REFERENCES


Jensen, M. C. and Meckling, W. H. Equity and Production Function: An Application to Labor-


REFLECTIONS

Introduction by the Editor

Routine obligations, we lament, leave too little time for reflection, independent of career goals. Ironically, as careers wind down, time becomes available, but distance from the reigning protocol makes the professional journals inaccessible to all but a select few.

Reflections, which will be an occasional feature in The Eastern Economic Journal starting with this issue, will offer such an opportunity, by invitation, to emeritus members of our profession. The Editors are particularly pleased that Professor Carolyn Shaw Bell, distinguished teacher of two generations of young women at Wellesley College, consented to write the first of these essays.

Carolyn Shaw Bell

Carolyn Shaw Bell, Katharine Coman Professor of Economics, taught at Wellesley College from 1950 to 1989. Before that she worked as economist for OPA during World War II and spent 1945-49 at the London School of Economics; her Ph.D. was completed in 1949. She was included in Who's Who in Economics on the basis of her published research: the first in the Quarterly Journal of Economics, 1951 and the most recent in the Monthly Labor Review, January 1994. She holds three honorary degrees and numerous awards for her work in various advisory roles for private, nonprofit, and governmental organizations. Currently she is revising a book on income distribution.

Data and the Economist

Economics, as a profession and a field of study, has grown and changed significantly over the past thirty to forty years. The activities of professional economists today bear very little resemblance to the pursuits of their predecessors. Much of this transformation is attributable to the stunning changes that have occurred in the methods economists use, particularly computer modelling of complex economic activities and analysis of large banks of economic data. Unfortunately, while the development of sophisticated econometric techniques has depreciated the profession, most economists give scant attention to the availability and quality of the data which form the foundation for their analyses.

HOW ECONOMISTS USE DATA

When explicit models of the economy were developed after World War II to accommodate empirical data and apply Keynesian theory to the real world, there was
little confidence that forecasting would be significantly fostered by such models. When the Brookings-SRRC model was formally introduced at the American Economic Association meetings in 1964, economic forecasting became confined to a handful of firms and a few government agencies. Their needs, it was thought, could be satisfied with quite crude analysis or gross figures, with no need for the sophisticated modeling techniques being presented.

Instead, of course, economic forecasting became a growth industry and the demand for economists expanded. Otto Eckstein founded DRI in 1963 and his guess about the market for such tailor-made forecasts was brilliantly correct. Over the next decades economists were increasingly accepted as influential participants in business and government and employment opportunities for economists proliferated.

Today economists in many different occupational settings use an extraordinary wealth of data without thinking twice about their availability, or about the achievement with which they would have been received in earlier times. It is difficult to overemphasize how rapid and profound this metamorphosis of the profession has been. A generation ago economists knew what multiple regression was but did not use it, outside the textbook. It remained an esoteric solution, not much use for a real contemporary economic problem. Today undergraduates learn hypothesis testing and multiple regression analysis in the beginning course in statistics and do original exercises with data from a myriad of different sources.

Developments in econometrics and mathematical economics, new survey techniques, and bigger and better computers now enable economists to obtain far more refined and detailed analysis of any issue than was possible thirty years ago. Computerization revolutionized data collection and tabulation in all the chief statistical agencies of government: the Bureau of Labor Statistics (BLS), the Department of Agriculture, the Federal Trade Commission, and above all the Bureau of the Census.

The Census of Population and Housing collected more information from more people, as did the various Economic Surveys: the Census of Manufacturing, the Census of Minerals, the Census of Agriculture, and so on. Computerization made it possible to keep accurate records of the expansion of workers covered by Social Security and private insurance and pension plans. Data from administrative records multiplied and Congress and bureaucrats realized how easy it was, once the citizen was required to fill out a form, to add another line (or more) to get more information.

In business, public opinion polling married to marketing grew quite quantitative market research and to specialists knowledgeable about survey methods, sampling design, and interview or questionnaire construction. Every statistical undertaking in the private sector was based, fundamentally, on the Federal effort. Federal data on the number and characteristics of workers, firms, and establishments, and on the amount of production, sales, tax revenues or whatever enabled the business statistician or economist to define the population or check the representativeness of a sample.

