

TESTS DES MODÈLES DE DÉCISION EN SITUATION DE RISQUE

Le cas des parieurs hippiques en France

Ariane Charpin

Presses de Sciences Po | « Revue économique »

2018/5 Vol. 69 | pages 779 à 803

ISSN 0035-2764

ISBN 9782724635560

Article disponible en ligne à l'adresse :

<https://www.cairn.info/revue-economique-2018-5-page-779.htm>

Distribution électronique Cairn.info pour Presses de Sciences Po.

© Presses de Sciences Po. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Tests des modèles de décision en situation de risque

Le cas des parieurs hippiques en France

Ariane Charpin*

L'un des résultats les plus solides de la littérature utilisant des données sur les paris hippiques est que les cotes des chevaux reflètent très bien leurs chances de gagner, à l'exception de celles des favoris qui sont sous-estimées et de celles des outsiders qui sont surestimées. La théorie de l'utilité espérée et les théories comportementales de la décision en situation de risque se font concurrence pour expliquer ce résultat. Cet article cherche à discriminer entre les deux types de modèles en testant celui qui est le plus adapté pour expliquer le comportement des parieurs observé dans les données. En utilisant un jeu de données unique des paris sur les courses de chevaux en France, je trouve que les théories comportementales sont mieux ajustées aux données que l'utilité espérée. Elles fournissent ainsi une meilleure représentation des choix que l'utilité espérée.

TESTING MODELS OF DECISION UNDER RISK: THE CASE OF HORSERACE BETTORS IN FRANCE

One of the most robust findings in the literature using data on horseraces bets is that odds associated to horses reflect their chances of winning very well, with the exception that favorites are underbet while outsiders are overbet. Expected utility theory and behavioral theories of decision under risk compete to explain this finding. This paper seeks to discriminate between the two classes of models by testing which is the most suited to explaining the behavior of bettors observed in the data. Using a unique dataset of bets on horseraces in France, I find that behavioral theories of decision under risk better fit my data than expected utility. This result shows that behavioral theories provide a better representation of choice behavior than expected utility.

Mots clés : prise de décision, utilité espérée, théorie des perspectives cumulées, utilité dépendante du rang, fonction de pondération des probabilités, aversion au risque, parieur représentatif, biais favori-outsider

Keywords : decision making, expected utility, cumulative prospect theory, rank-dependent utility, probability weighting function, risk-aversion, representative bettor, favorite-longshot bias

Classification JEL : D81, L83.

* École d'économie de Paris – Université Paris 1. Correspondance : Bureau R2-18, 48 boulevard Jourdan, 75014 Paris, France. Courriel : ariane.charpin@psemail.eu

Je remercie les participants au séminaire de Paris 1, au meeting QED, au Congrès de l'AFSE et aux Journées de microéconomie appliquée pour leurs commentaires.

INTRODUCTION

Les paris sur les courses de chevaux permettent d'étudier les décisions prises en situation de risque dans une situation réelle. Premièrement, parier sur un cheval implique de faire un choix entre des alternatives clairement identifiées, chaque alternative étant associée à un résultat monétaire. Deuxièmement, l'alternative qui se produit est observée publiquement après une courte période de temps. Troisièmement, les choix sont effectués de manière répétée par un grand nombre de participants. Quatrièmement, une information abondante et complète est disponible sur les probabilités associées aux différents résultats possibles. Les marchés des paris hippiques offrent donc l'opportunité de tester le cadre théorique de la décision en situation de risque dans une situation simple mais réelle. En particulier, ils partagent de nombreuses caractéristiques avec des marchés financiers très simples.

Un grand nombre d'articles ont profité de ces caractéristiques. Ils ont étudié si les prix associés aux chevaux (les cotes) reflètent leurs valeurs intrinsèques (chances de gagner dans une course donnée). L'une des découvertes les plus solides de cette littérature est que les cotes associées aux chevaux reflètent très bien leurs valeurs intrinsèques, sauf que les favoris (chevaux ayant de fortes chances de gagner) ont tendance à être sous-estimés tandis que les outsiders (chevaux ayant une probabilité de gagner relativement faible) sont surestimés (Sauer [1998]). En conséquence, les rendements espérés sur les outsiders sont plus faibles que sur les favoris. Une abondante littérature tente d'expliquer l'existence de cette régularité empirique, appelée le biais favori-outsider (voir Ottaviani et Sørensen [2008] pour une revue des principales explications). En particulier, deux théories de la décision en situation de risque rivalisent à cet effet.

D'un côté, la théorie standard du choix individuel en économie (théorie de l'utilité espérée de von Neumann et Morgenstern [1947]) peut rationaliser le biais en supposant que les parieurs ont, au moins localement, une fonction d'utilité convexe pour les résultats monétaires. D'un autre côté, les théories comportementales, qui sont plus générales que l'utilité espérée, sont capables d'expliquer le biais en ajoutant que les décideurs transforment les probabilités lors de l'évaluation de la valeur des perspectives risquées.

L'objectif de cet article est de discriminer entre ces deux classes de modèles. Il s'agit d'identifier le modèle qui explique le mieux le biais favori-outsider. L'utilité dépendante du rang et la théorie des perspectives cumulées sont des généralisations de l'utilité espérée, on s'attend donc à ce que ces modèles expliquent mieux les données que l'utilité espérée. Ce travail vise à déterminer si le gain apporté par les modèles comportementaux ne se produit pas au prix de trop de complexité.

L'utilisation d'un modèle trop restrictif ou erroné pourrait empêcher de comprendre des comportements couramment observés ou conduire à des prédictions fausses. Les comportements qui ne peuvent être expliqués qu'à l'aide de modèles comportementaux incluent par exemple l'énigme de la prise de risque en finance, le choix de certains contrats d'assurance ou l'offre de main-d'œuvre des chauffeurs de taxi en économie du travail.

Notre analyse s'appuie sur un jeu de données unique sur les paris hippiques de l'opérateur de paris français de 2013 à 2015. Les principales contributions de l'article sont les suivantes. D'abord, j'étudie l'existence du biais favori-outsider

en France, ce qui n'a pas encore été fait à ma connaissance. Deuxièmement, je teste les résultats de Jullien et Salanié [2000] dans un contexte et à une période différents. Troisièmement, je surmonte l'une des limites de l'article de Jullien et Salanié [2000] liée au fait qu'ils n'ont pas pu mettre en œuvre les tests statistiques appropriés pour comparer les modèles. Cette nouveauté est une amélioration majeure de l'article de Jullien et Salanié [2000], puisque la comparaison des modèles est l'objectif principal de l'article. Quatrièmement, j'assouplis certaines des hypothèses du modèle de Jullien et Salanié [2000] parce que le système français est un système de pari mutuel et non un système reposant sur des bookmakers. Enfin, je vérifie également la robustesse des résultats à certaines des hypothèses du modèle en utilisant les caractéristiques des courses.

