KEYNES'S *TREATISE*, STATISTICAL INFERENCE, AND STATISTICAL PRACTICE IN INTERWAR ECONOMICS IN THE UNITED STATES

BY

JEFF E. BIDDLE

In his Treatise on Probability *John Maynard Keynes criticized the tools of statistical inference derived from probability that were coming into use in the early twentieth century, and outlined an alternative approach to statistical inference based on the logic of induction. This essay argues that Keynes's ideas were embraced and echoed by several leading US economists during the 1920s and 1930s, including those developing and applying the most sophisticated statistical methods of the day. These economists expressed views regarding statistical inference that were quite similar to those found in Keynes's* Treatise*, often citing Keynes as an authority in support. Also, the inferential methods recommended and actually employed by these writers were consistent with Keynes's ideas about the proper methods of statistical inference.*

I. INTRODUCTION

John Maynard Keynes devoted the fifth and final part of his *Treatise on Probability* to "The Foundations of Statistical Inference," opening this part of the *Treatise* by making a distinction between two functions of the "Theory of Statistics." The first was the "descriptive" function, which "devises numerical and diagrammatic methods by which certain salient characteristics of large groups of phenomena can be briefly described." The second, and the one with which Keynes was exclusively concerned, was the "inductive" function, which "seeks to extend its description of certain characteristics of observed events to the corresponding characteristics of other events which have not been observed." This inductive function of statistics or "Theory of Statistical Inference," Keynes noted, was "closely bound up with the theory of probability" (Keynes 1921, p. 371).

One of Keynes's goals in this part of the *Treatise* was to criticize inferential methods based on the inversion of Bernoulli's theorem, for example, the use of the frequency with which an event had occurred in an observed sample of cases or time periods as a basis for specifying, with a quantifiable level of certainty, the range of possible values of the frequency with which that event would occur in another sample of as yet unseen cases or future time periods. Keynes's criticisms of these methods were quite harsh: The claims made for them were "foolish" and "preposterous"; applications of them amounted to "mathematical charlatanry" (Keynes 1921, pp. 418, 436, 443).

Keynes also sought to outline an alternative approach to statistical inference. For the most part the methods of inference he recommended were non-mathematical in nature, and reflected his belief that sound inferences from statistical data were best built using "the methods of Analogy and Induction" to which Part III of the *Treatise* had been devoted. Bradley Bateman (1990) establishes that Keynes's belief in the importance of applying the logic of induction to the process of statistical inference was a primary concern underlying his discussions of statistical method and his criticisms of the statistical and econometric work of others throughout his career.

John Aldrich (2008) concludes that almost all of the leading statistical theorists of the 1920s rejected Keynes's arguments concerning statistical inference. However, Keynes's ideas were embraced and echoed by several leading empirical researchers among US economists during the 1920s and 1930s, including those developing and applying the most sophisticated statistical methods of the day. This essay documents the prevalence of Keynes's views on statistical inference among empirical economists in the US.[1] I will show that leading empirical economists in the US expressed views regarding statistical inference that were quite similar to those found in Part V of Keynes's *Treatise*, often citing Keynes as an authority in support. I will also argue that the inferential methods recommended and actually employed by these writers were consistent with Keynes's ideas about the proper methods of statistical inference. What I cannot say with certainty is whether these writers formed their views as a result of reading Keynes's *Treatise* or summaries thereof, or whether they simply found in Keynes's arguments a lucid and authoritative articulation and defense of conclusions to which they had come based on their own experiences analyzing and attempting to generalize from statistical data.

II. KEYNES ON STATISTICAL INFERENCE

Much of what Keynes had to say in Part V of the *Treatise* was critical in nature. He considered it "crucial" that he attack the arguments of authorities such as Pierre-Simon Laplace and Karl Pearson that, given the frequency with which an event had occurred on a series of occasions, one could determine the probability that it would occur on a further occasion. Keynes himself did not believe that there was "any direct and simple method by which we can make the transition from an observed numerical frequency to a numerical measure of probability" (Keynes 1921, p. 418). The "mathematical" methods developed to do so were "invalid," and

---

[1] This essay draws heavily from material presented in Biddle (2017 and 2021).

> [t]o apply these methods to material, unanalysed in respect of the circumstances of its origin, and without reference to our general body of knowledge, merely on the basis of arithmetic and of those of the characteristics of our material with which the methods of descriptive statistics are competent to deal, can only lead to error and to delusion. (Keynes 1921, p. 438).

