Economics and Philosophy (2019), **35**, 423–459
doi:10.1017/S0266267118000469

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The Allais paradox: what it became, what it really was, what it now suggests to us

Philippe Mongin

GREGHEC, 1 rue de la Libération, F-78350 Jouy-en-Josas, France
Email: mongin@greg-hec.com

(Received 11 June 2018; revised 24 August 2018; accepted 27 August 2018; first published online 30 January 2019)

Abstract

Whereas many others have scrutinized the Allais paradox from a theoretical angle, we study the paradox from an historical perspective and link our findings to a suggestion as to how decision theory could make use of it today. We emphasize that Allais proposed the paradox as *a normative argument*, concerned with 'the rational man' and not the 'real man', to use his words. Moreover, and more subtly, we argue that Allais had an unusual sense of the normative, being concerned not so much with the *rationality of choices* as with the *rationality of the agent as a person*. These two claims are buttressed by a detailed investigation – the first of its kind – of the 1952 Paris conference on risk, which set the context for the invention of the paradox, and a detailed reconstruction – also the first of its kind – of Allais's specific normative argument from his numerous but allusive writings. The paper contrasts these interpretations of what the paradox historically represented, with how it generally came to function within decision theory from the late 1970s onwards: that is, as an empirical refutation of the expected utility hypothesis, and more specifically of the condition of von Neumann–Morgenstern independence that underlies that hypothesis. While not denying that this use of the paradox was fruitful in many ways, we propose another use that turns out also to be compatible with an experimental perspective. Following Allais's hints on 'the experimental definition of rationality', this new use consists in *letting the experiment itself speak of the rationality or otherwise of the subjects*. In the 1970s, a short sequence of papers inspired by Allais implemented original ways of eliciting the reasons guiding the subjects' choices, and claimed to be able to draw relevant normative consequences from this information. We end by reviewing this forgotten experimental avenue not simply historically, but with a view to recommending it for possible use by decision theorists today.

Keywords: Allais paradox; decision theory; expected utility theory; experimental economics; positive vs normative; rationality; 1952 Paris conference; Allais; von Neumann and Morgenstern; Samuelson; Savage

JEL Classification: B21; B31; B41; C91; D8

1. Introduction

One does not need to study decision theory for very long before stumbling across the *Allais paradox*, a neat finding by the French economist Maurice Allais that

perhaps had a greater impact than any of his wider theoretical constructions.[1] The paradox targets the classical hypothesis that decision under risk conforms to the rule of expected utility (EU). This high theoretical stake explains the hesitancy with which it was initially received. Discovered in 1952 and published in 1953, hence not long after von Neumann and Morgenstern (VNM) had first axiomatized the EU rule, the paradox was initially thought little more than an ingenious curiosity. Only after a lengthy delay did it begin to attract sustained interest, belatedly being accepted as a genuine *paradox*, i.e. as shaking confidence in seemingly uncontroversial ideas. Along with related counter-examples, it acquired a central role in motivating decision theory to shift from EU to alternative hypotheses – such as subjectively weighted utility, prospect theory, generalized expected utility, rank-dependent utility theory and others.

Whereas many others have scrutinized the Allais paradox from a theoretical and technical angle, our intention here is to study it historically, and then expand on our findings in the conceptual and methodological directions. A basic claim of the paper is that the paradox is a normative argument, being concerned with the 'rational man', not the 'real man', as Allais himself made clear. We offer this claim as a definitive interpretation of the paradox as regards its *historical origin*, thus rejecting the by-now rather common view that Allais meant it to be an empirical argument. We also offer this claim as an alternative interpretation of the paradox regarding its *possible use* by today's decision theory. When the paradox eventually moved from the periphery to the centre of decision theory, roughly by the late 1970s or early 1980s, it was taken to be a coarse but improvable schema of empirical refutation of the EU hypothesis, and more specifically of the crucial VNM independence condition underlying that hypothesis. With some relevant exceptions, to be discussed below, this empirical interpretation has prevailed and is still the received one today.[2] Although it has some serious advantages, which we will duly record, our objective is really to promote the forgotten initial view once again.

To do so, we expand on Allais's own argument, claiming that the paradox is concerned not so much with *the rationality of choices* as with *the rationality of the agent*. This claim might interest even those decision theorists who have recognized normative suggestions in the paradox, as it differs from these suggestions. We then expand on a proposal that, perhaps reading rather generously, we find in Allais's writings. The classical decision-theoretic treatment induces a divide between the normative assessment of the EU hypothesis and its rivals, and the empirical stage at which these hypotheses get confirmed or disconfirmed by observed behaviour. Instead of following such a dual approach, we suggest that it is both possible and instructive to *let experiments themselves determine which of the testable hypotheses are normatively sustainable*. This is how we understand Allais's striking expression, 'the experimental definition of rationality'. More precisely, the experiment should be so devised that, first of all, it provides

[1]Allais's broader contributions to economics will not be discussed here. For general overviews, see Drèze (1989), Grandmont (1989), Munier (1991) and Diemer *et al.* (2010).

[2]The most significant exception relates to the normative dynamic argument that some decision theorists, following Machina (1989), have identified in the Allais paradox. Others have normatively interpreted the so-called certainty effect that underlies the paradox. More on these connections below.

information on the reflective judgements of the subjects, not simply their choices, and second, it pays special attention to a subgroup of subjects who are identified as rational on the basis of this previous information. When some of these selected subjects endorse a decision rule, they confer on it a normative warrant which is by nature different from that conferred in the standard discussion of the rationality of choices.

As well as departing from today's standard experimental methodology, the above proposal involves a mixing of the normative and positive that will seem shocking to all those – many philosophers of science included – who honour the sharp separation between *is* and *ought*, value and fact, and similarly polarized ideas. We hope to convince the sceptics by revisiting the experiments that a handful of Allais's followers performed in the 1970s, for they developed original ways of capturing rationality within their experiments, and claimed to be able to draw normative consequences from the results so obtained. Imperfect as they are by today's technical standards, these forgotten studies can serve as proof – in an eating-the-pudding style – that the approach we recommend is neither unsound nor unpromising.

The paper develops through six subsequent sections. Section 2 reviews the basics of EU theory and the Allais paradox; we need this material for what is to come next. Section 3 presents the received view of the paradox as an empirical finding and explores some of its methodological implications. We include this view in the paper both for completeness and to highlight our contrast of interpretations, but we can be brief here, as others have specialized in this part of the account.[3] Section 4 places Allais's writings of 1952–1953, and specifically his statement of the paradox, in their original context – that of the 1952 Paris conference on decision under risk. This conference famously saw a clash between the American school led by Samuelson and Savage, and the French school led by Allais. Although this event belongs to the decision theorists' collective memory, no one thus far had studied it in full detail. We find that the debate between the two schools over the EU hypothesis was primarily normative in character, a contextual finding that reinforces our general interpretation.[4]

In Section 5, we reconstruct Allais's argument about the paradox as set out in 1952–1953, centring on what he meant by the 'rational man' *versus* the 'real man', and 'the experimental definition of rationality'. Section 6 covers the aforementioned experimental work by Allais's early followers; we argue that this work had absorbed the original intended sense of the paradox, and then draw out some methodological suggestions for new experiments in decision theory. Section 7 briefly concludes. The appendices deal with three specific arguments that historically connect with the Allais paradox. One of them, due to the French economist Morlat, contains

[3]Machina's (1983, 1987) overviews contain rich and important methodological insights on the EU hypothesis and alternative hypotheses (including his own) from the empirical perspective. Mongin (1988, 2009) provides some treatment based on philosophy of science. Both writers claim that the move from EU to non-EU theories represented theoretical as well as empirical progress. Detailed reviews of the experimental evidence appear in, among others, Schoemaker (1982), Hey and Orme (1994), Camerer (1995), Starmer (2000) and Wakker (2010).

[4]This part of the paper should be compared with the studies by Guala (2000) and by Jallais and Pradier (2005). More historical references appear below in connection with other issues.

a little-known anticipation of the Allais paradox, while the other two, due to Samuelson and Savage, are the sources of the classic attempts at dismissing it from normative consideration.

2. An elementary review of the Allais paradox and its VNM connections

The Allais paradox emerged against the background of von Neumann and Morgenstern's recently published work – specifically, the 1947 edition of their *Theory of Games and Economic Behaviour*, which complemented the first edition (1944) with a mathematical appendix on 'the axiomatic treatment of utility'.[5] The contribution of this appendix was limited to axiomatizing the EU hypothesis, which had of course been long familiar; the two authors had used it, unreflectingly as it were, when developing their concept of mixed strategies. Their axiomatic treatment differs from those which it inspired later, even if all systems eventually became lumped together under the heading 'VNM axiomatization'. Most systems include a prominent condition of 'VNM independence' that the 1947 appendix did not have. Fishburn and Wakker (1995) have explained this lacuna, and traced the gradual emergence of 'VNM independence' in the post-war years. Here, we will use a form of 'VNM axiomatization' based on Friedman and Savage (1952). As a matter of historical fact, Allais and the other participants of the Paris conference were not exposed to this version, but rather to an alternative one due to Samuelson. However, the former is still acceptable by today's standards, which is not the case with the latter, and importantly they both share the same attractively simple version of 'VNM independence'.[6]

Let $X = \{x_1, \ldots, x_n\}$ be the set of final outcomes, arranged in some fixed order, and $u = (u(x_1), \ldots, u(x_n))$ the vector of utility values assigned to them. The set L of uncertain options consists of all probability measures, or *lotteries*, on X; in the present framework, these simply reduce to probability vectors $p = (p_1, \ldots p_n)$, $q = (q_1, \ldots, q_n)$, and so on.[7] A lottery attributing probability 1 to an outcome is called *degenerate* and identified with that outcome. The EU hypothesis states that for all $p, q \in L$,

$$(*)\ p\,R\,q \textit{ if and only if } \Sigma\, p_i\, u(x_i) \geq \Sigma\, q_i\, u(x_i).$$

Here R stands for the agent's weak preference relation, which will serve as the primitive of the axiomatization. The strict preference relation P and the indifference relation I are defined from R in the natural way.

Friedman and Savage impose three axiomatic conditions:

(A1) R is a *weak order*, i.e. it is transitive and complete.
(A2) R is *continuous* in a suitable technical sense.

[5]Leonard (1995: 753) explains the genesis of this important appendix. It was originally Morgenstern's idea, although von Neumann was clearly responsible for its mathematical implementation.

[6]Neither Friedman nor Savage presented their 1952 axiom system at the Paris conference, and Samuelson's (1952a) alternative system turned out to be the only one presented at the conference. Shortly afterwards, he published it in still another form (Samuelson 1952b). The main other axiomatizations known at the time were by Friedman and Savage (1948) and Marschak (1950). Later, Herstein and Milnor (1953) and Luce and Raiffa (1957) provided other systems.

[7]We simplify Friedman and Savage's treatment by assuming a finite set of outcomes.

(A3) R satisfies *VNM independence*, i.e. if p is weakly preferred to q, and r is another lottery, the lottery that gives p with probability α and r with probability 1-α, is weakly preferred to the lottery which gives q with probability α and r with probability 1-α, and conversely. Formally, for all α strictly between 0 and 1, $p\ R\ q$ if and only if $\alpha p + (1-\alpha)r\ R\ \alpha q + (1-\alpha)r$.

Axiom (A3) involves *convex combinations* of lotteries (and so does (A2), which we did not state formally). By a standard calculation, based on the 'multiplication law of probability', convex combinations of elements of L are elements of L. Explicitly, $\alpha p + (1-\alpha)r$ is $(\alpha p_1 + (1-\alpha)r_1, \ldots, \alpha p_n + (1-\alpha)r_n)$, and $\alpha q + (1-\alpha)r$ is $(\alpha q_1 + (1-\alpha)r_1, \ldots, \alpha p_n + (1-\alpha)r_n)$. Thus, mathematics obliterates the intuitive difference between *compound lotteries*, i.e. those having other lotteries among their outcomes, and *simple lotteries*, i.e. those having only final outcomes. This break with intuition offers both advantages and disadvantages. On the one hand, it is convenient to have a single set L of options. On the other hand, adopting L forces identifications that agents may not be able or willing to make. One should at least be clear about the fact that these identifications occur, and for this purpose we will add to (A1), (A2) and (A3) an *informally* stated principle to the effect that they do. We call it the *principle of compound lotteries*. The received terminology is in fact the 'reduction of compound lotteries to simple lotteries', but this is slightly misleading since the identification goes both ways.[8]

Corresponding to any 'VNM axiomatization' there is a 'VNM representation theorem' which connects the preference conditions with the EU hypothesis. In the present instance, the theorem states that R satisfies (A1), (A2) and (A3) on L if and only if the equivalence (*) holds, and that the function u entering this equivalence is unique up to the origin and the unit of measurement. Our paper needs the former part – the existence result – and has little or no use for the latter – the uniqueness result.

Both VNM and their first followers[9] specialized the EU hypothesis to the case where an agent's preferences bear on uncertain prospects whose outcomes have stated probability values. The proof that, under relevant preference assumptions, the hypothesis also applies to prospects without preassigned probabilities had to await Savage's *The Foundations of Statistics* (1954), which founded subjective expected utility (SEU) theory. Still busy with this extension when the Paris conference took place, Savage presented it in draft form on this occasion. This appears to have been the first time the nascent community of decision theory came across it.[10] For a long while, they would not sharply separate the VNM and SEU branches, gliding in particular over the difference between VNM independence

[8]Friedman and Savage allude to the principle of compound lotteries in just one passage (1952: 467). All early VNM followers accept the principle, and all except for Samuelson (1952a, b) build it into the mathematical structure; more on this in section 3. It is only at a late stage that decision theory tried to disentangle the principle from the inherent properties of the L space. See especially Segal (1990) and Karni and Schmeidler (1991).