Je montre d'abord que le biais favori-outsider est présent en France. Je trouve, en outre, que les théories comportementales de la décision en situation de risque, c'est-à-dire la théorie des perspectives cumulées et la théorie de l'utilité dépendante du rang, sont mieux adaptées pour expliquer le comportement des parieurs observé dans les données que l'utilité espérée. Ce résultat fournit la preuve que les parieurs pondèrent les probabilités de façon non linéaire lorsqu'ils font des choix. En utilisant la théorie des perspectives cumulées, je trouve une pondération significative des probabilités dans le domaine des pertes et une pondération linéaire des probabilités dans le domaine des gains, ce qui est cohérent avec le résultat de Jullien et Salanié [2000] mais contredit les résultats des expériences qui trouvent une pondération similaire des probabilités dans les domaines des gains et pertes (Tversky et Kahneman [1992]; Abdellaoui [2000]).

Cet article s'inscrit dans une littérature théorique et expérimentale considérable motivée par l'observation que dans les expériences en laboratoire les individus font des choix systématiquement incompatibles avec la théorie de l'utilité espérée (Allais [1953]; Kahneman et Tversky [1979]). La théorie des perspectives cumulées est actuellement le modèle préféré de la littérature expérimentale. Des preuves que les décideurs pondèrent les probabilités de façon non linéaire comme dans la théorie des perspectives cumulées ont été fournies par de nombreuses expériences (voir Camerer et Ho [1994], Tversky et Kahneman [1992], Wu et Gonzalez [1996], Gonzalez et Wu [1999] et Abdellaoui [2000]). La théorie s'est également avérée capable de rationaliser les comportements observés dans les laboratoires qui ne pouvaient pas être expliqués par l'utilité espérée. Un exemple est l'assurance probabiliste (voir Wakker, Thaler et Tversky [1997]). Ce type de police d'assurance implique une faible probabilité (disons 1 %) que le consommateur ne soit pas remboursé. Selon la théorie de l'utilité espérée (et quelle que soit la concavité de la fonction d'utilité), les individus devraient être prêts à payer l'assurance probabiliste environ 99 % du prix de l'assurance complète. Mais les réponses expérimentales montrent que les individus sont prêts à payer beaucoup moins pour compenser les faibles chances de ne pas être remboursés. Ce comportement ne peut pas être expliqué par l'utilité espérée, mais est cohérent avec la surpondération des faibles probabilités de la théorie des perspectives.

Certains se sont demandé si les résultats de la littérature expérimentale se généralisaient aux données du monde réel (voir List [2003], Levitt et List [2008]). Ils pensent que les biais sont moins probables en présence de grands enjeux, d'expérience et de concurrence.

Les études existantes reposant sur des données du monde réel utilisent généralement des données d'assurance, de finance et de paris ou de jeux. Alors que Cicchetti et Dubin [1994] présentent des preuves que la décision de souscrire à

une assurance contre le risque de dysfonctionnement de la ligne téléphonique au domicile est compatible avec la théorie de l'utilité espérée, Barseghyan *et al.* [2013] montrent que la pondération non linéaire des probabilités joue un rôle central dans le comportement des ménages dans le choix de l'assurance auto et habitation. Kliger et Levy [2009] constatent également que la théorie des perspectives cumulées est mieux ajustée à leurs données que l'utilité espérée et l'utilité dépendante du rang en utilisant des données sur les options d'achat de l'indice S&P500. À l'aide de données provenant de jeux télévisés, Post *et al.* [2008] montrent que les préférences dépendent du point de référence, de sorte qu'elles présentent des caractéristiques du modèle de la théorie des perspectives cumulées. List [2003], [2004] fournit la preuve que, bien que les consommateurs inexpérimentés se comportent comme dans la théorie des perspectives, l'expérience du marché amène le comportement des traders expérimentés proche des prédictions néoclassiques. Au contraire, Pope et Schweitzer [2011] montrent que les golfeurs professionnels très expérimentés, confrontés à des gains élevés et à une concurrence intense, présentent une aversion à la perte, comme le prédit la théorie des perspectives.

Le présent article est étroitement lié à la littérature utilisant des données sur les courses de chevaux. Jullien et Salanié [2000] montrent, en se concentrant sur les paris simples au Royaume-Uni où le système de bookmakers prédomine, que la théorie des perspectives cumulées décrit mieux le comportement d'un agent représentatif que les théories de l'utilité espérée et de l'utilité dépendante du rang. Contrairement à la fonction de pondération des probabilités habituelle en S inversé, ils trouvent peu de preuves de l'existence d'un effet de certitude et d'un changement de concavité des fonctions de pondération de probabilité qu'ils estiment. Ils établissent également que l'utilité dépendante du rang n'apporte pas d'amélioration par rapport à l'utilité espérée.

Snowberg et Wolfers [2010] utilisent une impressionnante base de données de paris mutuels aux États-Unis pour tester les prédictions dérivées de l'utilité espérée et de la théorie des perspectives cumulées dans le cas particulier des paris simples sur des paris complexes. Leur approche repose sur le fait que les deux théories ont des implications différentes sur les prix des paris complexes, ce qui permet, en comparant les prédictions avec les prix réels, d'identifier le meilleur modèle. Ils trouvent que le modèle avec pondération non linéaire des probabilités fournit la meilleure description des données, ce qui suggère que la théorie des perspectives permet une meilleure description des données que l'utilité espérée.

Le présent article est organisé comme suit. La deuxième section explique comment les paris hippiques sont organisés en France. La troisième section décrit les données. La quatrième section présente le modèle et la procédure d'estimation. La cinquième section présente les résultats. La dernière section conclut.

LES PARIS HIPPIQUES EN FRANCE

Le marché des paris hippiques en France est exclusivement un système de pari mutuel. Le concept du pari mutuel consiste à regrouper tous les paris correspondant à une course et à un type de pari donnés, à retirer une part pour couvrir les taxes et les dépenses de l'opérateur de paris et à redistribuer le reste parmi les

parieurs gagnants proportionnellement à leurs paris. Les gains finaux dépendent donc exclusivement du total des mises, de la part conservée par l'opérateur de paris (le « *take* »¹) et des mises sur chaque cheval.

Plus un cheval attire de mises par rapport au total des mises, moins un pari sur ce cheval rapporte. Les gains sur les chevaux sont appelés cotes. Dans le type de paris le plus simple (qui sont les paris analysés dans cet article), qui consiste à trouver le gagnant d'une course donnée, une cote de 1,2 sur un cheval et une course donnée signifie qu'un pari gagnant de 1 sur ce cheval rapporte au parieur sa mise (1) plus 1,2. La cote correspond donc au rendement net d'un pari unitaire. Ainsi, un cheval ne peut pas avoir une cote inférieure à 0,1.

