation should enhance student understanding of all dimensions of a course, creating benefits beyond those attached solely to learning about data generation, sampling distributions, and Monte Carlo studies. Advanced estimation or testing techniques do not mean much to students if they do not understand the fundamental principles that lie behind them.

In my own courses, I have given up most mathematical derivations. Do we really want our students at the end of their course to be able to do technical things like derive the ordinary least-squares estimator and prove that it is BLUE? Such things have little meaning if the concept of a sampling distribution is not thoroughly understood.

IV. Conclusion

This paper urges fundamental change in the way in which undergraduate econometrics is taught. It argues that, at the beginning of the traditional undergraduate econometrics course (i.e., the course following introductory statistics), instructors should exposit carefully the concept of a sampling distribution, explain how Monte Carlo studies can be used to characterize sampling distributions, and, most important, follow this up throughout the course with assignments asking students to explain how to conduct Monte Carlo studies to investigate specific econometric questions.

Students find this approach intellectually very challenging because it forces them to think hard to understand an abstract concept. Julian Simon and Peter Bruce (1991 p. 29) make the same point in the context of their resampling approach to learning basic statistics, claiming that it "requires only hard, clear thinking. You cannot beg off by saying 'I have no brain for math!' " Some students do not rise to this challenge; they are more comfortable learning by rote a bunch of techniques and mathematical proofs and so find this approach more difficult.

An obvious corollary to the message of this paper should be emphasized. Instructors of beginning-level graduate econometrics courses need to realize that their students do not have a good understanding of the sampling-distribution concept and how it fits into econometrics. A story told by one of my former undergraduate students illustrates this. At the beginning of his graduate econometrics course he was struggling because he had to learn from scratch formulas and algebraic derivations that other students had encountered in their more traditional undergraduate econometrics courses. But he found that after the first month or so he had become comfortable with the math and suddenly leapfrogged far ahead of the rest of the class because, he said, he understood what was going on but they understood only the mathematics.

REFERENCES

Duggal, Vijaya G. "Coping with the Diversity of Student Aptitudes and Interests." *American Economic Review*, May 1987 (*Papers and Proceedings*), *77*(2), pp. 24–28.

Garfield, Joan. "How Students Learn Statistics." *International Statistical Review*, April 1995, *63*(1), pp. 25–34.

Garfield, Joan and Ahlgren, Andrew. "Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research." *Journal for Research in Mathematics Education*, January 1988, *19*(1), pp. 44–63.

Hubbard, Ruth. "Assessment and the Process of Learning Statistics." *Journal of Statistics Education*, March 1997, *5*(1), ⟨www.stat.ncsu.edu/info/jse/v5n1/hubbard.html⟩.

Kennedy, Peter E. "Using Monte Carlo Studies for Teaching Econometrics," in W. E. Becker and M. Watts, eds., *Teaching undergraduate economics: Alternatives to chalk and talk*. Aldershot, U.K.: Edward Elgar, 1998a (forthcoming).

———. *A guide to econometrics,* 4th Ed. Cambridge, MA: MIT Press, 1998b.

Simon, Julian L. and Bruce, Peter C. "Resampling: A Tool for Everyday Statistical Work." *Chance*, Winter 1991, *4*(1), pp. 22–32.

Sowey, Eric. "University Teaching of Econometrics: A Personal View," with comments by Jacques Dreze, Pascal Mazodier, Peter Phillips, Takamitsu Sawa, and Arnold Zellner. *Econometric Reviews,* 1983, *2*(2), pp. 255–333.

AMERICAN
ECONOMIC
ASSOCIATION