Bulletin de la Sabix

Société des amis de la Bibliothèque et de l'Histoire de
l'École polytechnique

66 | 2020

Maurice Allais (1911-2010, X1931)

Le paradoxe d'Allais : une notoriété aussi universelle que justifiée

Eric Barthalon

Édition électronique

URL : <https://journals.openedition.org/sabix/2812>

DOI : 10.4000/sabix.2812

ISSN : 2114-2130

Éditeur

Société des amis de la bibliothèque et de l'histoire de l'École polytechnique (SABIX)

Édition imprimée

Date de publication : 1 octobre 2020

Pagination : 117-134

ISSN : 0989-3059

Référence électronique

Eric Barthalon, « Le paradoxe d'Allais : une notoriété aussi universelle que justifiée », *Bulletin de la Sabix* [En ligne], 66 | 2020, mis en ligne le 21 janvier 2021, consulté le 15 février 2023. URL : <http://journals.openedition.org/sabix/2812> ; DOI : <https://doi.org/10.4000/sabix.2812>

Tous droits réservés

Le paradoxe d'Allais : une notoriété aussi universelle que justifiée

Eric Barthalon*

Un paradoxe, nous disent les philosophes, c'est « *ce qui est contraire à l'opinion généralement admise, à la prévision ou à la vraisemblance* ». ¹ On remarquera que cette définition décrit l'effet d'un paradoxe, qui est de contredire ce que Maurice Allais appelait volontiers une « vérité établie », plus que sa nature même, où se trouve la cause de cet effet. Il y a donc paradoxe et paradoxe. Ainsi, le paradoxe formulé par Zénon d'Élée, le paradoxe d'Achille et de la tortue, se présente-t-il sous la forme d'un raisonnement parfaitement logique, du moins sous certaines hypothèses, dont la conclusion est que, même si Achille court plus vite que ne se déplace une tortue placée devant lui, jamais il ne pourra la rattraper, car lorsqu'Achille atteint le point de départ de la tortue celle-ci a avancé jusqu'à un nouveau point de départ, et ainsi de suite. ² Ce paradoxe n'en est un que parce que Zénon écarte, par hypothèse implicite, le

fait que la distance entre Achille et la tortue puisse converger vers zéro. ³

Le problème d'Allais

Ce qu'il est convenu d'appeler le paradoxe d'Allais est d'une toute autre facture. Tiré de l'expérimentation, c'est une observation, un contre-exemple conçu pour montrer qu'il y a des cas de figure où n'est pas valide la théorie de l'utilité espérée, telle que von Neumann et Morgenstern l'ont, en 1944, formulée, puis, en 1947, axiomatisée. ⁴ Selon cette théorie, lorsqu'un décideur rationnel doit choisir entre plusieurs combinaisons de gains (ou de pertes) possibles, chacun plus ou moins élevé, chacun plus ou moins probable, il doit toujours opter pour la perspective aléatoire dont l'espérance mathématique de l'utilité est la plus élevée.

* Responsable de la recherche sur les marchés de capitaux, Allianz SE.

1. Lalande, A., (1972), *Vocabulaire technique et critique de la philosophie*, P.U.F, 10^e édition.

2. Université d'Oxford, (1989), *Dictionnaire de l'Antiquité*, Editions Robert Laffont, Paris, 1993.

3. On pense ici à l'observation d'Henri Poincaré selon laquelle les « hypothèses dangereuses [...] ce sont d'abord, ce sont surtout celles qui sont tacites et inconscientes. Puisque nous les faisons sans le savoir, nous sommes impuissants à les abandonner » in *La science et l'hypothèse*.

4. Von Neumann, J. and Morgenstern, O., (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey, 3rd edition. C'est dans la deuxième édition de ce livre, celle de 1947, que sont énoncés différents axiomes.

Or, si – comme l’a fait Allais en 1952 – on demande à des sujets réputés rationnels de choisir entre, d’une part, les perspectives aléatoires A_1 et B_1 :

A_1 : 100 % de chance de gagner 100
 B_1 : 1 % de chance de ne rien gagner ; 89 % de chance de gagner 100 ; 10 % de chance de gagner 500

et, d’autre part, les perspectives aléatoires A_2 et B_2 :

A_2 : 89 % de chance de ne rien gagner ; 11 % de chance de gagner 100
 B_2 : 90 % de chance de ne rien gagner ; 10 % de chance de gagner 500

on observe qu’une majorité d’entre eux préfèrent A_1 à B_1 et B_2 à A_2 . Précision très importante : il s’agit de choix uniques, la possibilité de les effectuer ne se présente qu’une seule fois. Il ne s’agit pas d’un jeu répétitif où l’on pourrait compter sur la loi des grands nombres pour que s’opère une convergence des fréquences observées vers les fréquences théoriques. Pour comprendre en quoi ces choix contredisent la théorie de l’utilité espérée, il n’est pas inutile de visualiser ces quatre perspectives aléatoires comme on le fait dans la Table 1.

Table 1

Gains	A_1	B_1	A_2	B_2
0	0.00	0.01	0.89	0.90
100	1.00	0.89	0.11	0.00
500	0.00	0.10	0.00	0.10

Suivant la théorie de l’utilité espérée, u étant la fonction d’utilité d’un décideur donné, nous avons, d’une part :

$$u(A_1) = 1.00u(100)$$

$$u(B_1) = 0.01u(0) + 0.89u(100) + 0.10u(500)$$

d’où il résulte que

$$u(A_1) - u(B_1) = 0.11u(100) - 0.01u(0) - 0.10u(500)$$

et d’autre part

$$u(A_2) = 0.89u(0) + 0.11u(100)$$

$$u(B_2) = 0.90u(0) + 0.10u(500)$$

d’où il découle que

$$u(A_2) - u(B_2) = 0.11u(100) - 0.01u(0) - 0.10u(500)$$

Quelle que soit la fonction d’utilité, on a donc :

$$u(A_1) = u(B_1) + \Delta$$

$$u(A_2) = u(B_2) + \Delta$$

Il apparaît ainsi que la préférence $A_1 \succ B_1$ implique $\Delta > 0$, alors que la préférence $B_2 \succ A_2$ implique le contraire, $\Delta < 0$. Si Δ contribuait de la même manière à l’utilité de A_1 et à celle de A_2 , on devrait observer soit les préférences $A_1 \succ B_1$ et $A_2 \succ B_2$ si $\Delta > 0$, soit les préférences $B_1 \succ A_1$ et $B_2 \succ A_2$ si $\Delta < 0$.⁵ Autrement dit, 1 % de chance en plus ou en moins de ne rien gagner, 11 % de chance en plus ou en moins de gagner 100 et 10 % de chances en plus ou en moins de gagner cinq fois plus contribuent à l’utilité totale d’une perspective aléatoire quelconque d’une manière qui dépend du contexte, c’est-à-dire de l’ensemble de la distribution de probabilités ou encore de ce que Bertrand Munier a

5. Plus directement, mais moins significativement, on peut observer que
 $A_1 \succ B_1 \Rightarrow 0.11u(100) > 0.01u(0) + 0.10u(500)$,
alors que
 $B_2 \succ A_2 \Rightarrow 0.11u(100) < 0.01u(0) + 0.10u(500)$

appelé la structure du risque.⁶ Maintes fois confirmée depuis 1952, cette observation expérimentale contredit l'axiome d'indépendance sur lequel s'appuie, entre autres postulats, la théorie de l'utilité espérée.

