JAMES HECKMAN
2000 Nobel Laureate in Economics

Honoring James Heckman's Contributions to Economics: Identification, Heterogeneity, and Economic Models

Steven D. Levitt

Given James Heckman's enormous contributions to the field of economics and the social sciences more generally, the question regarding the Nobel prize was not "if," but "when." The choice of Heckman as the 2000 Nobel laureate (joint with Daniel McFadden, another deserving winner) has been warmly received throughout the profession. Heckman and McFadden were both pioneers whose work helped to create modern empirical economics.

In this short essay, I begin by briefly reviewing the empirical transformation of the field of economics over the last three decades. I spend the remainder of the paper reflecting on a number of important themes that have characterized (and continue to be integral parts) of Heckman's research over the course of his career.[1] His approach has had a profound impact on how I and many other empirical economists of the younger generation think about research and also has a broad applicability to social science research in other fields as well.

Steven D. Levitt is a senior research scholar at the American Bar Foundation, Chicago, and a professor in the Department of Economics, University of Chicago.

1. My discussion is left intentionally general and nontechnical. For a more in-depth presentation of Heckman's specific contributions, see Heckman (2001).

0897-6546/02/2701-35$01.00

THE REVOLUTION IN EMPIRICAL ECONOMICS

Historically, economics has been a field dominated by theorists. While the great economists have always been motivated by important real-world problems, the mode of analysis was virtually always theoretical in emphasis for Adam Smith, Alfred Marshall, A. C. Pigou, John Maynard Keynes, Paul Samuelson, Robert Lucas, Gary Becker, and others of their intellectual rank.[2] The primary reason for the theoretical focus was technological: only in recent years have increased computing capacities and the collection of large-scale data sets made modern empirical work feasible.[3]

Heckman was among the first to appreciate the coming revolution in empirical economics and, with a series of important papers, provided both the tools necessary to analyze these new data sets and also leading examples of these tools in practice (e.g., Heckman 1978, 1979; Heckman and Honore 1990; Heckman and Robb 1985; Heckman and Singer 1984; Heckman and Hotz 1989; Heckman, Lochner, and Taber 1998).[4] The remainder of this essay is devoted to highlighting what I consider to be the three fundamental principles of Heckman's approach to empirical research—principles that, in my opinion, are not nearly as widely appreciated or applied as they should be.

THE PRINCIPLES UNDERLYING HECKMAN'S APPROACH

In the remainder of this article, I focus on Heckman's contributions to our understanding and appreciation of three different (but sometimes interrelated) issues: identification, heterogeneity, and the role of economic models in empirical research.

---

2. Milton Friedman, who combined empirical and theoretical insights, is a notable exception.

3. This theoretical emphasis in economics, it should be noted, stands in contrast to most other social scientific disciplines. Anthropology, for instance, has always been an empirical discipline, with luminaries such as Margaret Mead and Marshall Sahlins. Psychology, with its emphasis on laboratory experiments (which do not require extensive computing power by virtue of randomization), has also been dominated by empiricists for decades.

4. Precisely because empirical analysis in the other social sciences does not involve the intersection of large-scale data and necessarily complex analytical approaches, the impact of computing power on these other disciplines has been much less great.

Legal scholarship has generally been slow to embrace micro-empirical approaches. Surprisingly, law and economics historically has been and continues to be a field overwhelmingly dominated by theorists (e.g., Ronald Coase, Gary Becker, Guido Calabresi, Richard Posner). The law and society movement, of course, is a notable exception to the theoretical emphasis in law. Because of the historical divide between law and society and law and economics, Heckman's influence on the empirical study of the law is far less great than it should be, his own contributions notwithstanding (Donohue and Heckman 1991; Heckman and Payner 1989).

Understanding and Demonstrating the Importance of "Identification"

In economics, the term *identification* refers to the source of variation from which a parameter estimate is derived. Understanding the source and validity of an identification strategy is absolutely critical in determining the plausibility and applicability of empirical research.