Keynes's critical exposition focused mainly on procedures designed to draw inferences about the "true" probability of an event in some universe based on the frequency with which that event occurred in a random sample from that universe. But the critique also applied, and was understood by statisticians of the 1920s to apply, to other inferential procedures based on similar arguments from probability theory, such as the use of the standard error of a sample mean to make inferences about the range in which the true mean would be found, or the use of the "standard error of a regression" to predict the range of values that one variable would take on the basis of values taken by one or more other variables. In what follows, I will refer to such procedures as "inferential methods based on probability theory."

As the quoted passage above shows, the reasons for Keynes's rejection of these procedures went beyond his skepticism regarding the assumptions and proofs underlying them. A valid approach to inference, he believed, could not rely solely on applying arithmetic to those characteristics of the sample cases that were amenable to being counted. Sound inference required attention to potentially unquantifiable aspects of the sample, including those associated with its "circumstances of origin," and consideration of our "general body of knowledge" regarding the phenomena about which one wished to draw inferences. Much of Part V of the *Treatise* was devoted to describing characteristics of what Keynes believed would be a more fruitful approach to inference than the one centered on inferential methods based on probability theory.

The central theme in these passages and chapters was that the logic underlying good statistical inference was similar to the more familiar logic of universal induction, that is, the process of reasoning on the basis of multiple instances of observation to formulate and build confidence in generalizations like "all eggs taste good."[2] Both forms of induction relied on what Keynes called the "method of analogy." Inductive arguments in support of general statements were built on numerous instances of observation. The characteristics shared by a set of instances constituted the "positive analogy" of the set (each involved an egg, in each the egg tasted good) while differences in characteristics across the instances constituted the "negative analogy" (eggs of different colors, eggs laid by different hens, or eggs laid during different seasons of the year).

Keynes explained that it was through careful consideration of the positive and negative analogy in sets of observations that one refined and/or built confidence in inductive generalizations. Finding that a speckled egg tasted good, if one had not previously tasted a speckled egg, would, in Keynes's terms, "strengthen the negative analogy" of one's set of instances, and strengthen confidence in the generalization that "all eggs taste good"; finding that an egg that had been around the house for a while did not taste good would narrow the scope of the generalization that the set of observations could support ("all fresh eggs taste good"), and so forth.

---

[2] Bateman (1990) provides a more detailed and complete treatment of Keynes's analysis, in the *Treatise*, of the logic of universal induction and its extension to statistical induction.

Keynes's explication of the logic of universal induction is found in Part III of the *Treatise*; in Part V he explained how, with suitable modification, the same logic could be applied to the problem of statistical inference, that is, reasoning from samples of statistical data to probabilistic generalizations about events or relationships of the form "if A then a 20% chance of B." In statistical inference, one built general conclusions on sets of samples of statistical data rather than sets of individual observational instances. Each sample was like previous samples in some ways (the positive analogy) and unlike those samples in other ways. Further, any given sample was being used to draw conclusions about some "universe," with which it would share some characteristics but from which it would differ in certain ways. Building strong inferences on the basis of summary statistics calculated with sample data required careful attention to the many unique characteristics of each sample, and how they compared with the circumstances surrounding the phenomena about which one wished to draw conclusions. Keynes illustrated this point with the example of drawing an inference about the relationship between age and the probability of death from a sample of deceased individuals:

> We note the proportion who die at each age, and plot a diagram which displays these facts graphically. We then determine by some method of curve fitting a mathematical frequency curve which passes with close approximation through the points of our diagram. … But in determining the accuracy with which this frequency curve can be employed to determine the probability of death at a given age in the population at large, [the statistician] must pay attention to a new class of considerations and must display a different kind of capacity. He must take account of whatever extraneous knowledge may be available regarding the sample of the population which came under observation, and of the mode and conditions of the observations themselves. Much of this may be of a vague kind, and most of it will be necessarily incapable of exact, numerical, or statistical treatment. (Keynes 1921, p. 372)

Keynes's advice regarding effective inference did include one class of mathematical procedures that he called "tests of stability." If, for example, one calculated the frequency of some event in a sample, and wished to know how well it would perform as an estimate of the probability of an event in similar cases, it would be informative to divide the sample into subsamples and calculate the frequency of the event in each of those subsamples. If it was sufficiently similar across the subsamples, the sample frequency was "stable," and confidence in it as an estimate of the unknown probability was increased. This confidence would be further bolstered if the frequency remained stable under different subdivisions of the sample. Keynes discussed with approval the work of Wilhelm Lexis and his proposed mathematical measure of the stability of sample statistics.