Pour une course se déroulant un jour particulier, les paris ouvrent sur Internet vers cinq heures du matin le jour de la course. Pour un parieur qui préfère se rendre dans un magasin spécialisé, les paris démarrent à l'ouverture des magasins. À l'hippodrome, un parieur ne peut parier qu'une trentaine de minutes avant le début de la course. Le marché ferme juste avant le début de la course. En raison de la façon dont les cotes sont calculées dans le système du pari mutuel, les parieurs n'ont accès qu'à des cotes temporaires qui sont calculées avec l'état actuel des mises et sont mises à jour chaque minute en ligne.

DONNÉES

Les données ont été collectées sur le site pmu.fr entre avril 2013 et mai 2015. Le PMU (pari mutuel urbain) est l'opérateur principal de paris en France. En ligne, il rassemble 84,8 % du total des mises et en magasin, c'est un monopole légal. La base de données contient des informations sur les paris, les courses, les chevaux et les hippodromes pour les courses pour lesquelles des paris ont été proposés par le PMU. Elle porte sur 33 196 courses.

Pour chaque course, la base de données contient la cote finale de chaque cheval, son classement dans la course et de nombreuses caractéristiques. En plus de l'heure, de la date et de l'hippodrome, les courses sont également caractérisées par leur discipline, leurs types et les conditions de participation. Les données ne contiennent que les cotes de tous les chevaux de chaque course pour les paris simples, qui sont l'objet de cet article.

Cet article s'intéresse à la modélisation du processus de décision concernant le choix d'un cheval spécifique dans une course. Je supprime donc les 7 919 courses dans lesquelles deux chevaux ou plus dans une course donnée appartiennent à la même écurie, ce qui arrive quand les chevaux ont le même propriétaire ou le même entraîneur. Dans ce cas, tous les chevaux de l'écurie ont la même cote et si l'un d'eux gagne la course, un pari sur l'un des chevaux de l'écurie gagne également. Par conséquent, la cote d'un cheval qui fait partie d'une écurie ne reflète pas sa probabilité de gagner, mais plutôt la probabilité que tout cheval de l'écurie gagne.

1. En plus du « *take* », les opérateurs français bénéficient également du « *breakage* », qui est le gain provenant du fait d'arrondir les cotes à la baisse à la première décimale.

Je supprime également les courses dans lesquelles plusieurs chevaux arrivent en première position, car je modélise une course dans laquelle un seul cheval gagne. Je supprime les courses pour lesquelles les cotes sont incomplètes ou manifestement erronées. Cela comprend les courses qui ne sont pas enregistrées comme étant terminées, dans lesquelles au moins un cheval participant n'a pas de cote et les courses pour lesquelles la cote finale du cheval gagnant ne correspond pas au dividende indiqué par ailleurs dans les données. Il me reste 23462 courses.

Comme le montre le tableau 1, le nombre moyen de courses par jour s'élève à 32. Certains jours, 80 courses ont lieu alors que d'autres jours seulement 13. Le nombre moyen de chevaux dans une course donnée est de 12. Le minimum est de 2 et le maximum de 24. La moitié des courses comptent entre 9 et 14 chevaux.

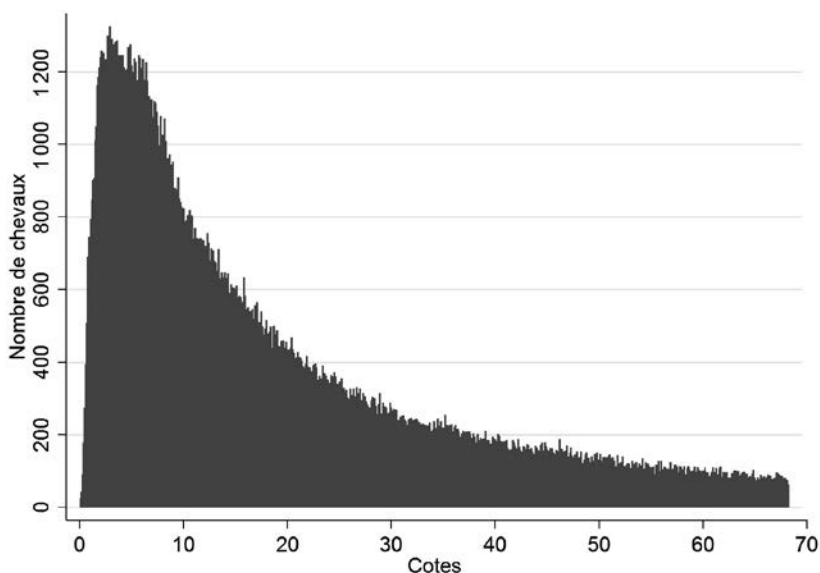
La distribution des cotes couvre un large intervalle. Le maximum atteint 998 alors que le minimum est de 0,1. La médiane s'élève à 15,4 et la moyenne à 27. Quatre-vingt-dix pour cent des cotes sont comprises entre 0,1 et 68,4.

Tableau 1. *Statistiques descriptives*

Variable	Moyenne	Écart type	Min.	Max.	N
Courses par jour	32	9,29	13	80	735
Chevaux par course	12	3,31	2	24	23464
Cotes	27	31,79	0,1	998	279 792

Le graphique suivant montre que l'échantillon contient des grands favoris, avec des cotes comprises entre 0,1 et 0,5 (0,2 % de l'échantillon) et des grands outsiders avec des cotes supérieures à 50 (19 % de l'échantillon).

Graphique 1. *Distribution des cotes (90 % des cotes)*



En utilisant ces définitions, 64 % des grands favoris et 0,6 % des très grands outsiders ont gagné leur course. Alternativement, en définissant les grands favoris comme des chevaux qui attirent deux fois plus de mises que le deuxième cheval attirant le plus de mises, les grands favoris gagnent 44 % du temps.

MODÈLE ET PROCÉDURE D'ESTIMATION

Modèle théorique

Le modèle représente la décision d'un parieur représentatif qui parie a dans une course donnée et qui est doté d'une richesse initiale notée M^2 . Le choix d'un cheval particulier dans la course dépend uniquement de sa probabilité de gagner et des cotes finales. Dans une course donnée r avec N chevaux, le parieur fait face à un menu de probabilités et de cotes $((O_1, p_1), (O_2, p_2), \dots, (O_N, p_N))_r$, les probabilités étant non négatives et leur somme valant un.

Je suppose que le menu est connu du parieur lorsqu'il effectue son choix. En pratique, les cotes finales ne sont pas connues avant le début de la course et le parieur n'a pas une parfaite connaissance des probabilités de gagner. Cependant, des études antérieures sur les parieurs ont montré que les parieurs ont une bonne connaissance des chances de gagner des chevaux³ (voir Sauer [1998]).

En notant H_i l'action de parier sur le cheval i , la valeur globale d'un pari sur le cheval i , $W(H_i)$, est un nombre tel que le parieur préfère le cheval i au cheval j ou est indifférent entre les chevaux i et j si et seulement si $W(H_i) \geq W(H_j)$. Le parieur représentatif est rationnel en ce qu'il parie sur le cheval avec la valeur globale la plus élevée.