[9]Except for Rubin (1949).

[10]In Fishburn and Wakker's (1995) bibliography, the oldest version of Savage's axioms is not dated earlier than 1952. His 1951 article puts forward minimax regret but not yet EU (Savage 1951).

and Savage's 'sure-thing principle', which are by no means mathematically equivalent. At the conference itself, even after Savage made his presentation, participants often overlooked this major difference. Savage himself does not rigorously attend to it in his 1952 paper with Friedman. We will see that this had some significant consequences for the assessment of the Allais paradox.

Conditions (A1), (A2) and (A3), and their related variants elsewhere, seem to enjoy intuitive plausibility, which contributed to the widespread success of the EU hypothesis in the post-VNM years. Here, however, is the counter-example, which would become famous, that Allais proposed in 1952–1953. Let two choices be made, firstly between p_1 and q_1, and then between p_2 and q_2 (the amounts indicating million French francs):

p_1: 100 with prob 1	p_2: 100 with prob 0.11; 0 with prob 0.89
q_1: 500 with prob 0.10; 100 with prob 0.89; 0 with prob 0.01	q_2: 500 with prob 0.10; 0 with prob 0.90

From the four pairs of choices, (p_1, p_2) and (q_1, q_2) respect the EU hypothesis, and (p_1, q_2) and (q_1, p_2) violate it. This is checked by inspecting the algebraic inequalities that express preference comparisons according to equivalence (*). For example, if the choices are (p_1, q_2), Allais's privileged pair, the preference comparisons are $p_1 \; P \; q_1$ and $q_2 \; P \; p_2$, which (*) transforms into a contradiction:

$$u(100) > 0.10\,u(500) + 0.01\,u(0) + 0.89\,u(100)$$

and

$$0.11\,u(100) + 0.89\,u(0) < 0.10\,u(500) + 0.90\,u(0).$$

This argument presumes that the *choices* made by the individual, and observed by the theorist, match this individual's *preferences*, which are the official object of the axiom system. Here we come across a possible discrepancy that the ordinary semantics of preferences in decision theory – 'revealed preference' – ignores. Allais and his opponents implicitly endorse this semantics, and we will in turn take it for granted here.

Allais (1953a, b) claims that very prudent individuals make the choices (p_1, q_2) and he produces various arguments to that effect. The choice of the term 'paradox' to designate either this claim or the choices themselves arises from the suggestion that the pre-existing state of opinion – the *doxa* – must be favourable to the EU hypothesis. Jallais and Pradier (2005) date the first systematic use of this term to the 1979 collection of papers assembled by Allais and Hagen, *Expected Utility Hypotheses and the Allais Paradox*. But to call it a paradox is merely a convenient tag, without any deep conceptual meaning being attached to the usage.[11]

By itself, the counter-example bears on the EU hypothesis as a whole, and not on one or other of its underlying axiomatic conditions. Allais relates it to (what is in

[11]Decision theory sometimes suggests a deeper sense of paradox – when it clashes not so much with a given state of opinion as with deeply rooted categories of practical reasoning. Newcomb's paradox is one such example.

effect) condition (A3), but this needs an argument, which we now spell out. First, introduce the auxiliary lottery l:

l : 500 with prob 10/11,
0 with prob 1/11,

and then restate the original choice problem as follows:

p'_1: 100 with prob 0.11; 100 with prob 0.89	p_2: 100 with prob 0.11; 0 with prob 0.89
q'_1: l with prob 0.11; 100 with prob 0.89	q'_2: l with prob 0.11; 0 with prob 0.89

With this restatement, the preference comparisons become $p'_1\ P\ q'_1$ and $q'_2\ P\ p_2$, and the violation now relates to (A3) specifically. Indeed, by (A3), it does not matter whether the common outcome is 100 or 0, so that $p'_1\ P\ q'_1$ is equivalent to 100 P l, and $q'_2\ P\ p_2$ to $l\ P$ 100. Actually, (A1) is needed to make P asymmetric; ignoring this detail, one imputes blame to (A3).

The principle of compound lotteries, which is a crucial assumption behind the previous reasoning, permits replacing q_1 by q'_1 and q_2 by q'_2 (and also p_1 by p'_1, perhaps not an innocent substitution either).[12] While never mentioning the principle, Allais alludes to it when he suggests that, among the individuals choosing p_1 and q_2, at least some are 'familiar with the probability calculus' (1953a: 524). This technical understanding and oblique acceptance of the principle were quite usual at the time. It is also typical that Allais directs the paradox at 'Savage's independence principle' (1953a: 525–527), and not VNM independence, as he should have. This is the first occurrence in this paper of the pervasive confounding we mentioned above. Just after stating the paradox, Allais (1953a: 528–530) objects to Samuelson's 'substitution principle', which comes close to (A3), by means of a new, specially devised counter-example. Ironically, despite being directed at the proper condition, this example has fallen into oblivion.

We have just presented the basics of the Allais paradox and its VNM connections, making sure not to preempt the interpretations now to come. Section 3 spells out the standard one, and in Sections 4 and 5 we set out our own.

3. What the Allais paradox became: an empirical refutation

As the standard view goes, the Allais paradox is a heuristic step towards the conclusion that the EU hypothesis, and more specifically the VNM independence condition, are empirically false. The standard view also describes it as a heuristic step towards establishing alternative empirical regularities of choice behaviour. Over and beyond these claims, we also identify the received view by what it does *not* say. It does not mention any other form of relevance for the paradox than its contribution to empirical knowledge, and in particular is mute on its possible normative relevance.

[12]The move from p_1 to p'_1 was long taken for granted, but this and the reverse move are now investigated under the labels 'event-splitting' and 'coalescing'; see Birnbaum (1998, 2004) and Birnbaum *et al.* (1992).

Most surveys of decision theory present the paradox this way, and so do most accounts of Allais's work.[13] We date the standard view so conceived to Kahneman and Tversky's (1979) classic paper. They reported experimental variations on the paradox that evoked a high prevalence of non-EU conforming choices; but other experimentalists had already done nearly as well, and what really sets their paper apart is what they *chose to ignore*. With them, the paradox became specialized in its role of disconfirming empirical hypotheses.[14] Section 5 will provide evidence of this historical break. In the present section, we briefly summarize what the received view contributes. There is no question that it has refined the empirical side of the paradox considerably; what we dispute is its lack of realization of a wider potential. In the rest of this section, we first follow the paradox on its road towards genuine experiments, and then explain the sense in which decision theorists have taken these experiments to reveal empirical regularities and deliver empirical refutations.

3.1 The Allais paradox on the road to experimentation

Allais himself turned to an empirical perspective shortly after inventing the paradox. Having first circulated it in the spring of 1952 under circumstances we explain below, later in the same year he embarked on a 'poll' (*sondage*), as he put it, of real individuals. For a long time after this early study, he would allude to it without detailing its results, except to say that they supported his objections to the EU hypothesis. It was as late as 1979, in a collection prepared with Hagen, *Expected Utility Hypotheses and the Allais Paradox*, that he provided the following explanations (Allais 1979: 447–449).

In 1952, he had distributed a postal questionnaire to selected correspondents, and also presented it, in a shorter version, to the members of a seminar he was conducting in Paris. From the responses then received, he had selected 101 – 53 from the first source and 49 from the second – with a view to analysing them statistically. At the time, he had to content himself with simply publishing the questionnaire.[15] Being involved in other theoretical pursuits, he was unable to begin the statistical analysis until the 1970s, and this explains why he published his results so late – in the same 1979 collection in which one can read his apology.[16]

By Allais's own admission, he could only provide 'selected findings'. His statistical analysis consisted in little more than statements of proportions and averages. Despite the decision theorists' occasional requests, nothing else on the 1952 'poll' filtered out from Allais's writings after 1979.

Inspection of the questionnaire shows that it included the paradox and other counter-examples, but did not place any emphasis upon them. Its main purpose was

[13]See fn 2 for examples of the former. Examples of the latter are Grandmont (1989: 31–32) and Munier (1991: 191–194).

[14]Although they attracted much less attention than Kahneman and Tversky's paper, two contemporaneous studies by Karmakar (1978, 1979) also initiated the restriction of the paradox to purely empirical concerns.

[15]The questionnaire appeared in a French statistical journal (Allais 1953a: 55–73). It was later summarized in English in Allais and Hagen (1979: 612–614).

[16]'Selected Findings of the 1952 Experiment', Appendix C in Allais and Hagen (1979: 611–654).

numerically to estimate the utility functions of selected subjects. Being a 'cardinalist' in the tradition of early neoclassicals, Allais was hoping to recover utility functions that would measure the subjects' preference intensities, and not simply represent their preference levels, and many of his questions were intended to fix numerical values for these intensities. For instance, he would ask: 'Is your preference for an inheritance of 10 million rather than no inheritance stronger than your preference for an inheritance of 150 million rather than 10 million?' (Question 651, in Allais 1979: 616). Unlike some VNM theorists, Allais denied that preference intensities could be read off from an estimate of the u function in the EU formula. Allais actually claimed that no sense could be made of an EU formula that did not take the form $\Sigma p_i v(x_i)$, where the utility function v is substituted for u and measures preference intensities in the *certainty* case. This claim also brought him into conflict with the other VNM theorists, for whom the EU formula involved no measurement of preference intensities at all. These controversies over the 'cardinality' of the EU formula are logically independent of the argument against VNM independence, and although they should be part of a fuller account of Allais's position, they are too complex to be explored here alongside the ramifications of the paradox.[17]

Given the overall orientation of the questionnaire, it comes as no surprise that the 'selected findings' devote so little space to the experimental assessment of the counter-examples. Allais gives just a dozen lines to them in a note to the main text (Allais and Hagen 1979: 636). Regarding the paradox itself, he mentions 46% violations of the EU hypothesis. This is a lower proportion than for other counter-examples included in the questionnaire, but he unfortunately does not pursue the interesting comparison.

The 'poll' has remained famous because of its early date, which is enough to secure it a place in the history of the experimental work on decision, but it is hard to see in it a prototype of what this field would later become. Apart from the sketchy statistical treatment, the sample did not obey any rigorous principle of selection, the two groups of respondents did not answer exactly the same questions, and the questions themselves were too remote from real choices, except for the few based on the counter-examples.[18]

The first rigorous experiments inspired by these counter-examples did not take place until the early 1960s, and they became more varied only during the 1970s. This strikes one as being an extended delay – the more so since, both shortly after and shortly before the 'poll', a significant number of experimental studies had come out on decision making, both under certainty and uncertainty.[19] Admittedly, these studies were primarily conducted by psychologists, who studied decision rules without attending to their axiomatic decomposition, unlike the economists and statisticians who were Allais's prime public. Relatedly, the prevailing statistical methodology favoured the measurement of numerical magnitudes, especially of utility functions, even when the purpose was to confirm or refute an abstract

[17]Fishburn (1989) and Moscati (2013, Forthcoming: ch. 10) cover the history of the 'cardinality' problem in VNM theory. Bouyssou and Vansnick (1990) and Munier (1995) reconstruct Allais's idiosyncratic position on this problem.

[18]Hey (1991: 72–75) has related criticisms.

[19]Schoemaker (1982) lists these older experiments on the EU hypothesis. The first in this genre, by Mosteller and Nogee (1951), clearly preceded the Allais paradox. Preston and Barrata's (1948) experiment has overall precedence, but it is not entirely related to EU.

hypothesis that could have been checked on merely qualitative implications. These contextual facts may explain why Edwards (1954: 394) – perhaps the most prominent decision psychologist of the time – dismissed Allais's questionnaire as being unusable.[20] Such direct counter-examples as the paradox, which merely involved conflicting preference statements, long remained the preserve of properly trained theorists. They dissected them according to their particular goals, paying little attention to their empirical relevance. This applies to Savage (1954 [1972]: 101–104), Luce and Raiffa (1957: 25), and at greater length Raiffa (1968: 80–86), all of whom tried to reinstate the EU hypothesis against Allais. Savage's reconstruction of the paradox is the topic of Appendix 3.

Five remarkable studies – by MacCrimmon (1968), Moskowitz (1974), Slovic and Tversky (1974), MacCrimmon and Larsson (1979) and Kahneman and Tversky (1979) – eventually gave the Allais paradox a genuine experimental status. The first three studies involve populations of 30 to 50 subjects, the last rather more (up to 72), and the penultimate one rather fewer (19).[21] In each case, paired choice problems are posed, using Allais's exact set-up (Slovic and Tversky), or a variant that preserves the probabilities but changes the outcomes (Moskowitz), or a variant that changes both (Kahneman and Tversky), or an even more remote form (MacCrimmon). In the study that is the most thorough, despite having the smallest number of subjects, numerical probabilities and money outcomes give way to *algebraic parameters* and multiple sub-experiments result from taking values in the ranges of variations of these parameters (MacCrimmon and Larsson). Across the studies as a whole, the percentage of violations of the EU hypothesis varies from 27% to 61%. The latter percentage appears in Kahneman and Tversky, who artfully calibrate their numerical values; one may consider this less representative than the others, which are concentrated between 27% and 42% (see MacCrimmon and Larsson's summary, 1979: 366–367).[22]

3.2 Empirical regularities and empirical refutations connected with the Allais paradox

The multiple replications in the 1970s have conferred upon the paradox an empirical robustness that is difficult to challenge. Thanks to their parametric method, MacCrimmon and Larsson also discovered some non-obvious general characteristics. On the one hand, violations do not depend on the absolute value of monetary

[20]Edwards (1954, 1962) inspired Kahneman and Tversky's (1979) prospect theory more than the latter acknowledge. He was scathing of Allais, but nonetheless opposed the EU hypothesis for reasons of his own.