As I shall document in the following sections, several prominent themes consistent with Keynes's arguments about sound vs. unsound approaches to inference can be found in the writings of leading empirical economists in the US in the 1920s. A first was that the inferential procedures based on probability theory were valid only given certain stringent assumptions, and it was important when using those measures to determine whether those assumptions were met by one's sample material. Of especial importance was the assumption that one had a random sample of some well-defined universe. Second, even if one had a random sample, it was necessary to determine the extent to which the universe about which one wished to make generalizations was similar to the universe

from which the sample was drawn.[3] Time series data posed a special problem, because it often seemed impossible to regard a time series as a random sample from any stable, well-defined "universe," and, further, hazardous to regard future time periods about which one wished to make generalizations as cases drawn from that same universe.

Third, a reliable, trustworthy inferential procedure would make use of information beyond numerical measures calculable from the sample data. For the US economists this meant, among other things, knowledge of the historical period from which the sample came and the economic and social institutions surrounding the activity recorded in the sample, along with a good grasp of the cause-and-effect relationships suggested by economic theory. Fourth, tests of the stability of the values of summary statistics across subsamples and the comparison of results of samples from similar universes were useful aids to inference.

III. KEYNESIAN IDEAS IN THE WRITINGS OF EMPIRICAL ECONOMISTS IN THE US DURING THE 1920s AND 1930s

Warren Persons's ASA Presidential Address

An early and influential endorsement by a US economist of Keynes's views on inference came from Warren Persons, in his 1923 Presidential Address to the American Statistical Association (ASA) (Persons 1924). In the 1920s, Persons was considered one of the leading economic statisticians in the United States, having developed a basic approach to time series analysis and an associated set of statistical procedures that were quickly and widely adopted by empirical economists. He joined the faculty of Harvard in 1919, and was the lead developer of the Harvard Business Barometer, one of the more respected economic forecasting methods of the 1920s.[4]

As of 1919, Persons was enthusiastic about inferential methods based on probability theory. In an article explaining the construction of his forecasting model, for example, he provided readers with a detailed explanation of the meaning and use of the probable error of the correlation coefficient, noting at one point that

> the problem of the determination of a "probability" in economics is similar to the problem of ascertaining the ratio between the unknown numbers of black and white balls in a bowl based upon a record of sample drawings of, say, ten balls at a time. In other words, just as we may estimate the relative number of black and white balls in a bowl from a record of experiments, so we may determine an "objective probability" from available economic statistics. (Persons 1919, p. 125)

However, Persons's address to the ASA indicates that some time between 1919 and 1923, two things happened: Persons read Keynes's *Treatise on Probability*, and he changed his mind about the usefulness of inferential methods based on probability theory. The address included a discussion of "the nature of statistical inference" that sounded some distinctly Keynesian themes. He argued that the statistician's approach to

---

[3] One can distil this second point from the discussion on pp. 472–473 of the *Treatise*.
[4] For more on Persons, his methods, and the Harvard Barometer, see Morgan (1990, pp. 56–63) and Lenel (forthcoming).

inference was that of ordinary induction. In forecasting, for example, the statistician would look for a good "analogy," a statistical series from a period of time as similar as possible to the present. His confidence in the statistical results of analyzing that period as a basis for forecasting would be increased if the same or similar results were found in subperiods of the sample, or in samples from other periods marked by different circumstances. And if the results seemed consistent with relevant non-statistical knowledge, better still.

Persons then had a hypothetical interlocutor raise the argument that in the realm of statistical inference, ordinary induction could be improved upon by the use of probability theory, using as an illustration an application of the probable error of the correlation coefficient similar to the one that Persons himself had offered in 1919. Persons's response to this argument in 1923 was that "the view that the mathematical theory of probability provides a method of statistical induction or aids in the specific problem of forecasting economic condition" was "wholly untenable," a thesis, Persons noted, that had been "developed with great skill" by Keynes (Persons 1924, pp. 6, 7).

