# From Wald to Savage: Homo Economicus Becomes a Bayesian Statistician

**— Source link** ↗

Nicola Giocoli

**Institutions:** University of Pisa

Related papers:

- On the Foundations of Decision Theory
- Decision and Choice: Bounded Rationality
- Rationality of Belief Or: Why Savage's axioms are neither necessary nor sufficient for rationality *
- Towards a Unified Decision Theory: A Non-Bayesian Approach
- Collectivist Foundations for Bayesian Statistics

# From Wald to Savage: homo economicus becomes a Bayesian statistician

Giocoli, Nicola

University of Pisa, Department of Economics

14 October 2011

# From Wald to Savage:

## *homo economicus* becomes a Bayesian statistician

Nicola Giocoli[*]

*Department of Economics, University of Pisa*

Bayesian rationality is the paradigm of rational behavior in neoclassical economics. A rational agent in an economic model is one who maximizes her subjective expected utility and consistently revises her beliefs according to Bayes's rule. The paper raises the question of how, when and why this characterization of rationality came to be endorsed by mainstream economists. Though no definitive answer is provided, it is argued that the question is far from trivial and of great historiographic importance. The story begins with Abraham Wald's behaviorist approach to statistics and culminates with Leonard J. Savage's elaboration of subjective expected utility theory in his 1954 classic *The Foundations of Statistics*. It is the latter's acknowledged fiasco to achieve its planned goal, the reinterpretation of traditional inferential techniques along subjectivist and behaviorist lines, which raises the puzzle of how a failed project in statistics could turn into such a tremendous hit in economics. A couple of tentative answers are also offered, involving the role of the consistency requirement in neoclassical analysis and the impact of the postwar transformation of US business schools.

---

# From Wald to Savage:

# *homo economicus* becomes a Bayesian statistician

Nicola Giocoli*

*Department of Economics, University of Pisa*

## Introduction

Rational behavior is the cornerstone of neoclassical economics. In the specific case of decisions under uncertainty, an agent can be termed rational if and only if she behaves as a Bayesian decision-maker, that is to say, if she makes choices according to the three main tenets of Bayesianism, namely, if she *i*) captures uncertainty by probability (whenever a fact is not known, the decision-maker should have probabilistic beliefs about it); *ii*) captures information by conditioning probabilities (the decision-maker should update her prior beliefs according to Bayes's rule as new information arrives); *iii*) follows the expected utility rule (the chosen alternative should maximize the weighted average of probabilities and utilities). In short, in the neoclassical paradigm economic rationality coincides with Bayesian rationality, as embodied in standard game and decision theory.[1] Yet, exactly *how, when and why* did the traditional notion of *homo economicus* as a self-interested utility maximizer came to be equated with the more sophisticated one of a Bayesian decision-maker?

The answer is seemingly straightforward. Credit should be given to the theory of decision under uncertainty developed by Leonard J. Savage in his classic 1954 volume, *The Foundations of Statistics* (Savage 1954 [1972]; *FS* henceforth). In the book Savage successfully combined a personalistic notion of probability with Bayes's rule and the axiomatic method to develop his

[1] Cf. Gilboa, Postlewaite & Schmeidler 2009, 287, "…within economic theory the Bayesian approach is the sole claimant to the throne of rationality". On the three tenets, see Gilboa, Postlewaite & Schmeidler 2004, who also add, as a fourth principle, that "a state should resolve all uncertainty", i.e., specify "all possibly relevant causal relationships and all that is known about the way information is obtained" (ibid., 5).

subjective expected utility theory (SEUT henceforth) which became the new orthodox characterization of economic rationality. Moreover, SEUT and Bayesianism played a key role in the late 1970s – early 1980s boom of game theory: in the hands of John Harsanyi and others, Savage's decision theory became the logical underpinning of the new orthodoxy in the field, Bayesian game theory, so much so that the latter may be considered the former's extension to a multi-agent setting (Myerson 1991, 5). Hence, both parametric and strategic rationality are nowadays founded upon Savage's SEUT and share the three main tenets of Bayesianism.

The "how, when and why" question may thus look like a no-brainer – a settled issue devoid of further historical interest.  Yet, historically speaking, the emergence of Bayesianism and SEUT as *the* characterization of economic rationality was hardly warranted, in view of Savage's  1954 real goal – transforming traditional statistics into a behavioral discipline – and actual achievement – the proof that the transformation was impossible. The aim of the present paper is to tell the story of that goal and that achievement, i.e., of how a self-recognized fiasco eventually became an unintended triumph. Acknowledging this story turns the "how, when and why" question into a meaningful one as it reveals that the metamorphosis of the traditional *homo economicus* into a Bayesian decision maker was far from inevitable, if only because its origins lie in a botched scientific endeavor. Thus, though the paper does *not* provide an answer to the "how, when and why" of Bayesianism in economics,[2] it will hopefully set the stage for viewing that question as a historically serious one.

In his 1946 *JPE* review of the *Theory of Games and Economic Behavior*, Jacob Marschak focused, as many other commentators, on John von Neumann's far-from-intuitive notion of mixed strategies and remarked that, by embodying that notion into his characterization of rational strategic behavior, von Neumann's theory ended up requiring that "to be an economic man, one must be a statistical man". However, in a footnote Marschak also noted that, in the same years, Abraham Wald was working at a new kind of statistical decision theory, according to which "being a statistical man implies being an economic man" (Marschak 1946, 109, text and fn.14). While von Neumann was requiring the *homo economicus* to make inferences like a proven statistician, Wald was suggesting that statisticians should embrace an economic way of reasoning. This passage in Marschak's review aptly captures the gist of the present paper. As it turns out, Savage's 1954 project was to reinforce, by way of the axiomatic method, the Wald's part of Marschak's remark, i.e., to teach statisticians how to behave as rational economic agents, but eventually, and

---

[2] A couple of tentative explanations will nonetheless be given in the last section.

unintendedly, ended up strengthening the von Neumann's part, transforming the economic men populating neoclassical models into fully-fledged Bayesian statisticians.

## §1. The pioneer: Wald's economic approach to statistics

The name of Abraham Wald is associated with the rise of behavioral statistics on account of his work on statistical decision theory.[3] His basic intuition was that statistical problems should be considered as special instances of general decision problems, where decisions have to be taken under conditions of uncertainty. Hence, from a decision-theoretic point of view, statistics should be defined as the science of making decisions in the face of uncertainty. A solution to a statistical problem must therefore instruct the statistician about *what to do*, i.e., what particular action to take, not just *what to say*. This approach was dubbed by Wald as *inductive behavior*, following the similar expression used by Jerzy Neyman in a 1938 paper (Neyman 1938).[4]

Generally speaking, a statistical decision problem (SDP henceforth) arises when a set of alternative decisions exists and the statistician's preference over them depends on an unknown probability distribution. The key insight relating this literature to modern economics is that the decision model developed by Wald also provides a setup for analyzing individual behavior in mundane problems of everyday life. The central issues in any SDP are, first, to choose what "experiment" to perform in order to extract information from the available data and, then, to choose what action to take (say, continuing experimentation or taking a final decision) given the experiment's outcome.

The basic decision model, for which credit must be given to Wald, is made of four components (see Ferguson 1976): *i*) the available actions; *ii*) the states of the world, one of which is the true, unknown one (so-called parameter space); *iii*) the loss function, measuring the loss to the statistician if he takes a certain action when the true state is a given one; *iv*) an experiment, whose

---

[3] For Wald's intellectual biography see Wolfowitz 1952 and Menger 1952. Leonard 2010 contains other information about Wald's troubled life.

[4] As a matter of fact, the true inventor of the economic approach to statistics, including the notions of loss function and experimental cost, was William S. Gosset, aka "Student" (see Ziliak 2008; Ziliak & McKloskey 2008). Wald was seemingly unaware of "Student"'s approach, if not through the mediation of Egon Pearson and Jerzy Neyman who, themselves influenced by "Student", had put forward similar ideas (see e.g. Neyman & Pearson 1933). It may be worthwhile to remember here that the most prominent statistician of the time, Ronald Fisher, and his disciples (including Harold Hotelling, who hired Wald as a statistician at Columbia University) fiercely opposed the decision-theoretic approach to statistics (cf. Fisher 1955, 75: "…in inductive inference we introduce no cost functions for faulty judgments... […] We make no attempt to evaluate these consequences and do not assume that they are capable of evaluation…").

goal is to help the statistician to reduce the loss and whose results (called observations) depend on the true state. A decision function is a rule associating an action to each possible experimental outcome. The available decision functions are evaluated according to the expected loss their adoption may cause under the various possible states. The statistician's task is then to choose the decision function capable of minimizing, in some specified sense, the expected loss.

Wald developed this setup in several works. The seminal one came in 1939 (Wald 1939) and already contained most of the main ideas, such as the definition of general decision problem and the notion of loss function. It also featured the two analytical principles which were to play a crucial role in the derivation of Wald's theory, namely, the minimax principle and Bayes's rule.[5] In the paper, the solution of a SDP – what he calls "the determination of the region of acceptance" of a given hypothesis – is made to depend on two circumstances: first, that not all errors (i.e., taking a certain action, like accepting a given hypothesis, when the true, yet unknown, state is a given one) are equal, and, second, that the statistician may have an a priori probability distribution over the parameter space (Wald 1939, 301). The first circumstance is accounted for by introducing a *weight function* which expresses the relative importance of the possible errors. Wald emphasizes that the choice of the specific shape of the weight function is not a matter of either mathematics or statistics. Often the importance of errors may be expressed in monetary terms, so much so that the function measures the monetary loss of taking a certain action when a certain state is the true one.[6] As to the second circumstance, Wald explicitly rejects Bayesianism as a philosophy of probability and mentions many objections which may be raised against the adoption of an a priori probability distribution. Solving a SDP must therefore be independent of the availability of an a priori distribution. Yet, the existence of such distribution is a very useful analytical tool: "The reason we introduce here a hypothetical probability distribution of [states of the world] is simply that it proves to be useful in deducing certain theorems and in the calculation of the best system of regions of acceptance" (ibid., 302). What we see here is the first instance of Wald's instrumental approach to Bayesianism, a recurring theme in his later works and a key point to understand the real goal of Savage's 1954 project.

Having defined the risk function as the expected loss of selecting a certain "region of acceptance", Wald proposes the mimimax principle as a general solution for a SDP, i.e., as a rule to

---

[5] Note that the paper came before von Neumann & Morgenstern's 1944 *Theory of Games and Economic Behavior*, where the minimax principle is cornerstone, though, of course, after John von Neumann's first proof of the minimax theorem (von Neumann 1928). When still in Europe, Wald was well acquainted with von Neumann's works: see Leonard 2010, Ch.8.

[6] That was precisely "Student"'s approach and what the Fisher school explicitly rejected: see above, fn.4.

choose the region of acceptance under a given weight function. He argues that, whenever we decide not to take into consideration an a priori probability distribution over the parameter space, "it seems reasonable to choose that [region of acceptance] which for which [the maximum risk] becomes a minimum" (ibid., 305). Thus, as early as 1939, he advocates the minimax as a reasonable solution criterion. Both the minimax and Bayes's rule are singled out for their expediency in deriving analytical results, but it is only for the minimax that Wald suggests that an explicit justification may be found for employing it as a practical SDP solution.

In short, though several elements were still missing – above all the idea that the design of the experiment be also part of a SDP[7] – the 1939 paper shows Wald's awareness that his ideas might be used to build a unified general theory of statistics, as well as to solve explicit statistical problems. This awareness allowed Wald to make further steps in his program following two crucial external events. First, the publication in 1944 of the *Theory of Games and Economic Behavior*, which suggested him the key insight of re-interpreting a SDP as a kind of von Neumann's two-person zero-sum games, and, second, his being presented with a specific instance of SDP in the form of quality control of warfare supplies.[8]

The latter event took place in early 1943, when a US Navy Captain, Garret L. Schuyler complained with economist and statistician Allen Wallis about the excessive size of the sample required for comparing percentages in ordnance testing.[9] As Wallis himself recounted many years later, Schuyler argued that, in the specific case of judging alternative methods of firing naval shells, "a wise and seasoned ordnance expert […] would see after the first few thousand, or even few hundred, [rounds] that the experiment need not be completed, either because the new method is obviously inferior or because it is obviously superior" (Wallis 1980, 325). Why, complained Schuyler, did statisticians go on designing "an experiment which is so long, which is so costly and which uses up so much of your shells that you've lost the war before you get the test over"? (see Warren Weaver's personal remiscences, quoted by Klein 2000, 47). Why not take a "more economic" approach to testing warfare equipment, i.e., one which could at the same time minimize experiment costs and ensure adequate sampling for proper quality control?

---

[7] The experiments in the 1939 paper are always single-stage. The breakthrough came with the analysis of multi-stage experiments, starting from Wald's work on sequential testing.

[8] According to Ferguson 1976, the limitations of Wald 1939, first of all the modelling of the action space as not independent of the state space, were perhaps due to Wald's willingness to encompass within his approach the standard methods of hypothesis testing and point/interval estimation. No surprise at that, given the mentoring role exercised on Wald's early forays in mathematical statistics by Harold Hotelling, the American leader of the Fisher school (see Wolfowitz 1952, 2-3).