Compte tenu des hypothèses énoncées, les parieurs continuent de parier dans une course donnée jusqu'à ce que les cotes les rendent indifférents entre un pari sur n'importe quel cheval de la course et pas de pari⁴. Donc à l'équilibre :

$$\forall i \in 1, \dots, N, W(H_i) = w, w \text{ constant.} \quad (1)$$

Soit a la somme mise par le parieur représentatif, M sa richesse initiale, O_i la cote associée au cheval i et p_i la probabilité de gagner du cheval i . Je définis aussi u , une fonction d'utilité continue et strictement croissante avec $u(0) = 0$ et j'écris

2. Les données ne contiennent pas d'information sur le montant parié (a) ou sur la richesse des parieurs (M).

3. L'hypothèse que les parieurs connaissent les probabilités de gagner des chevaux participant à une course a été testée en réestimant le modèle sur un échantillon de données limité aux courses dans lesquelles les chevaux sont âgés de 4 ans ou plus et ont déjà couru dans des courses. Dans ces courses, les parieurs sont susceptibles d'avoir une très bonne connaissance des probabilités de gagner des chevaux. Le fait de limiter la base de données à ce sous-échantillon n'a pas modifié les résultats présentés dans l'article.

4. L'hypothèse d'équilibre est plus forte dans Jullien et Salanié [2000], car les cotes sont choisies par des bookmakers (système de bookmakers) et non par le marché (système de pari mutuel). Ici mon hypothèse est simplement que les cotes s'ajustent de sorte que les parieurs sont indifférents entre tous les chevaux à la fin de la période des paris. Jullien et Salanié doivent supposer que les bookmakers fixent les cotes telles que les parieurs sont indifférents entre tous les chevaux en raison de l'hypothèse supplémentaire selon laquelle tous les chevaux de la course reçoivent un pari.

φ^+ la fonction de pondération des probabilités des gains et φ^- celle des pertes. Les deux fonctions de pondération des probabilités sont strictement continues et croissantes de $[0,1]$ à $[0,1]$ et satisfont $\varphi^+(0) = \varphi^-(0) = 0$ et $\varphi^+(1) = \varphi^-(1) = 1$. En termes généraux, la valeur globale d'un pari sur le cheval i peut être écrite :

$$W(p_i, a, M, O_i) = \varphi^+(p_i) * u(M + aO_i) + \varphi^-(1 - p_i) * u(M - a). \quad (2)$$

L'expression spécifique de la valeur globale d'un pari sur le cheval i dépend du modèle de décision étudié. Dans la théorie de l'utilité espérée, $\varphi^+(p_i) = \varphi^-(p_i) = p_i$. Dans la théorie de l'utilité dépendante du rang, $\varphi^-(1 - p_i) = 1 - \varphi^+(p_i)$ et dans la théorie des perspectives cumulées $M = 0$, ce qui revient à supposer que le point de référence est l'absence de pari⁵. Plus de détails sur chaque modèle sont présentés en annexe.

Le modèle est résolu en utilisant la procédure de Jullien et Salanié [2000] qui consiste à calculer d'abord w puis la probabilité de gagner du cheval qui a gagné la course (p_i), qui est à son tour utilisée pour calculer la fonction de vraisemblance. La procédure utilisée pour obtenir p_i est expliquée pour chaque modèle en annexe.

Forme fonctionnelle de la fonction d'utilité

Tout comme Jullien et Salanié [2000], je suppose dans tout l'article une fonction d'utilité avec aversion absolue au risque de la forme suivante :

$$u(x, \theta) = \frac{1 - e^{-\theta x}}{\theta}.$$

La forme fonctionnelle avec aversion absolue constante au risque permet d'estimer le niveau d'aversion au risque absolu θ en supposant qu'il est constant. Les parieurs aiment le risque si $\theta < 0$ et sont averses au risque si $\theta > 0$. L'expression de u retenue est pratique puisque M , qui n'est pas observé, s'annule dans l'expression de la probabilité utilisée dans la fonction de vraisemblance.

Formes fonctionnelles des fonctions de pondération des probabilités

Les formes fonctionnelles classiques présentées dans le tableau 2 sont estimées. Plus d'informations sur ces fonctions sont disponibles en annexe.

Soit Ω le vecteur des paramètres des fonctions de pondération des probabilités correspondant à chaque modèle. Ω , qui dépend de la forme de la fonction des probabilités à estimer et du fait que le modèle utilise une seule fonction de pondération des probabilités (utilité dépendante du rang) ou deux fonctions de pondération des probabilités (une pour les gains et une pour les pertes dans la théorie des perspectives cumulées), est défini dans le tableau 3.

5. Cette hypothèse est cohérente avec la littérature qui suppose couramment que le point de référence est le *statu quo*. Pour une discussion sur la façon dont les individus pensent en termes de gains et de pertes, voir Köszegi et Rabin [2006], [2007], [2009].

Tableau 2. Fonctions de pondération des probabilités

Nom	Expression	Restrictions sur les paramètres
Puissance	$\varphi(p) = p^\alpha$	$\alpha > 0$
Tversky et Kahneman [1992]	$\varphi(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}}$	$\gamma > 0$
Cicchetti et Dubin [1994]	$\frac{\varphi(p)}{1-\varphi(p)} = \left(\frac{p}{1-p}\right)^{a_1} * \left(\frac{p_0}{1-p_0}\right)^{1-a_1}$	$a_1 \geq 0, p_0 \in [0,1]$
Prelec [1998]	$\varphi(p) = e^{-\beta(-\ln p)^\alpha}$	$0 < \alpha < 1, \beta > 0$
Lattimore, Baker et Witte [1992]	$\varphi(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$	$\gamma > 0, \delta > 0$

Tableau 3. Définition du vecteur de paramètres Ω

Nom	Utilité dépendante du rang	Théorie des perspectives cumulées
Puissance	$\Omega = \alpha$	$\Omega = (\alpha, \beta)$
Tversky et Kahneman [1992]	$\Omega = \gamma$	$\Omega = (\gamma, \gamma')$
Cicchetti et Dubin [1994]	$\Omega = (a_1, p_0)$	$\Omega = (a_1, p_0, a'_1, p'_0)$
Prelec [1998]	$\Omega = (\alpha, \beta)$	$\Omega = (\alpha, \beta, \alpha', \beta')$
Lattimore, Baker et Witte [1992]	$\Omega = (\delta, \gamma)$	$\Omega = (\delta, \gamma, \delta', \gamma')$

Estimation

Soient θ le paramètre de la fonction d'utilité et Ω le vecteur des paramètres des fonctions de pondération des probabilités défini dans le tableau 3 pour chaque forme fonctionnelle de la fonction de pondération des probabilités et chaque modèle à estimer.

p_1^r est la probabilité de gagner du cheval qui a effectivement gagné la course r , avec $r = 1, \dots, J$. Pour chaque r , la vraisemblance d'observer le cheval 1 gagner la course est $l(\theta, \Omega; O_1^r) = p_1(O_1^r; \theta, \Omega)$.

