Milton Friedman: A Bayesian?*

Gerald P. Dwyer
Clemson University
University of Carlos III, Madrid

Abstract

Milton Friedman's empirical research is very different in tone and substance than the research in a typical journal article. This chapter points out that the difference in tone and substance is directly related to Friedman's views on the foundations of statistics. Instead of viewing statistics as a classical statistician might, Friedman viewed statistics through the lens of personal probability. This had substantial implications for his overall research agenda and the particulars of his empirical work. He emphasized agreement across various sets of data because such evidence was more likely to produce agreement among economists than tests on a single set of data. He put less emphasis on statistical significance of results at least partly because of the problems involved in interpreting multiple tests on the same data.

August 2014

* James R. Lothian brought Gianluigi Pelloni's papers to my attention, which I appreciate. Conversations with John Devereux about Milton Friedman's methodology also have been very helpful.

Responding to a question about rational expectations and learning, Robert E. Lucas mentions that Friedman is “very influenced by Savage and by this Bayesian way of thinking about probabilities. So when I [Lucas] talk about people ‘knowing’ a probability, he [Friedman] just can't reach that language.” (Klamer 1983, p. 40). As Friedman noted (Friedman and Friedman 1998, p. 199), Jimmie Savage “greatly influenced my understanding of the foundation of statistics – particularly the interpretation of the concept of probability.”[1]

Having taken Friedman’s classes on price theory and having been in the Money and Banking workshop for two years with Friedman, I can agree with Lucas’s statement that Friedman was very influenced by the Bayesian way of thinking. While not everyone had personal contact with him or took classes from him, two papers by Pelloni (1986, 1997) some time ago argued that Friedman was a Neo-Bayesian.[2] Pelloni’s evidence includes personal correspondence with Friedman. Despite these papers, this point seems to be unknown. For example, an excellent book on Friedman’s empirical research on money’s causal role in business fluctuations by Hammond (1996) includes neither the name “Bayes” nor any derivative words in the index, which is consistent with my examination of the text of the book which turned up no reference to Bayes. But as Friedman stated in a letter quoted by Ebenstein (1996, p. 68) in which he spoke of two major influences that underlied his methodology – Karl Popper and Jimmie Savage:

> ‘I have increasingly come to put greater emphasis on the second compared to the first. The thing that impressed me about Savage’s analysis in his *Foundations of Statistics* and elsewhere is the emphasis on statistics, i.e. all scientific investigation, not as a search for truth in some abstract sense but rather as a

[1] Leonard Jimmie Savage went by the name “Jimmie” and this usage seems most appropriate in this context, in part because Friedman calls him “Jimmie”.

[2] The term “Neo-Bayesian” reflects analysis in which all probabilities are personal and subjective (Pelloni 1987, p. 408). I do not follow this particular usage and just the term “Bayesian”.

> mechanism which will produce agreement among people starting out with initially different views.'[3]

As Friedman himself (Friedman and Friedman 1998, p. 216) says,

> To oversimplify, Jimmie would say 'The role of statistics is not to discover truth. The role of statistics is to resolve disagreements among people. It's to bring people closer together.'

A claim that Friedman was a Bayesian might seem odd. After all, his empirical research certainly was not Bayesian in the sense that Arnold Zellner's research was Bayesian. It is easy to forget though, that Zellner spent time in the 1970s working on numerical integration routines, far from Friedman's interests. Computation was a seriously limiting factor for empirical research in the 1950s and early 1960s, with regressions being run by hand on electromechanical calculators at the cost of significant time and effort. A complete Bayesian analysis of a regression would turn the computation into days of work – probably many days. In the later 1960s and 1970s, regressions could be run on expensive mainframes.