F , G et H étant les utilités de trois événements possibles, α étant une probabilité, cet axiome d'indépendance peut être formalisé de la manière suivante :

$$F \succcurlyeq G \Rightarrow \alpha F + (1 - \alpha)H \succcurlyeq \alpha G + (1 - \alpha)H$$

ce qui revient à postuler que quel que soit α , la valeur de $(1 - \alpha)H$ ne dépend ni de F , ni de G .⁷

Les utilités espérées des quatre perspectives aléatoires constituant le problème d'Allais peuvent être reformulées de la manière suivante :

$$u(A_1) = 0.89u(100) + 0.11u(100)$$

$$u(B_1) = 0.89u(100) + 0.01u(0) + 0.10u(500)$$

$$u(A_2) = 0.89u(0) + 0.11u(100)$$

$$u(B_2) = 0.89u(0) + 0.01u(0) + 0.10u(500)$$

Il devient alors visible que A_1 et B_1 ont en commun 89 % de chances de rapporter un gain de 100, alors que A_2 et B_2 ont en commun 89 % de chances de ne rien rapporter du tout. De ce fait, choisir entre les perspectives aléatoires A_1 et B_1 d'une part, A_2 et B_2 d'autre part, constitue – une fois celles-ci dépouillées de leurs conséquences communes respectives – un seul et même problème.⁸ Suivant l'axiome dit de la chose sûre, axiome – formulé par Savage – suivant lequel doit être ignoré ce qui, étant « hors sujet », rend une décision certaine, ce problème devrait rece-

voir une seule et même réponse. Or, nous venons de le voir, tel n'est pas le cas.

En posant d'une part, $r = 500$, $s = 100$, $t = 0$, et d'autre part, $p' = 1.00$ et $p'' = 0.11$, on donne au problème d'Allais une expression paramétrique qui permet d'évaluer l'effet des probabilités et des gains sur la fréquence des violations de l'utilité espérée. On obtient en effet les utilités espérées suivantes :

$$u(A_1) = p'u(s) + (1 - p')u(t)$$

$$u(B_1) = 0.10u(5s) + (p' - 0.11)u(s) + (1.01 - p')u(t)$$

et

$$u(A_2) = p''u(s) + (1 - p'')u(t)$$

$$u(B_2) = 0.10u(5s) + (p'' - 0.11)u(s) + (1.01 - p'')u(t)$$

En écrivant

$$u(A) = (p - 0.11 + 0.11)u(s) + (1 - p)u(t)$$

et

$$u(B) = 0.10u(5s) + (p - 0.11)u(s) + (1.00 - p + 0.01)u(t)$$

soit

$$u(A) = (p - 0.11)u(s) + (1 - p)u(t) + 0.11u(s)$$

et

$$u(B) = (p - 0.11)u(s) + (1 - p)u(t) + 0.01u(t) + 0.10u(5s)$$

on fait apparaître la conséquence commune à A et B

$$(p - 0.11)u(s) + (1 - p)u(t)$$

L'expérimentation a confirmé ce que l'intuition laisse supposer : les violations de l'utilité espérée sont d'autant plus fréquentes que la probabilité p se rapproche de la certitude,

6. Munier, B. (1998), *Two Stage Rationality under Risk: Experimental Results and Perspectives*, Rivista di Matematica Applicata al Science Economica.

7. Quiggin, J. (1993), *Generalized Expected Utility Theory*, Springer-Science + Business Media, B.V.

8. Ici, le terme conséquence est à entendre au sens d'événement.

c'est l'effet dit de certitude, et que le gain intermédiaire s'éloigne du gain minimal t .⁹

Certains auteurs ont été tentés de voir dans cet effet dit de « conséquences communes » une condition nécessaire de violation de la règle de l'utilité espérée. Il n'en est rien.

Pour s'en convaincre, il suffit de considérer les problèmes P3 et P4 de Kahneman et Tversky, où l'on a, d'une part:¹⁰

$$u(A) = 0.80u(4000) + 0.20u(0)$$

$$u(B) = 1.00u(3000)$$

et d'autre part:

$$u(C) = 0.20u(4000) + 0.80u(0) = 0.20u(4000) + 0.75u(0) + 0.05u(0)$$

$$u(D) = 0.25u(3000) + 0.75u(0)$$

Dans ce problème, si C et D ont bien une conséquence commune – 75 % de chances de réaliser un gain nul –, A et B n'ont en revanche rien de commun, et pourtant on observe les préférences $B \succcurlyeq A$ et $C \succcurlyeq D$. Seul entre en jeu l'effet de certitude, ce qu'Allais appelle la préférence pour la sécurité au voisinage de la certitude.

Rôle du contre-exemple dans la démarche scientifique de Maurice Allais

Ce qu'il aurait peut-être mieux fallu appeler le contre-exemple d'Allais, mais il est trop tard pour changer l'usage, illustre à merveille le rôle qu'Allais fait jouer aux faits, « aux données de l'observation » dans l'enrichissement de la connaissance scientifique. Certes, sa

vie durant, Maurice Allais n'a eu de cesse d'insister sur le fait que « la soumission aux données de l'expérience est la règle d'or qui domine toute discipline scientifique », que « toute théorie, quelle qu'elle soit, si elle n'est pas vérifiée par les données de l'expérience, n'a aucune valeur scientifique ».¹¹ Mais, si l'on s'en tenait à cela, on ne ferait qu'effleurer ce qui caractérise la dynamique scientifique, car « en fait trois principes dominent toute démarche scientifique :

- La vérification d'une théorie par certaines données de l'observation n'est pas une preuve de sa validité. Elle montre simplement qu'elle est compatible avec ces données de l'observation ;
- Quels que puissent être le nombre, la qualité et la précision des vérifications antérieures d'une théorie, elle se trouve infirmée si une quelconque de ses hypothèses essentielles ou si une quelconque de ses implications est infirmée par une expérience cruciale ;
- Pour qu'une expérience incompatible avec une théorie antérieure puisse être considérée comme cruciale, il faut et il suffit que les observations correspondant à cette expérience puissent être considérées comme scientifiquement incontestables ».¹²

On reconnaît dans ces lignes tout ce qu'Allais doit non seulement à sa passion pour la physique, mais aussi, et peut-être surtout, à sa connaissance de l'histoire de cette discipline. Comment ne pas se souvenir ici que c'est un écart de huit minutes seulement (2/15^{es} de degré !) entre la position calculée de la planète Mars et sa position observée qui a mis Kepler sur la piste de la forme elliptique de l'orbite des planètes ?¹³

9. MacCrimmon, K.R., Larsson, S., (1979), *Utility Theory: Axioms versus 'Paradoxes'* in Allais, M., et Hagen, O., eds (1979), *Expected Utility Hypotheses and the Allais paradox, Contemporary Discussions of Decision under Uncertainty with Allais's Rejoinder*, D. Reidel Publishing Company, Dordrecht.

10. Kahneman, D., Tversky, A., (1979), *Prospect Theory: An Analysis of Decision Under Risk*, *Econometrica*, Vol. 47, n° 2 (Mar., 1979), p. 263-291.