Economists use a number of different sources of identification in their work: aggregate time-series, cross-sectional, panel data, natural experiments, and randomized experiments are some of the more common approaches. It is impossible within the space constraints of this article to demonstrate the enormity of Heckman's contribution to the understanding of identification. Rather, I am reduced to a superficial accounting of some of the subjects he has touched upon. (For a more detailed discussion, see Heckman 2001). One side of Heckman's work in this area is laying bare the hidden assumptions and misinterpretations of particular sources of identification, most notably natural experiments (Heckman 1997; Heckman 2000; Heckman and Vytlacil 2001). His research has also shown the very important weaknesses of randomized experiments (Heckman and Smith 1995) and audit-pair studies (Heckman and Siegelman 1993)—two approaches that to the uninitiated would appear to be ideal sources of variation. On the other hand, some of Heckman's greatest contributions have been in demonstrating that economic models can provide plausible sources of identification in a variety of settings, for example when different types of economic actors self-select (Heckman and Honore 1990) or in a hedonic model (Ekeland, Heckman, and Nesheim 2001).

In my opinion, the question of identification is of paramount importance in empirical work and the greatest improvements in applied economics over the last three decades relate to an increased understanding and appreciation of identification. Heckman has been instrumental to this profound change in the profession. There is no other individual whose work has influenced empirical economists in this dimension as much as his has.

Understanding and Demonstrating the Importance of Heterogeneity

Perhaps more than for any other issue, Heckman has made his reputation by appreciating the impact that heterogeneity has in empirical contexts. Heterogeneity comes in many forms. A classic form of heterogeneity is self-selection. In the presence of self-selection, estimates are likely to be biased and to lead to flawed inference with respect to optimal public policy. For instance, if one naively estimates the impact of wages on the number of hours that women work using only the sample of women who work a

positive number of hours, the results may very well be misleading. To the extent that unobserved characteristics of women are correlated both with labor market participation and market wages, the estimates will be biased. A recurring theme in Heckman's research has been devoted to understanding how to overcome this sort of problem (Flinn and Heckman 1982; Heckman 1979; Heckman and Honore 1990).

A more subtle issue—one that is still not well appreciated by most empirical researchers—is the fact that in the presence of heterogeneity in response to treatment, extrapolating parameter estimates from one population to other parts of the distribution can be misleading. This problem is most extreme in natural experiment papers in which, typically, an estimate is based on a very small subset of the population. For instance, one could try to estimate the benefit of more years of schooling by looking at changes in college attendance in response to fluctuations in tuition. Alternatively, one might attempt to use changes in how binding mandatory school attendance laws are. One could easily imagine, however, that the returns to schooling will vary widely among students. Thus, even if one obtained an unbiased estimate of the returns to schooling among the population affected by a particular natural experiment, it may be impossible to usefully extrapolate that coefficient outside of the targeted sample, leading to faulty public policy recommendations. In a series of papers (Heckman 1997; Heckman 2000; Heckman, Lalonde, and Smith 1999; Heckman and Vytlacil 1999; Ekeland, Heckman, and Nesheim 2001), Heckman has demonstrated the potential for misinterpretation of parameter estimates in a range of contexts. Typical of his approach to research, however, he has not been content to merely criticize current practices. Rather, he has developed an array of methods that allow researchers to estimate the entire distribution of parameter estimates across individuals and to interpret those coefficients in a way that allows for meaningful public policy advice.

Understanding and Demonstrating the Importance of Economic Models to Empirical Work

Earlier in this essay I argued for the increasing influence of empirical research in the field of economics. One unfortunate side-effect of the rise of empiricism is that among many empirical economists of the younger generation, the connection between economic models and estimation has been severed.[5] In the most extreme view, it is argued that all that can be learned

5. The bifurcation of the field into those who are empirically focused and those who do theory is widely recognized and maligned within the profession. Perhaps the clearest example of this is within the field of econometrics itself. Most of the most prominent econometricians who came of age in the 1970s and 1980s were motivated by problems of applied work. While Heckman is perhaps the greatest example of this, there are many others who worked in this

is in the data. Any further structure or assumptions imposed by the researcher interfere with the lessons that can be learned from empirical analysis. For others, while there is not active hostility to economic modeling, there is a decided lack of interest and appreciation. Many empirical economists of my generation—some quite successful—have never derived a formal economic model.

The primary virtues of atheoretical empirical work are that (a) it is easy to do, (b) it is easy to evaluate. Given training in basic empirical approaches, it is possible for an army of researchers to churn out plausible parameter estimates on questions of interest. Knowing the basic empirical rules, judging the quality of such research is relatively straightforward.

Heckman's career, however, has been devoted to demonstrating that neither theory nor estimation can stand alone. Only when those two pieces of the puzzle are brought together do the greatest insights emerge. As was mentioned above, economic models can provide a means of identification when atheoretical approaches cannot (e.g., Flinn and Heckman 1982; Heckman 1979; Heckman and Honore 1990; Heckman, Lochner, and Taber 1998) and economic models can be immensely useful in allowing researchers to take available estimates and meaningfully extrapolate these estimates out of sample.