More fundamentally, what we now call Bayesian analysis was developed in the 1950s and later by Savage and others.[4] Savage was Friedman's co-author of two articles on expected utility (Friedman and Savage 1948; 1952) and a colleague at the Statistical Research Group (SRG) at Columbia during World War II.[5] Fienberg (2006) provides an excellent summary of the development of Bayesian statistical analysis up to about 1970. Feinberg's summary makes it

---

[3] Friedman is referring to Savage (1972), first published in 1954. The eventual agreement by those with different priors is a fairly general theorem in Bayesian analysis. Bernardo and Smith (2000, Section 5.3) provide an example of such analysis.

[4] Zellner (1985: 254) gives Friedman and Savage's papers a role in the development of Bayesian statistics by showing the relationship between economic utility theory and statistics.

[5] Friedman and Savage also co-authored statistics paper as part of the SRG's work (Friedman and Friedman 1998, p. 146).

This group developed sequential analysis, about which Friedman regarded himself as an expert for a time (Bayesian Heresy, 2007). Sequential analysis is inherently Bayesian in its approach to sampling.

clear that it would be anachronistic to require that Friedman compute Bayes factors to classify his empirical research as Bayesian rather than classical.[6]

On the contrary, it could be claimed that it is anachronistic to suggest that Friedman was a Bayesian when the term did not have the same connotations as today. This is fair enough up to a point. On the other hand, it is more misleading to think of Friedman as having a research agenda dominated by classical statistical criteria or the non-statistical frame of reference of much early National Bureau empirical analysis. Evaluating Friedman's empirical research by classical statistical criteria, as some do (Hendry and Ericsson 1991), is missing the point, even leaving aside Friedman's objections to the particular implementation of those criteria in that paper.

Personal Probability and Rational Expectations

Friedman assigned the papers by Friedman and Savage (1948; 1952) on personal probability and expected utility to graduate students in his price theory course until he retired. His lectures in class reflected that development of personal probability and, while it is personal recollection, his answer to one question in class was adamantly clear that he completely rejected the notion that people could not behave consistent with personal probabilities under some circumstances. Rather, as in his *Price Theory* text (Friedman 1976, p. 282), he relied on the point that people's behavior reflected their estimates of the probabilities and those always could be inferred. An outside observer might believe the relevant probabilities were different than the implied

[6] Pelloni (1987, p. 408, fin. 3) notes that there is some inconsistency between Friedman's theory about the foundations of probability and his statistical techniques, maybe partly because the former was influenced by Savage after his education in statistical techniques. "Thus the interpretation of Friedman's methodological framework through his empirical work is a delicate task to be entertained with caution if one does not want to be misled." While I agree with the conclusion, I think the reason is simpler: detailed Bayesian analysis was not particularly feasible or consistent with Friedman's research agenda at the time.

probabilities from observed behavior, but that did not make the implied probabilities wrong, merely different.

More than once, Friedman suggested that the whole notion of rational expectations was misleading because it assumed objective probabilities.[7] Comments about the free-silver episode in Friedman and Schwartz (1982, p. 630) are related to those in class and in the Money and Banking workshop, namely

> "that participants in whatever market is [sic] considered have 'correct' estimates of the probability distribution of outcomes (itself something that is difficult or impossible to define objectively), so that on the average anticipations are correct; and (2) that errors of forecast in successive time units are uncorrelated. This episode brings sharply the difficulty of giving a precise meaning to the first assumption, and the ambiguity of 'time unit' for the second assumption."[8]

As they say about the first assumption elsewhere (Friedman and Schwartz 1982, p. 557), "What meaning, if any, can be given to the assertion: 'Mr X's personal probability about a specified event was correct [or wrong]'?"

Friedman did not however reject the proposition of rational expectations, even if he was dubious about some aspects of its implementation. He states in *Price Theory* (Friedman 1976, p.

---

[7] It is interesting to note that Friedman in his *Price Theory* text (Friedman 2007, p. 282), revised a few years before *Monetary Trends* was last revised, does assign a meaning to *objective probability*: A probability is objective if people agree about the probability.