11. Allais, M. (1989), *Autoportraits, une vie, une œuvre*. Conférence Nobel. Clément Juglar.

12. Allais, M. (2005), *Albert Einstein, Un extraordinaire paradoxe*, Clément Juglar.

13. Caspar, M. (1959), *Kepler*, Dover Edition, Dover Publications Inc., New York.

La critique de la théorie de l'utilité espérée antérieure au contre-exemple d'Allais

Maurice Allais a dit comment sa propre expérience de parieur (sur les courses de chevaux) l'avait conduit, dès 1937, « à réfléchir sur un aspect essentiel de l'analyse des choix aléatoires : l'arbitrage entre l'espérance mathématique et la probabilité de ruine ».

Certes, Allais n'a pas été le premier à douter de la théorie de l'utilité espérée. Il n'a pas manqué d'observer que von Neumann et Morgenstern eux-mêmes s'interrogeaient sur le caractère incomplet ou réducteur de leur théorie. Dès le premier chapitre de leur livre, on peut en effet lire les interrogations suivantes :

« [l'axiome d'indépendance] affirme que les valeurs numériques de l'utilité se combinent (avec les probabilités) comme des espérances mathématiques ! Et cependant, le concept d'espérance mathématique a souvent été critiqué...¹⁴ Nos postulats n'introduisent-ils pas, de manière oblique, les hypothèses qui font intervenir l'espérance mathématique ? Plus précisément : ne peut-il exister chez un individu une utilité (positive ou négative) du simple fait de « tenter sa chance », de parier, utilité qu'ignore l'usage de l'espérance mathématique ? Comment nos axiomes ont-ils contourné cette possibilité ? Pour autant que nous puissions en juger, ils ne tentent pas de l'éviter... Le fait qu'une utilité numérique – avec une formule se résumant à faire usage des espérances mathématiques – puisse être construite sur [l'axiome d'indépendance], semble indiquer ceci : Nous avons pratiquement défini l'utilité numérique comme étant cette chose pour laquelle est légitime le

calcul des espérances mathématiques. Des concepts tels que l'utilité propre au pari ne peuvent être formulés dans ce cadre sans contradiction ».¹⁵

Et, dans les toutes dernières pages de la deuxième édition de leur livre, von Neumann et Morgenstern reviennent sur le sujet de manière encore plus incisive :

« C'est un problème autrement plus compliqué que de formuler un système dans lequel le fait de parier a dans tous les cas une utilité ou désutilité définie, un système où les utilités numériques permettant le calcul des espérances mathématiques ne peuvent être définies par aucun procédé, direct ou indirect. Dans un tel système, l'un ou l'autre de nos axiomes doit nécessairement être erroné. Il semble probable que l'axiome réellement critique soit l'axiome [d'indépendance]. Cet axiome formule la règle de combinaison d'éventualités aléatoires multiples, et il est plausible qu'il faille abandonner cette règle de combinaison pour que puisse exister une utilité ou désutilité propre au fait de parier. Un abandon ou au moins une modification radicale [de l'axiome d'indépendance] peut peut-être conduire à un calcul des utilités mathématiquement complet et satisfaisant, un calcul qui prenne en compte la possibilité d'une utilité ou désutilité propre au fait de parier. On espère qu'il sera trouvé une manière d'y parvenir, mais les difficultés mathématiques semblent considérables ».¹⁶

Si Allais a invoqué von Neumann et Morgenstern à l'appui de sa critique de leur théorie, il ne semble pas avoir remarqué les réflexions,

14. Ici, les auteurs citent Menger et Tintner.

15. Traduction de l'auteur.

16. Traduction de l'auteur. Parmi les efforts faits en ce sens, sont à remarquer ceux de Pope, R. Selten. R. (2009), *Risk in a Simple Temporal Framework for Expected Utility Theory and for SKAT, the Stages of Knowledge Ahead Theory*, Risk and Decision Analysis, Vol. 2, January 2009. Cet article synthétise plusieurs décennies de recherche.

pourtant voisines, de Keynes sur le même sujet.¹⁷

« Pour deux raisons, on peut douter de la doctrine suivant laquelle les espérances mathématiques de lignes alternatives d'action sont de justes mesures de nos degrés de préférence – premièrement, car elle ignore ce que j'ai appelé [...] les poids des arguments, à savoir la quantité d'observations sur la base de laquelle se fonde chaque probabilité; et deuxièmement, car elle ignore l'élément de risque et suppose que des chances égales de paradis ou d'enfer sont tout aussi désirables que la certitude d'atteindre un état de médiocrité. [...]

« La [...] difficulté a trait à la question de savoir si l'espérance mathématique de différentes lignes d'action mesure de manière précise ce que nos préférences devraient être, c'est-à-dire, savoir si la désirabilité d'une ligne d'action donnée diminue proportionnellement à toute augmentation de l'incertitude de son succès, ou bien s'il devrait être fait une part au risque, sa désirabilité diminuant alors plus qu'en proportion de son incertitude.¹⁸ [...]

« Les justifications [*de l'espérance mathématique*] les plus souvent avancées ressemblent à celles de Condorcet dans ses « *Réflexions sur la règle générale, qui prescrit de prendre pour valeur d'un événement incertain, la probabilité de cet événement multipliée par la valeur de l'événement en lui-même* », où il soutient – s'appuyant sur le théorème de [Jacques] Bernoulli [*la loi*

faible des grands nombres] – qu'une telle règle conduira à des résultats satisfaisants si l'on réalise un très grand nombre d'épreuves. Comme, cependant, [...] le théorème de Bernoulli n'est aucunement applicable dans tous les cas, cet argument est irrecevable en tant que justification générale.

« Néanmoins, dans l'histoire du sujet, la théorie de l'espérance mathématique n'a été que très rarement contestée. Comme D'Alembert a presque été le seul à émettre des doutes sérieux à son égard (et bien qu'il ait, ce faisant, seulement réussi à se discréditer [*aux yeux de ses contemporains, sur ce sujet du moins*]), le principal passage dans lequel il déclare son scepticisme mérite d'être cité:¹⁹

« « Il me sembloit » (*en lisant l'Ars Conjectandi de Bernoulli*) « que cette matière avoit besoin d'être traitée d'une manière plus claire; je voyois bien que l'espérance étoit plus grande, 1° que la somme espérée étoit plus grande, 2° que la probabilité de gagner l'étoit aussi. Mais je ne voyois pas avec la même évidence, et je ne le vois pas encore, 1° que la probabilité soit estimée exactement par les méthodes usitées; 2° que quand elle le seroit, l'espérance doive être proportionnelle à cette probabilité simple, plutôt qu'à une puissance ou même une fonction de cette probabilité; 3° que quand il y a plusieurs combinaisons qui donnent différents avantages ou différents risques (qu'on regarde comme des avantages négatifs) il faille se contenter d'ajouter simplement

17. Keynes, J.M., (1921), *A Treatise on Probability*, Dover Publications, Inc. (2004), Mineola, New York. Traduction de l'auteur. Cet ouvrage de Keynes figure dans une liste dressée par Allais de livres relatifs à la théorie des probabilités et à son histoire, mais Allais ne semble jamais avoir fait précisément allusion au passage cité ici.