[21]The subjects in these experiments were students, except in MacCrimmon (1968), who was able to address businessmen. One of the first articles on the Allais paradox, by Morrison (1967), is sometimes cited as making an experimental contribution, but it only briefly mentions a poll the author performed in class.

[22]Kahneman and Tversky (1979: 265–266) propose the following example (the monetary unit is the shekel):

p_1: 2400 with prob 1	p_2: 2400 with prob 0.34; 0 with prob 0.66
q_1: 2500 with prob 0.33; 2400 with prob 0.66; 0 with prob 0.01	q_2: 2500 with prob 0.33; 0 with prob 0.67

The subjects choose p_1 and q_2 at 82% and 83% rates, with the intersection of the two groups making 61%.

outcomes, but on their relative value, following a given rule. On the other, they still occur when the certain lottery – p_1 in Allais – is replaced by uncertain lotteries of a relevant type. Curiously, one feature that jumps out at any casual observer long remained unelaborated. Subjects who contravene the EU hypothesis align themselves with one of two violations, neglecting the other; thus, in Allais, most of them select (p_1, q_2) instead of (q_1, p_2). Conlisk (1989) seems to have been the first to explore the significance of this feature.

MacCrimmon and Larsson (1979: 360) used the term *common consequence* to designate the parametric form subsuming Allais's numbers, and by the same token the empirical regularity, or *effect*, covering the paradox.[23] Today's decision theory has kept the term but redefined the effect more generally and abstractly. In this new sense, well explained by Machina (1983, 1987), the effect occurs in the following pattern of choice:

p_1: x with prob p; M with prob 1-p	p_2: x with prob p; M' with prob 1-p
q_1: Q with prob p; M with prob 1-p	q_2: Q with prob p; M' with prob 1-p

Here x is a money amount, and M, M' and Q are lotteries; x lies between the worst and best amounts of M; and M *stochastically dominates* M'.[24] The subject makes the choices (p_1, q_2), which contravenes the EU hypothesis. Allais's numbers are recovered by setting $x = 1$ million, $M = 1$ million with prob 1, $M' = 0$ with prob 1, and $Q = l$ as defined in last section. This takes M and p_1 to be degenerate lotteries, which the common consequence effect *per se* does not require.

The paradox can be subsumed under another putative generality. For Kahneman and Tversky (1979: 245), it follows from a *certainty effect*; that is, the salient feature of the paradox is that p_1 is a degenerate lottery offering a satisfactory outcome. This diagnosis has obvious force, but needs balancing against the previous one. Unfortunately, the two effects compare poorly, because they are not subject to the same degree of theoretical elaboration. While the common consequence effect has received a precise definition, the certainty effect has basically remained at the intuitive level at which it was originally introduced. Here is the most technical statement we found in Kahneman and Tversky: 'the overweighting of outcomes obtained through certainty relative to outcomes that are only probable' (Tversky and Kahneman 1986: S265). This is too vague for making it possible to apply MacCrimmon and Larsson's method of parametric variations. All the same, Kahneman and Tversky's unsystematic variations on the certainty effect deliver substantial rates of EU violations, and without being able to say much more, decision theory has allowed the two effects to coexist. In sum, the Allais paradox, viewed as a merely empirical phenomenon, lives at the crossroads between the certainty effect and the common consequence effect. The decision theorists' views can be more entrenched, however: for example, Machina (1983, 1987) assigns the Allais paradox to the common consequence effect without discussing the certainty effect, while Munier (1991, 2011)

[23]'Consequence' is here a synonym for 'outcome'.

[24]By definition, lottery M stochastically dominates lottery M' if, for any possible outcome, the probability of obtaining at least this outcome is at least as great with M as with M'.

privileges the latter. The recent experimental treatment of the paradox often focuses on the certainty effect: see Weber (2008) and Andreoni and Sprenger (2009).[25]

Although sometimes overshadowed by the fame of the paradox, Allais penned another brilliant counter-example in 1952–1953. The violation of the EU hypothesis in the latter arises when one pair of choices results from the other by a homothetic reduction of the probability values of the best outcomes. MacCrimmon and Larsson (1979: 350) also gave this phenomenon algebraic form and studied it experimentally using differing numerical values for the parameters. They use the term *common ratio* to refer to either the parametric form or the associated effect. Today's decision theory – see again Machina (1983, 1987) – has taken up the expression, this time keeping the initial parametric form unchanged:

p_1: x with prob p; 0 with prob $1-p$	p_2: x with prob αp, 0 with prob $1-\alpha p$
q_1: y with prob q; 0 with prob $1-q$	q_2: y with prob αq, 0 with prob $1-\alpha q$

with $p > q$, $x < y$, and α (the homothetic factor) lying between 0 and 1. The subject makes the choices (p_1, q_2), which contravenes the EU hypothesis.[26] The many studies bearing on this effect, including MacCrimmon and Larsson's (1979: 359), show rates of violations clearly above those obtained for the common consequence effect. In general, these studies make the parametric restriction $p = 1$, which amounts to involving the certainty effect, and thus no doubt increases the number of violations. Both the common consequence effect and the common ratio effect have provided powerful heuristics for the theoretical development of non-EU theories, which we do not cover here (readers may consult the references in fn 3).

Related to the empirical phenomena discussed above is the observation that different numerical utility curves are obtained for the same subject, depending on the kind of choice data that the experimenter elicits from this subject. This non-uniqueness would not occur if the EU hypothesis regulated the choices. Allais noted the phenomenon while analysing some answers to his 'poll' (1979: 474), and Karmakar (1978) identified it more fully; it was repeatedly corroborated afterwards. Machina includes this *utility evaluation effect* in his classic list of effects that are damaging for the EU hypothesis.

The philosophy of empirical sciences treats an experiment as refuting its target hypothesis only if it has been properly replicated. The paradox met this condition in the 1970s and can thus be said to have refuted the EU hypothesis. As to turning the refutation against VNM independence specifically, this is another matter, a proper discussion of which exceeds the scope of the present paper. Logically, the experiments associated with the effects of this section are all ambiguous. One can attribute their results to a violation of VNM independence, a violation of

[25]Although widely studied experimentally, the certainty effect has attracted less theoretical attention than the common consequence effect; see however Gilboa (1988), Jaffray (1988), Rubinstein (1988), Cohen (1992), who deal with it indirectly, and Cerreia-Vioglio *et al.* (2015), who provide a direct axiomatization.

[26]Allais's original counterexample (1952: 316–317, and 1953b: 528–559) conforms to a more complex structure which, to our knowledge, has not been subjected to experimental research.

the principle of compound lotteries, or both violations occurring simultaneously. Decision theorists here face an *underdetermination problem* in the sense classically explained by the philosopher of physics Duhem.[27] A majority of them have persisted with Allais's claim that the refutation of the EU hypothesis reduces to that of VNM independence; some have even extended this claim to weaker forms than (A3) that emerge when the principle of compound lotteries is duly formalized. This overall conviction relies on an admixture of arguments, some logically articulated, others rough and intuitive, and still others pragmatic. The most influential argument may well be the pragmatic one that alternatives to the EU hypothesis are mathematically easier to develop by retaining the principle of compound lotteries, or at least some form of it. This is typical of the hesitant way in which the empirical sciences resolve the Duhem underdetermination problem (more on this line in Mongin 2009).[28]

4. What the Allais paradox was: a normative objection

The paradox as it is commonly presented today bears little relationship to what it was for Allais.[29] For a clearer view, we need to go back to the texts where it first appeared: the memoir that Allais wrote in 1952 and published in 1953, 'Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine', and the article that he extracted from the memoir in 1952 and published in *Econometrica* the following year, 'Le comportement de l'homme rationnel devant le risque'. Of the two sections of the memoir, the article retains the 'critique' in full, but the 'positive theory' only in part – a loss of content which Allais (1953b: 503) deplored. However, the two pieces are the same as far as the paradox goes, and since the article has become the universal reference, we will only refer to it here. We preface our analysis with a discussion of the 1952 Paris conference, which is one of the scholarly contributions of the present paper. These proceedings – in French and still not translated – are collected in a volume, here referred to as *Econométrie*, which also came out in 1953.

4.1 The 1952 Paris conference and the invention of the Allais paradox

From 12 to 17 May 1952, the Centre National de la Recherche Scientifique (CNRS) hosted a meeting in Paris of the leading representatives of the nascent discipline of decision theory; this is actually one of its founding moments. Dealing with the theory of risk in economics, the conference brought two groups together, one

[27]In today's philosophy of science, the *Duhem problem* is that of showing how empirical refutations can be determinate, given that they do not logically contradict isolated propositions, but only sets of propositions. Some prefer the term *Duhem-Quine problem*, but Quine's angle on indeterminacy is different (see e.g. Mongin 1988).

[28]Birnbaum and collaborators (as referenced in fn 11) are an exception to this standard resolution of the Duhem problem of EU theory. In view of their experimental evidence, they question 'coalescence' and 'event-splitting', which even critics of the principle of compound lotteries would take for granted. At the theoretical level, Segal (1990) is also an exception: he argues against this principle rather than VNM independence (in the weaker form that the formalization of the principle makes sufficient).

[29]In the methodological and historical literature, only Guala (2000) has clearly recognized this, and his sketch of a 'normative falsification' account anticipates what this section will explain more fully. By contrast, Heukelom (2015: 17 and 20) attributes to Allais a descriptive view of both the paradox and the VNM axioms.

defending the EU hypothesis, the other disputing it. The former included nearly all the American participants, from which came the name 'American School' that Allais would use from then on to designate it, but drew supporters from elsewhere as well. De Finetti also supported the EU hypothesis and Frisch was mildly favourable to it. The latter group was reduced essentially to three Frenchmen, Allais, Massé and Morlat. It is noteworthy that they had been trained as economic engineers, whereas most of their opponents were academics and not all were economists. Guilbaud, a mathematician who had recently converted to social sciences, put the proceedings together for publication.[30] He included both the papers and the ensuing discussions, making this volume a source of great documentary value.

On the pro-EU side, Savage's (1952) and Samuelson's (1952a) contributions attracted most attention, the former because his axiomatization of subjective expected utility (SEU) was of stunning novelty, and the latter because he provided a VNM axiomatization that properly emphasized the role of VNM independence and developed an energetic defence of the EU hypothesis on this basis. Samuelson had initially reacted negatively to von Neumann and Morgenstern's 1947 appendix (von Neumann and Morgenstern 1944), but subsequently had become an enthusiastic convert to their approach, and his axiomatic work was meant to present it in the best possible way. We discuss (and actually rebut) Samuelson's defence of VNM independence in Appendix 2.[31] Two other participants who might have fought on the pro-EU side kept a low profile at the conference, namely Friedman (despite his close association with Savage) and de Finetti (despite the clear connection between some of his pre-war work and the EU hypothesis).

The anti-EU minority was impressively active, bringing ever more objections to both the axioms and the hypothesis itself, sometimes in the form of counter-examples, sometimes arguing from alternative proposals. Massé and Morlat's paper is impressive, and a short comment made by the latter perhaps even more so, but Allais had a more prominent role in the debates, and Frisch's closing address registers this state of affairs by singling him out as *the* representative of the dissident position. He made no less than two presentations, the first extending general economic equilibrium to situations involving risk, the second devoted to the individual theory of choices involving risk – the one that concerns us here.[32] Moreover, he unceasingly tried to refocus the often diffuse discussion on the fundamental disagreement between him and the American school.[33]

[30]Guilbaud made a name for himself in the different field of judgement aggregation; see Mongin's (2012) account of his contributions.

[31]Heukelom (2014), Moscati (Forthcoming: ch. 10–11), and in complete detail Moscati (2016), follow Samuelson's change of mind through the unpublished letters he exchanged with Baumol, Friedman, Marschak and Savage in 1950.

[32]*Econométrie*: 81–109 ('Généralisation des théories de l'équilibre général et du rendement social au cas du risque') and pp. 127–140 ('Fondements d'une théorie positive des choix comportant un risque'). Despite their distinctive theoretical objectives, the two papers are linked by their common reticence about the EU hypothesis. Some brief 'Observations générales', also by Allais, appear in the proceedings just before Frisch's closing address, but we do not know whether they were actually presented at the conference.

[33]See *Econométrie*: 34–35 and pp. 37–40, responding to Savage; pp. 153–155, responding to Marschak and Savage; pp. 158–159, responding to Samuelson; pp. 161–163, responding to de Finetti and Frisch; and pp. 194–197, relating to Massé and Morlat.