[9] The episode is detailed in Wallis 1980 and Klein 2000.

As we learn from the historical note in Wald 1945, Wallis, together with future Nobelist Milton Friedman, tried to answer Captain Schuyler's challenge by conjecturing that there might exist a sequential test capable of controlling type I and type II errors[10] as effectively as the ordinary most powerful tests, while requiring a smaller expected number of observations. "It was at this stage that the problem was called to the attention of the author of the present paper. […] In April 1943 the author devised such a test, called the sequential probability ratio test." (Wald 1945, 121).[11]

A sequential test is defined by Wald as any kind of statistical procedure which gives a specific rule for taking, at any stage of the experiment, one of the following three actions: either accept the hypothesis being tested, or reject it, or continue experimentation by making an additional observation (Wald 1945, 118). The crucial feature of a sequential test is therefore that the number of observations is not predetermined, but is itself a random variable, given that at any stage of the experiment the decision to terminate the process depends on the result of previous observations. This was a big change with respect to standard test procedures, which required a fixed number of trials to be specified in advance and thus could at most be considered special cases of sequential procedures. The latter greatly economized on the number of observations, thereby answering Captain Schuyler's complaints.

The analytical goal of sequential testing is to minimize the number of observations required to reach a decision about acceptance or rejection of a given hypothesis under the desired test power. In the 1945 paper Wald did *not* manage to build an optimal test, i.e., one minimizing both the expected values of the number of observations required when either the given statistical assumption or its negation is true. Yet, he provided a substitute for the optimal test, and the proxy was, once again, offered by the minimax logic. He claimed that, when no a priori knowledge exists of how frequently the given hypothesis or its negation are true in long run, "it is perhaps more reasonable to minimize the maximum of [expected number of observations]..." (ibid., 124). The main tool developed by Wald under this logic was the sequential probability ratio (SPR) test, a testing procedure based on an expected number of observations considerably smaller than in standard most powerful tests for any desired level of control of type I and type II errors. Crucially for our story, the SPR test was explicitly founded upon Bayes's rule: it required updating any a priori probability the experimenter might entertain about the truthfulness of a given hypothesis

---

[10] Recall that, in statistical jargon, by type I and type II errors it is meant the possibility of, respectively, falsely rejecting a true hypothesis and failing to reject a wrong one.
[11] Wald adds that "the National Defense Research Committee considered these developments sufficiently useful for the war effort to keep the results out of the reach of the enemy…". Secrecy explains why the new method of sequential analysis was published in a scientific journal only two years later, in 1945.

with the new information arising from experimental observations. At any stage of the experiment the updating warranted that one of three actions be taken, namely, either accept or reject the hypothesis or continue with one more observation. Wald first argued that the SPR test might warrant any desired level of control over the two kinds of errors, while at the same time requiring much less information. Then, he explicitly distanced himself from endorsing Bayesianism and proved that the SPR test might work even in the absence of any a priori probability distribution.

Thus, by 1943 Wald had devised a brand new approach to the testing of statistical assumptions. The approach was based on strict economic logic from both the viewpoint of economizing over the experiment's costs and of asking the experimenter to act as an economic agent and take at each stage an optimal action, i.e., whether to endorse a certain hypothesis or to continue experimentation. In the latter respect, sequential testing was a direct application of the behavioral logic introduced in the 1939 paper – indeed, it proved the logic might bring operational results. Yet, the new procedure went beyond the seminal paper in that it got rid of the traditional single-stage-experiment constraint, explicitly allowing for multi-stage experimentation, where the behavioral element was even more crucial. In the following years Wald pursued three different research lines: *i*) the construction of usable sequential testing procedures; *ii*) the solution of specific SDPs; *iii*) the development of a general theory of statistical decision. It is this third, and most important, branch which matters to us.

The final ingredient in Wald's statistical decision theory came with acknowledging the formal overlap between his SDP and von Neumann's two-person zero-sum games (2PZSG). We know that in a SDP the experimenter wishes to minimize the risk function $r(F,\delta)$, i.e., the expected maximum loss that taking a certain decision $\delta$ might cause when the true distribution of the parameter space is F. Risk depends on two variables, but the experimenter can choose only one of them, $\delta$, but not the other, the true distribution F. This is chosen by Nature and the choice is unknown to the experimenter. Wald realized that the situation was very similar to a 2PZSG, with Nature playing the role of the experimenter's opponent. Thus, in Wald 1945a we find the first formulation of a SDP as a "game against Nature", an approach which will enjoy considerable popularity in the following years and will shape much of postwar decision theory.

As in von Neumann's games, the solution to the SDP-turned-2PZSG comes from the minimax logic: "Whereas the experimenter wishes to minimize the risk $r(F,\delta)$, we can hardly say that Nature wishes to maximize $r(F,\delta)$. Nevertheless, since Nature's choice is unknown to the experimenter, it is perhaps not unreasonable for the experimenter to behave as if Nature wanted to maximize the

risk." (Wald 1950, 27).[12] In this framework, "a problem of statistical inference becomes identical with a zero sum two person game" (Wald 1945a, 279). Yet again, Wald's commitment to minimax was not complete. What really mattered to him was that, even without endorsing their underlying logic,[13] both the theory of 2PZSG and the minimax solution were crucial for the analytics of statistical decision theory.

In a 1947 paper Wald provided the first complete and truly general formulation of a SDP. He also demonstrated the complete class theorem, the crucial result upon which most of his later theory is founded. Having defined a statistical decision function as a rule associating each sequence of observations with a decision to accept a given hypothesis about an unknown distribution (Wald 1947, 549), the theorem claims that the class of Bayesian decision functions – that is, of decision functions based upon the existence of an a priori probability over the unknown distribution and on the updating of that probability according to Bayes's rule – is complete (ibid., 552). This means that for any non-Bayesian decision function which can be used to solve a given SDP, there always exists a Bayesian decision function which, for all possible a priori distributions, is at least as effective at minimizing the risk function, i.e., for which the expected value of maximum loss is never larger.

With all the necessary ingredients at hand, Wald was eventually able to present in a compact form the outcomes of his decade-long research in the 1950 volume, *Statistical Decision Functions*. The book states from the beginning the motivation behind the whole project, namely, setting statistical theory free of two restrictions which marred it "until about ten years ago" (Wald 1950, v).[14] These were, first, that experimentation was assumed to be carried out in a single stage and, second, that decision problems were restricted to the two special cases of hypothesis testing and point/interval estimation. Wald boasts his theory free of both restrictions, as it allows for multi-stage experiments and general multi-decision problems. Any instance of the latter is treated in his new approach as a problem of *inductive behavior*: this because a statistician's choice of a specific decision function uniquely prescribes the procedure she must follow for carrying out her experiments and making a terminal decision (ibid., 10). Thus, the behavioral character of statistics

---

[12] Cf. the similar passage in Wald 1945a, 279, where the statistician is assumed to be "in complete ignorance as to Nature's choice".

[13] On von Neumann's minimax logic see Giocoli 2003, Ch.4; 2006.

[14] The book also contains a historical note where Wald reconstructs the thread leading to the new theory (ibid., 28-31). Credit for the first intuition of the decision-making character of statistical testing procedures is once again given to Neyman & Pearson 1933 and Neyman 1938.

is at the core of Wald's theory. The generalization allowed by the new, decision-theoretic – or, as Marschak would say, "economic"[15] – approach is truly remarkable.

The 1950 book is in many respect just an outgrowth of Wald's previous papers, though it is said to contain "a considerable expansion and generalization of the ideas and results obtained in these papers." (ibid., 31). To our aims, it is important to mention the very clear exposition of the basic ingredients of the general SDP and, above all, the way Wald presents the crucial issue of how the experimenter may judge the relative merit of any given decision function. Two elements need be evaluated in this respect: the experimenter's relative degree of preference over the various possible decisions when the true state of the world is known and the cost of experimentation (ibid., 8). The loss suffered by making a given terminal decision d when F is the true distribution of the parameter space is captured by the weight function W(F,d):[16] this function is always among the data of a SDP, but the big issue is in many cases how to attach values to it, i.e., how to measure losses. Experiment cost depends on the chance variable selected for observation, on the actual observed values and on the number of stages in which experiment has been carried out. The sum of the expected value of W(F,d) and the expected cost of experimentation gives the risk function r(F,$\delta$), where $\delta$ is the specific decision function adopted by the experimenter. Hence, the merit of any given decision function for purposes of inductive behavior may be entirely evaluated on the basis of the risk associated with it (ibid., 12). The complete class theorem of Wald 1947 then allows to conclude that, if an a priori distribution $\xi$ exists and is known to the experimenter, a decision function for which the average risk – the average being calculated using Bayes's updating rule – takes its minimum value may be regarded as an optimum solution. In fact, a DF $\delta_0$ which minimizes this average risk for all possible $\delta$ is called a Bayes solution relative to the a priori distribution $\xi$ (ibid., 16).

Once more, Wald explicitly distances himself from "real" Bayesianism, claiming that the a priori distribution $\xi$ may often either not exist or be unknown to the experimenter.[17] As an alternative, one may thus make recourse to the minimax solution: a decision function is a minimax solution of the SDP if it minimizes the maximum risk with respect to the distribution F. The author also points to the "intimate connection" between Bayes and minimax solutions: under fairly general conditions a minimax solution to a SDP is also a Bayes solution (ibid., 89 ff.). And while both the

---

[15] Cf. Savage 1961 [1964], 177, where inductive behavior is defined as "the economic analysis of statistical problems".
[16] W(F,d) = 0 thus means that d is the correct terminal decision when F is true.
[17] "To avoid any possibility of misunderstanding, it may be pointed out that the notions of Bayes solutions and a priori distributions are used here merely as mathematical tools to express some results concerning complete classes of decision rules, and in no way is the actual existence of an a priori distribution in postulated here." (Wald 1950a, 238).

Bayes and the minimax solution are again justified for their usefulness in deriving basic results about complete classes of decision functions, Wald does offer his customary, though timid, defense of the minimax, calling it "a reasonable solution of the decision problem" precisely for those cases where ξ does not exist or is unknown (ibid., 18).

In a concise presentation of his new theory for the 1950 International Congress of Mathematicians, Wald concluded by saying that "While the general decision theory has been developed to a considerable extent and many results of great generality are available, explicit solutions have been worked out so far only in a relatively small number of special cases. The mathematical difficulties in obtaining explicit solutions, particularly in the sequential case, are still great, but it is hoped that future research will lessen these difficulties and explicit solutions will be worked out in a great variety of problems" (Wald 1950a, 242). The airplane crash which killed Abraham Wald and his wife in December 1950 brought an abrupt end to this research. It will be up to other scholars to continue it and to one of them to turn it towards an unexpected direction.


**§2. The legacy: Savage on Wald**


Savage's statistical work began at the Columbia University's wartime Statistical Research Group, where he joined a stellar team of economists and statisticians, which included, among others, Friedman, Wallis and Wald. The SRG had been the receiver of Captain Schuyler's complaints and where Wald had developed his analysis of sequential testing.[18] The impact of Wald's new theory on the young statistician was considerable, as is demonstrated by a couple of papers that Savage co-authored while at the SRG.[19] Yet his plan was more ambitious.

At the 1949 meeting of the Econometric Society, two sessions were held under the common title "Statistical inference in decision making". The sessions featured 5 papers related to Wald's research, including one by Wald himself who chaired one of the sessions. Savage was among the other presenters, with a paper titled "The role of personal probability in statistics". That work has never been published, but its *Econometrica* abstract shows that by that early date Savage had already identified the core of his 1954 book. The key idea was, so to speak, "let's take Wald

---

[18] On the history of the SRG see Wallis 1980 and Klein 2000.
[19] Both Arnold, Girschick & Savage 1947 and Friedman & Savage 1947 deal with the design of experiments. Note that the papers were published only after the war, once again due to secrecy constraints about the SRG's activity. On the impact of Wald's sequential analysis upon the SRG see Wallis 1980, 325-8.

seriously". That is, if statistics must really become a behavioral discipline, if statistical inference is a matter of decision theory, if statisticians must behave as rational economic men, then it is necessary to characterize more rigorously what rational behavior amounts to. Only a full theory of rational behavior under uncertainty will provide – as Wald's project requires – the decision-theoretic foundations for a general theory of behavioral statistics (cf. Savage 1949).