18. Soulignés par nos soins, ces mots indiquent que les préoccupations de Keynes et de D'Alembert sont de nature normative, comme le seront celles d'Allais, ainsi que nous le verrons plus loin. Keynes pose ici le problème dit de « continuité des préférences ».

19. Et ici, chose étonnante cent ans plus tard, Keynes n'hésite pas à citer D'Alembert sans le traduire, c'est-à-dire dans le Français du XVIII^e siècle.

ensemble toutes les espérances pour avoir l'espérance totale. »²⁰

« Dans les cas extrêmes, il semble difficile de ne pas accorder quelque force à l'objection de D'Alembert ; et c'était d'ailleurs en référence à des cas extrêmes qu'il l'a lui-même soulevée. Est-il certain qu'un plus grand bien, s'il est extrêmement improbable, soit précisément équivalent à un moins grand bien proportionnellement plus probable ? On peut douter que les valeurs morales respectives d'actions spéculative ou prudente puissent être soupesées l'une par rapport à l'autre d'une manière arithmétique simple, de même que l'on a déjà pu douter qu'un bien dont la probabilité ne peut être déterminée que sur une mince base de données puisse être comparé au regard seulement de la valeur de cette probabilité avec un autre bien dont la vraisemblance se fonde sur une connaissance plus complète.

« Il semble, en tout cas, qu'il y ait beaucoup à dire en faveur de la conclusion suivant laquelle, toutes choses égales par ailleurs, est préférable la ligne d'action qui implique le moindre risque, et sur les résultats de laquelle nous avons la connaissance la plus complète. Dans les cas marginaux, donc, le poids des données et le risque ainsi que la probabilité sont des facteurs à prendre en compte dans notre conclusion. Il semble naturel de supposer qu'ils devraient exercer une influence dans d'autres cas aussi, la seule difficulté étant ici l'absence de principe de calcul du degré de cette influence. Un

poids élevé et l'absence de risque augmente d'autant la désirabilité de l'action à laquelle ils s'appliquent, mais nous ne pouvons mesurer l'intensité de cette augmentation. »

Cette revue de la littérature antérieure au paradoxe d'Allais ne prétend pas être exhaustive. Au risque de sortir des canons du discours académique, on pourrait d'ailleurs y ajouter l'un de nos proverbes, qui a mis le doigt, il y a sans doute fort longtemps, sur la problématique du paradoxe d'Allais. Ce proverbe, c'est celui qui nous dit qu'« *un tiens vaut mieux que deux tu l'auras* ». A vrai dire, il faut lui préférer sa version anglaise – “*one bird in the hand is worth two in the bush*” – qui, n'employant que le temps présent, ne risque pas de nous égarer hors sujet dans un arbitrage temporel.²¹ Si « *un oiseau en main en vaut deux dans le buisson* », c'est bien parce qu'il est probable que ni l'un, ni l'autre de ces deux oiseaux ne se laisse attraper.

On le voit, la critique de l'utilité espérée était dans l'air bien avant 1952. Le singulier mérite de Maurice Allais est de l'avoir adossé à une base expérimentale.²²

Le programme de recherche défini par Allais en 1952

Il ne suffit pas de constater que des agents réputés rationnels violent la règle de l'utilité espérée. Encore faut-il expliquer pourquoi et comment ils le font.

20. Le Rond D'Alembert, J. (1768), *Opusculs mathématiques*, vol. iv., (extraits de lettres), p. 284-285. Allais critique vivement un ouvrage antérieur du même auteur où se voit exposé ce qu'il appelle le principe de D'Alembert, à savoir le caractère négligeable des petites probabilités.

21. Strauss, E. (1998), *Concise Dictionary of European Proverbs*, Routledge, London.

22. La place manque pour le démontrer ici, mais le fait est qu'en condamnant la théorie de l'utilité espérée, le contre-exemple d'Allais condamne du même coup les nombreuses constructions intellectuelles qui en dépendent et sur lesquelles s'appuient des pans entiers de la théorie financière : par exemple, sans utilité espérée, point de coefficient d'aversion pour le risque d'Arrow-Pratt, et point d'*equity premium puzzle* (voir Wakker, P. (2010), *Prospect Theory for Risk and Ambiguity*, Cambridge University Press).

Dès 1952, Allais a suggéré que quatre facteurs sont à prendre en compte pour étayer une théorie de la décision face au risque :²³

- › La déformation psychologique des valeurs monétaires et la courbure de la satisfaction absolue;
- › La déformation subjective des probabilités objectives²⁴;
- › La pondération suivant les probabilités [subjectives] des valeurs psychologiques et la considération des espérances mathématiques (moment d'ordre 1) de la distribution des probabilités [subjectives] des valeurs monétaires;
- › La prise en considération de la forme des distributions de probabilités des valeurs psychologiques et en particulier de leur dispersion (moment d'ordre 2 ou plus) en première approximation.

Pour résumer, Allais propose donc de remplacer la formulation de l'utilité espérée

$$V(X) = \sum_{i=1}^n p_i u(x_i)$$

par la formulation suivante

$$V^*(X) = \sum_{i=1}^n p_i^* u(x_i)$$

où les p_i^* représentent les probabilités subjectives.

Pour illustrer sa pensée, il donne un exemple de cette formule dans le cas d'une perspective

aléatoire à deux éventualités, soit, en utilisant les notations de la présente contribution :

$$V^* = p_1^* u(x_1) + (1 - p_1^*) u(x_2)$$

Si seulement Allais avait, à l'époque, pris pour exemple une perspective aléatoire à trois au lieu de deux branches, il aurait encore plus clairement esquissé les modèles à dépendance de rang apparus trente ans plus tard. Mais, nous anticipons.

L'impasse des moments

Comment prendre en considération « la forme des distributions de probabilités des valeurs psychologiques et en particulier leur dispersion [autour de leur moyenne] », exigence formulée dès 1952 par Maurice Allais ?²⁵

Si Allais évoque à l'époque le « moment [central] d'ordre deux de la distribution des valeurs psychologiques », il indique aussitôt qu'il ne peut s'agir que d'« une première approximation ». Souvent en effet, on oublie que ce moment central d'ordre deux, qui – lorsqu'il est appliqué non pas aux valeurs psychologiques (absolues), mais aux gains monétaires (relatifs) – n'est autre que le carré de cette volatilité si chère aux théoriciens, praticiens et régulateurs des marchés financiers, ne permet de caractériser que certaines distributions, comme par exemple la distribution de Laplace-Gauss. Toujours à la recherche de la formulation la plus générale possible, Allais formule en 1977 l'hypothèse suivant laquelle la dispersion d'une distribution peut

23. Allais, M., (1953), *Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'École Américaine*. Rédigé en Français en 1952, ce texte fondateur a été publié une première fois en 1953 comme mémoire annexé au volume *Économétrie des Actes des Colloques Internationaux du Centre National de la Recherche Scientifique*, Vol. XL, Paris. La même année, il fût partiellement reproduit dans la revue *Econometrica*, mais toujours en Français : Allais, M., (1953), *Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine*, *Econometrica*, Vol. 21, n° 4, October 1953. Ce n'est que vingt-six ans plus tard, en 1979, que fut publiée sa première traduction intégrale en Anglais dans Allais, M., (1977), *The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School* in Allais, M., et Hagen, O., éd. (1979), *Expected Utility Hypotheses and the Allais paradox, Contemporary Discussions of Decision under Uncertainty with Allais's Rejoinder*, D. Reidel Publishing Company, Dordrecht.