The private sector was well aware of the essential part played by Federal data. When 1964 budget cuts left the BLS unable to publish data it had already collected, the Ford Foundation and a group of major firms financed a group located at the Wharton School of Finance and Commerce to complete the tabulation and analysis.

Perhaps it was the close acquaintance with a major data source which prompted several firms to conduct their own surveys in later years.

Computerization also allowed the private sector to keep far more extensive records, to record information and disseminate it more swiftly, and to create more complex ways of doing business.

Completely new data sources came into being: the reports of the banking industry to the Federal Reserve Board, the basic facts about ownership collected by national associations like those for Life Insurance and the New York Stock Exchange, data on the Fortune 500 and data compiled by Dan and Bradstreet, industry data compiled by trade associations for electricity, railroads, publishing, steel manufacturing, and food processing, to name a few. Facts could be extracted from accounting records filed with the Securities Exchange Commission or other regulatory commissions and regularly published data on unemployment compensation, road safety, airline travel, and many other topics joined the time-honored basic facts of exports and imports, taxes and revenues, employment, money, and credit.

The growth of the economics profession has been accompanied by an explosion of data and, more obviously, of complex analytical methods to handle information in bulk and with detailed refinements. For most economists, however, being acquainted with econometric developments has appeared more useful than being acquainted with data. The obligation to "keep up with the field" rarely means learning about each new statistical source or series that appears in different areas of economics. It is even permissible, unfortunately, to use data without having mastered the technical appendices explaining sources, definitions, and limitations on use. For whatever reason, most economists do not pay much attention to data.

THE CASE OF LONGITUDINAL STUDIES

Most surveys provide ample amounts of empirical data for comparative statics. It is up to the economist, of course, to interpret the economic meaning of the difference between one set of circumstances and another.

Thus, an economist discussing the effects of income on saving may state that "As the amount of income rises, the average amount saved also rises." Figures from surveys showing average saving of households classified by total income appear to bolster this argument, although in fact they do not. Survey data prove nothing about the effect of getting more income on saving, because they do not allow such an occurrence for one individual or household. All that cross-sectional data provide are figures on different people with different incomes and savings.

The same problem exists in many areas including the impact of a change in price on demand or supply or of age on unemployment. To collect empirical data on the effect of any change in circumstances over time requires a longitudinal survey, collecting data from the same individuals (who may be people, firms, or other economic entities) over time. The number of longitudinal studies in economics (or in any field) is very small because they cost more than cross-sectional data, and few economists find longitudinal data of compelling interest.
The first large-scale longitudinal data bank was born with the Social Security Administration. To keep track of contributions for workers so as to be able to calculate eligibility for benefits later, the agency created a separate file for each individual, recording the contributions of employers as they were received. During its first two years the agency tabulated numbers and dollars for 100 percent of the files but then switched to sampling. Beginning in 1941, 1 percent of the records comprise the Continuous Work History Sample. By adding information on the employer to each worker’s file, the Longitudinal Employer Employee Data, or LEED File has been created.

Although such longitudinal data banks stem from administrative records, others were designed and implemented specifically to yield data for economic research. An early example was the work in labor economics at Ohio State University, where samples of individuals were drawn to represent first, young people, and then other groups of workers. The National Longitudinal Surveys obtained initial data about each individual in the sample, and then revisited the sample at regular intervals. Data on the intervening period were recorded to provide a continuous history of employment.

The second major longitudinal study grew out of years of work at the Survey Research Center of the University of Michigan, dating from 1946. The Survey of Consumer Finances, published annually, collected data on household income, assets, debts, saving and major expenditures and published analyses of income and wealth distribution. The sample of consumers was also interviewed quarterly about their attitudes and expectations. A panel of respondents drawn from those who had been interviewed for the annual Survey in 1960 were reinterviewed in 1961 and 1962. Information about income change and the subsequent change in variables like saving or expenditure on selected items could, for the first time, be reported.