Persons asked his audience to consider the problem of attempting to forecast business conditions in 1924 based on a time series of economic data from the previous 100 years. Application of the theory of probability to this inferential problem would require that the set of 100 observations forming the sample was randomly drawn from some universe, that each member of that sample could be regarded as independent of the others, and that the year 1924 could be considered as another independent, random draw from that same universe. Persons then explained why this was obviously not so. To further support his argument, he pointed out to his audience that the correlation between pig iron production and the interest rate six months later in a sample covering the period 1903 to 1914 was 0.75, with the probable error of that correlation coefficient indicating that there was less than a one in ten million chance that the correlation between those two variables between 1915 and 1918 would be less than .5. Yet its actual value was 0.38—a shocking occurrence if you considered only inferential measures based on probability theory but not surprising at all when you considered the unusual economic circumstances associated with the war. Persons used a quotation from Keynes to drive home his point:

> In order to get a good scientific argument we still have to pursue precisely the same methods of experiment, analysis, comparison, and differentiation as are recognized to be necessary to establish any scientific generalization. These methods are not reducible to a precise mathematical form… . But that is no reason for ignoring them, or for pretending that the calculation of a probability which takes into account nothing whatever except the numbers of instances, is a rational proceeding. (Persons 1924, p. 8, quoting Keynes 1921, p. 391)

Textbooks on Statistical Methods for Economists

The Keynes/Persons view of statistical inference was promoted by several leading textbooks on statistical methods for economists published by US authors in the 1920s. The most prominent of these was Frederick Mills's *Statistical Methods Applied to Economics and Business* (Mills 1924), one of the most popular and highly esteemed

statistics texts of the interwar period.[5] In his chapter "Statistical Induction and the Problem of Sampling," Mills adopted Keynes's distinction between the descriptive and inductive functions of statistics. He also described the meaning of a representative sample, derived the formulas for the standard errors of various common statistical measures, and explained how they could serve as measures of the reliability of a sample statistic as an estimate of a characteristic of the population from which the sample was drawn. But he then strongly cautioned his readers against using these measures, arguing that the circumstances that justified their use were "rarely, if ever" met in economic data. Further, these measures were designed only to account for the problem of sampling error, which was far less important a consideration for those wishing to generalize from sample data than sampling biases and measurement error. He recommended that instead of inferential techniques based on probability theory, the statistician use "actual statistical tests of stability," such as the study of successive samples and the comparison of descriptive statistics across subsamples of the population—the same sorts of inferential procedures recommended by Keynes and Persons.

The general message was the same, but the debt to Keynes and Persons more explicit, in Edmund Day's 1925 textbook *Statistical Analysis.* In his chapter on the interpretation of statistical results, Day asserted that the theory of probability was inapplicable to most economic data, and that inferences should be based on "non-statistical tests of reasonableness" and the logic of analogy. In support of this, Day reproduced close to two pages of text from Persons's ASA presidential address, including passages in which Persons quoted Keynes's *Treatise* (Day 1925, pp. 378–382).

Horace Secrist's text, *An Introduction to Statistical Methods*, had first appeared in 1917 to generally positive reviews, with Edmund Day calling it "the best-balanced book in English on statistical methods as related to economic investigations" (Day 1918, p. 403). The book was popular enough that Secrist issued a second edition in 1925, which included new material both deriving and explaining probable error measures. Secrist also argued, however, with the help of a quote from Persons, that these measures were largely inapplicable to economic data.

In 1930, Mordecai Ezekiel published *Methods of Correlation Analysis,* an advanced text for students and practitioners wishing to understand and apply multiple correlation (now called "multiple regression") analysis, which at that time was an arcane and rarely used statistical technique. For at least the next fifteen years the book remained a standard reference for economists using this increasingly popular method. Ezekiel, an agricultural economist, was quite intrigued by ongoing work deriving and refining inferential measures based on probability theory, and Ezekiel's book was the first after Ronald Fisher's (1925) *Statistical Methods for Research Workers* to explain the method of significance testing with z-tests and t-tests. In 1929, as he was preparing his book for publication, Ezekiel corresponded with Fisher to make sure that he was properly presenting Fisher's method of deriving a likely lower bound on a population correlation coefficient based on the sample estimate of the coefficient (Aldrich 2000).