24. Pour Allais, une probabilité objective est une fréquence empirique.

25. Allais, M., (1953), *Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine*, *Econometrica*, Vol. 21, n° 4, October 1953.

être prise en compte grâce à une fonction de ses moments.²⁶

Mais alors, avant même de tenter de trouver la forme de cette fonction, se pose la question suivante : de combien de moments a-t-on besoin pour caractériser une distribution discrète quelconque, faite de n événements possibles ? En 1985, Allais répond à cette question en démontrant un théorème suivant lequel, pour caractériser une telle distribution, il ne faut pas moins de ses $2n - 1$ premiers moments.²⁷ Ainsi, si la fonction recherchée doit être un polynôme, il faudrait un polynôme de degré cinq pour prendre en compte la dispersion de la perspective aléatoire B_1 du problème d'Allais, perspective qui ne comporte pourtant que trois événements possibles. Peu engageant, ce résultat a sans doute incité Maurice Allais à orienter ses recherches dans une autre direction, ainsi que nous allons bientôt le voir. La planète finance, quant à elle, continue de se contenter d'un seul moment, le moment central d'ordre deux, pour caractériser la dispersion d'une distribution de rendements.

On a néanmoins observé que, dans les problèmes de choix posés par Kahneman et Tversky, les sujets préfèrent toujours la perspective dont le moment central d'ordre 3 des seuls gains monétaires est le plus élevé, que celui-ci soit négatif, positif ou nul.²⁸ Autrement dit, si les écarts positifs par rapport à

la moyenne sont suffisamment importants et fréquents pour générer une asymétrie positive, ils sont désirables.

L'absence de l'entropie

Il est une mesure de la dispersion d'une distribution discrète de probabilités qui est absente de la littérature sur la théorie de la décision en situation de risque. Cette mesure de dispersion, c'est l'entropie d'information que Claude Shannon a défini en 1948.²⁹ Si tel est le cas, c'est sans doute pour les deux raisons suivantes. D'une part, l'entropie ignore le montant des gains et des pertes associés aux différentes probabilités.³⁰ D'autre part, l'entropie ignore la notion de probabilité subjective. Si l'entropie ne pouvait, pour ces deux raisons, fournir une alternative à la théorie de l'utilité espérée, elle aurait pu contribuer à sa critique en soulignant l'importance des petites probabilités.

Le triangle de Marschak-Machina

En présence d'une perspective aléatoire se composant de trois branches au plus ($x_1 < x_2 < x_3$), un ingénieux exercice de géométrie analytique – connu sous le nom de triangle de Marschak-Machina – suggère une interprétation féconde du contre-exemple d'Allais.³¹ Cet exercice consiste à représenter une perspective aléatoire à trois branches par un point situé à l'intérieur ou aux limites d'un triangle rectangle isocèle.

26. Allais, M., (1977), *The So-called Allais Paradox and rational decisions under uncertainty* in Allais, M., et Hagen, O., eds (1979), *Expected Utility Hypotheses and the Allais paradox, Contemporary Discussions of Decision under Uncertainty with Allais's Rejoinder*, D. Reidel Publishing Company, Dordrecht.

27. Allais, M., (1985), *Sur la détermination d'une distribution discrète d'ordre n par la donnée de $2n - 1$ moments consécutifs*, Doc. C-4341 (11/09/1985) du Centre d'Analyse Économique. Texte reproduit dans Munier, B. (ed), (1988), *Risk, Decision and Rationality*, Dordrecht, Reidel Publishing.

28. Barthalon, E., (2014), *Uncertainty, Expectations and Financial Instability, Reviving Allais's Lost Theory of Psychological Time*, Columbia University Press, New York. Cette observation doit pouvoir s'expliquer par le fait que plus l'ordre d'un moment est élevé, plus il donne de poids aux gains ou pertes extrêmes d'une distribution. Or, à ces valeurs extrêmes correspondent les probabilités, généralement faibles, auxquelles les modèles de transformation de probabilité, que nous présenterons plus loin, font subir les accroissements les plus prononcés. Par une voie ou par l'autre, le poids des extrêmes se voit donc accru.

29. Shannon, C.E. (1948), *The Mathematical Theory of Communication*, (1998) University of Illinois Press, Chicago.

30. Machina, M., Rotschild, M. (2008), *Risk*, New Palgrave Dictionary of Economics, 2nd edition, éd. By Steven N. Durlauf and Lawrence E. Blume.

31. Marschak, J., (1950), *Rational behavior, uncertain prospects and measurable utility*, *Econometrica* 18 et Machina, M., (1982), "Expected Utility" *Analysis without the Independence Axiom*, *Econometrica* 50(2), p. 277-323, February 1982.

Il est d'abord tracé deux axes perpendiculaires gradués de 0 à 1 : un axe horizontal sur lequel on porte la probabilité de x_1 , le moins favorable des trois cas possibles ; un axe vertical sur lequel on inscrit la probabilité de x_3 , le plus favorable des trois cas possibles.

Figure 1

La droite joignant les extrémités de ces deux axes (l'hypoténuse du triangle) est le lieu des points tels que $p_1 + p_3 = 1$ (ou – ce qui revient au même – $p_2 = 0$). Lorsque l'on remonte le long de l'hypoténuse de l'Est vers le Nord, p_1 diminue alors que p_3 augmente. Les valeurs de x_1 et x_3 étant données, celle de l'utilité espérée $p_1u_1 + p_3u_3$ augmente donc elle aussi.³² Autrement dit, plus une perspective aléa-

toire est proche de l'angle Nord du triangle, plus elle est préférable.

Comme le montre la Table 2, il est aisé d'interpréter tout déplacement, à l'intérieur du triangle de Marschak-Machine, d'un point de coordonnées (p_1, p_3) vers un point de coordonnées (p'_1, p'_3) .

Table 2

Probabilités initiales	Orientation du mouvement	Nouvelles probabilités
p_3	Nord	$p_3 + dp_3$
p_1	Ouest	$p_1 - dp_1$
$1 - p_1 - p_3$	Nord-Ouest	$1 - p_1 - p_3 - (dp_3 - dp_1)$

32. Pour alléger l'écriture, posons $u(x_1) = u_1$, $u(x_2) = u_2$ et $u(x_3) = u_3$.

A un déplacement vers le Nord correspond une augmentation $+dp_3 > 0$ de p_3 , la probabilité du gain le plus élevé, au détriment de p_2 , la probabilité du gain intermédiaire, puisque p_1 reste par hypothèse la même. A un déplacement vers l'Ouest correspond une diminution $-dp_1 < 0$ de p_1 , la probabilité du gain le plus faible, au bénéfice de p_2 , puisque p_3 reste inchangée. A un déplacement vers le Nord-Ouest correspond donc à la fois une diminution de la probabilité du pire des cas et une augmentation de celle du meilleur des cas, la probabilité du cas intermédiaire étant inchangée, augmentée ou diminuée suivant que dp_3 est égal, inférieur ou supérieur à dp_1 . Dans les trois cas de figure, en vertu du principe de dominance stochastique du premier ordre (appelé axiome de préférence absolue par Allais), la situation (p'_1, p'_3) est préférable à la situation (p_1, p_3) , car l'on a toujours $p'_1 \leq p_1$ et $p'_3 \geq p_3$. La position de p'_2 par rapport à p_2 ne fait que renforcer ou atténuer cette dominance.