Reading through the proceedings, one is surprised to find that they neither contain nor even mention the paradox or any of the other counter-examples that made Allais famous. By contrast, Massé and Morlat's contribution contained one, and the aforementioned comment by Morlat another, which may be the better of the two (since it is difficult to find, we restate it in Appendix 1). Both already consist of paired choices that contradict the EU hypothesis and are offered as objections to VNM independence. The delayed appearance of the Allais paradox raises a curious historiographical problem, already discussed by Jallais and Pradier (2005). The paradox first appeared in 1953 in *Econométrie*, which has wrongly led to an impression that it was presented at the conference; but it actually belongs to Allais's memoir, which was for some reason included in *Econométrie* alongside the conference proceedings. It is known that Allais worked on the memoir during the summer of 1952, and this places an upper limit to the dating of the invention of the paradox. After carefully studying the whole sequence, Jallais and Pradier (2005) settle on a lower limit for the circulation – if not the invention – of the paradox. In an encounter on the fringes of the colloquium that has remained famous among decision theorists, Allais presented Savage with one or several choice problems that trapped him into violating the EU hypothesis. Jallais and Pradier accept the widespread view that the paradox was presented during this conversation, and, exploiting more evidence, add that it is unlikely to have circulated earlier than this, despite Allais's occasional suggestions to the contrary. As for the possible influence of Massé and Morlat's counter-examples on his own, Jallais and Pradier neither confirm nor reject it.[34] They deflate the issue by arguing that the two authors' conference examples were themselves nothing exceptional at the time. Others were already known, and to invent more of them was a common mathematical exercise among the French (the Americans, for their part, seem to have first come across this genre while in Paris).

These scholarly points are important, because they help us recognize what Allais's target might have been when he launched his counter-examples: they are like *a last argumentative resort*, for use once more traditional means of exposition are exhausted. During the colloquium, Allais had expressed himself at a higher level of abstraction and with more use of mathematics. Although little inclined to the axiomatic method himself, he was drawn to it by the need to engage his opponents on their own ground. But judging from the conference discussions, this highbrow strategy failed. Savage and Samuelson remained inflexible, and the less committed participants contented themselves with being polite. In writing the memoir, and then extracting from it an article that emphasized the counter-examples at the expense of pure theory, Allais presumably believed that he would at last shift the balance in his favour.[35]

[34]Neither Massé nor Morlat left any testimony on this point. After the Paris conference, they distanced themselves from the debate. Both were economic engineers with Electricité de France. Morlat remained there while teaching statistics, and Massé later moved to public administration, where he held important positions.

[35]At the beginning of the memoir, Allais says in effect that he wrote it to refute Savage more thoroughly than he had been able to do at the conference (*Econométrie*, Appendice: 257).

4.2 Allais's structured argument against the EU hypothesis

The *Econometrica* article, on which we now focus our comments, puts forward four theses intended as the steps of a demonstration. The logical structure is so clear that it is surprising that subsequent decision theorists have paid so little attention to it. The theses run as follows. (1) The EU hypothesis does not apply to 'the real man', and everyone today agrees on this, including in the American school. (2) The only remaining question is how the hypothesis relates to the 'rational man', and this cannot be settled by merely pointing out that it conforms to an axiom system, since the question would arise again in relation to this axiom system. (3) If one begins with an 'abstract definition' of individual rationality under risk, one cannot deduce the EU hypothesis from it, because the latter is far more constraining than any such definition would entail. (4) Another way of approaching individual rationality under risk is to study how rational persons act, but this 'experimental definition of rationality' is no more favourable to the EU hypothesis than is the preceding one. Allais does not state his definitive conclusion entirely clearly, and we can distinguish two possible variants for it. Either (5) there are circumstances in which individual rationality under risk is compatible with departing from EU (the *weak* conclusion), or (6) there are circumstances in which individual rationality under risk entails departing from EU (the *strong* conclusion).[36]

Step (1) already signals an open break with what would become the standard interpretation of the paradox. Why would Allais propose an empirical refutation of a hypothesis that he already took to be false? And one, moreover, that he thought everybody else took to be false? Strange as this seems in retrospect, Allais was only slightly exaggerating when he said that his colleagues shared his empirical dismissal of the EU hypothesis. With the single exception of Friedman, the conference participants were not prepared to fight over this ground. Since this is an important fact both in itself and for our contrast of interpretations, we adduce some evidence from the conference proceedings.

Marschak had set the stage of the conference by an early pronouncement contrasting the 'descriptive' with the 'normative' use of the EU hypothesis. Only in the latter use, he said, could the hypothesis be defended as such; in the former use, it was 'too rigid' and would have to enter a scheme of stochastic preference (*Econométrie*: 25–26). Later in the conference, Savage prefaced the formal statement of his axioms with the claim that they were 'principles of conduct or decision' that 'nobody would like deliberately to contravene'. It followed that agreement with real behaviour was 'sometimes' possible (*Econométrie*: 29). This weak defence is compatible with the admission that the EU hypothesis is empirically refuted as a generality. In commenting on the paper, Allais was quick to exploit the loophole: 'it is naturally not here a matter of the behaviour of real men since M. Savage freely concedes his theory is not applicable to them' (*Econométrie*: 38). Savage did not complain, and the debate then moved on to the normative issues. Only

[36]Allais sometimes implicitly distinguishes between the weak and strong conclusions. Here is one passage: 'It follows that: (1) to be rational, a given individual does not necessarily have to behave in the manner that Bernoulli would dictate; (2) that in reality, the consequences of the abstract definition of rationality being less restrictive than the axioms of the Bernoulli school, these axioms contain *something more which, in fact, could be irrational!*' (1953b: 522, our emphasis). Point (1) and (2) reflect the weak and strong conclusions respectively.

Friedman made a sustained attempt at shifting it towards 'empirical generalizations' (*Econométrie*: 71), but from the terse comments he attracted, his paper does not seem to have been a hit.[37] Then came Samuelson, trenchant as ever: 'so far as explaining the behaviour of real men on our planet is concerned, the [EU] hypothesis is rather insignificant' (*Econométrie*: 142). Having 'given up the empirical validity of the theory', he would defend 'its normative interest'. Samuelson's very clear statements channelled the discussion exactly as Allais wished, and the latter could emphatically comment: 'There is a well-posed problem: should a rational man behave according to the [EU] hypothesis, or should he not? The problem lies here and nowhere else' (*Econométrie*: 155).

These quotes from the conference proceedings buttress the contrast of interpretations on which the present paper revolves, and as a secondary purpose they help dispel a not uncommon misconception of the beginnings of decision theory. As the story goes, the early followers of von Neumann and Morgenstern privileged an empirical interpretation of the EU hypothesis and its underlying axioms, and it was not until they encountered Allais's counter-examples, also viewed from the empirical perspective, that they shifted to a normative interpretation. This 'normative retreat' story has no clear grounding in von Neumann and Morgenstern, and directly contradicts Marschak's and Samuelson's very early claims that the theoretical merit of the EU hypothesis could only lie in the normative value of the underlying axioms. Our quotes further show that this story runs afoul of the best evidence one can gather on the collective state of mind in 1952. The thin textual basis for 'normative retreat' has to do with Savage, who first supported a bluntly empirical interpretation of the EU hypothesis in his 1948 paper with Friedman, and later moved to the normative interpretation that permeates his 1954 *Foundations*. However, our evidence casts light even on Savage's attitudes. His conference paper already expresses doubts about the empirical prevalence of EU-conforming choice, and emphasizes rather the normative plausibility of his axiomatic conditions. If Savage did change his mind, this change was well advanced in 1952 and certainly did not have to await the Allais paradox to crystallize.[38]

We now return to the thread of Allais's reasoning. Although obvious, the point made in step (2) was justified contextually. Von Neumann and Morgenstern (1944: 9) had stressed the need to explore the formal meaning of 'rational behavior', and some of their followers, like Marschak (1950: 111–112, 134–135), had taken up this theme less cautiously, by identifying 'rational behavior' with the axiomatic content of VNM theory. By contrast, step (3) is very informative. For Allais (1953b: 522), the 'abstract definition' of individual rationality has only three components: the ordering property of choices, the use of objective probabilities if they are available, and an 'absolute preference axiom', which in today's words means

[37]For the main, Friedman's paper was a repetition of his 1948 work with Savage, in which both authors tried to reconcile some basic facts of preference under risk by manipulating the shape of the VNM utility representation.

[38]The 'normative retreat' thesis appears in Guala (2000), Jallais et al. (2008), Bourgeois-Gironde and Giraud (2008), and in a more subdued form in Heukelom (2015). It seems to belong to an oral tradition of decision theory; how it gained momentum is unclear. Moscati (2016) has joined in the present criticism of the 'normative retreat' thesis.

that preferences respect first-order stochastic dominance comparisons.[39] This purposefully short list is not enough to entail the EU hypothesis. In our chosen axiomatic decomposition, the first item corresponds to (A1), and the third is strictly weaker than (A3). Among von Neumann and Morgenstern's followers, some ugly confusion had prevailed on this logical relation, but Allais made no mistake.[40] Concerning the second item, Allais contents himself with the banal observation that to use objective probabilities is a condition for practical success and it is therefore rational to follow them if possible. This comment has no analogue in the formal treatment. Notice that frequentism is Allais's conception of an objective probability.

Step (4) is where the Allais paradox belongs. It illustrates the 'experimental definition' of rationality at work, and, being allegedly decisive, closes the whole argument against the EU hypothesis. The need for this step follows from the fact that step (3) cannot be sufficient. EU theorists could retort to Allais that his 'abstract definition' begs the question of what rationality is, no more and no less than their axiom systems do. Allais would of course resist the objection, arguing that his definition is not as specific as their axiom systems; but since specificity is a matter of degree, this comparison does not settle the matter. We explain the presence of step (4), hence of the paradox, by the conjecture that Allais expected this objection and thus felt the need to say something more against the EU hypothesis. Having located the paradox in Allais's structured argument, we proceed to discuss the normative force that Allais conferred on it.

5. The normative force of the Allais paradox

We identify two normative lines on the paradox in Allais, one focusing on *the rationality of choices in themselves*, and the other on *the rationality of the agent viewed as a person*. As will be seen, this new distinction connects with that of the two possible conclusions – either weak or strong – which Allais's structured argument might deliver. Once we have covered the two lines of thinking, we will expand on what may be Allais's most provocative insight concerning the paradox, i.e. that rationality can be an object of experimental study.

5.1 Choice-relative normative senses and their limitations

Among the reasons for selecting (p_1, q_2), the most obvious one is the certainty of gain in p_1. According to Allais (1953b: 527), a 'very prudent' person will tend to prefer p_1, which guarantees a not inconsiderable amount, to q_1, which can offer more but without an absolute guarantee. And this same person might well prefer q_2 to p_2, on account of a bolder criterion of choice – in the first instance the mathematical expectation of gain – because in this case both options are similarly

[39]Here Allais draws upon a statement of the axiom by Massé and Morlat (*Econométrie*: 167–168). Both his formulation and theirs fail to distinguish *statewise* dominance, a non-probabilistic concept that belongs to subjective expected utility theory, and *stochastic* dominance, which requires probability values, not states of the world.

[40]Friedman and Savage (1948) sadly fell prey to the confusion. Samuelson never did. From Moscati's (2016) evidence, his pre-1950 conception of preference among lotteries is not unlike Allais's 'abstract definition'.

uncertain. Another less obvious reason matters, although Allais presents it in connection not with the paradox itself but with the counter-example that comes next in his article. This says that the chances of outcomes in a lottery should not be evaluated in isolation from each other, since they exist in relations 'of complementarity (or of non-complementarity)' (Allais 1953b: 528). In terms of the lotteries in the paradox, the 11 chances of winning 100 million do not mean the same thing in p_1, where the 89 remaining chances concern the same gain, as in p_2, where they concern a zero gain. Similarly, the 11 chances of drawing lottery L do not mean the same thing in q_1, where the remaining 89 chances bear on a positive gain, as in q_2, where they bear on a zero gain. These comparisons require the principle of compound lotteries to hold.

Allais's second argument seems less natural than the first, since what it compares are lotteries *across* choice problems (here p_1 with p_2), as opposed to lotteries *within* the same choice problem as in the first (p_1 with q_1, or p_2 with q_2). Such reasoning is unlikely to enter individuals' deliberations, unlike the observation that p_1 differs from the other lotteries by being certain. However, reflective individuals may come across it when trying to rationalize their choices *ex post*. The main advantage of the second argument over the first is its higher generality. Subjective 'complementarities' or 'non-complementarities' can occur for various patterns of chances, in particular when chances do not coalesce into certainty and thus extreme prudence has nothing to recommend. The 1953 article privileges the first argument, and several of Allais's later texts do not even mention the second, which seems like a weakness from his own perspective.[41] There is a striking parallel between the two arguments and the certainty and common consequence effects of the empirical literature. What this literature offers as an explanation of choice behaviour now appears from the normative angle of what may justify it.

Some decision theorists have endowed Allais's paradox with normative force, these being exceptions to what in Section 3 we described as the received view. The most significant proposal made in this respect differs from Allais's suggestions both in terms of content and level of technicality. It takes up a *dynamic* argument that EU theorists have long made in favour of VNM independence. Broadly speaking, these theorists reformulate static choices, as in the paradox and related counter-examples, as if they were multi-stage choices for which new principles – relative to intertemporal decision – can be defined formally and assessed normatively. Along this way, an equivalence emerges between VNM independence and a conjunction of such principles. Then, a normative argument for this condition follows from the claim that these equivalent principles are sound. EU theorists make this claim, and use the resulting argument to dismiss the paradox and other counterexamples, granting the assumption that they involve a violation of VNM independence rather than any other condition. Whereas EU theorists seemed to have taken the upper hand with this argument, Machina (1989, 1991) countered it by denying that the dynamic principles equivalent to VNM independence were all sound, and he used the Allais paradox to illustrate his point, thus offering a new normative defence of it.

[41] Among these texts are those of the 1979 collection, a dictionary article on the paradox (Allais 1987) and a rather late review he provided of his decision-theoretic contributions (Allais 1994).