En supposant que les résultats des courses sont indépendants et que les expressions de p_1 dérivées du modèle sont correctes, la probabilité d'observer l'échantillon étudié est le produit des J densités individuelles, ce qui correspond à la fonction de vraisemblance suivante :

$$\prod_{r=1}^J p_1(O_1^r; \theta, \Omega) = L(\theta, \Omega | O_1^r). \quad (3)$$

D'où la fonction de vraisemblance en logarithme :

$$LL = \ln L(\theta, \Omega | O_1^r) = \sum_{r=1}^J \ln p_1(O_1^r; \theta, \Omega). \quad (4)$$

L'estimateur du maximum de vraisemblance a les propriétés asymptotiques habituelles. Il est cohérent, asymptotiquement normal, asymptotiquement efficace et invariant.

$$\hat{\theta} \sim N\left(\theta_0, [I(\theta_0)]^{-1}\right),$$

$$I(\theta_0) = -E_0\left(\partial^2 \ln L / \partial \theta_0 \partial \theta_0'\right) = -E_0\left(\partial \ln L / \partial \theta_0 * \partial \ln L / \partial \theta_0'\right). \quad (5)$$

Les écarts types sont calculés à l'aide d'une procédure de *bootstrap* et de la formule précédente.

RÉSULTATS

Le biais favori-outsider

Le biais favori-outsider est le constat que les paris sur les favoris (chevaux avec des cotes faibles) ont un rendement espéré plus élevé que les paris sur les outsiders (chevaux avec des cotes relativement élevées). Il a été montré dans un grand nombre d'articles, à commencer par Griffith [1949]. Il a été mis en évidence dans différents types de courses et à différentes époques en Amérique du Nord (McGlothlin [1956]; Weitzman [1965]; Ali [1977]; Snyder [1978]; Asch, Malkiel et Quandt [1982]; Snowberg et Wolfers [2010]) où le système de pari mutuel prévaut, au Royaume-Uni, dans les systèmes de pari mutuel et de bookmakers (Williams et Paton [1997]; Jullien et Salanié [2000]), en Australie dans le système de pari mutuel (Coleman [2002]) et le système des bookmakers (Bird, McCrae et Beggs [1987]) et en Nouvelle-Zélande dans le système de pari mutuel (Coleman [2002]; Gandar, Zuber et Johnson [2001]).

Le premier résultat de l'article est que le biais favori-outsider est présent également en France. Le rendement espéré pour un pari de 1 sur le cheval i est $R_i = \pi_i * O_i + (1 - \pi_i) * (-1)$, où π_i est la probabilité de gagner du cheval i et O_i correspond à sa cote finale. La probabilité de gagner d'un cheval, qui est la proportion de fois où le cheval gagnerait la même course répétée un nombre infiniment grand de fois, est inconnue, donc je calcule le rendement espéré en utilisant l'approche couramment adoptée dans la littérature (voir Coleman [2004]). Cette approche consiste à regrouper tous les chevaux de la base de données, soit par intervalles de cotes, soit par ordre de préférence (le favori est dans le premier groupe, le deuxième favori dans le deuxième groupe, etc.), et à calculer le pourcentage de gagnants et la cote moyenne dans chaque groupe.

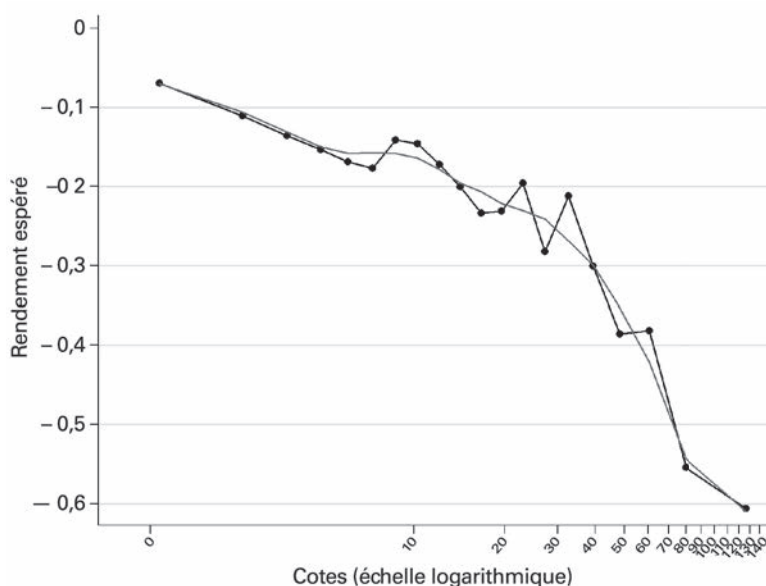
Les rendements espérés sont représentés graphiquement sur le graphique suivant, les chevaux ont été regroupés par centiles de cotes et les cotes sont représentées sur une échelle logarithmique.

Le graphique 2 montre que les rendements ne sont pas égaux pour les différentes cotes : les paris sur les favoris ont un taux de rendement plus élevé que les paris sur les outsiders. Le rendement espéré d'un pari sur un cheval avec une cote de 127 s'élève à -0,6 par unité mise, alors qu'il s'élève à -0,07 pour un cheval avec une cote de 1,43. Par conséquent, les cotes des favoris ne sont pas assez faibles pour compenser les probabilités élevées de gagner des favoris, ou, de manière

équivalente, les favoris sont sous-estimés par rapport à leur probabilité de gagner. Au contraire, les gains des outsiders ne sont pas assez élevés pour compenser leurs faibles probabilités de gagner, ou de manière équivalente, ils sont surévalués.

Par conséquent, dans un modèle simple avec des fonctions d'utilité et de pondération de probabilité linéaires, les données montrent qu'à l'équilibre, les rendements espérés ne sont pas égalisés entre les chevaux d'une même course. De nombreuses propositions ont été faites pour expliquer ce biais. L'une d'entre elles est que le modèle simple ne tient pas correctement compte des goûts et des croyances des parieurs.

Graphique 2. Rendement espéré des cotes



Tests des modèles de décision en situation de risque

Modèle d'utilité espérée

Les résultats des paramètres obtenus dans le cadre de l'utilité espérée, ainsi que la valeur maximale de la fonction de log-vraisemblance (LL) sont présentés dans le tableau 4.

Tableau 4. Modèle d'utilité espérée

	Modèle UE
$a\theta$	- 0,014*** (0,0011)
Max LL	- 45 977,315

$a\theta$ est statistiquement significatif, négatif et a une valeur absolue faible, ce qui signifie qu'au moins statistiquement et dans la fourchette de revenus dans

laquelle je les teste, les parieurs montrent un faible goût statistiquement significatif pour le risque. Le paramètre est plus petit que celui obtenu par Jullien et Salanié [2000] ($-0,055$) mais similaire à celui de Snowberg et Wolfers [2010] qui ajustent leurs données avec une fonction d'utilité avec aversion absolue au risque de paramètre $-0,017$.