Dans le cas où $p_1 + p_3 < 1$ (c'est-à-dire $p_2 > 0$), le point P de coordonnées (p_1, p_3) se trouve sur une droite horizontale qui coupe l'hypoténuse en un point de coordonnées $(p_1 + d, p_3)$. Puisque ce point d'intersection se situe sur l'hypoténuse, on a donc

$$p_1 + d + p_3 = 1 \Leftrightarrow d = 1 - p_1 - p_3 \Leftrightarrow d = p_2$$

Conduit verticalement, le même raisonnement aboutit à la même conclusion. A lui tout seul, ce point P représente donc p_1 , p_2 et p_3 , soit toute la distribution de probabilités considérée.

Sous l'hypothèse de l'utilité espérée, les valeurs de x_1 et x_3 étant fixes, il passe par ce point P une droite d'équation

$$p_3 = \frac{u_2 - u_1}{u_3 - u_2} p_1 + \frac{U - u_2}{u_3 - u_2}$$

équation que l'on obtient en remplaçant p_2 par $1 - (p_1 + p_3)$ dans l'équation de l'utilité espérée

$$U = p_1 u_1 + p_2 u_2 + p_3 u_3$$

La pente de cette droite est nécessairement positive. Son coefficient peut en effet s'écrire

$$\frac{u_2 - u_1}{u_3 - u_2} = \frac{u_3 - u_1}{u_3 - u_2} - 1$$

Or, puisque $x_3 > x_2 > x_1 > 0$, on a $u_3 > u_2 > u_1 > 0$ et donc $u_3 - u_1 > u_3 - u_2$ ce qui implique

$$\frac{u_3 - u_1}{u_3 - u_2} > 1$$

Tout point situé sur cette droite procurant donc la même utilité U , cette droite représente le lieu d'indifférence ou d'iso-utilité déterminé par le paramètre U . Si ce paramètre détermine l'ordonnée à l'origine de cette droite, il est en revanche sans influence sur sa pente. A des valeurs plus ou moins élevées de ce paramètre correspondent donc des droites parallèles d'iso-utilité, elles-mêmes plus ou moins élevées, si du moins l'hypothèse de linéarité en probabilités est valide.

Les quatre perspectives aléatoires du problème d'Allais forment un parallélogramme à l'intérieur du triangle de Marschak-Machina : le point B_1 est en effet par rapport au point A_1 dans la même position Nord-Est que le point B_2 par rapport au point A_2 . Révélée par les réponses au questionnaire d'Allais, la préférence orthodoxe $B_2 \succcurlyeq A_2$ implique que la droite d'iso-utilité passant par B_2 soit située au Nord-Ouest du point A_2 . Si l'hypothèse de linéarité en probabilités objectives est valide, la préférence $B_2 \succcurlyeq A_2$ devrait impliquer qu'une droite parallèle à celle passant par B_2 passe par B_1 et soit située au Nord-Ouest de A_1 , en sorte que B_1 devrait être préféré à A_1 .

Or, on sait que, « paradoxalement », c'est la préférence $A_1 \succ B_1$ que l'on observe. Autrement dit, la courbe d'indifférence passant par B_1 ne se situe pas au Nord-Ouest, mais plutôt à l'Est de A_1 . Quant à la courbe d'indifférence passant par A_1 , elle ne passe pas au Sud-Est, mais plutôt à l'Ouest de B_1 . Ces observations suggèrent que les courbes d'indifférence sont d'autant plus pentues que p_1 diminue et que p_3 augmente : au lieu d'être parallèles les unes aux autres, elles s'ouvrent à la manière d'un éventail. Autrement dit encore, tout semble se passer comme si les probabilités objectives étaient subjectivement transformées par les décideurs.

La piste des moments ayant été écartée, le triangle de Marschack-Machina suggérant quant à lui qu'il fallait transformer les probabilités objectives, il ne restait plus qu'à comprendre comment cette transformation s'effectue. Les premiers modèles proposés à cet effet à partir du milieu des années 1950, notamment en France (par Kreweras³³, Bernard³⁴, etc.), mais aussi, de manière plus médiatique, par Kahneman et Tversky en 1979, ont transformé les probabilités discrètes objectives³⁵. Mais on a vite observé que cette méthode pouvait avoir pour fâcheuse conséquence la violation du principe de dominance stochastique de premier ordre. Il a fallu attendre les années 1980 pour qu'Allais, Quiggin et Yaari trouvent simultanément et indépendamment les uns des autres la solution, à savoir une transformation de la distribution (dé)cumulée, solution pourtant intuitivement présentée et illustrée de manière éloquente par Allais dès 1952 dans le volume Écono-

métrie déjà cité, mais sur un exemple ne comportant que deux événements possibles, comme on l'a déjà évoqué.³⁶ Le modèle présenté en 1992 par Tversky et Kahneman ne fait qu'ajouter quelques paramètres à ce canevas.³⁷

Le modèle (u, θ) d'Allais

Afin d'alléger l'écriture, nous considérons une perspective aléatoire composée de trois éventualités seulement (x_1, x_2, x_3) , où $x_1 < x_2 < x_3$. A ces trois éventualités sont associées les probabilités objectives (p_1, p_2, p_3) , telles que $p_1 + p_2 + p_3 = 1$. L'intuition d'Allais est qu'un décideur ne va pas regarder et traiter ces données comme le ferait un mathématicien. Soucieux d'aller et venir entre probabilités discrètes (ou densité de probabilité) d'une part et fonction de répartition d'autre part, le mathématicien, habitué qu'il est à manipuler les intégrales, s'intéresse à la probabilité qu'une variable aléatoire soit inférieure à une valeur donnée. Dans le cas discret qui est le nôtre, les relations entre les probabilités discrètes $P\{X = x_i\}$ et la fonction de répartition $F(x) = P\{X < x\}$ s'écrivent

$$F(x) = \sum_{x_i < x} P\{X = x_i\}$$

et

$$P\{X = x_i\} = F(x_{i+1}) - F(x_i)$$

Dans le cas continu, on a

$$F(x) = \int_{-\infty}^x f(t)dt$$

33. Kreweras, G., (1961), *Sur une possibilité de rationaliser les intransitivités*, La Décision, Colloques Internationaux du CNRS, Paris, éd. p. 27-32.

34. Bernard, G., (1966), *Sur les fonctions d'utilité*, Revue Française de Recherche Opérationnelle, 41, p. 323-352.

35. Kahneman, D., Tversky, A., (1979), *Prospect Theory: An Analysis of Decision Under Risk*, Econometrica, Vol. 47, n° 2 (Mar., 1979), p. 263-291.

36. Allais, M. (1988), *The General Theory of Random Choices in Relation to the Invariant Cardinal Utility Function and the Specific Probability Function, the (u, θ) model: a General Overview* in Munier, B. (éd.) *Risk, Decision and Rationality (Selected papers of a conference held in Aix-en-Provence in 1986)*, Dordrecht/Boston, Reidel, p. 231-289.