Ezekiel's exposition of such cutting-edge inferential techniques, however, was accompanied by caveats. The assumptions required for them to be valid were described,

---

[5] An indication of the respect for and widespread circulation of Mills's book can be found in reviews of the revised (1938) edition, such as Brown (1939) and Yntema (1939).

with the warning that they "seldom completely obtained in practice," so that the measures supplied only a useful "rough estimate" of the reliability of the results as a basis for inference about what would happen in a similar sample. At one point, he warned that when using time series data, the standard error of estimate was valid for predictions of what might happen in a future period only if it were "definitely known" that "exactly" the same conditions prevailed in the future period as in the period covered by the sample data. Otherwise, it was merely suggestive of how the dependent variable might behave in the future (Ezekiel 1930, pp. 14–15, 18, 116).[6]

Agricultural Economists: Ideas about Statistical Inference

There existed in the US during the 1920s a large community of self-identified "agricultural economists," employed by the US Department of Agriculture (USDA), state departments of agriculture, and the many state-supported agricultural colleges. These economists maintained a group identity bolstered by their own professional organization and institutionalized linkages between government agencies and the agricultural colleges. Agricultural economists were much more likely than economists in general to do empirical research, and some education in statistical theory and method was considered essential for those in the field.[7]

As a group, the agricultural economists of the 1920s had little use for inferential methods based on probability theory. This comes through clearly, for example, in the 1928 report *Research Method and Procedure in Agricultural Economics,* sponsored by the Social Science Research Council and based on contributions from a number of active researchers. The report's editors commented that:

> economic statisticians … generally take the position that the mathematics of sampling and error and inference thus far developed, which holds rigorously only for pure chance and simple samples of entirely unrelated events, is inadequate for the needs of economic phenomena, and that there is little prospect of mathematical analysis soon being developed that will be adequate. Once the assumptions of pure chance are violated, inference has to proceed along other lines than those based on simple mathematical probability. (Advisory Committee, 1928, p. 38)

The message that the data of agricultural economics did not meet the assumptions required for probability-based inference, and that statistical inference required both a priori analysis and empirical knowledge beyond that provided by the sample, was repeated by several contributors to the report. Elmer J. Working, in the section of the report specifically devoted to statistical inference, questioned the usefulness of the inferential methods based on probability theory with the observation that the universe of which an economist's sample was truly representative was seldom the universe about which the economist wished to draw inferences. How, then, could an economist

---

[6] Ezekiel also warned that "in many types of problems, such as in time series, the observations can hardly be regarded as samples from a universe. In such cases, statistical measures of reliability have little meaning" (Ezekiel 1930, p. 264).

[7] See Biddle (2021). The 1928 survey of research methods in agricultural economics, discussed in the text, also testifies to the importance of statistical research in the field at this time, with roughly a third of its pages devoted to statistical methods, "the main resort of agricultural economists" (Advisory Committee 1928, p. 36).

determine whether a correlation found in a sample was likely to provide a reliable basis for inference about the relationship between variables in the universe that really interested him? Working argued that the tests of stability created by Wilhelm Lexis had a role to play—the relationship should at least be stable within the sample if it was to be trusted as an indication of what might be true beyond the sample. But one could have even more faith in a generalization based on a sample correlation if that correlation represented a true cause-and-effect relationship, and much knowledge of which correlations represented causal relationships could be provided by economic theory.[8] Louis Bean, a USDA economist writing on time series analysis, made similar points (Advisory Committee 1928, pp. 272–288).

Another USDA economist, Charles F. Sarle, devoted several pages of a technical bulletin on estimating crop yields to the subject of statistical inference both in general and as it applied to his topic, quoting Keynes's *Treatise* at four points (Sarle 1932a, pp. 12–38). Sarle explained, "The statistician's basis for assuming that a generalization concerning the average yield per acre of a crop from sample data will apply to the cases not included in the sample must be logically developed. The ordinary methods of inductive reasoning are used, basing the logical processes upon statistical data" (Sarle 1932a, p. 13). Measures of inference derived from probability theory were useful, but a complete analysis should consider also potential bias or lack of representativeness in the sample, and the sample results should be interpreted in light of basic knowledge of the phenomenon being studied.

Sarle's views on inference are of particular interest because he was also a statistics instructor for the USDA's graduate education program, which offered advanced courses for government employees.[9] A set of notes from Sarle's 1932 statistics course survives. Ezekiel's book was the text for the course, and Mills's book was also recommended. In an early session students were introduced to the concept of the standard error and the formulas for standard errors of various descriptive statistics. Emphasis was placed on the assumptions required for these measures to be valid, and during lectures devoted to the problem of statistical inference, students were cautioned to think about and if possible test those assumptions in their data. They were warned that the samples they worked with might not be representative of the universe of interest —indeed, "many surveys are made without any very clear idea of what the universe of inquiry is really supposed to be" (Sarle 1932b, p. 56).[10] Statisticians were encouraged to look for possible bias by comparing the distributions of variables in their samples to the distributions of the same variables in "check data," such as data from the periodic federal agricultural censuses. Likewise, when dealing with time series data, the universe about which one hoped to make generalizations had likely changed since the time the sample was collected, so one had to consider "the influence of factors that change materially with time" (Sarle 1932b, pp. 41, 57).