In my own view, however, the importance of economic models is greater than this. The success of economics as a discipline is in no small part due to the rigor that formal theory imposes (e.g., optimization, rationality, formal statement of assumptions). Moreover, this shared set of tools and beliefs makes it possible for economists to readily understand the ideas and insights being brought to bear in other parts of the field. The borrowing of ideas from one area, properly adjusted to fit a different problem, is a very efficient means of advancing knowledge. If empirical economists are not choosing questions based on theoretical puzzles and vice versa, then either one or both strands of research is likely to become irrelevant. Even if an empirical researcher cannot make seminal contributions to theoretical research (and as economists since Adam Smith have argued, the virtues of specialization often argue against being a generalist), appreciation of theory is important ingredient. And it is those rare few with talents so immense they can master both sides of the equation, on whose shoulders it falls to advance knowledge by great leaps rather than baby steps. Surely, James Heckman's career is a case study in how an enormous intellect, focused squarely on the questions of greatest importance, can make such a contribution.

---

vein: Lars Hansen, Jerry Hausman, Gary Chamberlain, and others. Most of these individuals both wrote econometric theory and made important empirical contributions. Among the younger cohort of econometricians, however, there has been a tendency toward increased specialization and the development of more complicated methods, which are rarely, if ever, adopted by those actually doing empirical research.

REFERENCES

Donohue, John, and James J. Heckman. 1991. Continuous vs. Episodic Change: The Impact of Affirmative Action and Civil Rights Policy on the Economic Status of Blacks. *Journal of Economic Literature* 29 (4): 1603–43.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim. 2001. Identifying Preferences and Technology in the Hedonic Model Using All of the Economic Implications of the Model. Manuscript. Chicago: University of Chicago Press.

Flinn, Christopher, and James J. Heckman. 1982. New Methods for Analyzing Structural Models of Labor Force Dynamics. *Journal of Econometrics* 18 (Jan.): 115–68.

Heckman, James J. 1978. Dummy Endogenous Variables in a Simultaneous Equations System. *Econometrica* 46 (4): 931–59.

———. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47 (Feb.): 153–61.

———. 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely Used Estimator. *Journal of Human Resources* 32 (3): 441–61.

———. 2000. Causal Parameters and Policy Analysis: A 20th Century Retrospective. *Quarterly Journal of Economics*, February, 45–97.

———. 2001. Microdata, Heterogeneity, and the Evaluation of Public Policy. *Journal of Political Economy* 109 (4): 673–748.

Heckman, James J., and Bo Honore. 1990. The Empirical Content of the Roy Model. *Econometrica* 58 (5): 1121–49.

Heckman, James J., and Joe Hotz. 1989. Choosing among Alternative Methods of Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association* 84:862–74.

Heckman, James J., Robert LaLonde, and Jeffrey Smith. 1999. The Economics and Econometrics of Active Labor Market Programs. In *Handbook of Labor Economics*, ed. Orley Ashenfelter and David Card, ch. 31. Amsterdam: North Holland.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. General Equilibrium Treatment Effects: A Study of Tuition Policy. *American Economic Review* 88 (2): 381–86.

Heckman, James J., and B. Payner. 1989. Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina. *American Economic Review* 79 (1): 138–77.

Heckman, James J., and Richard Robb. 1985. Alternative Methods for Evaluating the Impact of Interventions: An Overview. *Journal of Econometrics* 30 (1–2): 239–67.

Heckman, James J., and Peter Siegelman. 1993. The Urban Institute Audit Studies: Their Methods and Findings. In *Clear and Convincing Evidence: Measurement of Discrimination in America*, ed. Michael Fix and Raymond Struyk. Washington, D.C.: Urban Institute Press.

Heckman, James, and Burton Singer. 1984. A Method for Minimizing the Impact of Distributional Assumption in Econometric Models for Duration Data. *Econometrica* 54 (Mar.): 271–320.

Heckman, James, and Jeffrey Smith. 1995. Evaluating the Case for Randomized Social Experiments. *Journal of Economic Perspectives* 9 (Spring): 85–100.

Heckman, James, and Edward Vytlacil. 1999. Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. *Proceedings: National Academy of Sciences* 96:4730–34.

———. Forthcoming 2001. Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica*.