Quiggin, J. (1982), *A Theory of Anticipated Utility*, Journal of Economic Behavior and Organization, Vol. 3, p. 323-343.

Yaari, M. (1987), *The Dual Theory of Choice under Risk*, Econometrica, Vol. 55, p. 95-115

37. Tversky, A., Kahneman, D., (1992), *Advances in prospect theory: Cumulative Representation of Uncertainty*, Journal of Risk and Uncertainty, Vol. 5, p. 297-323.

Pour Allais, le décideur s'intéresse au contraire à la probabilité qu'une variable aléatoire soit supérieure à une valeur donnée, soit la distribution décumulée

$$F'(x) = P\{X > x\} = 1 - P\{X < x\}$$

Car, face à la perspective aléatoire $(x_1, x_2, x_3; p_1, p_2, p_3)$, le décideur observe que, dans le pire des cas, il est certain de gagner au moins x_1 puisque $p_1 + p_2 + p_3 = 1$. Il observe ensuite que la probabilité de gagner $x_2 - x_1$ en plus de x_1 est égale à $p_2 + p_3$. Enfin, la probabilité de gagner $x_3 - x_2$ en plus de x_2 est égale à p_3 . Quant à la probabilité de gagner quoi que ce soit en plus de x_3 , elle est nulle. Arrivé à

ce stade, nous pouvons constater que nous n'avons fait que reformuler l'utilité espérée comme suit :

$$V(X) = \sum_{i=1}^3 p_i u(x_i) = (p_1 + p_2 + p_3)u(x_1) + (p_2 + p_3)[u(x_2) - u(x_1)] + p_3[u(x_3) - u(x_2)]$$

Allais suppose alors que le décideur transforme les probabilités objectives décumulées $p_1 + p_2 + p_3$, $p_2 + p_3$ et p_3 en probabilités subjectives. Cette transformation est opérée par une fonction θ , qui doit à l'évidence au moins satisfaire les deux conditions suivantes : $\theta(1) = 1$ et $\theta(0) = 0$. La table 3 ci-dessous reprend les différentes étapes du modèle (U, θ) .

Table 3

Probabilités objectives	Utilité	Utilité additionnelle	Probabilités décumulées	Probabilités décumulées transformées
p_1	u_1	u_1	$p_1 + p_2 + p_3$	$\theta(p_1 + p_2 + p_3)$
p_2	u_2	$u_2 - u_1$	$p_2 + p_3$	$\theta(p_2 + p_3)$
p_3	u_3	$u_3 - u_2$	p_3	$\theta(p_3)$

Il ne reste plus alors qu'à récupérer les probabilités subjectives en écrivant :

$$p_1^* = \theta(p_1 + p_2 + p_3) - \theta(p_2 + p_3)$$

$$p_2^* = \theta(p_2 + p_3) - \theta(p_3)$$

$$p_3^* = \theta(p_3) - \theta(0)$$

et à calculer la valeur psychologique de la perspective aléatoire considérée, soit

$$V^*(X) = \sum_{i=1}^3 p_i^* u(x_i)$$

Chacune des probabilités subjectives ainsi définies dépend bien des autres éléments de la distribution, puisque l'on peut écrire

$$p_1^* = \theta(1) - \theta(p_2 + p_3)$$

$$p_2^* = \theta(1 - p_1) - \theta(p_3)$$

$$p_3^* = \theta(1 - p_1 - p_2) - \theta(0)$$

Que ce soit sous l'appellation d'utilité espérée à dépendance de rang, de théorie généralisée de l'utilité espérée ou bien encore de *cumulative prospect theory*, l'article fondateur de 1952 et son rejeton, le modèle (U, θ) ont fait florès. Les solutions aujourd'hui proposées par les différents auteurs se distinguent les unes des autres par les formes, somme toute assez peu différentes, qu'ils donnent à leur fonction d'utilité (ou de valeur) d'une part, et d'autre part à leur fonction de transformation des probabilités objectives.

Ainsi, la fonction θ s'est-elle vue donner par d'autres qu'Allais les formulations suivantes $\theta(p) = p^\gamma$, ou $\theta(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)$, ou encore $\theta(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)^{1/\gamma}$ où p est une probabilité décumulée.

Suivant la valeur du paramètre γ , ces fonctions sont convexes ou concaves, en forme de S, symétrique ou non. Quant à la fonction θ d'Allais, elle a la formulation suivante :

$$\theta(p) = \frac{m'p}{a-p} \times \frac{[1 + (m' - 1)a](a - p) - a(a - 1)(mm' - 1)(1 - p)}{[1 + (m' - 1)a][1 + (m' - 1)p] - a(a - 1)(mm' - 1)(1 - p)}$$

où m' et m sont la pente de θ en 0 et 1 respectivement, avec $m > 1$, $0 < m' < m$, $mm' > 1$ et $1 < a < m/(m-1)$. Comme le montre la Figure 2, les paramètres m et m' déterminent la courbure, éventuellement asymétrique, de la fonction θ ; le paramètre a détermine, quant à lui, sa hauteur.

Figure 2

Comme le montre la Figure 3 dans le cas de la fonction d'Allais, toutes ces fonctions produisent le même effet, à savoir une sur-pondération des événements situés aux

extrémités des distributions, sur-pondération contrebalancée par une sous-pondération des événements centraux.

Figure 3

La fonction d'utilité exprime, quant à elle, trois idées. Premièrement, la valeur d'un gain ou d'une perte dépend du capital initial du décideur: ce sont les pertes ou les gains relatifs qui comptent.³⁸ Deuxièmement, le désagrément occasionné par une perte de 10 % est plus intense que l'agrément généré par un gain de 10 %. Troisièmement, l'utilité marginale des gains est décroissante, mais ne devient jamais négative.³⁹ De là découle la

forme convexe des fonctions d'utilité proposées par les différents auteurs. Sans emporter l'adhésion, Allais a présenté comme « invariante » sa fonction d'utilité cardinale des gains x' par rapport au capital du sujet (voir Figure 4). Pour des pertes relatives allant jusqu'à - 60 %, elle s'approxime par la fonction $0.5223 \ln(1 + x')$; pour des gains relatifs allant jusqu'à 150 %, elle s'approxime par la fonction $0.1865 \ln(1 + x')$.

38. 1 000 euros de plus ou de moins comptent plus pour l'homme de la rue que pour un milliardaire, qui ne remarque même pas qu'il les perd ou les gagne.

39. Contrairement à ce que rend possible la fonction quadratique d'utilité que Markowitz (1959, *Portfolio Selection, Efficient Diversification of Investments*, Blackwell, Cambridge, MA) a placé au cœur de sa théorie du portefeuille, hypothèse dont Borch (1969, *A Note on Uncertainty and Indifference Curves*, Review of Economic Studies, Vol. 36), grâce au paradoxe éponyme, a montré l'absurdité.