---

[8] Working continued to argue throughout the 1930s against the unreflecting application of probability-based methods of statistical inference, and the need to combine statistical results with non-statistical information. See, e.g., Working (1930, pp. 122–125; or 1937, pp. 149–150).

[9] From 1927 to 1929, Sarle also taught courses for statisticians working in the USDA's state offices (USDA 1969, p. 72).

[10] Page references to Sarle (1932b) refer to the downloadable PDF version. The original is not consistently paginated.

Malcolm Rutherford (2011) provides additional information about statistics instruction at the USDA Graduate School in the interwar period. Given that G. Udny Yule's (1911) *Introduction to the Theory of Statistics* was a standard text for the year-long course in the 1920s, it is reasonable to assume that students were exposed to Yule's treatment of the theory of sampling and to formulas for the standard errors of basic sample statistics. Ezekiel began teaching the course in 1926, and his book was a standard text for the course after it appeared. One can only speculate on how the ambivalence towards inferential methods based on probability theory that Ezekiel displayed in his book manifested itself in his lectures. It is clear, however, that as the 1930s wore on, interest in those methods was growing among USDA economists and statisticians. In 1933 Edward Deming introduced a new course at the Graduate School in which he compared Ronald Fisher's new methods of fiducial inference with previous attempts to use probability theory as an aid to statistical inference. In 1936, Fisher gave three lectures at the Graduate School on statistical inference, and in 1937 Jerzy Neyman gave a series of lectures that included discussions of hypothesis tests and confidence intervals (Rutherford 2011). Sarle's response in the late 1930s to the development and refinement of inferential methods based on probability theory was to strongly advocate for new data collection methods at the USDA that would produce samples meeting the assumptions required by the new methods.[11]

IV. THE PRACTICE OF STATISTICAL INFERENCE IN AGRICULTURAL ECONOMICS

The Keynesian approaches to statistical inference recommended by the agricultural economists of the 1920s and 1930s were also those that they practiced in their own statistical work. In Biddle (2021) I support this assertion with examples from the crop and livestock forecasting program of the USDA's Bureau of Agricultural Economics (BAE); here I will do so with reference to the bureau's price analysis program.

The BAE's price analysis program, launched in the early 1920s, involved the statistical analysis of factors influencing the prices of agricultural commodities. A price analysis study often involved the application of multiple regression analysis to time series data on the price of a commodity and other measurable factors thought to affect supply or demand in the market for the commodity, with the goal of creating a statistical model that would produce accurate forecasts of the commodity's price. It was believed that farmers, if supplied with such forecasts, could make more profitable decisions regarding what and how much to plant, when to market their output, and so on. Inefficient overproduction would be avoided, and disruptive price fluctuations in agricultural markets would be smoothed.[12] Statistical forecasting, of course, is a form of statistical inference in Keynes's

---

[11] See, e.g., Sarle (1939). The growing interest in adopting sampling procedures that would justify use of the new inferential methods was also evident in other parts of the federal government at this time (USDA 1969, p. 85).

[12] See Banzhaf (2006) for more on price forecasting at the USDA in the 1920s.

sense, that is, using information from an observed sample to make generalizations about phenomena beyond the sample.

During the 1920s and 1930s, the BAE regularly issued price forecasts for a wide range of commodities, typically based on regression models developed in-house but never published (Ezekiel 1940). However, some price analysis studies that were particularly innovative or exemplary were published in journals or USDA bulletins, and there one sees the inferential methods used to go from regression results to forecasts of future prices.

For example, in the mid-twenties Bradford B. Smith worked on forecasting cotton prices. One step in his procedure was to estimate a regression explaining current cotton production as a function of past cotton prices. Smith had the necessary data covering the years 1901 to 1924, but the simple price-production correlation prior to 1907 was greater than for the subsequent years, and a graphical analysis showed a trend break prior to 1913, so Smith omitted the pre-1913 data from the analysis. To support this move Smith also cited several historical developments affecting US cotton production that would lead one to expect key relationships affecting the size of the cotton crop to have been in flux prior to the mid-teens before stabilizing. When discussing his regression explaining fluctuations in cotton prices, Smith observed that the sample period covered a panic, war and inflation, and postwar depression, and commented that "[t]he fact that the various relationships set forth have consistently held true through a wide range of economic circumstances … is evidence in itself of a measure of stability in them and thus encourages acceptance of them as approximately the true quantitative relations among these factors," thus appealing to what Keynes would have called the "strong negative analogy" provided by his sample (Smith 1925, pp. 38, 44; 1928, pp. 53–54).