[8] They also say: "The formalization in the theory of rational expectations of the ancient idea that economic actors use available information intelligently in judging future possibilities is an important and valuable development." (Friedman and Schwartz 1982, p. 630).

This objection could be interpreted as reflecting a "peso problem", an expected event not observed in the data, which Friedman also raised in the Money and Banking Workshop as the peso problem for Mexico. The point is more general though and refers to the whole notion of objective probabilities.

I will not distinguish between statements in pieces written by Friedman alone and those by Friedman and Schwartz because the statements would not have been made if they did not agree. This is not just a sop to Anna, as anyone acquainted with her knows.

230) that John Muth's original article was "important" and that it is not possible to take seriously the notion that people will persistently form expectations in a way that is inconsistent with the process actually generating inflation.

A Theory of the Consumption Function

Friedman's book on the consumption function is remarkable in many ways, including the almost complete absence of tests of statistical significance in the book. His explanation is more informative than a paraphrase by me. He (Friedman 1957, pp. ix-x) states that

> The origin of the book may explain some features of it, in particular the extensive reliance on secondary sources for data and the almost complete absence of statistical tests of significance. An hypothesis like the one presented below is typically a by-product of original empirical work; so it is in this case, but the original work was Mrs. Brady's and Miss Reid's, not my own. What systematic empirical work I did came after the development of the hypothesis, not before, and was directed at bringing together as wide a variety of data as I could with which to confront the hypothesis. It is a defect, of this confrontation that I make so little use of objective statistical tests of significance. There are several reasons for this defect. First, many of the data do not lend themselves readily to such tests: For example, it would be necessary in some cases to go back to individual observations rather than to be content, as I have been, with means of groups. Secondly, sampling fluctuations seem to me a minor source of error, particularly in interpreting family budget data for rather large samples, compared to both biases in the samples and inadequacies for my particular purpose in the definitions used and the kind of information collected. In consequence, I have preferred to place major emphasis on the consistency of results from different studies and to cover lightly a wide range of evidence rather than to examine intensively a few limited studies.

There is much that can be made of this remarkable statement and the book. It is clear that the treatment of error variances and their sources is careful and nuanced. It also is clear that substantial judgement enters into the empirical analysis. The purpose of the work is to be convincing that the existing evidence including differences in results across papers and sets of data can be explained by the permanent-income hypothesis.

Friedman examines a large variety of data and finds that virtually all of it is consistent with the permanent-income hypothesis. For analysis using group means, he eschews tests of statistical significance because the underlying number of observations is sufficiently large that sampling error is not an issue. In the text, he uses "intuitive judgements" about whether a particular result might be due to sampling error (Friedman 1957, p. 214). He goes on to suggest that "it would be highly desirable to have such judgements supplemented by formal tests of statistical significance whereever possible." (Friedman 1957, p. 215).

Statistical significance appears primarily in one context: the statistical significance of constant terms in time-series regressions relating consumption and income (Friedman 1957, 138-141). Here, the residual variance reflects omitted factors and errors in the data and a satisfactory standard error is available. The sole other test of statistical significance in the entire book is a test whether estimated variances across cities are solely due to sampling error (Friedman 1957, p. 175).

The analysis overall can be read as showing a classical statistician confronted by data available on group means and an implied uninformativeness of tests. From a strictly classical standpoint, though, the issue becomes "Why bother with analyzing data for which no statistical analysis is feasible?" It is extremely unlikely that the data are exactly consistent with the hypothesis. From Friedman's standpoint though, and the standpoint of most of his readers, the tests are informative because the point estimates can be examined for consistency with the hypothesis relative to their economic significance. Each of the computations is an additional piece of information based on available data. While it would be a non-trivial endeavor to formulate likelihood functions to use in a Bayesian model of learning across these data, the overall result is based on accumulating evidence across a large amount of data and combining the evidence to form a posterior distribution. The purpose of the analysis is to present evidence which overwhelms various priors and leads to tentative agreement that the permanent-income hypothesis is a useful way of thinking about consumption behavior.