Yet, in order to do so one has to go beyond "the tendency of modern statisticians to countenance no other than the frequency definition of probability". According to Savage, the frequentist view is indeed responsible for "insurmountable obstacles" preventing the development of behavioral statistics – obstacles that even Wald's minimax theory has been unable to overcome, but that "may be bypassed by introducing into statistical theory a probability concept, which seems to have been best expressed by Bruno de Finetti…". The latter has argued that "plausible assumptions about the behavior of a 'reasonable' individual faced with uncertainty" imply that "he associates numbers with the [uncertain] events, which from the purely mathematical point of view are probabilities." Moreover, these probabilities, which Savage calls "personal probabilities", are "in principle *measurable by experiments*" on the individual and their interpretation offers a well-defined characterization of how the individual should act in the face of uncertainty "in view of the von Neumann – Morgenstern theory of utility". Unfortunately, Savage notes, de Finetti's theory "compares unsatisfactorily with others (in particular Wald's theory of minimum risk)" because it neither predicts nor demands that a crucial feature of modern statistical analysis, deliberate randomization, be undertaken by the decision-maker. Thus, "both Wald's and de Finetti's theories are incomplete descriptions of what statistical behavior is and should be", so much so that "*we may look forward to their unification into a single more satisfactory theory*".[20]

As it turns out, all the essential ingredients of *FS* are already here: the rejection of frequentism, though, note well, not of frequentist-based statistical techniques; the praise of Wald's minimax; the personalistic view of probability as the numerical evaluation of uncertainty entertained by a "reasonable" decision-maker; the idea that personal probabilities can be elicited observing the agent's behavior under uncertainty; the idea that these probabilities fully characterize that behavior according to von Neumann's expected utility; the explicit normative penchant of the analysis; above all, the intuition that combining Wald and de Finetti may represent the most promising path towards a general theory of statistics as "acting in the face of uncertainty". Also

---

[20] All quotes are from Savage 1949, emphasis added.

noteworthy is what is *not* in Savage's 1949 sort-of-manifesto. No reference is made, in fact, to Bayesianism as a general philosophy of probability,[21] nor to the idea of upturning consolidated statistical techniques. Even the goal of developing a general theory of decision-making under uncertainty, to be used as a guide to rational behavior beyond the boundaries of statistical work, is conspicuously absent. The continuity of Savage's manifesto with respect to Wald's work was therefore quite strong. As in Wald, he wanted to apply (what today we call) Bayesian techniques to provide traditional statistics with more solid decision-theoretic foundations. As in Wald, his aim was to offer a guide to statisticians in their daily work. Where he wished to improve upon Wald was in the characterization of what it actually meant *for a statistician* to behave rationally in the face of uncertainty.

The continuity is felt even more strongly in Savage's review of Wald 1950. The paper – which appeared in the *Journal of the American Statistical Association* only after Wald's tragic death but had been written before it – had been commissioned as more than a simple review. The goal assigned to Savage by *JASA* Editor was to give an informal exposition of Wald's new approach to statistics. This explains why, as Savage himself observes, "the paper is no longer exclusively a review" (Savage 1951, 55, fn.1). Indeed, it is made up of three distinct parts, each of which of great historical relevance: *i*) a presentation of the decision-theoretic approach to statistics; *ii*) an introduction to the state-act-consequence model as a method to characterize an agent's decision under uncertainty; *iii*) a critical exposition, plus a possible defense, of Wald's minimax rule. Much like the 1949 paper, what we have here is a kind of manifesto, or, more properly, a work-in-progress report, of the *FS* theoretical edifice.

The review begins with the remark that the traditional statistical problem is to draw inferences, i.e., to make assertions on the basis of incomplete information. A related problem is the design of experiments permitting the strongest inference for given expenditure. But Wald's theory is about statistical *action*, not inference, i.e., about deciding what behavior to take under incomplete information. As prominent examples Savage mentions quality control and experiment design, but his main point is that "all problems of statistics, including those of inference, are problems of action" (ibid., 55). Thus, statistics must be reinterpreted as a discipline concerned with behavior, rather than assertions. As he will put it in FS, statistics is about what to do, rather than what to say: "The name 'statistical decision' reflects the idea that inductive inference is not always, if ever,

---

[21] Indeed, as showed by Feinberg 2003, the term itself "Bayesianism" was hardly used before the 1950s. Even the term "frequentist" seems to have become popular only in the 1950s, following its use as a polemic target by (whom today we call) Bayesians.

concerned with what to believe in the face of inconclusive evidence, but that at least sometimes it is concerned with action to decide upon such circumstances." (FS, 2).[22]

Having affirmed the behavioral content of statistics, Savage endeavors to explain what "a course of action" or, more simply, "an act" actually is. The notion must be understood "in a flexible sense", as even the whole design of a complicated statistical program may be regarded as a single act. More than that, "in a highly idealized sense", even an agent's entire existence may be thought as involving only one, single-act life decision, such as, say, "the decision to conduct himself according to some set of maxims envisaging all eventualities" (Savage 1951, 56). Of course, such an idealized view of an act "is manifestly far-fetched", but he believes it may call attention to the appropriateness in the new approach of considering "very large decision problems as organic wholes." (ibid.). This is a crucial passage in the paper and, possibly, in Savage's overall intellectual project, as it is precisely at this stage that his analysis parts company with Wald's.

We know that the latter had already presented a generalized version of SDP. Yet, Wald's SDPs were still expressed in the technical jargon of probability distributions, set theory and parameter spaces, with no concession to the reader in terms of simplified, possibly non-statistical examples. Having been assigned the task to offer a cut down exposition of Wald 1950, Savage elects to present the basic decision problem in a more straightforward way: "Acts have consequences for the actor, and these consequences depend on facts, not all of which are generally known to him. The unknown facts will often be referred to as states of the world" (ibid., 56). The rhetorical power of the state-act-consequence (SAC) language can hardly be downplayed. It brings the reader the message that SDPs are really just like any other kind of decision problems. This is reinforced by the first example chosen by Savage to illustrate the power of the new language (ibid., 56-7), namely, the problem of deciding… whether to carry an umbrella under uncertain weather conditions!

Provided each consequence can be assigned "a numerical score such that higher scores are preferred to lower scores" (monetary income being of course the most intuitive way to measure these scores) and provided agents may assign probabilities to the various states, it is possible to calculate the expected value associated with an action. The decision-maker will then follow von Neumann and Morgenstern's utility theory and choose the action maximizing this expected value (ibid., 57-8). However, if the agent does *not* assign probabilities to states "this trivial solution does not apply and the problem is newer". Indeed, the main theme of "modern, or un-Bayesian,

---

[22] In 1954 Savage will call the standard approach to statistics the *verbalistic outlook* (FS, 159), to be contrasted with the new *behavioralistic outlook*. A few years later the latter will be re-christened as the *economic-theoretic* approach: Savage 1961 [1964], 177.

statistical theory", as Savage calls it, has been precisely that of dealing with uncertainty when probability does not apply to unknown states of the world (ibid., 58).

Reading the 1951 review, one may not escape a sense of discontinuity in Savage's exposition. While the paper's first page, dedicated to explaining the new behavioral approach to statistics, might have been written by Wald himself and is fully pertinent to the general SDP issue, the introduction of the SAC terminology and, even more, the umbrella example bring the reader away from the realm of statistics and into that of economics, i.e., into the world of the theory of decision under uncertainty. The jump is of course intentional, as Savage wished to promote an economic approach to statistics, i.e., to bring forward the view that statisticians should behave as rational economic men. The message is: if we can devise a rule to effectively solve the umbrella dilemma, the very same rule can be applied to *any* kind of decision problems, including complicated statistical ones, because they all are amenable to treatment according to the SAC formalism and because their solution always involves the selection of an act among a set of alternatives.

Having defined the notions of dominant and mixed acts, Savage proceeds to state "the general rule by which the theory of statistical decision functions tentatively proposes to solve all decision problems" (ibid.), that is, Wald's minimax rule. Again, the rhetorical device is remarkable. Wald's minimax is presented as a way out from the stalemate caused by the unavailability of probability values to be attached to states, a stalemate – which in *FS* Savage will call "complete ignorance" – that is said to always affect statistical problems within the "modern, or un-Bayesian" approach. Let $I(a,s)$ be the expected income if act $a$ is chosen when the true state is $s$ (both $a$ and $s$ belong to finite sets). Let loss $L(a,s)$ be the difference between the most that can be earned by choosing any act when state $s$ obtains and what is actually earned by choosing $a$ when $s$ obtains, that is to say, $L(a,s) = max_{a'} I(a',s) – I (a,s)$. Thus, as in Wald, the loss measures "how inappropriate the action $a$ is in the state $s$" (ibid., 59). Wald's minimax principle then states: choose an action $a$ such that the maximum loss which may be suffered for any possible state $s$ is as small as possible, i.e., that minimizes the maximum loss. This principle, which Savage credits as being "the only rule of comparable generality proposed since Bayes' was published in 1763",[23] is said to be "central to the theory of statistical decision functions, *at least today*" (ibid., added emphasis). The emphasized words elucidate Savage's attitude with respect to Wald's minimax. In the rest of the review he will, first, offer a possible argument to justify the minimax, then criticize the criterion and, finally, argue

---

[23] Note that never in the review Savage mentions the circumstance that Wald also employed Bayes's rule, if only as a technical device.

that Wald's book – on account of its reliance on so arguable a criterion – is just a preliminary step towards a more complete reconstruction of statistics on behavioral basis, an endeavor that, after Wald's death, it will be up to Savage himself to accomplish.

Savage's defense of the minimax as a rule for making statistical decisions under uncertainty is ingenious. In a clear anticipation of *FS*'s SEUT, he argues that in the case of individual decisions, the criterion is *not* required because, following de Finetti's personalistic view of probability, individuals are never in a situation of "complete ignorance": a single agent can always rank the consequences of her actions because she always holds probabilistic beliefs about the states of the world. Yet, the minimax becomes indispensable whenever decisions must be taken by *a group*: "If, however, the actor is a group of individuals who must act in concert with a view to augmenting the income of the group, the situation is radically different, *and the problems of statistics may often, if not always, be considered of this sort*." (ibid., 61, emphasis added). The reason is intuitive: whose probability belief about the states of the world should be given priority in deciding what action to take? Absent any reason for privileging one belief over the others, adopting the minimax criterion as if no such beliefs exist sounds appealing, because it "means to act so that the greatest violence done to anyone's opinion shall be as small as possible" (ibid., 62). Thus the minimax commends itself as "a principle of group action", a compromise solution which may avoid causing undue losses to any of the group's members. As the previously emphasized words clarify, Savage believes that group action is ubiquitous in statistics – indeed, in the whole of science.[24] Moreover, it is often the case that, "under the minimax rule the same amount will be given up by each member of the group", again reinforcing the compromise character of the minimax choice. Finally, the minimax also offers a way out from another, potentially troublesome issue, namely, that of selecting who is entitled to be part of the decision-making group: under the minimax all "reasonable" opinions may be considered, without having to decide beforehand whose opinion is legitimate and whose is not (ibid.).[25]

This strong defense of the minimax does not hinder Savage from criticizing Wald's theory. The first motive of dissatisfaction is precisely Wald's inability to offer a valid justification for the

---

[24] "…one might idealize at least part of the activity of academic science in terms of decisions reached jointly by the whole group of people interested and competent in the science in question." (ibid., 62).

[25] Although determining what qualifies as a "reasonable" opinion is still "a real difficulty of the theory" (ibid., 65). Savage highlights another plus of the minimax rule in case of group action. The fact that the addition of a new act to the list of available ones may lead to the selection of a different minimax choice even if the latter assigns zero probability to the newly available act has brought many to attack Wald's principle as paradoxical – indeed, as a violation of the independence axiom. Yet, in case of group decisions, there is no paradox in admitting that the addition of a new act to the list may lead the group to achieve a different compromise which now takes into account the circumstance that some of the group's members may actually prefer the newly available act (ibid., 64).

criterion, which to him was merely an analytic device. Moreover, Wald's definition of the loss function is unsatisfactory and exposes the minimax to the easy critique of being an unjustifiably ultra-pessimistic rule. Indeed, Wald failed to distinguish his notion of loss *L(a,s)* from that of negative income − *I(a,s)*. The two notions coincide only if zero is the maximum value of *I(a,s)*, that is, if the most the decision-maker may earn by guessing the right state and selecting the right action is the absence of any loss. In such a case it is *L(a,s) = − I(a,s)*, but this is truly ultra-pessimistic and "no serious justification for it has ever been suggested"; actually, it may even bring to the absurd conclusion that, in some cases, no amount of experimentation will bring the agent to behave differently than as if he were under complete ignorance (ibid., 63). It is this potential irrelevance of observations for the agent's choice which makes Wald's negative-income version of the minimax rule untenable for statistics.[26]

Savage concludes the review by saying that the 1950 book, a "difficult and irritating" one to read, is at best a sort of intermediate report of Wald's research, albeit of "great scholarly value" and with "inestimable" possible influence on future statistics. Clearly, the project of instructing statisticians to behave as economic actors was still incomplete: the minimax rule was itself far from perfect and, above all, no results had been achieved about either the applicability of the rule to instances of concrete statistical problem or the possibility to encompass the traditional inference techniques under the minimax umbrella. Yet the review already offers the guidelines for Savage's own attempt to fulfill Wald's project – what in the next § will be called the *FS*'s game plan. A theoretical breakthrough will nonetheless still be needed to accomplish the task, namely, the replacement of Wald's minimax with a new decision criterion which, in turn, will require a novel characterization of probability as subjective belief.