Some decision theorists have pursued the same line, while EU theorists have tried to reinforce their initial position.[42] This is an important debate, but it does not easily connect with Allais's own defence of the paradox. Remarkably, at the Paris conference, Samuelson presented what may be the first historical defence of VNM independence by dynamic principles, but Allais brushed it away as being irrelevant (see Appendix 2 for details).[43]

We now return to Allais's above two arguments, asking whether they are capable of underwriting a judgement of rationality in favour of the paradoxical choices. The notion of rationality applicable to choices, as distinguished from the agent who performs them, is consistency of choice. Rationality viewed in this way only imposes *formal* conditions upon choice, whereas the justifications above determine it *substantially*. A comparison might bring this point home. Standard economics, which epitomizes the notion of rationality just mentioned, treats the existence of a utility function and the property that choices maximize it as being *desiderata* of rationality, on the ground that choice inconsistencies arise if they are not satisfied. By contrast, it treats special properties of a utility function, such as monotonicity or concavity, as being neither rational nor irrational, because choice consistency says nothing here. The attitudes of letting the certainty of an outcome weigh heavily, or of paying attention to chance complementarities, are broadly analogous to the special properties of a utility function, i.e. neither rational nor irrational by the consistency-of-choice rationality criterion. Thus, the answer to the question that begins this paragraph must be negative. The same answer would actually follow from step (3) by a less roundabout argument. If the 'abstract definition' of rationality is too weak to entail the EU hypothesis, it should consistently not entail any violation of this hypothesis either.

To take stock, we have found genuine reasons for making the paradoxical choices, but are unable to endow them with full normative force. If we stopped the analysis here, only the weak conclusion (5) would emerge. One would be entitled to discard the objection of irrationality against the paradoxical choices, but not to turn it against some EU-respecting choices, as the strong conclusion (6) requires. Modest in itself, this upshot would stand out only by virtue of the historical circumstances. At the Paris conference, Samuelson had daringly claimed that the VNM axioms derived from the very notion of rationality under risk, and Savage had made related – though not so sweeping – statements. Against such strong contentions, Allais only needed to establish (5). Actually, step (3) would already fulfil the argumentative task if the audience were prepared to endorse Allais's weak 'abstract definition' of rationality. Only against a recalcitrant audience does step (4) enter play, and this further limits the impact of the paradox. However, there is more to be said for its normative force.

[42]Among many valuable contributors, Cubitt *et al.* (1991) provide a synthetic account of this debate leaning towards a non-EU position, and Wakker (1999) provides another account leaning in the opposite direction.

[43]It emerges from the theoretical work on the certainty effect (see the works mentioned in fn 25, in particular by Cohen, Gilboa and Jaffray) that decision theorists have sometimes kept closer to Allais's inspiration by conferring normative value on that effect. This is not so much the case with the theoretical work on the common consequence effect or the common ratio effect, to the extent that they are properly distinguished from the certainty effect (which is not always the case).

5.2 An agent-relative normative sense and its promises

Reading between the lines of the 1953 article, we discern another argument, this one concerning the rationality of the agents themselves as against the rationality of their choices. Few – if any – readers of Allais have noted that he discusses his counter-examples in terms of *agent-relative properties*. These are: being 'prudent' or 'very prudent' (Allais 1953b: 524–525, 527–529); being 'very well aware of the probability calculus' (1953b: 524); being 'rational by general consent' (1953b: 525, 527–529); being 'both reasonable and prudent' (1953b: 533); and of course, being 'rational' (throughout and in the title). What characterizes the 'rational man' in contradistinction to the 'real man' appears to be the ability to choose in an all-things-considered, fully reflective manner that eschews any later change of mind, even in the presence of an opponent (see 1953b: 539, concerning Allais himself as an example). This ability should apply not in isolation, but across multiple choices (see 1953b: 534 and elsewhere). Extrapolating from these hints, we propose to say that, for Allais, *the 'rational man' strives to conform his particular choices to general rules of conduct that he is prepared to answer for.*[44]

Such a conception fits in with rationality as *consistency*, provided the latter idea is appropriately broadened. Consistency in the standard economics sense is a property of choices alone, and thus independent of the agents' motivations. However, in a different sense, which surfaces in various philosophical writings, it is the property that an agent displays of being able to subsume his or her actions – choices being a particular case – under general rules of conduct. Thus construed, consistency becomes a property of the agent's motivations. This notion of rationality as *internal* consistency must now be related to the other agent-relative properties that the last paragraph mentioned. Let us consider these properties in turn.

It is easy to dispose of the property of being rational by common consent. One should clearly view this only as a (fallible) criterion for recognizing rational people, not as a component of rationality itself. Prudence raises a more subtle problem. According to the traditional definition, prudence means the ability to conduct the affairs of one's life wisely; this would make it partly coincidental with instrumental rationality, and some philosophers including Kant have in fact equated the two ideas. However, on a more popular conception, prudence is the attitude that consists in forestalling possible harms by relevant precautionary measures. Viewed in this way it is a psychological property, which rational individuals may or may not exhibit. Allais's own understanding of prudence is not entirely clear, but his conjunctive formula – 'both reasonable and prudent' – does suggest the latter rather than the former sense.[45] By using a term of ordinary language, he might have created an ambiguity, but he at least avoided another one, since he thus made clear he

[44]The conception attributed here to Allais bears some analogy with that defended by Gilboa (2011: 18); see also Gilboa and Schmeidler (2001: 17–18) and Gilboa *et al.* (2010). However, we emphasize the role of generalities in the agent's reflective process more than these authors do.

[45]Jallais *et al.* (2008) relate Allais's remarks to older ideas of prudence as pragmatic wisdom, and concomitantly of public recognition as a touchstone of prudence. Although this is an interesting direction to take, we doubt that the textual evidence supports it. The philosophical contrast between the 'reasonable' and the 'rational' should not be overdone either; it seems as if Allais uses both words interchangeably.

did not mean 'risk-aversion' in the technical sense associated with VNM utility functions.[46]

The ability to make probabilistic judgements raises few substantial problems. If one adheres to the notion of rationality as reflective compliance with general rules of conduct, one has no reason to make probabilistic ability a necessary part of rationality. As to Allais's own position, it is rather unclear. He uses a conjunctive expression again – 'rational men *and* very well aware of the probability calculus' (1953b: 524, our emphasis). This suggests he endorses the view that probabilistic ability is extraneous to rationality. However, as we have seen, his 'abstract definition' of rationality does refer to the probability calculus, at least in one interpretation (the objective one, which Allais favours). We have no easy solution to offer to this exegetical conundrum and simply gloss over it.[47]

At any rate, Allais needs to endow his individuals with some probabilistic ability, since the 'complementarities' or 'non-complementarities' he deems relevant to their choices can only be discovered by those who know how to operate with the multiplicative rule of probability. He also needs this ability to turn his counterexamples specifically against VNM independence. The nagging question of compound lotteries surfaces once again – although now in a normative context of discussion.

With the notion of agent-relative rationality to hand, we can reinforce the previous arguments on choice-relative rationality. Being detached from any personal features, these arguments provided only *possible* reasons for the paradoxical choices. But if the agents align their choices with general rules of conduct, and these rules agree with the arguments in question, then the possible reasons become actual and the normative status of the arguments is enhanced. The strong conclusion (6) is now within view. This needs a further condition, however: it must be the case that the paradoxical choices are *uniquely* compatible with the adopted rules of conduct. If the rules encompass both paradoxical and non-paradoxical choices, as in many non-EU construals of the ensuing literature, no argumentative progress towards (6) can take place. It is dubious that Allais went through all these points, but he was clearly tempted by the strong conclusion (6).[48]

5.3 The 'experimental definition of rationality'

Whether one targets the strong conclusion or just the weak one, it is not enough to say abstractly what rational individuals are; one also needs to be able to spot them.

[46]Friedman and Savage (1948) had proposed to represent risk attitudes in terms of the curvature properties of VNM utility functions, and Marschak (1950) had pursued the same line, introducing the terminology of 'risk-aversion' on this occasion. Given this context, Allais had to make very clear that he was not concerned with cautiousness in the VNM sense.

[47]Following one possible line, the 'abstract definition' entered Allais's argument only rhetorically, i.e. as an *ad hominem* argument against EU theorists, and Allais would give substantial importance only to the broader notion of rationality we have identified here. But this interpretation lacks clear textual evidence.

[48]This is perhaps clearest from the following: 'far from being rational, certain implications [of VNM independence] can to the contrary be entirely irrational in some psychological situations (for example, in the case above of a very prudent individual, for whom the psychological value increases still significantly around 100 million)' (1953b: 527–528).

Allais's expression, 'the experimental definition of rationality', raises this problem.[49] We now propose several solutions to it, each taking the form of a *possible experimental procedure*.

- Procedure (P1) involves setting particular rationality criteria in advance, using them to select the participating subjects, and then running the experiment. The objective is to find out what choices these presumably rational subjects make.
- Procedure (P2) does not select subjects in advance, but devises an experimental stage that makes selection possible. The objective is again to find out who the rational subjects are and what choices they make, but this time the selection is endogenous to the experiment.
- Procedure (P3) runs like (P2), and pursues the same objective, while also trying to learn more about what rationality substantially constitutes; for the data may raise questions that point beyond the conception of rationality as consistency, which still underlies (P2). Hence, this time it is rationality itself that the experiment contributes to determining.

(P1) is Allais's own procedure. By appealing to his intuition about the 'rational man', he preselected the recipients of his questions, first in informal asides at the Paris colloquium, and then more systematically with the famous 'poll'. Those whom he approached were primarily decision theorists, and as such capable of formulating rules of conduct for risky choices. Presumably, Allais expected nearly as much from his École des Mines students, another sophisticated audience. But it is of course dubious to correlate rationality with theoretical reputations and even just the passing of exams, and Allais's poll has been mocked – not without reason – for manifesting the prejudices of a meritocratic *polytechnicien*.

(P2), an immediate improvement, uses the experiment itself to select rational individuals. Since they are disposed to relate any particular choice they make to the rules they endorse, it is sufficient to put this disposition to the test. The protocols can be summarized as follows (more detail in the next section). They will combine stages of choices among lotteries, with reflective stages in which subjects evaluate the rules the experimenter submits to them for consideration. Subjects who endorse rules that run counter to these choices, or disapprove of rules favouring them, and nonetheless persist in making such choices, exclude themselves as being irrational.

(P3) is linked to (P2), but extends its explanatory aim. Here the notion of rationality not only serves to identify the choice-making subjects, but itself becomes an object of study. The protocols of (P2) may introduce several operative notions of consistency, and one might then try to investigate their coincidence or lack of same.[50] This inquiry remains within the purview of rationality as consistency, but the data produced by implementing (P2) may suggest more remote questions. For instance, the experiment might employ diverse stimuli for rational reflection – the experimentalist might intervene in person, or not, subjects might

[49]Allais's readers usually neglect this expression, presumably because it seems to involve a contradiction in terms. However, Roth (1995: 8) notices it.

[50]The consistency of choice with endorsed rules is the essence of (P2), but the procedure can also test the mutual consistency of rules that the subject endorses; the next section will give an example of this.

run through their reflections in isolation, or collectively, and so on. If these stimuli operate differently, a comparison is in order. It may also be that definite psychological characteristics emerge among the subpopulation of rational subjects – e.g. they might display more or less prudence than average. Such correlations would enrich the understanding of rationality with substantial, not merely formal, elements. All these questions belong more to the cognitive sciences than to decision theory as traditionally conceived. (P2) is already an ambitious and provocative procedure by ordinary standards, and the next section will focus on it, leaving (P3) as a mere sketch of an alternative perspective.

6. Experimental implementations of rationality

As documented in Section 3, it was in the early 1970s that the Allais paradox reached the stage of being seen as an empirical refutation of the EU hypothesis. In saying merely this, we do not yet do justice to the experimental studies of the time, and one objective of this section is to investigate them more thoroughly. These experiments not only produced quantitative evidence and explanatory arguments on violations of the hypothesis, but also aimed to probe the normative strength of these violations by somehow including the subjects' rationality in the experiments themselves. This purpose is the common thread running through the studies of MacCrimmon (1968), Moskowitz (1974), Slovic and Tversky (1974) and MacCrimmon and Larsson (1979). Kahneman and Tversky (1979) would eventually break that thread, obliterating the memory of their predecessors to whom we now propose to return. That Kahneman and Tversky opened a new era in decision-related studies is widely recognized, and the recent work on the history of behavioural economics has only reinforced this point.[51] We substantiate it here in a specific way: the break represented by the 1979 paper becomes most obvious by comparing it with the earlier ones in the same temporal sequence, and notably with Slovic and Tversky's, which was oriented very differently. A little-known reflective paper by Tversky (1975) is further evidence of a dramatic change of mind.

6.1 Early attempts at experimenting on rationality

MacCrimmon's study (1968) calls for special attention, because it influenced the others and made the connection with Allais's normative perspective.[52] This study proposes nothing less than probing 'the normative implications of the postulates' of decision theory (1968: 4). Five postulates are dealt with, each one paired with a specific experiment; among them, (A3) has a variant of the Allais paradox as its experimental counterpart. The chosen variant is questionable, however, as

[51]See Heukelom (2014), Hands (2015), Moscati (Forthcoming). Hands rightly stresses that Kahneman and Tversky did not content themselves with specializing in the empirical issues of decision making, but also delegated competence on normative issues to EU theorists. This 'normative turn', as Hands labels it, should of course not be confused with the alleged 'normative retreat' discussed in section 4.