In 1968 the Center for Survey Research selected some 5,500 families to make up the Panel Study of Income Dynamics (PSID); all members of the original sample were followed through time, for twenty-one years. For the first time, empirical data existed to answer questions like "Does growing up in a welfare family increase the chances of going on welfare?" or "What is the chief cause of income increases among families?"

The logical successor to the two-decade operation of the PSID was a national probability sample, considerably larger than the Michigan study, to assess longitudinal data as a matter of routine. In 1982 the Census began the Survey of Income and Program Participation (SIPP) to collect information on each member of some 20,000 households periodically for two and one-half years. New samples of households have been added yearly. Continuing detailed information on employment, income, eligibility for income assistance, and special topics like assets, pensions, and health have been provided over the ten years of SIPP’s life.

Other specialized longitudinal studies exist in the field of health and education. The Framingham Study has followed the same individuals until their death in order to have lifetime histories of illness and health and associated behaviors. The National Center for Health Statistics collects data by interview and reinvestigation for several different surveys. The National Center for Education Statistics collects data from high school students and then reinterviews them in subsequent years.

Each of the longitudinal data sets suffers its own problems but none has received much attention from economists. Economists go right on using static data from the Census, BLS, or Gallup polls to make assertions about changes in the economy or in the economic welfare of its inhabitants. Economists do not loom large among users of the new SIPP data even while discussions are underway to rely on this survey for income data, rather than the CPS March survey.

ECONOMISTS AND THE SUPPLY OF DATA

Few economic studies explore the suitability of the data to test the specific hypothesis under consideration. All too frequently the data shape the hypothesis. Investments in human capital have been equated with expenditures on education or years of schooling because data on both these variables exist. Neither, of course, measures the acquisition of knowledge, which is presumably the human capital acquired by the students involved. And if years of schooling must be used, what is the economic meaning of the official definition for "school?"

Economists know that any measure of poverty is controversial, and therefore the numbers of poor people may be contested, but they rarely discuss the major assumption of the official definition. Economists know that the supply of money can be measured by M1, M2, M3, and so forth and that the amount of unemployment can also be measured with different definitions. But non-specialists rarely mention the defects in collecting data or the unreliability of samples.

Along with this general neglect by the profession, the second major problem with economic data has been the steady deterioration, over the past fifteen years, in the quality of many statistics and statistical systems. The deficiencies of data, due almost entirely to underfunding, have been deplored in Congressional hearings, in a few brilliant expositions, and by the Council of Economic Advisers.

People who care about data have their favorite horror stories:

Item: The Standard Industrial Classification was not revised between 1972 and 1987, a period in which technological change led to the growth of entirely new industries and the reshaping of others. How to measure the "computer industry" typifies the first problem, how to calculate the output of "medical services" the second.

Item: The political repercussions of the Census underline get considerable attention, but technical details of why it is difficult to correct do not.

Item: The United States has been out of step with the rest of the world for over a generation in emphasizing Gross National Product rather than Gross Domestic Product. In 1991 the great decision was made to follow suit, and do all economists now recognize that although the GNP deflator no longer exists there is no GDP deflator either?
THE FUTURE OF ECONOMICS

I believe that economics will grow and that economists will continue to benefit the country only if they pay more attention to economic data. The profession ought to change its lackadaisical attitude to the data which are the foundation of economic analysis. Economists should descend en masse on Congress to insist on adequate funding for the statistical agencies on which Congress depends. Every beginning economics course should examine at least one data series, or one statistical statement, so that students learn the most basic work involved in producing empirical facts. They should be able to identify the work performed by the Census enumerator or the business bookkeeper filing FICA records and count such work as equally important as the work of calculating variance or degrees of freedom. (The exercises would probably benefit most instructors, too.) The activist stance I am recommending to the profession does not mean that everyone must become a "data freak" or fret about time series or data collection in the midst of speculation about mergers or terms of trade. It merely asks every economist to be responsible, from now on, for the quality and availability of data.

The 1992 Council of Economic Advisers points out that in many cases improving the data requires basic economic research, not just technical refinements. Timely and adequate measures of the money supply may not be helped by instantaneous reporting as much as by a better understanding of how households and businesses regard money. As non-monetary compensation grows, workers become less responsive to changes in wages, yet little is known about individual preferences for different types of compensation.

