# INTRODUCTION TO AN ONLINE HANDBOOK FOR THE USE OF UP-TO-DATE ECONOMETRICS IN ECONOMIC EDUCATION RESEARCH

William E. Becker
Professor of Economics, Indiana University, Bloomignton, Indiana, USA
Adjunct Professor of Commerce, University of South Australia, Adelaide, Australia
Research Fellow, Institute for the Study of Labor (IZA), Bonn, Germany
Fellow, Center for Economic Studies and Institute for Economic Research (CESifo, Munich Germany)

The Council for Economic Education (first named the Joint Council on Economic Education and then the National Council on Economic Education) together with the American Economic Association Committee on Economic Education have a long history of advancing the use of econometrics for assessment aimed at increasing the effectiveness of teaching and learning economics. As described by Welsh (1972), the first formal program was funded by the General Electric Education Fund and held in 1969 and 1970 at Carnegie-Mellon University, and was a contributing force behind the establishment of the *Journal of Economic Education* in 1969 as an outlet for research on the teaching of economics. As described in Buckles and Highsmith (1990), in 1987 and 1988 a second econometrics training program was sponsored by the Pew Charitable trust and held at Princeton University, with papers featured in the *Journal of Economic Education*, Summer 1990. Since that time there have been major advances in econometrics and its applications in education research -- most notably, in the specification of the data generating processes and computer programs for estimating the models representing those processes.

Becker and Baumol (1996) called attention to the importance that Sir Ronald A. Fisher assigned to "random arrangements" in experimental design.[i] Fisher designed experiments in which plots of land were randomly assigned to different fertilizer treatments to assess differences in yield. By measuring the mean yield on each of several different randomly assigned plots of land, Fisher eliminated or "averaged out" the effects of nontreatment influences (such as weather and soil content) so that only the effect of the choice of fertilizer was reflected in differences among the mean yields. For educational research, Fisher's randomization concept became the ideal: hypothetical classrooms (or students) are treated as if they could be assigned randomly to different instructional procedures.

In education research, however, the data are typically not generated by well-defined experiments employing random sampling procedures. Our ability to extract causal inferences from an analysis is affected by sample selection procedures and estimation methods. The absence of experimental data has undoubtedly provided the incentive for the econometricians to exert their considerable ingenuity to devise powerful methods to identify, separate out, and evaluate the magnitude of the influence exercised by each of the many variables that determine the shape of any economic phenomenon. Econometricians have been pioneers in the design of techniques to deal with missing observations, missing variables, errors in variables, simultaneous relationships and pooled cross-section and time-series data. In short, and as comprehensively

reviewed by Imbens and Wooldridge (2009), they have found it useful to design an armory of analytic weapons to deal with the messy and dirty statistics obtained from opportunistic samples, seeking to extract from them the *ceteris paribus* relationships that empirical work in the natural sciences is able to average out with the aid of randomized experiments. At the request of the Council for Economic Education and the American Economic Association Committee on Economic Education, I have developed four modules that will enable researchers to employ these weapons in their empirical studies of educational practices, with special attention given to the teaching and learning of economics.

## MODULE ONE

Although first employed in policy debate by George Yule at the turn of the 20th century, economic educators and general educationalists alike continue to rely on least-squares estimated multiple regressions to extract the effects of unwanted influences.[ii] This is the first of the four modules designed to demonstrate and enable researchers to move beyond these basic regression methods to the more advanced techniques of the 21st century using any one of three computer programs: LIMDEP (NLOGIT), STATA and SAS.

This module is broken into four parts. Part One introduces the nature of data and the basic data generating processes for both continuous and discrete dependent variables. The lack of a fixed variance in the population error term (heteroscedasticity) is introduced and discussed. Parts Two, Three and Four show how to get that data into each one of the three computer programs: Part Two for LIMDEP (NLOGIT), Part Three for STATA (by Ian McCarthy, Senior Consultant, FTI Consulting) and Part Four for SAS (by Gregory Aaron William Gilpin, Assistant Professor of Economics, Montana State University). Parts Two, Three and Four also provide the respective computer program commands to do least-squares estimation of the standard learning regression model involving a continuous dependent test-score variable but with the procedures to adjust for heteroscedastic errors. The maximum likelihood routines to estimate probit and logit models of discrete choice are also provided using each of the three programs. Finally, statistical tests of coefficient restrictions and model structure are presented.