### §3. The strategy: *FS*'s game plan

Savage's project in *FS* followed Wald's steps and aimed at re-founding statistics by transforming it into a fully-fledged behavioral theory. Yet, differently from Wald, the discipline's new foundations were to lie in a subjectivist notion of probability (FS, 4-5). Any investigation of the 1954 book must

---

[26] The distinction between loss and negative income in the general case where $max_{a'} I(a',s) \neq 0$ will be emphasized in the *FS*, where the loss will be, though only half-heartedly, re-christened as the *regret function* (*FS*, 163-4) and where the point of the potential absurdity of Wald's account of the rule will be made even more forcefully (ibid., 170-1).

therefore keep in mind these two facts, i.e., that Savage targeted statistics, not economics, and that the two key ingredients in his book were behaviorism and subjectivism.

"The purpose of this book, and indeed of statistics generally, [is] to discuss the implications of reasoning for the making of decisions." (FS, 6). These words are placed by Savage at the beginning of *FS*'s Ch.2 to illustrate the first essential ingredient of his analysis, namely, the reinterpretation of statistics as a discipline concerned with the making of decisions, rather than the statement of assertions. As he puts it, "[t]he name 'statistical decision' reflects the idea that inductive inference is not always, if ever, concerned with what to believe in the face of inconclusive evidence, but that at least sometimes it is concerned with action to decide upon such circumstances." (FS, 2). The traditional verbalistic outlook of statistics, where statistical problems concern deciding what to say, need be replaced by Wald's behavioralistic outlook, according to which the object of statistics is to recommend wise action in the face of uncertainty (FS, 159).

The second key ingredient in FS is the subjectivist, or personalistic, view of probability. Following de Finetti's pioneering work, probability is defined as an index of a person's opinion about an event, i.e., a measure of "…the confidence that a particular individual has in the truth of a particular proposition…" (FS, 3). In this view, "…personal probability […] is […] the only probability concept essential to science and other activities that call upon probability." (FS, 56). Savage acknowledges that the 20th-century boom of statistical research – carried on by what he calls the British-American school of statistics – had taken place entirely in the objectivist field. Yet the frequentist view suffers from a series of weaknesses. Some are well known (like the fact that objective probability applies only to the very special case of repetitive events or the circularity of the frequentist definition, which depends on the existence of infinite sequences of independent events whose definition requires in turn a notion of probability), but Savage adds a further critique, namely, that objective probability cannot be used as a measure of the trust to be put in a proposition: "…the existence of evidence for a proposition can never, on an objectivistic view, be expressed by saying that the proposition is true with a certain probability. […] if one must choose among several courses of action in the light of experimental evidence, it is not meaningful, in terms of objective probability, to compute which of these actions […] has the highest expected income." (FS, 4). Hence, objective probability is unsuited as a basis for a truly behavioral approach

to statistical problems,[27] although any effort to rebuild the whole of statistics cannot overlook the bulk of invaluable results produced by the British-American School.

Savage's general goal of re-founding statistics explains why the book is divided in two distinct parts. In the first (Chs 2-7), he develops a rigorous theory of rational behavior under uncertainty which provides the required normative benchmark for behavioral statisticians. In the second (Chs 9-17), he endeavors to reinterpret the statistical methods of the British-American School according to the new decision theory in order to demonstrate that standard inference techniques can work even in a subjectivist/behaviorist framework. As can be read in the book's general outline: "It will, I hope, be demonstrated thereby that the superficially incompatible systems of ideas associated on the one hand with a personalistic view of probability and on the other with the objectivistically inspired developments of the British-American School do in fact land each other mutual support and clarification" (FS, 5). Drop the "subjectivist" component and what you have in this second part of *FS* is once again Wald's general project for a behavioral statistics.

Only by keeping in mind this two-part program it is possible to disentangle the complex structure of the 1954 book, which otherwise may look messy or, to be generous, as though it had been written by two different, sometimes even opposed, authors.[28] Indeed, the program may be best appreciated by bringing to light the logical sequence followed by Savage to implement it, what I call the *FS*'s "game plan": first, develop subjective probability theory and the new theory of rational behavior (SEUT); second, present and defend Wald's minimax rule; third, show that the minimax rule may also be given a subjectivist interpretation; fourth, apply the minimax rule to demonstrate that orthodox inference techniques may be translated into behaviorist terms; fifth, replace in the latter task the minimax rule with the new SEUT. At the end of this sequence, a brand new kind of statistics would emerge, the traditional verbalistic and frequentist approach having been transformed into a behavioral discipline governed by a rigorous subjective decision theory. A very ingenious game plan, but, alas, also an unsuccessful one.

As we already know, the *FS*'s second part did not fulfill its goal, so much so that the book is universally praised only because of its first seven chapters, while the remaining ones are usually neglected as ill-founded. Savage himself recognized the failure. By 1961 he had already realized

---

[27] Note that Savage acknowledges that all the three main approaches to probability (objective, subjective and logicist) accept the axiomatic, or measure-theoretic, notion developed by Kolmogorov in the 1930s, the differences only amounting to the extra-mathematical interpretation given to the same formal concept of probability measure. Yet, Savage, as de Finetti before him (cf. von Plato 1994, 276-8), refrains from using in *FS* the measure-theoretic notion of probability on account of the dubious axiomatic legitimacy of the key postulate of countable additivity (FS, 42-3).

[28] Just think of the crucial role played in the second part by Wald's minimax rule, a criterion based upon an objectivist notion of probability: see below, §5.

that Wald's project was a dead alley: "…the minimax theory of statistics is […] an acknowledged failure. […] …those of us who, twelve or thirteen years ago, hoped to find in this rule an almost universal answer to the dilemma posed by abstinence from Bayes' theorem have had to accept disappointment." (Savage 1961 [1964], 179).[29] In the preface to the 1972 edition of the book, he was even more explicit and admitted that the promised subjectivist justification of the frequentist inferential devices had not been found – quite to the contrary, what he had proved in *FS* was that such a justification could *not* be found! Indeed, his late 1950s-early 1960s "Bayesian awakening" came, at least in part, because he recognized the theoretical fiasco of the *FS*'s second part. As he put it in 1972, "Freud alone could explain how [such a] rash and unfulfilled promise […] went unamended through so many revisions of the manuscript." (FS, iv).

Actually, it was not a matter of psychoanalysis, but rather of the author's not-yet-mature awareness of the potential of the Bayesian approach. It is again Savage himself who tells us that in the early 1950s he was still "… too deeply in the grip of the frequentist tradition […] to do a thorough job" (Savage 1961 [1964], 183).[30] At the time of writing *FS* Savage had just converted from his initial objectivist stance to the subjectivist creed,[31] as is confirmed by the book's goal to complete Wald's program on the behaviorist foundations of orthodox statistical techniques. To this aim, subjectivism – both in terms of the use of personal probability and in terms of SEUT – was just a useful tool, not the outcome of any deep philosophical commitment. Indeed, it is even legitimate to ask whether Savage was a Bayesian in 1954. Of the three main tenets of Bayesianism – assess empirical claims via subjective probabilities; use Bayes's rule to evaluate new evidence; make decisions according to the expected utility rule – he fully endorsed in *FS* just the last one. This is not to mean that he disregarded the first two, but to underline that the book is rid with the author's doubts and critiques about them and includes features one would never expect in a "truly" Bayesian analysis, like the ample space dedicated to the analysis of vagueness in probability claims and to choice under complete ignorance (see below, §5).[32]

The fact is that Savage 1954 aimed neither at revolutionizing statistics, nor at describing how agents (let alone, *economic* agents) really behave. As to the latter, his new decision theory was

---

[29] The reference to "twelve or thirteen years ago" is probably to Savage 1949.

[30] Remember once again that in the postwar years the Fisher School dominated US statistics, mainly through the overwhelming influence of Harold Hotelling.

[31] "The book shows him in a transition period. He had accepted the personalistic view of probability but had not abandoned minimax or other ideas of the frequentist school." (Lindley 1980, 7).

[32] Also note that while the first two tenets are accepted under every possible definition of Bayesianism, it is more debatable that reliance on the expected utility rule should deserve the same status. As it happens, the inclusion of EU as the third fundamental tenet of Bayesianism is often justified on pragmatic grounds, while it might well be rejected at a non-pragmatic level of analysis: see Mongin 2011.

explicitly normative: the goal was to teach statisticians how they *should* behave to solve their statistical problems, not to describe how they – or any other decision-maker – actually behave. As to the former, we should not be deceived by the book's foundational emphasis. Yes, Savage wanted to rebuild statistics on behavioral foundations, but this was hardly a novel project (actually, it was Wald's one) and its most immediate implications were, to Savage, far from revolutionary. He openly admired and wanted to preserve orthodox inferential techniques – actually, he tried to strengthen them by providing the logical and behavioral underpinnings which they lacked in their standard frequentist rendition. Bayesian revolution, in both statistics and economics, was conspicuously absent from Savage's 1954 radar.[33] As his disciple and co-author Dennis Lindley observed, in *FS* Savage was an unconscious revolutionary, so much so that the paradigm change (in Kuhnian sense) the book brought to statistics was, at least initially, an unintended one (Lindley 1980, 5-7). Yet, a revolutionary he was, and he was soon to recognize it: suffices to compare the timid, hands-off tone with which he presented personal probability in 1954 with the self-confident, almost arrogant way he used time and again the expression "we Bayesians" (when not "we *radical* Bayesians") in the 1961 paper.[34]

## §4. The achievement: FS's SEUT

Savage's place in the history of science is unquestionably due to the axiomatic development of subjective expected utility theory (SEUT) in the first seven chapters of *FS*. In the crucial Chapter 5 he demonstrates what in decision-theoretic jargon is called a *representation theorem*, that is, a theorem showing that, given an evaluation criterion for determining a preference relation over a set of options and given a set of axioms that the decision-maker's preferences satisfy, the preferences about the options determined according to the evaluation criterion always coincide with the decision-maker's preferences.[35] In the case of Savage's SEUT, where the options are acts, the evaluation criterion is the expected utility formula and the axioms are those listed in *FS*'s Chapters 2 and 3, the theorem shows that there exist a unique utility function and a unique

---

[33] This is confirmed by Savage himself: "Savage (1954) […] was not written in the anticipation of radical changes in statistical practice. The idea was, rather, that subjective probability would lead to a better justification of statistics as it was then taught and practiced, without having any urgent practical consequences." (Savage 1962, 9).

[34] Savage has claimed that his definitive conversion to Bayesianism was due to the likelihood principle: "I came to take Bayesian statistics seriously only through recognition of the likelihood principle" (Savage 1962).

[35] For a survey on representation theorems see Gilboa 2009, Chs. 9-15, and Wakker 2010, Part I.

subjective probability function such that the decision-maker's preferences always conform with – i.e., are captured by – the expected utility formula. This means that an agent prefers act *f* over act *g* if and only if the expected utility associated with act *f* and calculated using those unique utility and probability functions is no less than that associated with act *g* (cf. FS, 79, Theorem 1). A rational agent, that is, an agent whose preferences satisfy Savage's axioms, makes her decisions according to the expected utility formula. The latter is therefore the criterion for rational decision-making under uncertainty.

There are three distinguishing features in Savage's SEUT: subjectivism, consistency and behaviorism. First of all, SEUT incorporates a personalistic view of probability, where uncertainty about the states of the world is captured in terms of probabilistic beliefs over those states. Second, the theory embodies a consistency view of rationality: the subjective beliefs are not entirely free, but need be constrained by a set of consistency requirements (the axioms). Third, SEUT partakes of the behaviorist methodology, as the unknown beliefs are elicited by observing the agent's behavior in the face of simple choices under uncertainty (more on this below).

Savage did not develop his theory in a vacuum. He availed himself of the foundations provided by von Neumann's EUT and explicitly targeted the latter's main weakness, namely, the not-so-well-specified nature of the probability values. In the *Theory of Games* the derivation of numerical utilities had been based on the "…perfectly well founded interpretation of probability as frequency in long runs." (von Neumann & Morgenstern 1944 [1953], 19), but the authors themselves had left the door open to a subjectivist restatement (ibid., fn.2). That opportunity was seized by Savage: "A primary and elegant feature of Savage's theory is that no concept of objective probability is assumed; rather a subjective probability measure arises as a consequence of his axioms" (Luce & Raiffa 1957, 304). Hence, the first feature of SEUT, the subjectivist approach to probability, finds a rationale in the author's willingness to improve upon von Neumann's theory.

As to the notion of probability itself, Savage defined it as "the confidence that a particular individual has in the truth of a particular proposition" (FS, 3) and exploited de Finetti's characterization of subjective beliefs as purely formal objects, constrained by consistency requirements and elicited by observing the decision-maker's behavior under uncertainty (see de Finetti 1937 [1964]). This established a strong connection between the subjectivist side of SEUT and the other two main components, i.e., axiom-based consistency and behaviorism. It is indeed the latter's combination which give analytical content to the former.

The second feature of SEUT, axiom-based consistency, warrants the theory's desired normativeness (don't forget that Savage's main goal was to *instruct* statisticians about how to behave rationally). Inconsistent beliefs are those which violate the axioms and thus entail irrationality. Conversely, an agent may label herself rational if and only if her beliefs, as well as her preferences, obey the axioms. The role of Bayes's rule is to police the consistency of beliefs by allowing their updating as long as new information arrives.