[52]MacCrimmon (1968: 9) briefly acknowledges Allais for the leading idea in his article, which summarizes a doctoral dissertation that has otherwise remained unpublished (MacCrimmon 1965).

MacCrimmon's later paper with Larsson would indirectly recognize. Instead of considering only probability values, MacCrimmon assumes states of the world that realize them in particular ways, a translation that changes the meaning of (A3). This is the same confusion of frameworks we already noticed in early discussions of the EU hypothesis (see also Appendix 3).

This weakness compromises MacCrimmon's results, but not his general approach, which was to pay attention to the subjects' rationality within the experiment. In the case of the Allais paradox, subjects had first to choose between (analogues of) p_1 and q_1, and p_2 and q_2; they then received two 'prepared answers' in written form, each stating both a pair of choices and reasons justifying this pair. Allegedly, these answers came from other subjects in a comparable experiment. One of them had (q_1, q_2) and a defence paraphrasing VNM independence (or rather what takes its place here), the other had (p_1, q_2) with an argument drawing on the certainty effect, and subjects had to state in writing which one they selected and why they did so. This whole process was repeated twice over with minor variations. Finally, the experimenter conducted face-to-face interviews to check whether subjects had understood the experiment properly, and make them evaluate the competing justifications (as well as others that might have occurred to them).

This protocol corresponds to MacCrimmon's aim of probing 'the normative implications of the postulates' if one grants the Allaisian conception that the judgement of rational *individuals* can serve as a touchstone for the rationality of *choices*. By prompting a deliberation process among subjects, it operationalizes the idea of the rational individual within the experiment, and thus belongs to the (P2) type of our taxonomy. However, MacCrimmon does not implement (P2) properly. When stating his results, he emphasizes that the interview stage witnessed a shift towards conformity with (A3), but he does not say how many of the subjects who manifest the shift are rational in the desired sense of achieving consistency between their choices and reasons. That is, MacCrimmon uses his protocol to enhance the rationality of his subjects, but finally refrains from turning it into a selection device, as (P2) more specifically requires. The other papers of the sequence show the same vacillation except, to an extent, for the one by MacCrimmon and Larsson (1979), which is the last and most thorough of them all.[53]

The interview plays an equivocal role in MacCrimmon's protocol. It verifies that the subjects have understood the 'prepared answers', possibly identifies their other reasons, and collects their considered final positions. But clearly, the subjects might take as a persuasive argument something which the experimentalist thinks of being only as a semantic elucidation. Struck by this objection, Slovic and Tversky (1974) replaced the interview with another round of choices, while preserving a general *ternary* structure: a stage of initial choices, a stage of presenting reasons, and a final stage that encapsulates the normatively relevant information. The authors also modified MacCrimmon's first two stages by returning to a formulation of the paradox in terms of probability values alone, a change that could shift the third

[53]There is also some (P1) element in this early experiment. The subjects were business people, hence supposedly experts at making decisions. The choice problems were phrased as stylized financial examples.

stage in the paradoxical direction. Indeed, at this stage, they found that 17 out of 29 subjects chose in the Allais way.[54]

Not stopping there, Slovic and Tversky thought up another experimental sequence consisting of two stages only: the first devoted to evaluations, and the second to choices. This experiment uses the same choice problems and the same statement of reasons as before, but different subjects. In the first stage, Slovic and Tversky introduced the reasons to the subjects, and then asked them to evaluate those reasons on a numerical scale. The second stage consisted of choices. The choices turned out to be severely discrepant with the numerical marks given to the reasons (the reasons tending to favour Allais, the choices clearly conforming to EU). Slovic and Tversky drew no conclusions from the study overall because, quite obviously, they did not know how to reconcile this experiment with the previous one. The scepticism they finally expressed connects with the shift that Tversky would soon make, giving up any further experimental implementation of rationality, and turning instead to purely empirical research on behaviour. This was a pity, since the conflict between the two experiments called for performing yet others (admittedly, more in line with (P3) than (P2) according to our taxonomy).

Likewise starting out from MacCrimmon, Moskowitz (1974) had also arrived at a ternary structure, i.e. a stage of presenting reasons placed between two stages of choice. His study includes two notable innovations. First, he compares different visual modes of presentation for identical choice problems, thus foreshadowing the 'framing effects' of later decision psychologists.[55] Second, he includes in the intermediate stage a free discussion among the subjects. At the final stage, subjects moved towards the EU hypothesis, but a significant proportion remained faithful to the paradoxical choices.

The last study, by MacCrimmon and Larsson (1979), which we brought to the fore in Section 3, needs to be reconsidered here, since it takes up and actually pushes much further MacCrimmon's initial project of testing normative statements of decision theory.[56] Here the protocol returns to a *binary* structure: a stage of choices followed by a stage presenting the reasons for choice; perhaps as a result of intervening criticisms, the authors eschew the discussion with the subjects. They do not have a second round of choices either, and this may be explained by the sheer complexity of the first two stages, each of which is much richer than in any previous study. In the first stage, MacCrimmon and Larsson test their parametric representation of the common consequence and common relation effect, which leads them to investigate whole ranges of numerically defined choice problems. In the second stage, they introduce an extremely well-stocked inventory of reasons for the possible choices. For the first time, they present these as *general rules of conduct*, the previous practice having merely involved *ad hoc* rationalizations for the numerically defined choice problems (as in MacCrimmon's 'prepared answers'). Rules are generally formulated

[54]Furthermore, Slovic and Tversky transformed the second stage into a contradictory trial. Subjects who had chosen according to the EU hypothesis got an answer following an Allais-style argument, and those who had chosen against the hypothesis got an answer following a Savage-style argument.

[55]Allais's paradox has often been approached from this angle; see e.g. Bierman (1989).

[56]Compare MacCrimmon and Larsson (1979: 346, 350, 403) with MacCrimmon's (1968) claim to be probing 'the normative implications of the postulates'.

in the imperative mood. Some reproduce axioms from decision theory (like VNM independence in several variants), others are rules of thumb contradicting the axioms (like the one about preferring certainty in some circumstances), others reproduce weak necessary conditions for the axioms, and there are yet others relating to utility representations. Once subjects have been introduced to the rules, they are requested to mark them from 0 (complete disagreement) to 10 (complete agreement). This use of a numerical scale accords with Slovic and Tversky's work, but is more extensive here.

We have already covered the results of MacCrimmmon and Larsson's first stage. Concerning those of the second stage, the average marks show a lower agreement with EU-style rules than with the Allais-style rule of preferring certainty. This holds for both the common ratio effect (MacCrimmmon and Larsson 1979: 358) and the common consequence effect (1979: 368). However, the value of this finding is diminished by the observation that the average marks are rather high (above 5), suggesting a bias towards acceptance. The most interesting results exploit all the information *by cross-referencing choices and marks*. When dealing with the common ratio effect, MacCrimmon and Larsson divide their 19 subjects into three groups: those whose choices are entirely consistent with EU-style rules (3), those whose choices are only partly consistent with them (4), and those whose choices are clearly inconsistent with them (12). The average mark is very slightly higher for EU-style rules in the first group, and notably higher for the Allais-style rule in the third group. These data indicate some form of consistency between choices and approved reasons, but the data for the common consequence effect are unfortunately chaotic. As MacCrimmon and Larsson discover, subjects whose choices are entirely consistent with EU-style rules (10 out of 19) rank them on average *lower* than they do the Allais-style rule.

With this cross-comparison of the two stages of the experiment, MacCrimmon and Larsson come close to an endogenous selection of rational subjects, as recommended by (P2), but stop short of this objective. Thanks to the number of choice problems offered at the first stage, they are able, for each effect, to screen out subjects who choose irregularly (e.g. the 4 subjects of the second group of the common ratio effect). Quite rightly, they limit their tentative normative conclusions to regular subjects (e.g. the $3 + 12$ of the first and third groups of the same effect). However, we have just seen that, in respect to these, they produce only *averages*, and this is clearly the wrong kind of data. Only *individual* data would make it possible to screen out those regular subjects who approve rules that clash with their choices or disapprove rules that agree with these choices – that is, irrational subjects according to (P2). This is the more disappointing since their small population of subjects – 19 as against 38 in MacCrimmon (1968) and 29 and 39 in Slovic and Tversky's (1974) two experiments – made the collection of individual data a relatively easy task.

It is thus hardly surprising that MacCrimmon and Larsson can only offer a very broad qualitative conclusion: 'many careful, intelligent decision makers do seem to violate some axioms of expected utility theory, even upon reflection on their choices' (1979: 403). Logically, the strong conclusion (6) against the EU only needs an *existential* claim; in Allais's terms cited above, 'certain implications can to the contrary be entirely irrational in some psychological situations'. The problem was to establish this claim beyond doubt, which Allais had not yet managed to

do. MacCrimmon and Larsson fill the gap differently from Slovic and Tversky (in their first experiment). The latter have EU-contradicting answers in their second round of choices, which is arguably more rational than the first; the former had EU-contradicting pronouncements, and this carries special weight among rational men, in the (P2) sense of rationality. The two sets of results support each other, which is satisfactory from the viewpoint of establishing the target claim. However, given the high sophistication of MacCrimmon and Larsson's experiment, one would have expected more. In particular, by using their data to the full, they would have been able to quantify the proportion of rational subjects, as well as the sub-proportions of EU and Allais supporters among them.

MacCrimmon and Larsson offer interesting material for researchers who would like to pursue the study along (P3) lines. Given the large number of rules considered, it is possible to observe the subjects' patterns of endorsement and investigate how these take into account the logical interrelations of the rules. This criterion of *internal* consistency could be investigated both for itself and in relation to the *external* criterion of consistency between endorsed rules and choices, which is the one considered in (P2). The comparison might even be extended to a third notion of consistency that underlies MacCrimmon and Larsson's results, i.e. *regularity*, in the sense of making choices that are always pro or contra a given rule or set of rules (we have just seen how they use this notion in reporting on common ratio experiments). Such investigations would drag the experimental treatment of choice in the direction of today's cognitive sciences.

6.2 Some lessons for today's experimental economics

Table 1 summarizes the methodological features of the four reported studies and introduces a proposal of our own, which is meant to improve on the experimental implementations of rationality that guided the studies. The main technical innovation of this proposal is to combine Slovic and Tversky's ternary structure (choices followed by a presentation of reasons followed by choices) with MacCrimmon and Larsson's restatement of reasons in terms of general rules of conduct. At the same time, we would recommend pursuing the strategy of selecting rational subjects within the experiments more consistently than was done in any of the previous studies. Once again, this requires collecting individual data (very much like the panel data of micro-econometrics).

With the proposed ternary structure, the consistency between choices and endorsed rules will be assessed on two different sets of choice data. It is appropriate to give more weight to choices at the last round since the reflective stage influences them. With this in mind, we submit some guidelines to select subjects. Two kinds of experimental sequences would suggest an 'irrational man' classification:

- initial choices – endorsed rules of conduct in agreement with choices – final choices disagreeing with these reasons (hence also with the initial choices);
- initial choices – endorsed rules of conduct in disagreement with choices – final choices disagreeing with these reasons (for instance similar to the initial choices).

Table 1. Summary of the methodological features of the four reported studies

	Experimental sequence	Stage of reasons	Selection of rational subjects	Normative conclusion
MacCrimmon (1968)	Ternary (choices/ presentation of reasons/ interview)	'Prepared answers' (suggested choices with reasons for making them) *Accepted or rejected by subjects*	None (except perhaps exogenously)	Paradox dissolved
Slovic and Tversky (1974)	Ternary (choices/ presentation of reasons/ choices)	'Prepared answers' *Simply read*	None	Paradox upheld
	Binary (presentation of reasons/ choices)	'Prepared answers' *Scored by subjects*		Unclear
Moskowitz (1974)	Ternary (choices/ presentation of reasons/ choices)	'Prepared answers' *Read. Discussed between subjects*	None	Paradox upheld
MacCrimmon and Larsson (1979)	Binary (choices/ presentation of reasons)	General rules *Scored by subjects*	Some form of endogenous selection (based on the consistency of initial choices with endorsed rules)	Paradox upheld
Our proposal	Ternary (choices/ explanations/ choices)	General Rules *Scored by subjects. Possibly also discussed*	Endogenous selection (based on the consistency of final choices with endorsed rules)	To be determined

Conversely, two kinds of experimental sequences would suggest a 'rational man' classification:

- initial choices – endorsed rules of conduct in agreement with choices – final choices in agreement with these reasons (hence also with the initial choices);
- initial choices – endorsed rules of conduct in disagreement with choices – final choices agreeing with these reasons (hence not with the initial choices).

Subjects exhibiting the last sequence convey their effort at rationality by revising their initial position, but those exhibiting the penultimate sequence may be suspected of being simply inert. However, the second round of choices involves choice problems that differ from those of the first round, which means that subjects who consistently apply the same rule of conduct in the two rounds must have made

an effort of thought. At least, they are less likely to be inert than if they were simply asked whether, on reflection, they approve or disapprove of their previous choices.[57] While the proposed screening is no doubt imperfect, it suggests that one can nonetheless implement (P2). This would also require increasing the number of subjects compared with the initial studies. We only know of one current attempt to explore the avenue opened up here.[58]

7. Conclusion

Allais's paradox, and more generally his critique of the EU hypothesis, became famous at the cost of becoming impoverished. While we tried to do justice to the project of empirically refuting and superseding the EU hypothesis, we meant primarily to change this standard perspective on the paradox and related counterexamples. Our recurring claim, i.e. that Allais made them part of a *normative* argument against the hypothesis, appears to be supported both by the historical account, especially concerning the 1952 Paris conference, and the conceptual analysis we appended to this account. The two main ideas around which Allais's argument revolves are that of the 'rational man' *versus* the 'real man', and that of 'the experimental conception of rationality'. Filling the gaps in Allais's allusive treatment, we have clarified the former idea, which neither fully breaks nor exactly coincides with the economist's view of rationality as consistency; and we have expanded at length on the latter idea, which is more obviously heterodox. To make sense of it, we have returned to intellectual history, revisiting early and long-forgotten works that attempted to capture individual rationality *within* the experiment. That they address the initial sense of the paradox is precisely the reason why they have been wiped out of the collective memory. Despite being imperfect, they point to interesting avenues that today's experimental work on decision could fruitfully explore.