Monetary Trends

The final volume of Friedman and Schwartz's analysis of money and income *Monetary Trends in the United States and the United Kingdom* (1982) raised different empirical issues than the earlier work on the consumption function. Here, all the data are time-series data but the data have various degrees of measurement error with which Friedman and Schwartz were well acquainted. The original draft of *Monetary Trends* was completed in 1966 when the National Bureau of Economic Research reading committee suggested the addition of evidence for the United Kingdom (Hammond 1996: 188-9). The result was a delay of publication until 1982. By that time, statistical methods had changed dramatically. Friedman did not have a particularly high opinion of some of these methods in any case, but the book itself did not confront the issue directly, instead apparently pursuing a course plotted many years before publication.[9] This generated a critique by Hendry and Ericsson (1991) and an informative reply by Friedman and Schwartz (1991).

[9] Hammond (1996: 191) has a very nice quote from a letter by Friedman to Charles Goodhart indicating his lack of confidence in commonly used time-series techniques. On a personal note, I confronted this directly when I made the mistake of presenting at the Workshop in Money and Banking some research on the Gibson Paradox using Box-Jenkins analysis.

The overall method pursued in *Monetary Trends* is similar in many respects to that pursued in *A Theory of the Consumption Function*. Rather than running a sequence of tests on one common set of data, they examined hypotheses and built up a set of observations that were consistent with the data when examined in various ways. Friedman always was concerned about pre-test bias when a sequence of tests was run on the same data. Even if statistical theory did not raise this concern, his experience as part of the Statistical Research Group with multiple regression certainly was consistent with being concerned about any regression that results from repeated tests on the same data (Friedman and Schwartz 1991, Appendix.)[10]

The title of Friedman and Schwartz's response is an excellent one for my purpose: "Alternative Approaches to Analyzing Economic Data." As Friedman and Schwartz note, they have "literally hundreds of regressions" in their book. They have not avoided the problem of estimating more regressions than they have observations, or if so, not by much. Instead the book builds up simple hypotheses, examines them from various standpoints and draws conclusions. Friedman and Schwartz (1982) build up their hypotheses combining economics, prior information on institutions and statistical results including careful, close examination of the data.[11]

In their critique, Hendry and Ericsson regard it as a virtue of their analysis that the residual variation of their estimated equation is similar before and after World War I. Friedman and Schwartz regard this as evidence "against, not for, their [Hendry and Ericsson's] equations."

[10] The issue of repeated tests of different sorts on a set of time series is the same in principle as the issue of repeated tests in the (I hope) totally discredited stepwise-regression procedure. When the entire set of tests is done, there is little assurance of the appropriate p-value for any test statistic, other than being sure that the p-value overstates the statistical significance of the test statistic. I say this in full awareness that I have more than a few papers using a set of time-series data with sequences of tests which are not likely to be orthogonal.

[11] Hammond (1996, Ch. 10) emphasizes the same point although he does not state it in terms of priors.

(Friedman and Schwartz 1991, p. 41). It is not obvious how to interpret this in a classical setup, in which constancy of parameters is consistent with simplicity. In a Bayesian framework, it is easy to interpret the statement. Friedman and Schwartz had a strong prior, based on the extensive interpolation of the data before World War I, that measurement error was more important before World War I. If a statistical procedure shows constant residual variance, an explanation is necessary to be consistent with the interpolation and the resulting prior of a higher error variance in the earlier period.

As Friedman summarized his overall research strategy in “Money and the Stock Market” (Friedman 1988, p. 230, fn. 12) in response to two referees’ comments about serial correlation of residuals, he preferred to emphasize the congruence of evidence from a number of different sources with due attention to the quality of the data. He thought this was more informative than emphasizing the statistical significance of the “’best’” estimates among those examined. Such congruence is simplest to contemplate in terms of posterior probabilities after examining the data, which is quite consistent with Friedman’s orientation.