Un pari de 20 euros sur un cheval avec une cote de 10 et une probabilité de gagner de 25 % équivaut à la loterie gagner 200 euros avec une probabilité de 25 % et perdre 20 euros avec une probabilité de 75 %. Le paramètre d'attitude face au risque estimé rend un parieur indifférent entre cette loterie et le montant certain de 38 euros. Un parieur neutre au risque serait indifférent entre la même loterie et la somme certaine de 35 euros, donc le comportement des parieurs présente un certain goût pour le risque.

Modèle d'utilité dépendante du rang

Les résultats pour toutes les formes fonctionnelles des fonctions de pondération des probabilités testées sont présentés dans le tableau 5 et représentés sur le graphique 3.

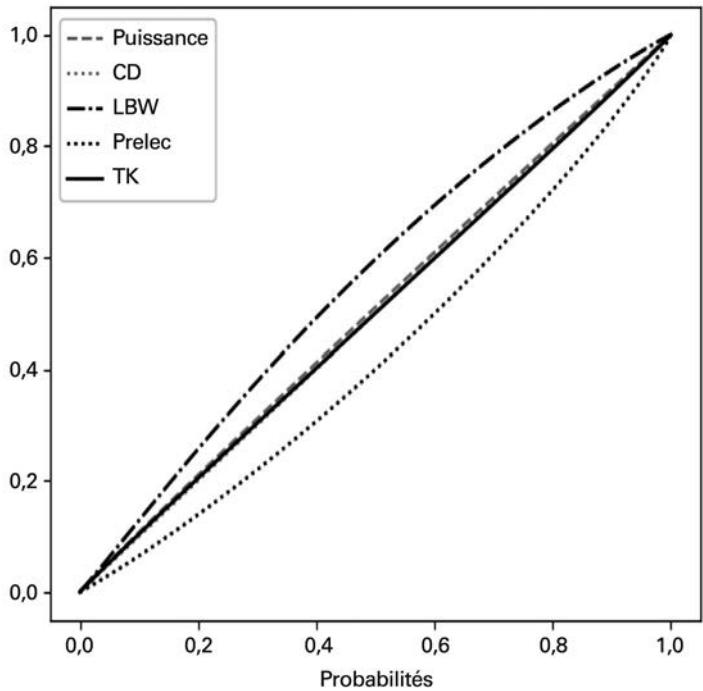
Tableau 5. Modèles d'utilité dépendante du rang

	Puissance	CD	LBW	Prelec	TK
$a\theta$	$-0,011^{***}$ (0,0019)	$-0,011^{***}$ (0,0026)	$-0,016^{***}$ (0,0029)	$-0,013^{***}$ (0,0038)	$-0,011^{***}$ (0,0018)
α	$0,97^{***}$ (0,014)			$0,91^{***}$ (0,045)	
a_1		$0,99^{***}$ (0,026)			
p_0		$0,99^{***}$ (0,22)			
γ			$1,05^{***}$ (0,035)		$0,98^{***}$ (0,011)
δ			$1,48^{***}$ (0,237)		
β				$1,28^{***}$ (0,077)	
Max LL	$-45\,975,289$	$-45\,974,628$	$-45\,972,658$	$-45\,972,481$	$-45\,975,363$

Le paramètre qui représente l'attitude face au risque de la fonction d'utilité est négatif, statistiquement significatif et faible dans chaque spécification du modèle d'utilité dépendante du rang.

Le coefficient de la fonction puissance est statistiquement significatif, diffère de et est légèrement inférieur à 1 (col. 1 du tableau 5) donc $\varphi(p) > p$, ce qui reflète l'optimisme. Dans la fonction Cicchetti et Dubin (col. 2), a_1 n'est pas significativement différent de 1 bien qu'il soit précisément estimé, donc p_0 n'est pas identifié et nous revenons au modèle d'utilité espéré. Dans la fonction Lattimore, Baker et Witte (col. 3), γ ne diffère pas significativement de 1 alors que δ est élevé et statistiquement différent de 1 au niveau de confiance 5 %.

Graphique 3. Utilité dépendante du rang –
Fonctions de pondération des probabilités estimées



La fonction est concave et supérieure à la droite à 45 degrés, ce qui suggère de l'optimisme. Dans la fonction Prelec (col. 4), α est proche mais statistiquement différent de 1, β est élevé et aussi statistiquement différent de 1. La fonction est convexe et inférieure à la droite à 45 degrés, ce qui montre du pessimisme. Dans la fonction Tversky et Kahneman (col. 5), γ est proche mais statistiquement différent de 1.

Le modèle d'utilité espérée est imbriqué dans tous les modèles. Les tests du rapport de vraisemblance entre ces derniers et les modèles d'utilité dépendante du rang sont présentés dans le tableau 6.

Tableau 6. Comparaisons des modèles d'utilité dépendante du rang et utilité espérée : tests statistiques

	UE	UDR Puissance	UDR CD	UDR LBW	UDR Prelec	UDR TK
AIC	91957	91955	91955	91951	91951	91955
BIC	91965	91971	91979	91976	91975	91971
Rapport de vraisemblance vs UE		4,1	5,4	9,3	9,7	3,9
Valeur critique		3,8	6	6	6	3,8

Les tests du rapport de vraisemblance montrent que le modèle d'utilité dépendante du rang est statistiquement significativement mieux adapté aux données que le modèle d'utilité espérée dans quatre spécifications (puissance ; Lattimore, Baker et Witte ; Prelec ; Tversky et Kahneman). On remarque que les rapports de vraisemblance sont cependant très proches des valeurs critiques associées pour les fonctions puissance et Tversky et Kahneman. La première conclusion de l'analyse est que le modèle d'utilité dépendante du rang représente statistiquement mieux les données que le modèle d'utilité espérée. Cette conclusion diffère de Jullien et Salanié [2000] qui ont conclu que les modèles d'utilité dépendante du rang n'apportaient pas d'amélioration par rapport au modèle d'utilité espérée, parce qu'ils ont trouvé que seule la spécification Prelec s'ajustait mieux à leurs données que l'utilité espérée.