The characterization of rationality as consistency was the trademark of postwar decision theory, which in turn represented the culmination of an almost century-long endeavor undertaken by marginalist and neoclassical economists who wished to model rational behavior. While the economists' efforts had focused on the notion of utility, SEUT extended the consistency requirement to another kind of unobservable entities – the decision-maker's beliefs about "the truth of a particular proposition". Just like Gerard Debreu did for the agent's preferences (Debreu 1959) and Paul Samuelson for the agent's choices (Samuelson 1948), Savage transformed those beliefs into tightly-constrained theoretical objects, suitable for formal modeling.[36] This was a major innovation with respect to the previous literature on expectations and conjectures, where – with the only exception of non-economists de Finetti and Frank Ramsey – economists had treated agents' beliefs as loose introspective entities, implicitly delegating their analysis to less rigorous disciplines, such as psychology or sociology.

Note that Savage's theory contains two elements which seemingly evoke introspection, namely, the property of cardinality of the EU function and, obviously, the subjective character of probability. Thus, it might seem natural to interpret SEUT, and the whole Bayesian program, as a step back towards a "metaphysical" view of rationality, imbued with unobservable mental variables and a naïve psychologism.[37] Yet, nothing could be farther from Savage's project. His goal was to obtain a characterization of rational behavior devoid of any psychological contamination by extending *logic* "to bear more fully on uncertainty" (FS, 6) – the kind of logic he referred to being the axiomatic one. It is indeed by focusing on the key property of consistency that the misunderstanding can be avoided. As we know, SEUT's main claim is that if agents' preferences and beliefs are consistent – in the sense specified by the axioms – then these preferences may represented by the expected utility formula. The axioms' role is, therefore, mainly normative, that

---

[36] On the history of choice and decision theory, see Giocoli 2003, 2005 and, for a more technical reconstruction, Moscati 2007. Note that Debreu (1959, 102, fn.1) explicitly acknowledged Savage's influence.

[37] This reaction was quite common at that time: see e.g. Howard Raiffa's recollection of the harsh critiques he received for his conversion to Bayesianism (Fienberg 2008, 141).

is, "to police my own decisions for consistency and, where possible, to make complicated decisions depend on simpler ones." (FS, 20). Axioms, and more generally logic itself, should be viewed as "…a set of criteria by which to detect, with sufficient trouble, any inconsistencies there may be among our beliefs and to derive from the beliefs we already hold such new ones as consistency demands." (ibid.). Far from being a retreat to old-style introspection or psychologism, Savage's theory is first and foremost a normative guide to the formation of consistent beliefs: a rational agent – better, a rational *statistician* – is defined as someone who checks whether her beliefs satisfy the axioms and, in the negative case, revises them according to Bayes's rule.[38]

The requirement of consistency is also the yardstick to evaluate the admission to the axioms' list. Here Savage followed again de Finetti to argue that only those properties whose violation entails an inconsistency should be legitimately listed as postulates (FS, 43). This is another noteworthy statement in view of the economists' subsequent adoption of SEUT as a *description* of agents' behavior. Take for instance Savage's first postulate, P1, namely, that the preference relation be complete and transitive (preference as a simple ordering among acts). Economists before and after *FS* would oscillate between a descriptive and a prescriptive interpretation of the postulate, but Savage was much more clear-cut: P1 is simply a normative rule of consistency (FS, 18-19). Viewing that axiom as a prediction of how real agents evaluate their market alternatives is, to say the least, to betray its original purpose of a rule to be obeyed by rational statisticians.[39]

The third main feature of SEUT is behaviorism. Savage had a technical problem to solve: how to make the decision-maker's probabilistic beliefs operational? In other words, how to "extract" these beliefs from the agent's mind and plug them into an empirically verifiable choice criterion, without making recourse to unacceptable introspection? His answer is in the observable notion of choice: it is choice which reveals the agent's preference for one alternative over another and it is still choice which reveals the agent's belief on the different probability of two events. Introspection is explicitly rejected as an "…especially objectionable [solution], because I think it of great importance that preference, and indifference, between [two events] be determined, at least in principle, by decisions between acts and not by response to introspective questions." (FS, 17). Yet, even with respect to behaviorism, Savage's 1954 views were hardly extreme. Once again, one

---

[38] Cf. de Finetti 1937 [1964], 103: "It is precisely [the] condition of coherence which constitutes the sole principle from which one can deduce the whole calculus of probability…".

[39] The pre-condition of what may legitimately count as a postulate also explains why the famous sure-thing principle, at least in its basic formulation – if an agent would not prefer act *f* to act *g* either knowing that event B obtained or knowing that event not-B obtained, then he does not prefer *f* to *g* (FS, 21) – was *not* included in Savage's axioms' list, but only left as a loose principle: "The sure-thing principle cannot appropriately be accepted as a postulate in the sense that P1 is…" (FS, 22).

gets the impression that his commitment to this aspect of SEUT was motivated more by a technical necessity than by a methodological endorsement.[40]

That choice occupy a central position in Savage's analysis is confirmed by his discussion in Chapter 3 of the possible methods through which beliefs may be elicited. The first method is that of simply asking the agent which of two events he considers more probable. Not surprisingly, Savage rejects it with the same argument used against introspection: "[W]hat could such interrogation have to do with the behavior of a person in the face of uncertainty, except of course for his verbal behavior under interrogation? If the state of mind in question is not capable of manifesting itself in some sort of extraverbal behavior, it is extraneous to our main interest." (FS, 27–8). The second method is called "behavioral interrogation", i.e., the observation of the agent's actual choices under uncertainty. It might surprise the reader to learn Savage was skeptical about this approach as well, but a moment's reflection reveals that the reasons for his skepticism – the method's cost in terms of time, money and effort and the practical impossibility to identify any two really identical choice situations (see FS, 29) – are pretty understandable given the theory's elected field of application, the solution of statistical decision problems. Savage's preference explicitly then went to a third, intermediate method, "a compromise between economy and rigor", which he considered "just the right one", at least in the normative interpretation of his theory (FS, 28). The method entails a kind of intellectual experiment, and consists of, first, questioning the agent not about her generic feelings concerning the probability of two events, but about how she would behave in a very specific choice situation, and, second, using the answer to draw inferences about her probabilistic beliefs. As remarked by Karni & Mongin (2000, 246-7), with his intermediate method Savage was drawing the line between unacceptable purely epistemic statements, on the one side, and acceptable choice statements, on the other. An agent's answers, provided they are expressed as choices among alternatives, constitute admissible evidence to elicit probabilistic beliefs in his theory.

Regardless of the specific method employed, the general procedure by which probabilities are elicited in FS resembles Samuelson's technique in revealed preference theory (Samuelson 1948; Wong 1978). Those subjective probability numbers that did not feature among the initial data do emerge from extending von Neumann's EU theorem to the preferences held over a properly constructed probabilistic space. With his new representation theorem Savage proves that any choice made under conditions of uncertainty can be represented in terms of the EU rule and, in

---

[40] This differs from his commitment to behaviorism in statistics which, as we know, was very strong.

24

particular, as if the agent's decision is guided by the attribution to the states of the world of a well-defined (i.e., numerical) subjective probability. Thus, we may speak of "revealed" subjective probabilities, though this should not be taken to imply that they emerge from a collection of real choice data, rather than from the answers to a merely intellectual experiment.

Still in Chapter 5 Savage applies his theorem to classes of acts of increasing generality (FS, 76 ff.). In its most general version the result is extended to bounded acts, i.e., acts whose utilities are bounded random variables. Adding that the axioms allow the replacement of consequences with utility numbers, so much so that any act $f$ may be re-interpreted as a real-valued random variable of the kind featuring in statistical problems, Savage's SEUT may be summarized as follows: given an event B and two bounded acts $f$ and $g$, $f$ is not preferred to $g$ if and only if, under event B, the expected utility of $f$ is less than that of $g$ (FS, 82). This is the version of the representation theorem which is employed in the rest of *FS* as the cornerstone for the new foundations of statistics.

## §5. The challenge: *FS*'s second part

Armed with the brilliant results of Chapter 5, Savage begins the following chapter declaring that: "With the construction of [SEUT] the theory of decision in the face of uncertainty is, in a sense, complete." (FS, 105). The time was ripe, then, to tackle his real target, statistical decision problems (SDPs). Thanks to Wald, postwar statistics already conceived of the latter in terms of the general structure described above, §1. Take for instance Kenneth Arrow's 1951 *Econometrica* survey, where the typical SDP is described *à la* Wald as follows: first, it is known that one out of a number of different hypotheses about a given situation is true; second, one out of a number of different experiments, or observations, must be performed in order to obtain an outcome that may be interpreted as a random variable whose probability distribution depends on which of the hypotheses is correct; third, on the basis of that outcome, an action has to be selected (Arrow 1951, 409-10). Thus, the difference among the various approaches to statistical decision theory hinged upon the alternative ways to process the additional knowledge, if any at all, which could be obtained via experiments or observations. Remarkably, in the second part of his book Savage stuck to Wald's solution to this issue, despite its being founded upon an objectivistic view of probability. Indeed, one of the key passages in *FS*'s game plan was precisely that of reconciling Wald's objectivist minimax rule with Savage's subjectivist view of probability (see above, §3). The

reconciliation was essential in view of the desired re-interpretation in subjectivist and behaviorist terms of orthodox inference techniques. This explains why Savage spent so much effort to achieve it, albeit with meager results.

Technically speaking, the bridge between the two approaches is granted by the notion of *partition problem*.[41] When expressed in objectivist terms, a partition problem is a special kind of decision problem where the consequence of each act depends on which of several possible probability distributions applies to an observed random variable. Here again Savage follows Wald in arguing that such a setup is so pervasive in statistics that "…no other type of problem is systematically treated in modern statistics." (FS, 120). In subjectivist terms, a partition problem is composed by *i*) a set of possible acts, *ii*) an unobservable random variable whose values uniquely characterize the consequences of each act (i.e., each realization of the random variable allows replacing the act with one of its consequences), and, *iii*) an observation, whose outcome affects the subjective probability attributed to the realizations of the random variable, so much so that the expected utility of each act is modified by the availability of the observation (ibid.). The crucial insight is that, since the unknown states of the world may well play the role of the random variable, any *basic* SDP[42] may be reformulated as a partition problem.

The reconciliation between Savage's and Wald's approaches comes in the opening remarks of Chapter 8, that is, at the very beginning of *FS*'s second part.[43] There the author identifies the two worst drawbacks of existing decision theory – and thus the two main obstacles preventing the full development of behavioral statistics – as, respectively, the inevitable vagueness of the decision-maker's beliefs and the absence of a solution to multi-person decision problems with heterogeneous agents. It is exactly here that we find the key passage introducing the reader to the second part of Savage's project: "From the personalistic point of view, statistics proper can perhaps be defined as the art of dealing with vagueness and interpersonal difference in decision situations." (FS, 154). Surprisingly enough, Savage claims that the solution to both issues lies in the application of Wald's minimax rule.

Savage had already dealt with vagueness in Chapter 4, when discussing the traditional objections against subjective probability. The standard critique is that an agent's beliefs may

---

[41] The use of partitions to represent information had been known in economics and statistics since at least von Neumann's exposition in the *Theory of Games* (see von Neumann & Morgenstern 1944 [1953], 60-73).

[42] What Savage calls observation problems, i.e., problems where the statistician makes an observation and then selects her act according to the outcome (FS, 105). These are a reduced version of general SDPs, which also include, as in Wald, the decision about what observations to make.

[43] The chapter's title is, not casually, "Statistics Proper" (FS, 154).

hardly be assigned a precise quantitative value and that such vagueness is even more serious in the case of a normative theory, like that in *FS*, because it makes virtually impossible to formulate specific behavioral prescriptions (FS, 59). To counter the critique, Savage offers a fervent defense of… Wald's *objectivist* minimax rule! To start with, he underlines that the key ingredients of his theory of subjective probability – i.e., the notions of states, events, acts and consequences – apply as well to Wald's theory, "from which they were in fact derived" (FS, 158). The dividing line between the two approaches is the postulate that the decision-maker's preferences establish a simple ordering among *all* acts. This is an assumption that no frequentist could accept because, when combined with neutral (in terms of probabilistic views) consistency requirements, it would entail, as demonstrated in Chapter 5, the existence of a subjective probability distribution.

Three years after the publication of *FS*, Duncan Luce and Howard Raiffa surveyed the different approaches to decision theory in Chapter 13 of their highly influential *Games and Decisions*. The survey lists the alternatives available in a period when SEUT was still far from dominant. Not surprisingly, Luce & Raiffa rank Savage's approach as intermediate between the two extreme cases of risk and complete ignorance. The former refers to situations, like those in von Neumann's EUT, where a probability distribution over the set of states of nature is either known or taken by the decision-maker as if it were known. The latter refers to cases where the decision-maker has no information about the states of the world (Luce & Raiffa 1957, 277-8). Between the two extremes, Luce & Raiffa identify the vast territory of decision making under *partial* ignorance, i.e., when an agent holds the subjective feeling that a state be more plausible than another. Here the decision-maker is assumed to possess at least some vague information about the true state of the world, the real point being how such vague information can be processed (ibid., 299). It is in this territory that Luce & Raiffa place SEUT.