A Natural Experiment

Subsequent to *Monetary Trends*, Friedman wrote another empirical paper addressed to fellow economists which is a natural for my purpose (Friedman 2005). There is no test of statistical significance in “A Natural Experiment in Monetary Policy…”. Instead, the paper focuses on three episodes of rapid economic growth associated with the emergence of new industries and increases in stock prices followed by subsequent crashes in stock prices. In a well-defined sense, each of these episodes is an observation. There are only three observation but they are quite striking. The observations were chosen for their similarity before their respective stock market crashes: the United States in the 1920s; Japan in the 1980s; and the United States in the 1990s. The evidence, graphs and tables examine the nominal quantity of money, nominal stock prices and nominal GDP. The Great Contraction in the United States was associated with a substantial decrease in the nominal stock of money and nominal GDP. These substantial decreases did not occur in Japan or the United States. The decreases in stock prices are interpreted as reflecting "the inner dynamics of a collapsing bubble" (Friedman 2005: 149) in the stock market until about a year after the peak. At that time, economic developments due to the behavior of the nominal stock of money become important. The behavior of nominal GDP in all three episodes mirrors the behavior of the nominal stock of money, with the Great Contraction in the 1930s standing out with its continued decreases in stock prices and GDP.

Friedman's conclusion emphasizes the support of the data for his hypothesis:

> The results of this natural experiment are clear, at least for major ups and downs: what happens to the quantity of money has a determinative effect on what happens to national income and to stock prices. The results strongly support Anna Schwartz's and my 1963 conjecture about the role of monetary policy in the Great Contraction. They also support the view that monetary policy deserves much credit for the mildness of the recession that followed the collapse of the U.S. boom in late 2000.

The emphasis on support for a hypothesis, as opposed to a failure to reject, is itself consistent with a Bayesian interpretation of the data.

Conclusion

As the introduction indicated, there is no reason to expect Friedman's empirical work to reflect a detailed Bayesian analysis of data as appears in journals today. The tools and the hardware were not there.

Friedman's economic theory and his use of statistics reflected a common background theory: personal probability, which is Bayesian and the basis of Savage's explicitly Bayesian theory of statistics. The commonality is not an accident: People use personal probability to assess various possible states of the world, whether assessing the prospect of finding a job or the informativeness of an hypothesis. Personal probability is the basis of Friedman's expected utility theory and of his friendly criticisms of rational expectations.

Personal probability also is the basis of Friedman's empirical work. From a Bayesian standpoint, the purpose of empirical analysis is to assess the support for a theory and parameter values, something that is lacking in classical statistics with its emphasis on rejection and the failure to reject. The differences in tone and substance of Friedman's empirical analyses from the empirical analysis in a typical journal article could not be more startling. Those differences are directly related to Friedman's views on the foundations of statistics. Instead of viewing statistics as a classical statistician might, Friedman viewed statistics through the lens of personal probability. This Bayesian view of the foundations of statistics had dramatic implications for his overall research agenda and the particulars of his empirical work. He put less emphasis on the statistical significance of a single result because the common use of multiple tests on the same data can render such p-values meaningless. He emphasized agreement across various sets of data partly because such evidence was more likely to produce agreement among economists than a series of tests on a single set of data.

Friedman took very seriously the problems with sequential tests on a single data set and emphasized those problems more than once when discussing his preferred empirical strategy. This emphasis at least partly resulted from his personal experience with the Statistical Research Group and a model he estimated which fit extraordinarily well in sample and predicted terribly. Friedman preferred to accumulate evidence from a variety of sources and examine their consistency with a theory rather than find a p-value of five per cent at the end of numerous tests on a set of data. His book *A Theory of the Consumption Function* (1957) is remarkable for the almost complete absence of tests of statistical significance. While this by itself need not suggest a Bayesian approach, the accumulation of evidence across sets of data is consistent with using personal probabilities that can produce agreement amongst economists after many sets of data are examined. While economists cannot conduct controlled experiments to obtain replications of tests of many economic propositions, the use of data for different periods provides similar support when experiments are not feasible. Data for different periods are nicely used in Friedman's paper on a "natural experiment," which also does not contain a test of statistical significance. Instead, data are compared across three different episodes and are interpreted in terms of their consistency with the effects of the nominal quantity of money on the economy.