Figure 4

Pour une définition expérimentale de la rationalité

La quasi-totalité des propos relatifs au paradoxe d'Allais abrègent ou développent ce que l'on a pu lire jusqu'ici dans la présente contribution. Or, comme l'a montré Philippe Mongin, co-lauréat du Prix Maurice Allais 2019, c'est là une vision fort réductrice de l'apport de Maurice Allais à la théorie de la décision rationnelle face au risque.⁴⁰ Pour s'en convaincre, il suffit d'observer que le contre-exemple d'Allais n'apparaît qu'à la trente-neuvième des cinquante pages du texte fondateur de 1952. Et encore n'y apparaît-il qu'au côté d'autres exemples de comportements face au risque qui, sans être conformes à l'hypothèse de l'utilité espérée, sont pourtant ceux d'individus dont tout laisse penser qu'ils sont rationnels. Ces autres contre-exemples sont les suivants : l'achat simultané de billets de loterie et de polices d'assurance, le comportement de gens très prudents dans les

choix aléatoires mettant en jeu des petites sommes, le comportement des entrepreneurs lorsque de grandes pertes sont possibles. Les choix aléatoires au voisinage de la certitude mettant en échec l'axiome d'indépendance ne sont donc qu'un argument empirique parmi d'autres.

En revanche, dès la première page du même texte, juste après avoir énoncé les quatre points de ce que nous avons appelé plus haut le programme de recherche de 1952, Allais affirme qu'« *il est reconnu par tout le monde que l'homme réel ne se comporte pas suivant le principe de Bernoulli* [le principe de l'utilité espérée] ». Et il ajoute aussitôt que « *seule existe une profonde différence de vue sur ce que devrait faire l'homme rationnel* ». ⁴¹ On retrouve ici l'ambition normative déjà présente chez Keynes, et avant lui chez D'Alembert. C'est dans la méthode expérimentale

40. Mongin, P. (2019), *The Allais paradox: what it became, what it really was, what it now suggests to us*, Economics and Philosophy, Vol. 35, p. 423-459, Cambridge University Press.

41. Le terme « devrait » est ici souligné par nos soins.

qu'Allais retient pour définir la rationalité que se trouve son originalité.

Pour lui, la rationalité ne saurait se définir comme une obédience axiomatique, quand bien même les axiomes en question seraient formulés par d'éminentes autorités intellectuelles. Pourquoi ? Parce qu'on ne peut jamais exclure que de tels axiomes négligent un aspect important de la réalité et soient de ce fait inappropriés.⁴² En l'occurrence, l'axiome d'indépendance néglige le phénomène psychologique bien réel d'appétence ou d'aversion à l'égard de la dispersion plus ou moins grande des éventualités.

Allais propose alors deux voies concurrentes, mais éventuellement complémentaires, l'une abstraite, l'autre expérimentale, pour donner de la rationalité une définition qui soit scientifiquement utile.

De manière abstraite, est rationnel un comportement qui obéit au « *critère généralement utilisé dans les sciences sociales de non-contradiction avec soi-même, critère qui implique la cohérence des fins poursuivies et l'emploi de moyens appropriés pour atteindre ces fins* ». Jusqu'ici rien de nouveau, comme Allais le souligne lui-même. Outre l'utilisation de probabilités objectives lorsqu'elles existent, ce principe de non-contradiction ne fait qu'impliquer « *l'axiome de préférence absolue* [le principe de dominance stochastique du premier ordre] *qui exprime que de deux situations l'une est certainement préférable si, dans toutes les éventualités, elle comporte des gains au moins équivalents et, pour au moins une éventualité, des gains supérieurs* ». On le voit, en matière d'axiomes,

l'approche d'Allais est minimaliste. On l'a critiqué pour cela.

Si tel est le cas, c'est sans doute parce que, pour Allais, les données de l'observation sont à la fois la clé de voûte et le garde-fou de la connaissance scientifique. Aussi, selon lui, est également rationnel « *ce que font les personnes dont on a des raisons de penser, par ailleurs, qu'elles se conduisent de manière rationnelle* » et qui, ajoute-t-il plus loin, sont « *tout à fait au courant du calcul des probabilités* ».

Il est tentant d'objecter que cette définition expérimentale de la rationalité est tautologique, entachée d'un vice de circularité. Mais en fait, en plusieurs endroits du texte de 1952, Allais précise sa pensée en soulignant que ces personnes réputées rationnelles s'efforcent de minimiser le risque de ruine, ou ce qui revient au même, elles cherchent à maximiser leur probabilité de survie économique (et financière).⁴³ On retrouve cette problématique dans la distinction que fait Vernon Smith entre rationalité constructiviste et rationalité écologique.⁴⁴ Ces personnes réputées rationnelles se refusent donc à « *appliquer la loi des grands nombres dans les cas où elle ne s'applique pas* », c'est-à-dire quand elles n'ont pas la possibilité de donner suffisamment de temps au temps pour observer la convergence des fréquences empiriques cumulées vers les probabilités objectives.

Le paradoxe d'Allais, c'est donc d'abord un contre-exemple expérimental d'une remarquable fécondité. Il est le point de départ d'une dynamique scientifique internationale dépassant la simple critique, Français contre Américains, de l'utilité espérée, pour

42. Allais citait souvent Pareto : « L'histoire de la science est l'histoire des erreurs des hommes compétents ». On peut aussi évoquer Corneille : « Pour grands que soient les rois, ils sont ce que nous sommes/Ils peuvent se tromper comme les autres hommes ».

43. Le détenteur d'un portefeuille d'actions ne peut, par exemple, engranger le surcroît de rendement offert à long terme par cette classe d'actifs, la fameuse « *equity risk premium* », que s'il peut survivre, psychologiquement et financièrement, aux inévitables fluctuations boursières de court terme.

44. Smith, V.L. (2008), *Rationality in Economics, Constructivist and Ecological Forms*, Cambridge University Press, New York, NY.

aboutir à une théorie plus générale capable de faire ce que toute nouvelle théorie doit faire : conserver la théorie ancienne dans ce qu'elle a de juste (la préférence $B_2 \succcurlyeq A_2$) ; améliorer sa capacité explicative, et donc prédictive, là où elle n'y parvenait pas (la préférence $A_1 \succcurlyeq B_1$). Mais, le paradoxe d'Allais, c'est aussi une approche expérimentale de la rationalité face au risque. À ce double titre, il n'a pas fini de faire parler de Maurice Allais, non seulement au sein du monde académique, mais aussi dans les très nombreux domaines, publics comme privés, où il peut

et devrait à la fois éclairer et inspirer l'action. Qu'il s'agisse – pour ne citer que quelques exemples – de défense nationale, d'approvisionnement énergétique, d'environnement, de santé publique, de régulation financière, d'endettement, de trésorerie ou d'investissement... une politique rationnelle peut-elle se fonder uniquement sur le « scénario central » et ignorer l'importance psychologique des événements extrêmes, fussent-ils de faible probabilité?⁴⁵ Le paradoxe d'Allais nous dit que non.

45. Sur ce que les modèles à dépendance de rang devraient apporter aux épargnants et à leurs conseillers, grâce à des protocoles de profilage psychologique (appétence ou aversion pour le risque, tolérance des pertes) plus rigoureux que ceux préconisés par MIFID-II, voir, par exemple, Munier, B., Menk-Bertrand, E. (2019), *Comment faire émerger le vrai coût du risque financier en gestion de fortune ? Une critique de MIFID-II*, Revue d'Economie Financière, Vol. 1, n° 133.