Keeping in mind Luce & Raiffa's classification, we can better understand Savage's reconciliation of his own theory with Wald's one. He argues that "…in practice the theory of personal probability is supposed to be an *idealization* of one's own standards of behavior; [and] that the idealization is often imperfect in such a way that *an aura of vagueness* is attached to many judgments of personal probability…" (FS, 169, emphasis added). Thus, the minimax criterion may be invoked as a rule of thumb to be applied whenever the notion of "best choice" is "impractical" – impractical for the frequentist statistician because it would require the use of subjective probabilities; impractical for the subjectivist statistician in case of an overwhelming vagueness (FS, 206). Conversely, every time a coherent notion of "best choice" can be pursued – that is, every time a

simple ordering among all available acts may be established – the statistician should follow the consistency axioms and therefore apply SEUT. In short, Savage defends the minimax as a pragmatic rule which may step in whenever we face a failure of SEUT's tight requirements – so tight, indeed, that Savage had to invent the (highly debatable) notion of *small worlds* to warrant them a sort of "protected reserve" (see FS, 9 and 82 ff.).

It goes without saying that the reason mentioned by Savage for the "impracticality" of SEUT, namely, the vagueness of beliefs, would be almost anathema to "real", die-hard Bayesians. Savage's acknowledgment of a complete ignorance environment, and thus of the pragmatic validity of the minimax solution, actually goes even further. He admits that, in many instances, "I do not feel I know my own mind well enough to act definitely on the idea that the expected loss for [act] *g* really is L; but […] of course, feel perfectly confident that [act] *f* cannot result in a loss greater than [minimax value] L*" (FS, 169). Given that almost all SDPs allow for the possibility of a (relatively) inexpensive observation, and given that it may always exist an act conditional on such observation and capable of warranting a (relatively) small maximum loss L*, it turns out that the minimax criterion represents a desirable solution in every SDP where the statistician is unable to *confidently* select the SEU-maximizing act (ibid.). Thus, far from being incompatible with, or alternative to, SEUT, Wald's minimax looks like a welcome complement to Savage's approach to tackle those cases where vagueness prevails. This conclusion is, as we already know, a crucial step in *FS*'s game plan.

The second drawback of existing decision theory, SEUT included, is the inability to solve multi-person decision problems with heterogeneous agents (MPPs henceforth). Confirming his 1951 views, Savage repeats that statistics has to do first and foremost with the behavior of groups, rather than with isolated decision-makers: "Multipersonal considerations constitute much of the essence of what is ordinarily called statistics." (FS, 154; also see FS, 105). Two kinds of MPPs need be distinguished: those arising out of differences in tastes and those arising out of differences in judgment (FS, 155-6). The former can well be tackled by SEUT, which explicitly allows for heterogeneous preferences, but the latter cannot since agents diverge in their subjective probability assessments. The MPPs issue is thus a very serious one. Savage observes that statisticians of the objectivist school have begged it, on account of their denial of any role for personal judgments. Even subjectivists have coped with this difficulty only by self-limiting their analysis, i.e., by postulating that statistics be "largely devoted to exploiting similarities in the judgments of certain classes of people and in seeking devices, notably relevant observation, that

tend to minimize their differences." (FS, 156). Why then did Savage put so much emphasis on such an intractable – and seemingly lateral – issue?

The first obvious answer is that, within a truly behaviorist approach, statistics *must* deal with all kinds of probabilistic judgments as essential ingredients of action, regardless of their possible heterogeneity. Secondly, the emphasis on MPPs may be viewed as another outgrowth of Savage's normative perspective, more specifically of his desire to offer a recipe for rational behavior valid for a group of statisticians, be they a team of experimenters facing a given observation problem or the members of a scientific community summoned to sanction a published result. Finally, but crucially, MPPs have a role to play in *FS*'s game plan, namely, that of allowing the desired reinterpretation of the main results of the British-American school of statistics in subjectivist terms (see above, §3). Savage acknowledges the standard view that subjective probability may at best apply to individual decision-making, but never to MPPs. The scientific method itself is said to rely exclusively on objective notions, the only ones which may effectively be shared inter-personally. It follows that "[t]he theory of probability relevant to science […] ought to be a codification of universally acceptable criteria", so much so that no room is left for "personal difference" in probability judgments (FS, 67). As a counter-argument, Savage claims that "the personalistic view incorporates all the universally acceptable criteria for reasonableness in judgment", although these criteria "do not guarantee agreement on all questions among all honest and freely communicating people" (ibid.). The key insight to uphold this claim is his reinterpretation of a partition problem as a MPP where agents, who differ in their a priori probability distribution on a random variable, share the same set of acts and the same a posteriori (i.e., post-observation) distribution (FS, 122). This reinterpretation provides the basis to demonstrate what is allegedly the most important result of the entire second part of *FS*, namely, that, having properly defined a subjectivist variant of MPP, we may solve it by invoking once again Wald's minimax rule.

In Chapter 10 Savage offers a model of group decision (FS, 172 ff.). Take a group of agents, indexed by *j*. Assume they share the same utility function, but differ in their personal beliefs. It is this difference – explicitly denied by the British-American approach – which captures the subjectivist element of the analysis. The group is called to choose, by acting in concert, an act *f* out of a set F of alternatives. Such a setup – epitomized by the functioning of a jury, but actually "…widespread in science and industry" (FS, 173) – defines a *group decision problem* (GDP). Savage recognizes that while no decision rule may have general validity for every GDP under all

circumstances,[44] rules of thumb do exist which may lead the group to reach an *acceptable compromise*. Indeed, the possibility of using mixed acts – one of the main tools in Wald's box – can be given a natural interpretation in a GDP as a way to avoid the impasse and achieve the required compromise (ibid.). Although the members may differ in their personal judgments, there are plenty of instruments, such as coins, dice, cards, allowing the group to "objectively" mix, in any desired proportion, the individually preferred acts. Hence, it may well be assumed that the set F of available alternatives in any GDP always contains all the possible mixtures of its primary elements. This in turn allows a straightforward application of Wald's criterion, in the novel form a *group minimax rule* (FS, 174). The rule states that, as a compromise solution, the group should adopt the act such that the largest personal loss faced by any member of the group is as small as possible. That is, the act *f ′* should be chosen such that, for any agent j,  $max_j L(f ′; j) = L^* = min_f max_j L(f ; j)$.[45]

As we know, the notion of minimax as an "acceptable compromise" was already in Savage 1951 (see above, §2). Yet, it cannot be expected that the group minimax rule will or should be accepted by every member of the group. As in the standard, single-agent case, what may be said is just that, *if* the minimax loss L* is small, the group may find it reasonable to follow the rule and select act *f ′* because none of its members should consider this choice unacceptable nor could suggest any alternative capable of limiting every member's loss to, at most, *L\**. Moreover, as he also noted in the 1951 review, a symmetry argument may be invoked, whereby in case of symmetric GDP the minimax rule would impose exactly the same loss L* on every member: an element of fairness which, according to Savage, would further augment the rule's appeal in GDPs (ibid.). In short, far from being intractable, multi-person statistical problems, explicitly formulated in behaviorist terms and explicitly allowing for the heterogeneity of the members' subjective beliefs, may find a pragmatic solution in the minimax criterion. The appeal to Wald's pioneering contribution allows Savage to also get rid of the second major drawback of behavioral statistics.

## §6. The fiasco: FS's abrupt ending

The continuity between the second part of *FS* and Savage's review of Wald 1950 should by now be apparent. This reinforces our point that the 1954 endeavor was no more, and no less, than

---

[44] Nowhere in *FS* Savage mentions Arrow's social choice theory or the search for "best" voting procedures.
[45] The formula is identical with the usual minimax rule, but its elements now have a different meaning. Also recall that the idea of the minimax as an "acceptable compromise" was already in Savage 1951: see above, §2.

Savage's try at completing Wald's project, with the only, though crucial, addition of the SEUT theorem. The continuity is confirmed by Savage's critical attitude towards the minimax rule, and even more towards its "group" version – again a feature that went almost unchanged from the 1951 paper to the 1954 book. Yet, within the theoretical edifice of *FS*, the importance of the group minimax rule does *not* hinge upon its practical acceptability as a decision-making criterion to be used in real MPPs. What the rule brings home for him is the desired synthesis between the subjectivist view of probability – here captured by the members' differences in their beliefs about events – and Wald's behaviorist approach – here epitomized by the asserted ubiquity of, explicitly behavioral, GDPs and MPPs in modern statistics. Hence, the achievement of *FS*'s Chapter 10, weak and debatable as it may be, is pivotal in Savage's game plan.

The rest of the book validates my assertion. The final chapters contain Savage's attempt to reinterpret the three most important inference techniques – respectively, point estimation (Ch.15), the testing of hypothesis (Ch.16) and interval estimation (Ch.17) – in behavioral terms[46] by applying Wald's objectivist minimax rule. Granted such a reinterpretation, the *FS*'s game plan would then require to reformulate it in subjectivist terms, exploiting the newly established group minimax rule. Finally, for those special decision problems where vagueness is *not* a problem, the minimax rule would be replaced by the SEU rule, i.e., by the criterion allowing statisticians to select not just an "acceptable compromise" solution, but the "best" solution. This would accomplish the game plan.

Unfortunately, Savage's ingenious strategy ended in a fiasco. While he partially succeeded in giving a behavioral reinterpretation of point estimation (see FS, 229 ff.),[47] doing the same for interval estimation proved impossible. In a behaviorist framework, the statistician's problem is to make her "best" choice of an unknown parameter. Yet, following the orthodox doctrine of "accuracy estimation", which only allows to know how "good" such a "best" choice is with respect to an interval of values, is of no avail for such a problem and "…could at most satisfy [the statistician's] idle curiosity…" (FS, 257-8). In other words, when the decision-maker is called to choose an act, i.e., to produce a point estimate, interval estimation is simply nonsensical. Thus, regardless of the fact that "most leaders in statistical theory have a long-standing enthusiasm for

---

[46] As we know, this reinterpretation had been tried by Wald himself (see above, §1). The topic was a hot one in 1950s decision theory, the peculiarity of Savage's attempt being its subjectivist foundation. For a compact statement of the whole issue, see Luce & Raiffa 1957, 309-24.

[47] The trick being that of reinterpreting point estimation as a decision problem to be solved via Wald's minimax rule, assigning a zero loss to correct estimates.

[interval estimation]", the technique may at most be used as a rough and informal device, while its usefulness for statistical decision-making is nil (FS, 260-1).

The impossibility to offer a behaviorist reinterpretation of so crucial a tool as interval estimation was a fatal blow to Savage's project of reconstructing mainstream statistics in behavioral and subjectivist terms. The simple truth emerging from his analysis was that behaviorism, subjectivism and orthodox inference techniques could *not* be reconciled. The collapse of *FS*'s game plan is testified by the somehow abrupt conclusion of the book, which features no concluding remarks, no summary of the main results, no recapitulation of the analytical trajectory followed thus far. It really looks like the author had suddenly realized that his entire program was doomed and thus decided to quit the work. As we know, in the preface to the second edition Savage invoked Freud to explain how the book's project managed to survive in the face of such a dire conclusion (FS, iv). Fortunately, it took him only little time to react against the rout and become the most ardent supporter of a wholly new approach to statistics.

In a 1961 re-assessment of his book Savage looks already aware of the revolution *FS* helped to trigger and wholeheartedly praises it.[48] Far from orthodox methods, inference is defined by Bayesians as "changes in opinion induced by evidence on the application of Bayes's theorem" (Savage 1961 [1964], 178). The frequentist approach is chastised as being *more* subjective than the personalistic one on account of the unconstrained nature of experiment design (ibid.). As an attempt to get rid of this subjective element of traditional statistics, Wald's minimax criterion is deemed a failure – this despite "the strongest apology for the rule" given in *FS* (ibid., 179). The point is that subjectiveness is an unavoidable element of statistical analysis. By omitting empty questions about what "nice properties" a certain statistical procedure should have, but rather asking the decision-maker to choose between alternative procedures subject to strict consistency conditions (ibid., 179-80), the Bayesian approach ends up being more objective – i.e., *more scientific* – than the frequentist one, in that it imposes greater order on the subjective element of statistical decision-making (ibid., 181).