Acknowledgements. This paper evolved from a French version on which the author received useful feedback during presentations at conferences and seminars. The author particularly benefitted from the comments and suggestions made on this earlier version or the present one by Mohammed Abdellaoui, Jean Baccelli, Richard Bradley, Patrick Colyvan, Vincent Eli, Itzhak Gilboa, Raphaël Giraud, Dorian Jullien, Edi Karni, Jean-Sébastien Lenfant, Ivan Moscati, Bertrand Munier, Samir Okasha, Peter Wakker and Paul Weirich. Ben Young's linguistic assistance considerably improved the present text.

Financial support. The author also thanks the Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) for supporting this research.

[57]Gilboa (2011: 17–18) also entertains a scheme for 'experimentally delineating the scope of rationality'. This scheme differs from ours in being binary, with one round of choices followed by an 'analysis' presented to the subjects, and also in being liberal concerning the way this 'analysis' works (we specifically require that rules of conduct be presented to the subjects). Gilboa recognizes the possibility that inert people may be classified as rational. Among other purposes, the second round of choices partly restricts this possibility.

[58]Eli's (2017) doctoral thesis has one experimental implementation. The results identify a subgroup of rational subjects, among which those endorsing an Allais-style rule based on the certainty effect are more numerous than those endorsing an EU-style rule. This conclusion gives some quantitative flesh to MacCrimmon and Larsson's merely qualitative claim. It needs to be double-checked using a wider set of rules, and compared with the results obtained when the experimenter steers the subjects' reflections just by asking them to perform several choices in succession (on this strand of literature, see van de Kuilen and Wakker 2006).

References

Fondements et applications de la théorie du risque en économétrie, Colloques internationaux du Centre National de la Recherche Scientifique, XL Econométrie, Paris, Centre National de la recherche scientifique, 1953. (Proceedings of the conference held in Paris from May 12th to 17th, 1952.) Cited here as *Econométrie*.[59]

Allais M. 1952. Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine. *Econométrie* 257–331.

Allais M. 1953a. La psychologie de l'homme rationnel devant le risque. La théorie et l'expérience. *Journal de la Société de Statistique de Paris* **94**, 47–73.

Allais M. 1953b. Le comportement de l'homme rationnel devant le risque. *Econometrica* **21**, 503–546.

Allais M. 1979. The so-called Allais paradox and rational decisions under uncertainty. In M. Allais and O. Hagen (eds), *Expected Utility and the Allais Paradox*, Dordrecht: D. Reidel, pp. 437–683.

Allais M. 1987. Allais's paradox. In *The New Palgrave. A Dictionary of Economics*. London: Macmillan.

Allais M. 1994. An outline of my main contributions to Risk and Utility Theory: theory, experience, and applications: general overview. In B. Munier and M. J. Machina (eds), *Models and Experiments in Risk and Rationality*. Dordrecht: Kluwer, pp. 173–221.

Allais M. and O. Hagen eds. 1979. *Expected Utility and the Allais Paradox*. Dordrecht: D. Reidel.

Andreoni J. and C. Sprenger 2009. *Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena*. Manuscript, University of California, San Diego.

Bierman H. 1989. The Allais paradox: a framing perspective. *Behavioral Science* **34**, 46–52.

Birnbaum M. H. 1998. The paradoxes of Allais, stochastic dominance, and decision weights. In J. C. Shanteau, B. A. Mellers and D. Schum (eds), *Decision Research from Bayesian Approaches to Normative Systems: Reflections on the Contributions of Ward Edwards*. Norwell: Kluwer, pp. 27–52.

Birnbaum M. H. 2004. Causes of Allais common consequence paradoxes: an experimental dissection. *Journal of Mathematical Psychology* **48**, 87–106.

Birnbaum M. H., G. Coffey, B. A. Mellers and R. Weiss 1992. Utility measurement: configural weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance* **18**, 331–346.

Borch K. and J. Mossin eds. 1968. *Risk and Uncertainty*. London: Macmillan.

Bourgeois-Gironde S. and R. Giraud 2008. Le tournant cognitif en économie de la décision et des comportements. In B. Walliser (ed.), *Economie et Cognition*. Paris: Ophrys, pp. 55–91.

Bouyssou D. and J. C. Vansnick 1990. Utilité cardinale' dans le certain et choix dans le risque. *Revue économique* **41**, 979–1000.

Camerer C. 1995. Individual decision making. In A. E. Roth and J. H. Kagel (eds), *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press, pp. 587–703.

Cerreia-Vioglio S., D. Dillenberger and P. Ortoleva 2015. Cautious expected utility and the certainty effect. *Econometrica* **83**, 693–728.

Cohen M. 1992. Security level, potential level, expected utility: a three-criteria decision model under risk. *Theory and Decision* **33**, 101–134.

Conlisk J. 1989. Three variants on the Allais example. *American Economic Review* **79**, 392–407.

Cubitt R. P., C. Starmer and R. Sugden 1991. Dynamic choice and the common ratio effect: an experimental investigation. *Economic Journal* **108**, 1362–1380.

Diemer A., J. Lallement and B. Munier 2010. *Maurice Allais et la science économique*. Paris: Clément Juglar.

Drèze J. H. 1989. Maurice Allais and the French Marginalist School. *Scandinavian Journal of Economics* **91**, 5–16.

Edwards W. 1954. The theory of decision making. *Psychological Bulletin* **51**, 380–417.

Edwards W. 1962. Subjective probabilities inferred from decisions. *Psychological Review* **69**, 109–135.

Eli V. 2017. *Essays in Normative and Descriptive Decision Theory*. Doctoral thesis, Université Paris-Saclay and HEC Paris.

[59]We use the date 1952 to refer to the papers in this collection, in order to distinguish them more clearly from other published versions.

Fishburn P. C. 1989. Retrospective on the utility theory of von Neumann and Morgenstern. *Journal of Risk and Uncertainty* **2**, 127–158.

Fishburn P. C. and P. Wakker 1995. The invention of the independence condition for preferences. *Management Science* **41**, 1130–1144.

Friedman M. and L. J. Savage 1948. The utility analysis of choices involving risk. *Journal of Political Economy* **56**, 278–304.

Friedman M. and L. J. Savage 1952. The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy* **60**, 463–474.

Gilboa I. 1988. A combination of expected utility and maxmin decision criteria. *Journal of Mathematical Economics* **32**, 405–420.

Gilboa I. 2011. Questions in decision theory. In K. R. Apt (ed.), *Proceedings of Tark XIII*. ACM Publications, pp. 17–27.

Gilboa I. and D. Schmeidler 2001. *A Theory of Case-Based Decisions*. Cambridge: Cambridge University Press.

Gilboa I., F. Maccheroni, M. Marinacci and D. Schmeidler 2010. Objective and subjective rationality in a multiple prior model. *Econometrica* **78**, 755–770.

Grandmont J. M. 1989. Report on Maurice Allais's scientific work. *Scandinavian Journal of Economics* **91**, 17–28.

Guala F. 2000. The logic of normative falsification: rationality and experiments in decision theory. *Journal of Economic Methodology* **7**, 59–93.

Hagen O. 1972. A new axiomatization of utility under risk. *Teorie e Metoda* **4**, 55–80.

Hands D. W. 2015. *Normative Rational Choice Theory: Past, Present and Future*. University of Puget Sound Working Paper.

Herstein I. N. and J. Milnor 1953. An axiomatic approach to measurable utility. *Econometrica* **21**, 291–297.

Heukelom F. 2014. *Behavioral Economics. A History*. New York, NY: Cambridge University Press.

Heukelom F. 2015. A history of the Allais paradox. *British Journal for the History of Science* **48**, 147–169.

Hey J. D. 1991. *Experiments in Economics*. Oxford: Blackwell.

Hey J. D. and C. Orme 1994. Investigating generalizations of expected utility theory using experimental data. *Journal of Risk and Uncertainty* **62**, 1291–1326.

Jaffray J. Y. 1988. Choice under risk and the security factor. *Theory and Decision* **24**, 169–200.

Jallais S. and P. C. Pradier 2005. The Allais paradox and its immediate consequences for expected utility theory. In P. Fontaine and R. Leonard (eds), *The 'Experiment' in the History of Economics*. London: Routledge, pp. 25–49.

Jallais S., P. C. Pradier and D. Teira 2008. Facts, norms and expected utility functions. *History of the Human Sciences* **21**, 45–62.

Kahneman D. and A. Tversky 1979. Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–291.

Karmakar U. S. 1978. Subjectively weighted utility: a descriptive extension of the expected utility model. *Organization Behavior and Human Performance* **21**, 61–72.

Karmakar U. S. 1979. Subjectively weighted utility and the Allais paradox. *Organization Behavior and Human Performance* **24**, 67–72.

Karni E. and D. Schmeidler 1991. Atemporal dynamic consistency and expected utility theory. *Journal of Economic Theory* **54**, 401–408.

Leonard R. 1995. From parlor games to social science: von Neumann, Morgenstern and the creation of game theory, 1928–1944. *Journal of Economic Literature* **33**, 730–761.

Luce R. D. and H. Raiffa 1957. *Games and Decisions*. New York, NY: Wiley.

MacCrimmon K. R. 1965. *An Experimental Study of the Decision-Making Behavior of Business Executives*. Ph.D. dissertation, University of California, Los Angeles.

MacCrimmon K. R. 1968. Descriptive and normative implications of the decision-theory postulates. In K. Borch and J. Mossin (eds), *Risk and Uncertainty*. London: Macmillan, pp. 3–23.

MacCrimmon K. R. and S. Larsson 1979. Utility theory: axioms versus 'paradoxes'. In M. Allais and O. Hagen (eds), *Expected Utility and the Allais Paradox*. Dordrecht: D. Reidel, pp 333–409.

Machina M. J. 1983. Generalized expected utility analysis and the nature of observed violations of the independence axiom. In B. Stigum and F. Wenstop (eds), *Foundations of Utility and Risk theory with Applications*. Dordrecht: Reidel, pp. 263–293.

Machina M. J. 1987. Choices under uncertainty: problems solved and unsolved. *Journal of Economic Perspectives* **1**, 121–154.

Machina M. J. 1989. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature* **27**, 1622–1668.

Machina M. J. 1991. Dynamic consistency and non-expected utility. In M. Bacharach and S. Hurley (eds), *Foundations of Decision Theory*. Oxford: Blackwell, pp. 39–91.

Marschak J. 1950. Rational behavior, uncertain prospects and measurable utility. *Econometrica* **18**, 111–141.

Mongin P. 1988. Problèmes de Duhem en théorie de l'utilité espérée. *Fundamenta Scientiae* **9**, 289–317.

Mongin P. 2009. Duhemian themes in expected utility theory. In A. Brenner and J. Gayon (eds), *French Studies in the Philosophy of Science*. New York, NY: Springer, pp. 303–357.

Mongin P. 2012. Une source méconnue de la théorie de l'agrégation des jugements. *Revue économique* **63**, 645–657.

Morrison D. G. 1967. On the consistency of preferences in Allais' paradox. *Behavioral Science* **12**, 373–383.

Moscati I. 2013. How cardinal utility entered economic analysis. *European Journal of Economic Thought* **20**, 906–939.

Moscati I. 2016. How economists came to accept expected utility theory: the case of Samuelson and Savage. *Journal of Economic Perspectives* **30**, 219–236.

Moscati I. (Forthcoming). *Measuring Utility: From the Marginal Revolution to Neuroeconomics*. Oxford: Oxford University Press.

Moskowitz H. 1974. Effects of problem representation and feedback on rational behavior in Allais and Morlat-type problems. *Decision Sciences* **5**, 225–241.

Mosteller F. and P. Nogee 1951. An experimental measurement of utility. *Journal of Political Economy* **59**, 371–404.

Munier B. 1991. The many other Allais paradoxes. *Journal of Economic Perspectives* **5**, 179–189.

Munier B. 1995. Fifty years of Maurice Allais' economic writings: seeds for renewal in contemporary economic thought. In B. Munier (ed.), *Markets, Risk and Money. Essays in Honor of Maurice Allais*. Dordrecht: Kluwer, pp. 1–50.

Munier B. 2011. Maurice Allais, précurseur et devancier de l'analyse du risque contemporain. In A. Diemer, J. Lallement and B. Munier (eds), *Maurice Allais et la science économique*. Paris: Clément Juglar, pp. 75–94.

Preston M. G. and P. Barratta 1948. An experimental study of the auction-value of an uncertain outcome. *American Journal of Psychology* **61**, 183–193.

Raiffa H. 1968. Decision analysis. In *Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley.

Roth A. E. 1995. Introduction to experimental economics. In A. E. Roth and J. H. Kagel (eds), *Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press, pp. 3–109.

Rubin H. 1949. The existence of measurable utility and psychological probability. Cowles Commission Discussion Paper: Statistics 332.