For the profession to grow and progress in the years ahead will require a shift of attention, away from "pure" theory or method to messy data and statistical series.

NOTES

1. Thinking about what has happened over this time period causes thinking about my own professional life. I read Keynes's General Theory as part of an honors seminar in my senior year and graduated from college in 1941 with an economics major. During the war years I did economics at the Office of Price Administration, after which I studied at the London School of Economics for a Ph.D., awarded in 1949. I then joined the faculty of Wellesley College, making frequent forays into both private sector and public sector counseling and advising assignments, as well as serving on a variety of commissions and boards over the years. My professional services included membership of the AEA Executive Committee, Nominating Committee, and Census Advisory Committee, as well as membership of the BEA Executive Committee, Nominating Committees, and Committee to Review the Journal. I've published in most major economic journals as well as written a few books and contributed to non-academic publications like the New York Times, the Wall Street Journal, and the Boston Globe, to which I currently contribute a regular column.

2. The Brookings-SSRC Quarterly Economic Model of the U.S. Economy, James D. Douglas, Gary Frenn, Lawrence Klein, and Edin Koh, editors, was published by Brookings and North Holland in 1969; Frenn and Klein presented an explanatory description of the model at the meetings, which was discussed by Edin, Max Nordhaus, and R. A. Gordon.

3. In 1964 NSF identified over 12,000 economists known to some professional society such as the AEA, with a Master's degree or more (equivalent experience was also accepted) and designating some field of economics as his (her) field of highest competence (American Economic Review, December 1968, supplement). Forty-two percent were employed by educational institutions, 33 percent in business or industry, and 19 percent by the Federal government. In 1970 the NSF register held 92,500 names and in 1986 164,005 names, 60 percent were employed in industry and 15 percent in teaching (National Science Foundation-Profile Economics, Washington, D.C. NSF, 1987).

4. Basically, the private sector stepped in to correct a misguided attempt at cutting Federal expenditures by economizing on statistical data. The Bureau of Labor Statistics (BLS) itself is more careful. The basic data presented in those various statistical tables was collected by the BLS in its Survey of Consumer Expenditures in 1956, which had as its primary purpose the revision of expenditure weights in the Consumer Price Index. After this immediate objective was achieved the Bureau was not in a position to tabulate the survey results in a manner similar to that followed in the comparable survey series so that the data would be available for the great variety of uses for which they were wanted. ( footwear, in Study of Consumer Expenditures, Bureau, and Savings, tabulated by the Bureau of Labor Statistics for the Wharton School of Finance and Commerce, University of Pennsylvania, 1956, p.)

5. Other continuous history files developed at the Social Security Administration include a 30 percent sample of those receiving disability benefits and a continuous record of claims and benefits for a sample of individuals covered by Medicare.

6. As about the same time market researchers were trying to improve the accuracy of consumer buying habits. Several major firms set up panels of consumers who kept a record, particularly of small items frequently purchased, and the results of an expenditure diary was refined. These procedures were major issues for the BLS in designing its survey of consumer expenditures.

7. People realized that before the first collection of data in the SIEP survey it was not possible to count the number of people receiving income subsidies. In the country. Obviously such programs -- Food Stamps, AFDC, SSI, Medicaid -- could provide tabulations of its recipients, but these totals could not be summed into people could receive benefits from two or more programs.


10. Such problems ponerous many areas of statistics. The revision of the producer price index does not occur frequently enough to keep up with changes in the product mix, and the only reason for such inadequacies is lack of funding by Congress (Interview with Janet L. Norwood, "Good Policy Builds Reliable Data," Challenge Magazine, Fall-2003, p. 20).

11. See, for example, Alan Greenspan, Chairman of the Federal Reserve Board, "The increased number of financial markets, the rapid changes in them, and the increased pace of market responses to shocks made possible by improved computer and communications technologies, challenges the ability of central bankers to monitor closely developments in the financial system and react in a timely manner when necessary," Speech Prepared for Changing Capital Markets Implications for Monetary Policy, a conference held at Jackson Hole, Wyoming, August 19, 2000.)