What did Savage retain of his old game plan? The answer lies in his remark that: "One of the most valuable outgrowths of the frequentist movement is the behavioralistic (or one might say *economic-theoretic*) approach to statistical problems" (ibid., 177, emphasis added). Jerzy Neyman and, above all, Abraham Wald deserve credit for having launched the "economic analysis of statistical problems", what they called "inductive behavior". Thus, the truly valuable part of Wald's

---

[48] Note that, while the term "Bayesian" never featured in FS, at the turn of the 1960s its use in statistical jargon had already become commonplace. See Feinberg 2003, 16-17.

project – and, as a consequence, also of Savage's own one which aimed at fulfilling it – is its behaviorist and decision-theoretic character, first and foremost its conception of the statistician as an economic man. Under this respect, *FS* was indeed successful, in that it crowned Wald's efforts by developing a theory for decision-making under uncertainty which could be prescribed as a rational criterion for statistical problems. Yes, the other (possibly, the main) part of Wald's project, and of *FS* as well, namely, the reconciliation of inductive behavior with orthodox inference via the minimax rule, was a failure. But, no, *FS* had not been written in vain, as it contained a new, powerful tool, SEUT, for rebuilding statistics as a wholly behavioral discipline. In this sense, Wald's project *had* been accomplished, though by 1961 that rebuilding was still far from complete. The 1954 book might be said to have honored its title, by laying solid *foundations* of statistics – though not of the statistics Wald and the early Savage were thinking of, but of a brand new kind of it.

## §7. The puzzle: why in economics?

An instance of unintended consequences, the spread of Bayesian statistics helped popularize Savage's SEUT. In itself, the theory did represent a remarkable achievement, and a source of perennial glory for its author. By his masterful use of the axiomatic method, Savage elegantly succeeded to fulfill the economists' almost century-old pursuit of a general criterion for rational decision-making. Yet, a big historiographic puzzle remains: how could SEUT so quickly and so wholeheartedly be endorsed by neoclassical economists? How to explain its success in economics?

   To understand that a puzzle it is, consider what we have learned so far. First, Savage's book was not addressed to economists, but to statisticians. Second, the book's overall project was a remarkable fiasco as it ended up demonstrating exactly the opposite of what it was intended to. Third, *FS*'s most significant contribution, SEUT, was an explicitly normative theory aimed at guiding the statisticians' work, not a positive one describing how real economic agents behave. Two further circumstances should be taken into account. On the one hand, Bayesian statistics, while spreading after the *FS* at a considerable pace in a small number of academic strongholds,[49] was, and still is, a minority approach. As a matter of fact, most of the discipline has kept using traditional inference methods, though recognizing their foundational weaknesses, while

---

[49] For an overview of the early years of Bayesianism in terms of academic success, see Feinberg 2003.

Bayesianism itself has been constantly and severely criticized ever since.[50] Hence, Savage's revolution was, at most, only a partial success. In no sense any part of it could be considered mainstream by the time neoclassical economics embraced SEUT. On the other hand, SEUT itself was hardly the only available option in mid-1950s decision theory: as the list of criteria in Chapter 13 of the influential Luce & Raiffa 1957 makes clear, several alternatives did exist. Some had been even devised with the specific goal of encompassing situations of complete ignorance (i.e., Savage's vagueness) as a way out from one of most vexed issues of theoretical economics (viz., Knightian or Keynesian uncertainty). Moreover, all these criteria commended themselves as lying upon rigorous axiomatic foundations – indeed, following Chernoff 1954, Luce & Raiffa's goal in their 1957 survey had precisely been that of offering a general axiomatic characterization of all available decision rules, including those working under complete ignorance.[51]

In view of all these circumstances, I reiterate the claim that explaining the postwar triumph of SEUT in neoclassical economics really is a historiographic puzzle. To reword Marschak's 1946 dictum, why "to be an economic agent implies being a *Bayesian* statistician"? How, when and why did the agents populating economic models come to be modeled as rigorous followers of Savage's rationality prescriptions? As I said in the Introduction, I have no proper answer to this question. The paper's goal was, as announced, just that of showing that the question itself is an interesting one for the history of postwar microeconomics. Yet, a couple of conjectures, whose consistency awaits further validation, may be offered here as concluding remarks.


## §8. The conjectures: consistency and business schools


My first tentative explanation hinges upon a history of economic analysis argument. The catchword is "consistency", that is to say, the circumstance that, as I have showed at length elsewhere,[52] the history of the 20th-century neoclassical notion of rational behavior may be told in terms of the progressive replacement of the traditional maximization approach, where rationality consisted of the reasoned pursuit of one's own self interest, in favor of the consistency approach,

---

[50] The literature is huge. The controversy about Savage and Bayesianism started immediately (Kadane 2001, 5). For a modern rendition and rebuttal of the critiques, see Gelman 2008 (paper and discussion), and, on more philosophical grounds, Gelman & Shalizi 2011. For a detailed critique of SEUT, also see Shafer 1986 and the enclosed discussion.
[51] Note that, as a testimony to *FS*'s achievement, Luce & Raiffa adopted Savage's state-act-consequence setup in their analysis. For an assessment of Luce & Raiffa's survey vis-a-vis Savage's contribution, see Giocoli 2005a.
[52] See Giocoli 2003; 2005.

where rationality means the satisfaction of a formal requirement of consistency and where the notion of agency becomes an all-purpose concept, valid for real individuals as well as for groups or machines. The process of distilling the formal essence of the notion of rationality in order to make it as general and rigorous as possible found its milestones in the works of authors such as Paul Samuelson, John von Neumann, Kenneth Arrow and Gerard Debreu, and culminated in Savage's SEUT.[53] Thus, we might say that accepting to portray the *homo economicus* as a Bayesian statistician was just the inevitable, and almost "automatic", conclusion of an intellectual journey neoclassical economists had begun long before.

Modern historiography has showed that many top-notch theoretical economists working in the postwar era had been happy to embrace a new standard of representation of subjective variables in economic models based upon a (variable) combination of axiomatic and behavioral methods. At the turn of the 1960s, the standard had already been successfully applied to the modeling of the agents' *preferences* and, therefore, of their choices under conditions of either certainty or risk. What remained to be done was to apply the same standard to the other source of subjectivism in economic behavior, namely, the agent's *beliefs*, which mattered in case of choice under uncertainty. From this point of view, Savage's SEUT was just the right theory at the right time. It may be thus surmised that what made it so attractive for postwar economists was the demonstration that even the agent's beliefs could be subjected to tight rationality (viz., consistency) constraints, so much so that they could satisfy the new rigorous standard of theorizing and thus become proper ingredients of formalized economic models. In short, SEUT enabled neoclassical microeconomics to boast "mission accomplished", at least as far as parametric rationality was concerned.[54]

My second explanation is based on a wholly different argument, involving institutional and sociological factors. The catchword is "business schools", that is to say, the conjecture that economists came to appreciate Bayesian decision theory via their increasing involvement in major MBA programs, where the new approach was explicitly taught.

In a recent paper, Marion Fourcade and Rakesh Khurana have reconstructed the 20th-century evolution of teaching in US business schools, from their being dominated by mere practitioners devoid of any academic legitimacy to their becoming the largest employers of disciplinary trained

---

[53] Anscombe & Aumann 1963 might also be included to the list.

[54] Even with respect to strategic rationality, it may be argued that the popularity of the Bayesian approach was instrumental, via the work of Reinhard Selten, John Harsanyi and Robert Aumann, in pushing economists towards Nash's non-cooperative approach after its substantial neglect in the 1950s and 1960s: see Giocoli 2005; 2009.

social scientists. The new era for business schools began after WWII, when the rise of modern management science met the demand for a more scientific kind of managerial capabilities originating from the postwar conception of corporate "command and control" stemming from the war economy experience. In the new conception, "managers were increasingly described as 'systems designers', 'information processors', and 'programmers' involved in regulating the interfaces between the organization and its environment and bringing rational analysis to bear on a firm's problems" (Fourcade & Khurana 2011, 18). What Fourcade & Khurana call the "scientization" of business disciplines was triggered by a small group of scholars and academic officers who, working in newly established programs like Carnegie's Graduate School of Industrial Administration (GSIA), managed to transplant into business teaching the decision-making techniques they had crafted doing operations research on behalf of the military during the war, thus paving the way to the new "scientific approach" to the managing of US corporations (ibid., 5).

The GSIA did represent a radical departure from existing practices. Business education was seen at Carnegie "as an extension of the social sciences, rooted in quantitative analysis and the behavioral disciplines." (ibid., 15). With the help of few funding patrons, chief among them the Ford Foundation, the landscape of US business schools underwent in a few years a major transformation, with a whole generation of economists, statisticians and other social scientists, all with formal PhD training and serious research interests, replacing the old instructors who frequently lacked academic credentials. Thus, "MBA courses were to be taught by disciplinary trained scholars steeped in the latest quantitative methods to study various phenomena of business. […] Business schools were to restructure their own doctoral programs by grounding students in the basic social science disciplines and direct their research toward more fundamental theory" (ibid., 20).

Crucially for our story, Fourcade & Khurana argue that the fertilization was not unidirectional: while the new recruits brought the rigor of scientific methods to the teaching of would-be managers, the encounter with real world business problems affected, and sometimes re-directed, their own research agenda, with deep implications for the evolution of contemporary social science and, in particular, of post-1960s economics. As the authors put it, business schools have become increasingly intertwined with the progress of economics over the second half of the 20<sup>th</sup> century, "as *both recipients and agents* of scientific and intellectual change" (ibid., 30, emphasis

added).[55] Of special import in this cross-fertilization was the Chicago Graduate School of Business (GSB), whose rapid ascendance in MBA rankings marked at the same time the triumph of mainstream price theory in business teaching and a major transformation in the subject matter and analytical orientations of neoclassical economics itself. While Fourcade & Khurana mention topics such as firm theory, principal/agent model, rational expectations theory and financial economics, I surmise that a similar argument may be constructed for Bayesian decision theory.

The clues are indeed significant. All the three business schools the authors single out in their narrative, Carnegie, Chicago and Stanford, featured in the (rather short) list of early Bayesian strongholds in US statistics community. In particular, the dean of Chicago GSB was that Allen Wallis whom we already met as the first recipient of Captain Schuyler's problem of sequential testing (see above, §1). Chicago was also the place where, at the recently founded Department of Statistics, "an amazing confluence of statisticians" materialized around Savage's scholarship (Feinberg 2003, 16). The list included faculty members Harry Roberts and David Wallace, graduate students Morris de Groot (who later helped establish Carnegie Department of Statistics) and Roy Radner, visiting scholars Dennis Lindley, Frederick Mosteller (later the first chair of Harvard Department of Statistics) and John Pratt (who also became a Harvard professor of statistics). Thus, there seems to be a substantial degree of overlap between Fourcade & Khurana's story and the spread of Bayesian thinking in the US.

Another clue comes from the events of statistical teaching at Harvard Business School. There, independently of Wald and Savage, Howard Raiffa and Robert Schlaifer developed their own version of Bayesian analysis. In a recent interview, Raiffa has told the story of his "almost religious conversion" to Bayesianism, after learning about the sure-thing-principle in Chernoff 1954 and understanding that "push[ing] the axiomatics […] it made sense to assign a prior probability distribution over the states of the problem and maximize expected utility" (see Raiffa interview in Feinberg 2008, 141). Note that Raiffa was neither a statistician nor an economist – he knew nothing about Savage and had never taken a course in economics – but his background in game theory and operations research made it natural to him to think of statistics in terms of decision problems. Hired in 1957 jointly by the newly established Harvard Department of Statistics and by Harvard Business School (HBS), he spent most of his time working in the latter. It was at HBS that Raiffa met Schlaifer, another accidental statistician who, trained as a historian but unexpectedly

---

[55] That business schools have become formidable players within economic science is attested by the list of Nobel Prizes awarded to business school economists, seven of which were granted to GSIA scholars: see Fourcade & Khurana 2011, 4, 21.

assigned to teaching basic statistics to HBS students, had quickly immersed himself in the frequentist literature, only to conclude that "standard statistical pedagogy did not address the main problem of a businessman: how to make decisions under uncertainty". Thus, "he threw away the books and invented Bayesian decision theory from scratch" (Raiffa interview, in ibid., 143-4).[56]

Both Schlaifer and Raiffa believed that frequentist inference methods were simply unsuited for making real business decisions: when a manager had to make a choice, traditional Fisher or Neyman/Pearson techniques were of little help, as she often had neither objective probabilities at her disposal nor the possibility to avail herself of a properly constructed sample. Yet, what she could do was to apply her experience to formulate an informed probabilistic belief and then exploit any further information to consistently revise her a priori belief. In short, Bayesian techniques looked like the most proper set of decision tools a manager could employ in her day-to-day business. The striking similarity with Wald's original intuition for tackling sequential testing problems (see above, §1) should not go unnoticed.