Rubinstein A. 1988. Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory* **46**, 145–153.

Samuelson P. A. 1952a. Utilité, préférence et probabilité. *Econométrie*, 141–150. English version: Utility, preference, and probability. In *Collected Economic Papers, I*, 127–136. Cambridge, MA: MIT Press, 1966.

Samuelson P. A. 1952b. Probability, utility, and the independence axiom. *Econometrica* **20**, 670–678.

Savage L. J. 1951. The theory of statistical decision. *Journal of the American Statistical Association* **46**, 55–67.

Savage L. J. 1952. Une axiomatisation du comportement raisonnable face à l'incertitude. *Econométrie*, 29–33.

Savage L. J. 1954. *The Foundations of Statistics*. New York, NY: Dover (2nd edn, 1972).

Schoemaker P. J. H. 1982. The expected utility model: its variants, purposes, evidence and limitations. *Journal of Economic Literature* **20**, 529–563.

Segal U. 1990. Two-stage lotteries without the reduction axiom. *Econometrica* **58**, 349–377.

Slovic P. and A. Tversky 1974. Who accepts Savage's axioms? *Behavioral Science* **19**, 368–378.

Starmer C. 2000. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* **38**, 332–382.
Tversky A. 1975. A critique of expected utility theory: descriptive and normative considerations. *Erkenntnis* **9**, 163–173.
Tversky A. and D. Kahneman 1986. Rational choice and the framing of decisions. *Journal of Business* **59**, S251–S278.
Tversky A. and D. Kahneman 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5**, 297–323.
van de Kuilen G. and P. Wakker 2006. Learning in the Allais paradox. *Journal of Risk and Uncertainty* **33**, 155–164.
von Neumann J. and O. Morgenstern 1944. *The Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press (2nd edn, 1947).
Wakker P. 1999. *Justifying Bayesianism by Dynamic Principles*. Mimeo: Leiden University Medical Center.
Wakker P. 2010. *Prospect Theory*. Cambridge: Cambridge University Press.
Weber B. J. 2008. The effects of payout and probability magnitude on the Allais paradox. *Memory and Cognition* **5**, 1013–1023.

Appendix 1. Morlat's counter-example

At the 1952 Paris conference, Morlat proposed the following counter-example (*Économétrie*: 156–157). The agent has to choose between p_1, and q_1, and between p_2 and q_2:

p_1: you are guillotined with prob 0.01 you win 1000 F with prob 0.49 you win 2000 F with prob 0.50	p_2: you are guillotined with prob 0.49 you win 1000 F with prob 0.01 you win 2000 F with prob 0.50
q_1: you are guillotined with prob 0.02 you win 1000 F with prob 0.48 you win 1,000,000 F with prob 0.50	q_2: you are guillotined with prob 0.50 you win 1,000,000 F with prob 0.50

The typical choice pair (p_1, q_2) violates the EU hypothesis, and, given the principle of compound lotteries, VNM independence. Morlat proposed his counter-example during the joint discussion of Allais's and Samuelson's papers. In the same normative spirit as most discussants, he writes: 'many men thought to be reasonable do prefer p_1 to q_1, but q_2 to p_2, because, in this case, the risks of death seem very similar, whereas if they were saved, they would be richer than in q_2. ... This is the outcome I have had from many people I have asked, and I do not think for myself this position to be unreasonable.'

While the Allais paradox illustrates both the certainty and consequence effects, Morlat's counter-example formally relates only to the latter; however, the very small probabilities for death in p_1 and q_1 are evocative of the former. The counter-example has vanished from collective memory, but the early authors saw it well enough. Savage (1954 [1972]: 101) cited it in his well-known passage on the Allais paradox, Moskowitz (1974) used it in an experiment, and MacCrimmon and Larsson (1979) took account of it. In Borch and Mossin (1968: 28), one of MacCrimmon's discussants mentions a 'test' performed by Morlat in 1952, which probably refers to this counter-example. To the best of our knowledge, Allais never mentions it.

Appendix 2. Samuelson's argument for the independence condition

In his conference paper, Samuelson (1952a: 147) stated a VNM independence condition ('strong independence') that is exactly (A3), and he added the following justificatory argument for this condition. Assume an individual has already formed a preference between two lotteries p and q, and is now comparing the two compound lotteries $\alpha p + (1-\alpha)r$ and $\alpha q + (1-\alpha)r$. The preference between these lotteries cannot depend on what they offer if the event of probability $1-\alpha$ realizes, since this is r in either case. If the preference only depends on what the event of probability α brings about, it must be the same as

that existing between p and q. This argument supports (A3) in the 'only if' direction, but is easy to adapt to the 'if' direction.

Of all the participants at the colloquium, only Allais (*Économétrie*: 159) emphasized that Samuelson's problem was *dynamic* in nature, and that this changed the nature of the solutions. Samuelson reconstructs the individual's *ex ante* preferences in terms of *ex post* preferences (i.e. those following the first resolution of uncertainty) and this is questionable from a rationality point of view. In particular, the individual could *ex ante* take into account the complementarity of r with either p or q. Here Allais clearly echoes one of his two normative arguments for the paradoxical choices (see Section 5). In the *Econometrica* paper, Allais takes up the same point: 'There is a considerable difference between an *ex ante* reasoning and an *ex post* reasoning, as the *ex post* reasoning eliminates an essential component, to wit the composition operation of aleatory prospects' (1953b: 538). With some generosity, one may read this passage as an ancestor of Machina's (1989) dynamic argument for making paradoxical choices. More generally, one may see the brief discussion between Allais and Samuelson as marking out the debate that was to come on the contrasting dynamic principles that underlie EU and non-EU hypotheses.[60]

This is not all that needs saying concerning Samuelson's argument. Surreptitiously, he had changed the VNM framework of uncertainty into a state-of-the-world framework that makes the defence of the EU hypothesis easier. In Appendix 3 we will explain this shift while reviewing a famous argument by Savage, in which it also occurs.

Appendix 3: Savage's response to Allais and two representations of the Paradox

In a well-known passage of *The Foundations of Statistics*, Savage (1954[1972]: 101–104) revisits Allais's paradox, in a version that slightly changes the monetary amounts but keeps the probability values.[61] As he reports, he had first been trapped in the paradoxical choices (p_1, q_2), but on deeper reflection irrevocably chose (p_1, q_1), one of the two EU-conforming pairs. Savage relates his revised judgement to a general procedure for dealing with a 'normative theory' (1954[1972]: 102). This consists in replacing the theory in those concrete situations where it seems to 'lead [one] astray', and deciding for each such case 'whether to retain [the] initial impression, or to accept the implications of the theory' (1954[1972]: 102).

This sketch follows the principle of 'reflective equilibrium', which moral philosophy, and even everyday moral reflection, has long applied. It is noteworthy that Savage's adherence to this procedure links decision theory to other disciplines that are unequivocally normative. The experimental work covered in Section 6 also resorts to 'reflective equilibrium', although in a complex form that needs to be contrasted with the simple form that Savage illustrates. First, the 'reflective equilibrium' there is not introspective. The experimenter induces subjects to perform a 'reflective equilibrium' of their own, and faces the problem of deciding from the outside whether they have effectively achieved this result. Second, in the experimental work, subjects first make choices without generally being aware of the rules of conduct they obey or violate, whereas the simple form presupposes a knowledge of the rules of conduct to be put to the test. Third, this simple form is concerned with one rule of conduct at a time, which is not the case in the experimental work.

The fact that Savage answered the Allais paradox along normative lines has led some to conclude that it took Allais's paradox for Savage to adopt a normative attitude concerning the VNM axioms and his own new postulates. This is a flawed conclusion, since the evidence of Section 4 shows that Savage already adopted a normative attitude before discovering the Allais paradox. The dubious 'normative retreat' construction, which we discuss in Section 4, largely depends on a misapprehension of what Savage meant before publishing the *Foundations*.

In our view, Savage's answer to Allais matters more for a technical than a methodological reason. It illustrates a confounding of the paradox, which is formulated in terms of *probability values*, as befits its objective of undermining VNM theory, with a related but different paradox formulated in terms of *states of the world*, which is not relevant to VNM theory as such. Savage's train of thought suggests to him 'one way in which [Allais's gambles] could be realized' (1954[1972]: 103). There are 100 equiprobable

[60]Samuelson (1952b: 672) concisely reproduces the argument he made at the conference and does not mention Allais's objection.

[61]We cite the 1972 edition, but the 1954 edition already had the argument.

tickets, and the four lotteries of the paradox now correspond to four allocations of money amounts to these tickets.

	Ticket 1	*Tickets 2–11*	*Tickets 12–100*
p^*_1	\$500,000	\$500,000	\$500,000
q^*_1	0	\$2,500,000 \$	\$500,000
p^*_2	\$500,000 \$	\$500,000 \$	0
q^*_2	0	\$2,500,000	0

When comparing p^*_1 and q^*_1, Savage only considers tickets 1–11, since tickets 12–100 do not distinguish these two from each other, and by a similar argument when comparing p^*_2 and q^*_2 he only considers tickets 1–11. But truncating the options in this way makes the two problems of choice identical, and the only possible choice pairs are thus (p^*_1, p^*_2) or (q^*_1, q^*_2), among which Savage prefers the former.

As Hagen (1972) first pointed out, this line of reasoning is specious, since other allocations of prizes to tickets also respect the probabilistic data of the paradox, while not leading to identical choice problems by truncation. In the next example, p^*_1, p^*_2 and q^*_1 are as before, but q^*_2 is changed into q^{**}_2, which moves the 10 chances of gain to tickets 91–100 instead of 2–11:

	Ticket 1	*Tickets 2–11*	*Tickets 12–90*	*Tickets 91–100*
p^*_1	\$500,000	\$500,000	\$500,000	\$500,000
q^*_1	0	\$2,500,000	\$500,000	\$500,000
p^*_2	\$500,000	\$500,000	0	0
q^{**}_2	0	0	0	\$2,500,000

After pruning p^*_1 and q^*_1, and p^*_2 and q^*_2, of their common outcomes on tickets 12–90, one does *not* get identical pairs, unlike what happened in Savage's example. Either strategically or inadvertently, he had selected the kind of special cases that makes it possible to unfold his normative argument. Allais (1979: 533–536) took up Hagen's analysis when belatedly responding to Savage's response to his paradox.

MacCrimmon (1968) implicitly follows Savage in his first experiment. He considers an urn containing 100 equiprobable balls and offers the following choices to his subjects:

	Balls 1–10	Ball 11	Ball 12–100
Investment p^*_1	500% yield	loss	5% yield
Investment q^*_1	5% yield	5% return	5% yield
Investment p^*_2	500% yield	loss	loss
Investment q^*_2	5% yield	5% return	loss

At the stage of initial choices, MacCrimmon found that about 40% of subjects made paradoxical choices. At the reflective stage of the experiment, the results were mixed, with some inconsistency between the subjects' initial choices and their endorsement of 'prepared answers'. One of these 'prepared answers' amounted to spelling out Savage's argument to the effect that pairwise choice should neglect common outcomes. It is useful to register that even when the conditions for applying this argument hold, the subjects' initial choices are far from all being in agreement with the EU hypothesis, but MacCrimmon's reflective stage is disappointingly uninformative.[62]

[62]Tversky and Kahneman (1992) tested a variant of the Allais paradox using states of the world, just as MacCrimmon did in 1968, and the significant violation rate they found reinforces the latter's finding on the subjects' initial choices. Wakker (2010: 134) refers to these experiments when he claims that 'the Allais paradox is relevant to uncertainty as much as it is to risk'.

By not recognizing that their reformulation of the Allais paradox changes its content, Savage and MacCrimmon participate in a confusion which this paper has repeatedly identified, i.e. between the probability-based and state-based axiomatizations of the EU hypothesis. We found Allais himself falling prey to this confusion (see Section 2). In due course, decision theorists would learn how to distinguish the two axiomatizations, and correspondingly the two conditions of VNM independence and Savage's 'sure-thing principle'. MacCrimmon and Larsson (1979: 344–345) competently identify the two conditions, and they make relevant use of the difference between them at the reflective stage of their experiments.

The literature has discussed Savage's and Samuelson's arguments separately, whereas they share the same logical structure and weakness. Indeed, regardless of Allais's objections against Samuelson's postulated equivalence between a static and dynamic statement of VNM independence, Samuelson, like Savage, can be blamed for selecting an all-too-convenient states-of-the-world structure among all those fitting the probabilistic data. Instead of placing lotteries p and q on an event with probability α, and lottery r on the complementary event, one can relate the three lotteries to the following partition of four events:

	E of prob α-ε	F of prob ε	G of prob ε	H of prob $1-\alpha-\varepsilon$
Lottery $\alpha p + (1-\alpha)r$	p	r	p	r
Lottery $\alpha q + (1-\alpha)r$	q	q	r	r

Philippe Mongin is *Directeur de recherche* (Senior Research Professor) at Centre National de la Recherche Scientifique (C.N.R.S.) and a member of the Economics and Decision Sciences team of the HEC School of Management, Paris. He has contributed several books and more than one hundred articles on normative economics and collective decision theory, the philosophy of economics and the social sciences, and philosophical logic. He was the general editor of *Economics and Philosophy* from 1995 to 2000, and he has repeatedly published in this journal. His current work is concerned with conceptual and technical issues in the foundations of decision theory.

Cite this article: Mongin P (2019). The Allais paradox: what it became, what it really was, what it now suggests to us. *Economics and Philosophy* **35**, 423–459. https://doi.org/10.1017/S0266267118000469