In 1927, Ezekiel published a paper on "Two Methods of Forecasting Hog Prices." The "empirical formula" method involved estimating a regression equation explaining prices in a current month based on variables that could have been observed several months earlier, then using current data on the independent variables to forecast price several months ahead. The "demand curve" or "synthetic" method involved first using regression analysis to estimate the elasticity of price with respect to hog supplies. One then made the forecast using this elasticity estimate, an estimate of future supply based on current information, qualitative and quantitative, and any material that might suggest a change in demand conditions. Ezekiel compared the accuracy of several months' worth of forecasts made with each method, and commented that a combination of the two methods—that is, using a formula estimate to start with, then factoring in knowledge of specific current conditions—might be the best approach. He also advised that examination of past periods when the regression model fit the data badly "may be of great assistance in working out corrections to be applied to subsequent forecast" (Ezekiel 1927, p. 29). The Keynesian tenor of Ezekiel's approach to forecasting is obvious. Likewise, John Hopkins estimated a regression for explaining cattle prices using monthly data from 1922 to 1926, then tested it for stability on years outside that sample range. He explained that when forecasting, the regression would be most useful when "considered as one source of information along with various others" (Hopkins 1927, p. 446).

In 1940 Ezekiel offered an assessment of the record of the price analysis program. He argued that one of the best strategies for improving the program going forward

would be to update old studies, sometimes modifying them in light of institutional changes since they were originally done, and to conduct comparative analyses of the successes and failures of past studies covering different times and/or different commodities. Work of this sort was already being done, he observed, and should be expanded. In short, Ezekiel, like Keynes, was commending an inductive process based on a comparison of more or less similar samples, a process of "experiment, analysis, comparison, and differentiation … recognized to be necessary to establish any scientific generalization" (Keynes 1921, p. 446).

V. EPILOGUE

Keynes's discussions of statistical inference in the *Treatise on Probability* had two important elements: rejection of a set of inferential tools based on probability theory that were just beginning to be adopted by practicing statisticians, and a description of a positive approach to statistical inference based on the logic of induction and analogy. During the 1920s and 1930s, a number of the leading US economists adopted an approach to statistical inference consistent with both these elements of Keynes's view, and understood themselves to be doing so.

The professional consensus among US economists regarding the limited usefulness of inferential methods based on probability theory was strongly challenged beginning in the 1940s, and gradually eroded over the next few decades, to be replaced eventually by the view that those methods were essential to the practice of statistical research in economics.[13] Indeed, most modern economists understand the phrase "statistical inference" simply to mean the application of those methods to statistical results, in the form of formal statistical hypothesis tests, measures of statistical significance, and so on.

Keynes's exhortations about the use of the traditional logic of induction to build strong generalizations on the basis of samples of statistical data are no longer part of the formal pedagogy of econometrics. Inferential practices consistent with that advice, however, remain a part of econometric practice, and are commonly employed to support or challenge conclusions drawn from statistical data. Researchers who, in the process of making generalizations based on their statistical results, consider the circumstances surrounding the generation of their sample data and the ways in which their sample is and is not representative of the universe about which they are generalizing, and who make use of relevant knowledge evidence from beyond their sample—quantitative, qualitative, and theoretical—are admired by their colleagues and imitated by students. Still, in modern empirical research these non-probabilistic approaches to inference play a secondary role to the methods rejected by Keynes, in this sense: a relationship found in sample data must be certified as "statistically significant" by an acceptable test derived from probability theory before a generalization based on that relationship will be considered credible, regardless of how much other evidence and reasoning a researcher might offer in support of the generalization.

---

[13] Biddle (2017) documents this gradual but dramatic change.

REFERENCES

Advisory Committee on Economic and Social Research in Agriculture of the Social Science Research Council. 1928. *Research Method and Procedure in Agricultural Economics*. Two volumes. New York: Social Science Research Council.

Aldrich, John. 2000. "Fisher's 'Inverse Probability' of 1930." *International Statistical Review* 68 (2): 155–172.