It is hard to find consistency of Friedman's empirical analysis with standard classical statistics and it is easy to interpret his theoretical and empirical work as being consistent with Bayesian personal probability. While Hammond (1996: 208) surely is correct that Friedman partly carried out "distinctive National Bureau business-cycle work", there is a substantial Bayesian statistical foundation underlying that work.

References

Bayesian Heresy. 2007. "Milton Friedman the Statistician." Available at http://bayesianheresy.blogspot.com/2007/08/milton-friedman-statistician.html.

Berndado, José M., and Adrian F. M. Smith. 2000. *Bayesian Analysis*. Chichester, England: John Wiley & Sons, Inc.

Ebenstein, Lanny. 2007. *Milton Friedman*. New York: Palgrave Macmillan.

Fienberg, Stephen E. "When Did Bayesian Inference Become 'Bayesian'?" *Bayesian Analysis*, 1, 1-40.

Friedman, Milton. 1976. *Price Theory*. New Brunswick, New Jersey: Transaction Publishers, 2007 reprint.

Friedman, Milton. 2005. "A Natural Experiment in Monetary Policy Covering Three Episodes of Growth and Decline in the Economy and the Stock Market." *Journal of Economic Perspectives* (4, Autumn), 145-50.

Friedman, Milton. 1988. "Money and the Stock Market." *Journal of Political Economy* 96 (2, April), 221-45.

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Friedman, Milton, and Rose D. Friedman. 1998. *Two Lucky People.* Chicago: The University of Chicago Press.

Friedman, Milton, and Leonard J. Savage. 1952. "The Expected-Utility Hypothesis and the Measurability of Utility." *Journal of Political Economy* 60 (6, December), 463-74.

Friedman, Milton, and Leonard J. Savage. 1948. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy* 56 (4, August), 279-304.

Friedman, Milton, and Anna J. Schwartz. 1991. "Alternative Approaches to Analyzing Economic Data." *American Economic Review* 81 (1, March), 39-49.

Friedman, Milton and Anna J. Schwartz. 1982. *Monetary Trends in the United States and the United Kingdom.* Chicago: The University of Chicago Press.

Hammond, J. Daniel. 1996. *Theory and Measurement: Causality Issues in Milton Friedman's Monetary Economics*. Princeton: Princeton University Press.

Hendry, David F., and Neil R. Ericsson. 1991. "An Econometric Analysis of U.K. Money Demand in Monetary Trends in the United States and the United Kingdom by Milton Friedman and Anna J. Schwartz." *American Economic Review* 81 (1, March), 8-38.

Klamer, Arjo. "Robert E. Lucas, Jr." 1983. In *Conversations with Economists* edited by Arjo Klamer, pp. 29-57. Totawa, New Jersey: Rowman and Allanheld, Publishers.

Pelloni, Gianluigi. 1996. "De Finetti, Friedman, and the Methodology of Positive Economics." *Journal of Econometrics* 75, 33-50.

Pelloni, Gianluigi. 1987. "A Note on Friedman and the Neo-Bayesian Approach." *The Manchester School*. 55 (4, December), pp. 407-18.

Savage, L.J. (1972). *The Foundations of Statistics*. Second revised edition. New York: Dover Publications, Inc.

Tukey, J.W. (1991). "The Philosophy of Multiple Comparisons". Statistical Science, 6(1): 100-16.

Zellner, A. (1985). "Bayesian Econometrics". *Econometrica*, 53(2): 253-69.