Soon Raiffa and Schlaifer were to join forces and teach together an elective course in statistics at HBS. It was as a pedagogical move within this course that Raiffa transformed the game tree employed in extensive form games into his famous decision tree, where probability distributions, subjectively assessed by the decision-maker, are attached to chance nodes – a landmark tool for modern decision theory. And it was still in support of their shared course that the two published in 1961 *Applied Statistical Decision Theory* (Raiffa & Schlaifer 1961). The book, which featured Raiffa's decision tree as the prototype of statistical decision problems, was to become the main reference for the teaching of Bayesian decision theory in US business schools. In the mid-1960s, Schlaifer's course became compulsory for the hundreds of HBS students in the Managerial Economics program. As Raiffa put it, that course was "like an existence theorem: it demonstrated that decision analysis was relevant and teachable to future managers" (Raiffa interview, in Feinberg 2008, 148). Even more relevant were two further initiatives launched by the duo. First, from 1961 to 1964 they ran the HBS Decision Under Uncertainty weekly seminar, which offered a stage for the presentation of advances in decision theory. The seminar was a real turning point in

---

[56] The outcome of this dissatisfaction and reinvention was Schlaifer 1959, a book that was praised by Savage in the preface to the second edition of *FS* (see FS, v). In the Introduction, Schlaifer defined his book "a nonmathematical introduction to the logical analysis of practical business problems in which a decision must be reached under uncertainty. […] a really complete theory to guide the making of managerial decisions" (Schlaifer 1959, v-vi). The exposition was "based on the modern theory of utility and […] the 'personal' definition of probability. […] when the consequences of various possible courses of action depend on some unpredictable event, the *practical* way of choosing the 'best' act is to assign values to consequences and probabilities to events and then to select the act with the highest expected value." (ibid., v, original emphasis). I reiterate the point that these words, which might as well be written in the *FS*, were put on paper by Schlaifer without any knowledge whatsoever of Savage's theory.

that it testified that Raiffa and Schlaifer were starting to see themselves as fully-fledged decision theorists, rather than "just" statisticians. This new self-characterization of their own work enormously broadened the field for applying Bayesian analysis. Second, in 1960-61 Raiffa organized a 11-month program for 40 professors teaching in management schools who felt the need to learn more mathematics. The measure of the program's success is given by the circumstance that future deans of Stanford, Harvard and Northwestern business schools were all graduates of the program. In it, Raiffa taught subjectivist statistics with a heavy decision theory orientation, thereby ensuring that "the gospel radiated outward in schools of management" (Raiffa interview, in ibid., 147). Remarkably, the program was financed by the Ford Foundation, that is, by the same patron which, by generously supporting both the GSIA and the GSB, as well as other important MBA programs, such as Stanford's,[57] favored the "scientization" revolution in US business schools narrated by Fourcade & Khurana.

Summing up, we may say that Bayesianism was at least "in the air" in the Department of Statistics of places like Carnegie, Chicago and Stanford which also happened to host the leading business schools of the new, "scientific" kind. Harvard Business School may also be added to the list, thanks to the teaching of scholars like Raiffa, Schlaifer and, later, John Pratt. While of course the conjecture awaits further confirmation (say, by examining MBA curricula and reading lists, or by investigating personal relations between economists and Bayesian statisticians in those very universities), it may provisionally be concluded that the dramatic change underwent by management teaching in the second half of the 20th century has probably played a major role in the spreading of Bayesianism in general, and of SEUT in particular, within contemporary economics. To reword Marschak's dictum once again, it was not that *homo economicus* directly became a Bayesian statistician, but, possibly, that, first, *homo managerialis* was taught to behave that way, and, then, that economic agents came to be modeled as Bayesian corporate managers.

---

[57] As far as Stanford is concerned, it may be added that Robert Wilson, the giant of contemporary game and decision theory who has been at Stanford Business School since 1964, learned his Bayesian skills under Raiffa at HBS. The list of Wilson's students and colleagues at Stanford reads like a who's who in the application of Bayesian methods to modern economics. See Giocoli 2009 for more details.

**References**

ANSCOMBE F. AND AUMANN R.A. 1963, "A definition of subjective probability", *Annals of Mathematical Statistics*, 34 (1), 199-205.

ARNOLD K.J., GIRSCHICK M.A. & SAVAGE L.J. 1947 "Abandoning an experiment prior to completion", in: Eisenhart C., Hastay M.W. & Wallis A.W. (eds.), *Techniques of Statistical Analysis*, McGraw-Hill, New York, 1947, 353-362.

ARROW K. 1951, "Alternative approaches to the theory of choice in risk-taking situations", *Econometrica*, 19 (4), 404-437.

CHERNOFF H. 1954, "Rational selection of decision functions", *Econometrica*, 22, 422-443.

DEBREU G. 1959, *Theory of Value*, New York: Wiley.

DE FINETTI B. 1937, "Foresight: its logical laws, its subjective sources", in: Kyburg H.E. and Smokler H.E. (eds), *Studies in Subjective Probability*, Wiley, New York, 1964, 93-158.

FERGUSON T.S. 1976, "Development of the decision model", in: Owen D.B. (ed.), *On the History of Statistics and Probability*, Dekker, New York, 1976, 335-346.

FIENBERG S.E. 2003, "When did Bayesian inference become 'Bayesian'?", *Bayesian Analysis*, 1 (1), 1-41.

FIENBERG S.E. 2008, "The Early Statistical Years: 1947-1967. A Conversation with Howard Raiffa", *Statistical Science*, 23 (1), 136-149.

FISHER R.A. 1955, "Statistical methods and scientific induction", *Journal of the Royal Statistical Society, Series B (Methodological)*, 17 (1), 69-78.

FOURCADE M. & KHURANA R. 2011, "From Social Control to Financial Economics: The Linked Ecologies of Economics and Business in Twentieth Century America", *Harvard Business School Working Paper*, No. 11-071.

FRIEDMAN M. & SAVAGE L.J. 1947, "Planning experiments seeking maxima", in: Eisenhart C., Hastay M.W. & Wallis A. (eds.), *Techniques of Statistical Analysis*, McGraw-Hill, New York, 1947, 363-372.

GELMAN A. 2008, "Objections to Bayesian statistics", *Bayesian Analysis*, 3 (3), 445-450.

GELMAN A. & SHALIZI C. 2011, "Philosophy and the practice of Bayesian statistics in the social sciences", in: Kincaid H. (ed.), *Oxford Handbook of the Philosophy of the Social Sciences*, OUP, Oxford, forthcoming.

GILBOA I. 2009, *Theory of Decision under Uncertainty*, CUP, Cambridge.

G<small>ILBOA</small> I., P<small>OSTLEWAITE</small> A. & S<small>CHMEIDLER</small> D. 2004, "Rationality of Belief, or: Why Savage's Axioms are neither Sufficient nor Necessary for Rationality", *Synthese*, forthcoming.

G<small>ILBOA</small> I., P<small>OSTLEWAITE</small> A. & S<small>CHMEIDLER</small> D. 2009, "Is it always rational to satisfy Savage's axioms?", *Economics and Philosophy*, 25, 285-296.

G<small>IOCOLI</small> N. 2003, *Modeling Rational Agents: from Interwar Economics to Early Modern Game Theory*, Elgar, Cheltenham.

G<small>IOCOLI</small> N. 2005, "Modeling rational agents: the consistency view of rationality and the changing image of neoclassical economics", *Cahiers d'économie politique*, 49, 177-208.

G<small>IOCOLI</small> 2005<small>A</small>, "Savage vs. Wald: was Bayesian decision theory the only available alternative for postwar economics?", papers.ssrn.com/sol3/papers.cfm?abstract_id=910916 .

G<small>IOCOLI</small> N. 2006, "Do prudent agents play lotteries? Von Neumann's contribution to the theory of rational behavior", *Journal of the History of Economic Thought*, 28 (1), 95-109.

G<small>IOCOLI</small> N. 2009, "Three alternative (?) stories on the late 20th-century rise of game theory", *Studi e Note di Economia*, 14 (2), 187-210.

K<small>ADANE</small> J. 2001, "Jimmie Savage: an introduction", *ISBA Bulletin*, March, 5-6.

K<small>ARNI</small> E. & M<small>ONGIN</small> P. 2000, "On the determination of subjective probability by choices", Management Science, 46 (2), 233-248.

K<small>LEIN</small> J.L. 2000, "Economics for a client: the case of statistical quality control and sequential analysis", in: Backhouse R.E. & Biddle J. (eds.), Toward a History of Applied Economics, Duke University Press, Durham, 2000, 27-69.

L<small>EONARD</small> R. 2010, *Von Neumann, Morgenstern and the Creation of Game Theory*, CUP, Cambridge.

L<small>INDLEY</small> D.V. 1980, "L.J. Savage – His work in probability and statistics", *Annals of Statistics*, 8 (1), 1-24.

L<small>UCE</small> R.D. <small>AND</small> R<small>AIFFA</small> H. 1957, *Games and Decisions*, New York: Wiley.

M<small>ARSCHAK</small> J. 1946, "Neumann's and Morgenstern's new approach to static economics", *Journal of Political Economy*, 54 (2), 97-115.

M<small>ENGER</small> K. 1952, "T<small>HE FORMATIVE YEARS OF</small> A<small>BRAHAM</small> W<small>ALD AND HIS WORK IN GEOMETRY</small>", *Annals of Mathematical Statistics*, 23 (1), 14-20.

M<small>ONGIN</small> P. 2011, "What the decision theorist could tell the Bayesian philosopher", Lecture presented at the 14th International Congress of Logic, Methodology and Philosophy of Science (CLMPS), July 19–26, 2011, Nancy (France).

MOSCATI I. 2007, "History of Consumer Demand Theory 1871-1971: A Neo-Kantian Rational Reconstruction", *European Journal of the History of Economic Thought*, 14 (1), 119-156.

MYERSON R.B. 1991, *Game Theory. Analysis of Conflict*, Cambridge, Mass.: Harvard UP.

NEYMAN J. 1938, "L'estimation statistique, traitée comme un problème classique de probabilité", *Actualités Scientifiques et Industrielles*, 739, 25-57..

NEYMAN J. & PEARSON E.S. 1933, "The testing of statistical hypothesis in relation to probabilities *a priori*", *Mathematical Proceedings of the Cambridge. Philosophical Society*, 29 (4), 492-510.

RAIFFA H. & SCHLAIFER R. 1961, *Applied Statistical Decision Theory*, Division of Research, Graduate School of Business Administration, Harvard University, Boston.

SAMUELSON P.A. 1948, "Consumption theory in terms of revealed preference", *Economica*, November, 243-253.

SAVAGE L.J. 1949, "The role of personal probability in statistics", Report of the Boulder Meeting, August 29 – September 2, 1949, *Econometrica*, 18 (2), 183-184 (abstract).

SAVAGE L.J. 1951, "The theory of statistical decision", *Journal of the American Statistical Association*, 46 (253), 55-67.

SAVAGE L.J. 1954, *The Foundations of Statistics*, Wiley, New York, 2$^{nd}$ edn. 1972.

SAVAGE L.J. 1961, "*The Foundations of Statistics* reconsidered", in: Kyburg H.E. & Smokler H.E. (eds), *Studies in Subjective Probability*, Wiley, New York, 1964, 173-188.

SAVAGE L.J. 1962, "Subjective probability and statistical practice", in: Savage L.J. et al. (eds.), *The Foundations of Statistical Inference: A Discussion,* Methuen, London, 1962, 9-35.

SCHLAIFER R. 1959, *Probability and Statistics for Business Decisions*, McGraw-Hill, New York.

SHAFER G 1986, "Savage revisited", *Statistical Science*, 1 (4), 463-485.

VON NEUMANN J. 1928, "On the Theory of Games of Strategy", in: Luce D. & Tucker A.W. (eds.), *Contributions to the Theory of Games*, Vol. IV, PUP, Princeton, 1959, 13-42.

VON NEUMANN J. AND MORGENSTERN O. 1944, *Theory of Games and Economic Behavior*, PUP, Princeton, 3$^{rd}$ edn. 1953.

VON PLATO J. 1994, *Creating Modern Probability*, CUP, Cambridge.

WAKKER P. 2010, *Prospect Theory. For Risk and Ambiguity*, CUP, Cambridge.

WALD A. 1939, "Contributions to the theory of statistical estimation and testing hypotheses", *Annals of Mathematical Statistics*, 10, 299-326.

WALD A. 1945, "Sequential tests of statistical hypotheses", *Annals of Mathematical Statistics*, 1945, 16 (2), 117-186.

W<small>ALD</small> A. 1945<small>A</small>, "Statistical decision functions which minimize the maximum risk", *Annals of Mathematics*, 46 (2), 265-280.

W<small>ALD</small> A. 1947, "An essentially complete class of admissible decision functions", *Annals of Mathematical Statistics*, 18 (4), 549-555.

W<small>ALD</small> A. 1950, *Statistical Decision Functions*, Wiley, New York.

W<small>ALD</small> A. 1950<small>A</small>, "Basic ideas of a general theory of statistical decision rules", in: Wald A., *Selected Papers in Statistics and Probability*, McGraw-Hill, New York, 1955, 656-668.

W<small>ALLIS</small> A.W. 1980, "The Statistical Research Group, 1942-1945", *Journal of the American Statistical Association*,75, 320-330.

W<small>OLFOWITZ</small> J. 1952, "Abraham Wald (1902-1950)", *Annals of Mathematical Statistics*, 23 (1), 1-13.

W<small>ONG</small> S. 1978, *The Foundations of Paul Samuelson's Revealed Preference Theory*, Routledge, London.

Z<small>ILIAK</small> S.T. 2008, "Guinessometrics: the economic foundation of 'Student's' *t*", *Journal of Economic Perspectives*, 22 (4), 199-216.

Z<small>ILIAK</small> S.T. & M<small>CCLOSKEY</small> D.N. 2008, *The Cult of Statistical Significance*, The University of Michigan Press, Ann Arbor.