———. 2008. "Keynes Among the Statisticians." *History of Political Economy* 40 (2): 265–316.

Banzhaf, H. Spencer. 2006. "The Other Economics Department: Demand and Value Theory in Early Agricultural Economics." *History of Political Economy* 38 (ann. suppl.): 9–30.

Bateman, Bradley W. 1990. "Keynes, Induction, and Econometrics." *History of Political Economy* 22 (2): 359–379.

Biddle, Jeff E. 2017. "Statistical Inference in Economics, 1920–1965: Changes in Meaning and Practice." *Journal of the History of Economic Thought* 39 (2): 149–173.

———. 2021. "Statistical Inference in Economics in the 1920s and 1930s: The Crop and Livestock Forecasts of the US. Department of Agriculture." *History of Political Economy* 53 (ann. suppl.): 149–173.

Brown, Frederick. 1939. "Review of Mills, F.C. *Statistical Methods Applied to Business and Economics*, revised edition." *Economica New Series* 6 (24): 468–469.

Day, Edmund. 1918. "Review of *An Introduction to Statistical Methods*." *American Economic Review* 8 (2): 401–404.

———. 1925. *Statistical Analysis*. New York: Macmillan.

Ezekiel, Mordecai. 1927. "Two Methods of Forecasting Hog Prices." *Journal of the American Statistical Association* 22 (157): 22–30.

———. 1930. *Methods of Correlation Analysis*. New York: John Wiley and Sons.

———. 1940. "Price Analyses, Wars, and Depressions." *Journal of Farm Economics* 22 (4): 673–679.

Fisher, Ronald A. 1925. *Statistical Methods for Research Workers*. Edinburgh: Oliver & Boyd.

Hopkins, John A. 1927. "Forecasting Cattle Prices." *Journal of Farm Economics* 9 (4): 433–446.

Keynes, John Maynard. 1921. *A Treatise on Probability*. London: Macmillan and Co.

Lenel, Laetitia. Forthcoming. "Searching for a Tide Table for Business. Interwar Conceptions of Statistical Inference in Business Forecasting." *History of Political Economy* 53 (ann. suppl.).

Mills, Frederick. 1924. *Statistical Methods Applied to Economics and Business*. First edition. New York: Henry Holt & Co.

Morgan, Mary. 1990. *The History of Econometric Ideas*. Cambridge, UK: Cambridge University Press.

Persons, Warren M. 1919. "II. The Method Used." *Review of Economics and Statistics* 1 (2): 117–139.

———. 1924. "Some Fundamental Concepts of Statistics." *Journal of the American Statistical Association* 19 (145): 1–8.

Rutherford, Malcolm. 2011. "The USDA Graduate School: Government Training in Statistics and Economics, 1921–1945." *Journal of the History of Economic Thought* 33 (4): 419–447.

Sarle, Charles F. 1932a. *Adequacy and Reliability of Crop-Yield Estimates*. Technical Bulletin 311. Washington, DC: United States Department of Agriculture.

———. 1932b. *Statistics.* Bound typescript held by Albert R. Mann Library, Cornell University. Digitalized version accessed at https://babel.hathitrust.org/cgi/pt?id=coo.31924013703693. Accessed August 21, 2021.

———. 1939. "Future Improvement in Agricultural Statistics." *Journal of Farm Economics* 21 (4): 838–845.

Secrist, Horace. 1925. *An Introduction to Statistical Methods*. Revised edition. New York: Macmillan.

Smith, Bradford B. 1925. "Forecasting the Acreage of Cotton." *Journal of the American Statistical Association* 20 (149): 31–47.

———. 1928. *Factors Affecting the Price of Cotton*. Technical Bulletin 50. Washington DC: United States Department of Agriculture.

United States Department of Agriculture. 1969. *The Story of U.S. Agricultural Estimates*. Washington DC: USGPO.

Working, Elmer J. 1930. "Evaluation of Methods Used in Commodity Price Forecasting." *Journal of Farm Economics* 12 (1): 119–133.

———. 1937. "New Indices of Agricultural Supplies and Carry-Over." *Review of Economics and Statistics* 19 (3): 144–153.

Yntema, Theodore O. 1939. "Review of Mills, F.C. *Statistical Methods Applied to Business and Economics*, revised edition." *Journal of Political Economy* 47 (3): 449–450.

Yule, G. Udny. 1911. *Introduction to the Theory of Statistics*. London: Griffin.