Les résultats ne permettent pas de conclure sur l'attitude globale vis-à-vis du risque des parieurs, qui combine à la fois le paramètre d'attitude face au risque de la fonction d'utilité et la forme de la fonction de pondération des probabilités. Comme le suggère le graphique 3, la fonction puissance et la fonction Tversky et Kahneman sont extrêmement proches de la diagonale. Le tableau 5 montre que l'attitude envers le risque est similaire dans le modèle d'utilité espérée et dans les modèles de l'utilité dépendante du rang avec les fonctions de pondération des probabilités puissance et Tversky et Kahneman. La différence entre le modèle d'utilité espérée et ces modèles ne change donc rien au comportement des parieurs : les parieurs ont un faible goût pour le risque. Dans Lattimore, Baker et Witte [1992], le paramètre d'attitude face au risque estimé suggère que les parieurs ont un goût pour le risque, ce qui est renforcé par la surpondération des probabilités. Dans la spécification Prelec, le paramètre d'attitude face au risque suggère un goût pour le risque mais les parieurs sous-pondèrent les probabilités, de sorte qu'à la fin leur comportement présente une aversion au risque. À titre d'illustration, les paramètres estimés dans le cas de la fonction Prelec rendent un parieur indifférent entre la loterie qui consiste à gagner 200 euros avec une probabilité de 25 % et à en perdre 20 avec une probabilité de 75 % et le montant certain de 22 euros. Dans la spécification Lattimore, Baker et Witte, ils rendent le parieur indifférent entre la même loterie et le montant certain de 54 euros. Étant donné que la neutralité au risque correspond à la somme certaine de 35 euros, le parieur présente une aversion au risque dans le premier cas et un goût pour le risque dans le second. Il n'est pas possible de discriminer entre les modèles estimés dans le cadre de la théorie dépendante du rang, car le critère d'information d'Akaike et le critère d'information bayésien, ainsi que les tests de Vuong entre les modèles, ne sont pas concluants⁶.

Théorie des perspectives cumulées

Les résultats pour toutes les formes fonctionnelles des fonctions de pondération des probabilités testées sont présentés dans le tableau 7. Les fonctions de pondération des probabilités pour les gains sont représentées sur le graphique 4 et les fonctions de pondération des probabilités pour les pertes sur le graphique 5.

Le paramètre d'attitude face au risque de la fonction d'utilité est négatif, statistiquement significatif et petit dans chaque spécification du modèle de la théorie des perspectives cumulées.

6. Je peux seulement établir, en utilisant un test de rapport de vraisemblance, que le modèle Prelec, dans lequel le modèle puissance est imbriqué, est meilleur que la fonction puissance.

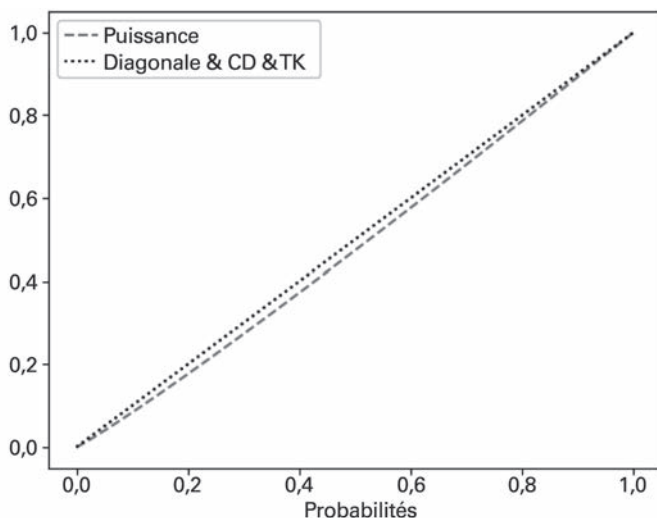
Tableau 7. *Modèles de la théorie des perspectives cumulées*

	Puissance	CD	LBW	Prelec	TK
$a\theta$	-0,017*** (0,0027)	-0,016*** (0,0048)	-0,015*** (0,0038)	-0,0089** (0,0044)	-0,014*** (0,0021)
α	1,078*** (0,036)			0,75*** (0,26)	
β	0,474*** (0,123)			1,79 (1,35)	
α'				0,99 (2,94)	
β'				0,084 (0,40)	
a_1		1,05*** (0,15)			
p_0		0,99*** (0,25)			
a'		0,74 (0,85)			
p'_0		0,99*** (0,31)			
γ			1,05*** (0,085)		0,99*** (0,012)
δ			0,39 (1,00)		
γ'			0,96 (2,35)		0,77*** (0,072)
δ'			16,94 (123,00)		
Max LL	-45 972,178	-45 971,818	45 970,667	-45 970,320	-45 972,613

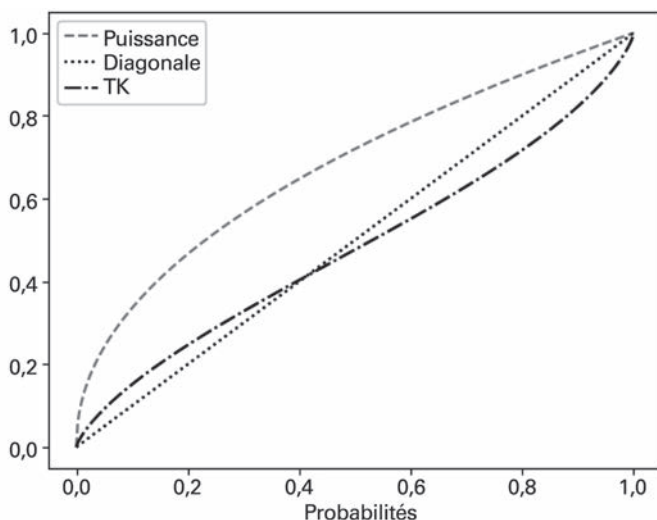
Le coefficient de la fonction puissance est statistiquement significatif, diffère de et est légèrement supérieur à 1 dans la fonction puissance de pondération des probabilités des gains et inférieur à 1 dans la fonction puissance de pondération des probabilités des pertes (col. 1 du tableau 7). Dans le cas des gains, il est très proche de la diagonale, ne reflétant presque aucune pondération des probabilités. Dans le cas des pertes, il est bien au-dessus de la droite à 45 degrés, ce qui montre clairement que les parieurs surpondèrent les probabilités de pertes, ce qui reflète le pessimisme.

Dans la fonction Tversky et Kahneman (col. 5), γ ne diffère pas de 1 dans la fonction de pondération des probabilités des gains, mais diffère de 1 dans la fonction de pondération des probabilités des pertes, avec $\gamma' = 0,77$. La fonction surpondère légèrement les faibles probabilités et sous-pondère les probabilités élevées, étant donc en forme de S inversé. La courbure de la fonction est moins prononcée que dans les études expérimentales telles que Camerer et Ho [1994] qui ont estimé une fonction de pondération des gains de 0,56, Tversky et Kahneman [1992] qui ont trouvé 0,61 pour les gains et 0,69 pour les pertes, Wu et Gonzalez [1996] qui ont trouvé 0,71 pour les gains et Abdellaoui [2000] qui a trouvé 0,60 pour les gains et 0,70 pour les pertes. Je trouve donc plus de sensibilité à des variations de probabilités loin de 0 et 1 que ces études et des effets de certitude et de possibilité moins prononcés.