## MODULE TWO

The second of four modules is devoted to endogeneity in educational studies. Endogeneity is a problem caused by explanatory variables that are related to the error term in a population regression model. As explained in Part One of Module Two, this error term and regressor dependence can be caused by omitted relevant explanatory variables, errors in the explanatory variables, simultaneity (reverse causality between $y$ and an $x$), and other sources emphasized in subsequent modules. Endogeneity makes least squares estimators biased and inconsistent. The uses of natural experiments, instrumental variables and two-stage least squares are presented as means for addressing endogeneity. Parts Two, Three and Four show how to perform and provide the commands for two-stage least squares estimation in LIMDEP (NLOGIT), STATA (by Ian McCarthy) and SAS (by Gregory Gilpin) using data from a study of the relationship between multiple-choice test scores and essay-test scores.

**MODULE THREE** (co-authored by W. E. Becker, J.J. Siegfried and W. H. Greene)

As seen in Modules One and Two, the typical economic education empirical study involves an assessment of learning between a pretest and a posttest. Other than the fact that testing occurs at two different points in time (before and after an intervention) , there is no time dimension in this model of learning. Panel data analysis provides an alternative structure in which measurements on the cross section of subjects are taken at regular intervals over multiple periods of time. Collecting data on the cross section of subjects over time enables a study of change. It also opens the door for economic education researchers to look at things other than test scores that vary with time.

This third in the series of four modules provides an introduction to panel data analysis with specific applications to economic education. The data structure for a panel along with constant coefficient, fixed effects and random effects representations of the data generating processes are presented. Consideration is given to different methods of estimation and testing. Finally, as in Modules One and Two, contemporary estimation and testing procedures are demonstrated in Parts Two, Three and Four using LIMDEP (NLOGIT), STATA (by Ian McCarthy) and SAS (by Gregory Gilpin).

**MODULE FOUR** - to be added, but tentatively to be done by W. E. Becker and W. H. Greene and to include comments on

a. Sample selection problems and alternative modeling and estimation
b. Heckman selection and correction
c. Quasi-random experiments
d. Regression discontinuities
e. Propensity score matching

**ALL MODULES AVAILABLE AT**

http://www.vanderbilt.edu/AEA/AEACEE/Econometrics_Handbook/index.htm

**RECOGNITION**

As with the 1969-1970 and 1987-1988 workshops, these four modules have been made possible through a cooperative effort of the National Council on Economic Education and the American Economic Association Committee on Economic Education. This new work is part of the Absolute Priority direct activity component of the *Excellence in Economic Education* grant to the NCEE and funded by the U.S. Department of Education Office of Innovation and Improvement. Special thanks are due Ian McCarthy and Greg Gilpin for their care in duplicating in STATA and SAS what I did in LIMDEP (in Modules One, Two and Three). Comments and constructive

criticism received from William Bosshardt, Jennifer (gigi) Foster, Peter Kennedy, Mark Maier, KimMarie McGoldrick, Gail Hoyt, Martin Shanahan, Robert Toutkoushian and Michael Watts in the Beta testing must also be thankfully acknowledged. Finally, as with all my writing, Suzanne Becker must be thanked for her patience in dealing with me and excellence in editing.

## REFERENCE

Becker, William E. and William J. Baumol (1996). *Assessing Educational Practices: The Contribution of Economics,* Cambridge MA.: MIT Press.

Buckles, Stephen and Robert Highsmith (1990). "Preface to Special Research Issue," *Journal of Economic Education*, (Summer): 229-230.

Clogg, C. C. (1992). "The impact of sociological methodology on statistical methodology," *Statistical Science* , (May): 183-96.

Fisher, R. A. (1970). *Statistical methods for research workers*, 14th ed. New York: Hafner.

Imbens, Guido W. and Jeffrey M. Wooldridge (2009). "Recent development in econometrics of program evaluation," *Journal of Economic Literature*, (March): 5-86.

Simpson, E.H. (1951). "The interpretation of interaction in contingency tables", *Journal of the Royal Statistical Society,* Series B, 13: 238-241.

Stigler, Stephen M. (1986). *The History of Statistics: The measurement of Uncertainty before 1900*. Cambridge: Harvard University Press.

Welsh, Arthur L. (1972). *Research Papers in Economic Education*. New York: Joint Council on Economic Education.

**ENDNOTES**

i. In what is possibly the most influential book in statistics, *Statistical Methods for Research Workers*, Sir Ronald Fisher wrote:

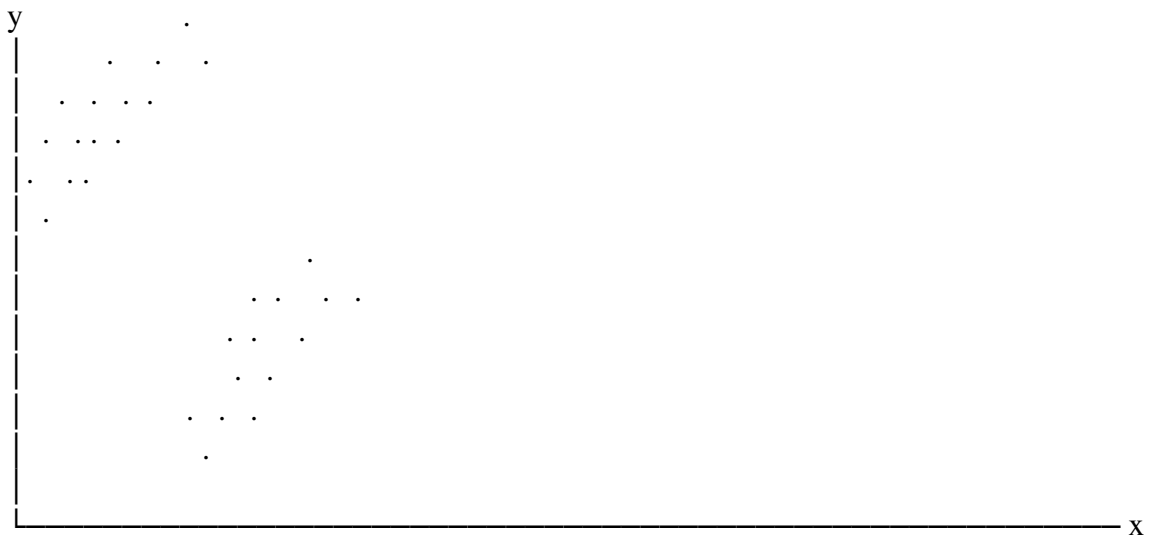
> The science of statistics is essentially a branch of Applied Mathematics, and may be regarded as mathematics applied to observational data. (p. 1)
>
> Statistical methods are essential to social studies, and it is principally by the aid of such methods that these studies may be raised to the rank of sciences. This particular dependence of social studies upon statistical methods has led to the unfortunate misapprehension that statistics is to be regarded as a branch of economics, whereas in truth methods adequate to the treatment of economic data, in so far as these exist, have mostly been developed in biology and the other sciences. (1970, p. 2)

Fisher's view, traceable to the first version of his book in 1925, is still held today by many scholars in the natural sciences and departments of mathematics. Econometricians (economists who apply statistical methods to economics), psychometricians, cliometricians, and other "metricians" in the social sciences have different views of the process by which statistical methods have developed. Given that Fisher's numerous and great contributions to statistics were in applications within biology, genetics, and agriculture, his view is understandable although it is disputed by sociologist Clifford Clogg and the numerous commenters on his article "The Impact of Sociological Methodology on Statistical Methodology," *Statistical Science*, (May 1992). Econometricians certainly have contributed a great deal to the tool box, ranging from simultaneous equation estimation techniques to a variety of important tests of the validity of a statistical inference.

ii. George Yule (1871-1951) designed "net or partial regression" to represent the influence of one variable on another, holding other variables constant. He invented the multiple correlation coefficient $R$ for the correlation of $y$ with many $x$'s. Yule's regressions looked much like those of today. In 1899, for instance, he published a study in which changes in the percentage of persons in poverty in England between 1871 and 1881 were explained by the change in the percentage of disabled relief recipients to total relief recipients (called the "out-relief ratio"), the percentage change in the proportion of old people, and the percentage change in the population.

> Predicted change percent in pauperism = – 27.07 percent + 0.299(change percentage in out-relief ratio) + 0.271(change percentage in proportion of old) + 0.064(change percentage in population).

Stigler (1986, pp. 356-7) reports that although Yule's regression analysis of poverty was well known at the time, it did not have an immediate effect on social policy or statistical practices. In part this meager response was the result of the harsh criticism it received from the leading English economist, A.C. Pigou. In 1908, Pigou wrote that statistical reasoning could not be rightly used to establish the relationship between poverty and out-relief because even in a multiple regression (which Pigou called "triple correlation") the most important influences, superior program management and restrictive practices, cannot be measured quantitatively.

Pigou thereby offered the most enduring criticism of regression analysis; the possibility that an unmeasured but relevant variable has been omitted from the regression and that it is this variable that is really responsible for the appearance of a causal relationship between the dependent variable and the included regressors. Both Yule and Pigou recognized the difference between marginal and partial association.

Today some statisticians assign credit for this identification to E. H. Simpson (1951); the proposition referred to as "Simpson's paradox" points out that marginal and partial association can differ even in direction so that what is true for parts of a sample need not be true for the entire sample. This is represented in the following figure where the two separate regressions of *y* on the high values of *x* and the low values of *x* have positive slopes but a single regression fit to all the data shows a negative slope. As in Pigou's criticism of Yule 90 years ago, this "paradox" may be caused by an omitted but relevant explanatory variable for *y* that is also related to *x*. It may be better named the "Yule-Pigou paradox" or "Yule-Pigou effect." As demonstrated in Module Two, much of modern day econometrics has dealt with this problem and it occurrence in education research.