Graphique 4. *Théorie des perspectives cumulées – Fonctions de pondération des probabilités estimées pour les gains*



Graphique 5. *Théorie des perspectives cumulées – Fonctions de pondération des probabilités estimées pour les pertes*



Dans la fonction Cicchetti et Dubin (col. 2), a_1 n'est pas significativement différent de 1 dans le cas des gains, donc p_0 n'est pas identifié et le modèle est équivalent à l'utilité espérée. Les paramètres estimés de la fonction de pondération des probabilités des pertes Cicchetti et Dubin ont des écarts types très élevés, de sorte que je ne peux tirer aucune conclusion de leurs valeurs. Les paramètres sont également estimés de manière trop imprécise dans les spécifications Lattimore, Baker et Witte (col. 3) et Prelec (col. 4) pour tirer une quelconque

conclusion. En examinant les coefficients estimés, les fonctions de pondération des probabilités Prelec et Lattimore, Baker et Witte présentent la même forme que la fonction puissance et la fonction Cicchetti et Dubin est proche de celle de Tversky et Kahneman. Cependant, la signification statistique des paramètres ne permet pas de rejeter l'utilité espérée, même si les tests du rapport de vraisemblance favorisent ces spécifications par rapport à l'utilité espérée.

Les tests du rapport de vraisemblance entre le modèle d'utilité espérée et les modèles de la théorie des perspectives cumulées sont présentés dans le tableau 8.

Tableau 8. Comparaisons des modèles de la théorie des perspectives cumulées et de l'utilité espérée : tests statistiques

	UE	TPC Puissance	TPC CD	TPC LBW	TPC Prelec	CPT TK
AIC	91 957	91 950	91 954	91 951	91 951	91 951
BIC	91 965	91 975	91 994	91 992	91 991	91 975
Rapport de vraisemblance vs UE		10	11	13,3	14	9,4
Valeur critique		6	9,5	9,5	9,5	6

Les tests du rapport de vraisemblance montrent que les modèles de la théorie des perspectives cumulées s'ajustent nettement mieux aux données que l'utilité espérée dans toutes les spécifications, de sorte que l'utilité espérée est clairement rejetée. Ce résultat est cohérent avec les conclusions de Jullien et Salanié [2000].

Les tests de Vuong entre les modèles de la théorie des perspectives cumulées utilisant les fonctions de pondération puissance et Tversky et Kahneman et les modèles d'utilité dépendante du rang tombent dans la région non concluante, sauf le modèle d'utilité dépendante du rang avec la fonction de pondération des probabilités Prelec, pour lequel je trouve que les modèles correspondant à la théorie des perspectives cumulées sont plus proches du vrai modèle. Il y a donc de faibles preuves que la théorie des perspectives cumulées donne de meilleurs résultats que la théorie de l'utilité dépendante du rang, mais je ne peux formuler aucune conclusion claire à ce sujet.

Jullien et Salanié [2000] trouvent que leurs données ne montrent pas un changement de concavité dans les fonctions de pondération des probabilités et que la fonction de pondération des probabilités des pertes est concave, tandis que la fonction de pondération des gains est linéaire. Camerer [2000] interprète ce résultat comme une nouvelle explication du biais favori-outsider : « Les parieurs aiment les outsiders parce qu'ils ont une utilité convexe et pondèrent leurs chances élevées de perdre et leurs chances faibles de gagner approximativement linéairement. Mais ils détestent les favoris parce qu'ils aiment parier ($u(x)$ est convexe), mais ont une peur disproportionnée de la faible chance de perdre quand ils parient sur un grand favori. »

Je trouve, comme Jullien et Salanié [2000], une nette différence entre la fonction de pondération des probabilités des gains, qui est quasi linéaire, et la fonction de pondération des probabilités des pertes, qui s'écarte de la droite à 45 degrés.

La fonction puissance donne le même résultat que Jullien et Salanié [2000] : la fonction puissance qui pondère les probabilités associées à l'utilité des pertes

est concave et bien au-dessus de la diagonale. J'estime également la fonction Tversky et Kahneman, qui n'a pas été estimée par Jullien et Salanié [2000], et qui suggère une histoire différente puisqu'elle est en forme de S inversé et proche de la diagonale.

Je ne suis pas en mesure de faire la distinction entre les deux modèles. L'AIC et le BIC des modèles de la fonction puissance et de la fonction Tversky et Kahneman sont très proches (cf. tableau 8). De plus, la statistique de Vuong pour les modèles non imbriqués se situe dans la région non concluante. Je ne peux donc pas conclure sur la façon dont les parieurs pondèrent les probabilités. Je ne peux pas non plus conclure sur l'attitude des parieurs face au risque puisque les deux modèles ont des implications différentes en termes de comportements des parieurs. Pour illustrer, les paramètres estimés dans le modèle Tversky et Kahneman rendent un parieur indifférent entre la loterie gagnant 200 euros avec probabilité 25 % et perdant 25 euros avec probabilité 75 % et le montant certain de 40 euros. Dans le modèle correspondant à la fonction puissance, ils rendent le parieur indifférent entre la même loterie et le montant certain de 31 euros. Dans le premier cas, le parieur montre donc un léger goût pour le risque. Dans le second cas, il est clairement averse au risque.

CONCLUSION

Le présent article s'appuie sur le modèle de Jullien et Salanié [2000] pour comparer l'ajustement des théories de l'utilité espérée, de l'utilité dépendante du rang et des perspectives cumulées aux données françaises sur les paris hippiques. Il montre que le biais favori-outsider existe en France. Il établit en outre que l'utilité dépendante du rang et la théorie des perspectives cumulées sont mieux adaptées pour expliquer les données que l'utilité espérée, ce qui suggère que les parieurs pondèrent les probabilités de façon non linéaire lorsqu'ils font des choix. Concernant l'utilité dépendante du rang, mes résultats contredisent ceux de Jullien et Salanié [2000] qui ne trouvent aucune amélioration de l'ajustement aux données avec ce modèle. Dans la théorie des perspectives cumulées, mes résultats confirment ceux de Jullien et Salanié [2000].

L'analyse souffre cependant d'une limitation principale. Je n'ai pas de données sur les parieurs individuels, donc je dois étudier le comportement d'un parieur représentatif. Cela est potentiellement problématique puisque les parieurs peuvent différer en termes d'attitude face au risque et de croyances.

Deux voies de recherche intéressantes pourraient être poursuivies pour compléter cet article. Premièrement, je ne teste qu'une forme de la fonction d'utilité et me concentre uniquement sur la fonction de pondération des probabilités. Jullien et Salanié [2000] ont ouvert cette voie possible de recherche en testant une fonction d'utilité avec aversion absolue hyperbolique au risque. Cependant, comme ils ne disposaient pas de données sur la richesse des parieurs, ils n'ont pas pu aller très loin et ont finalement choisi d'utiliser une fonction d'utilité avec aversion absolue constante au risque pour plus de commodité. Je fais face aux mêmes limitations dues aux données. Deuxièmement, mes données ne permettent pas de tester l'aversion aux pertes, qui est l'une des principales caractéristiques de la théorie des perspectives cumulées. Des données sur le montant des paris pourraient permettre de suivre cette voie dans les recherches futures.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ABDELLAOUI M. [2000], « Parameter-Free Elicitation of Utility and Probability Weighting Functions », *Management Science*, 43 (11), p. 1497-1512.
- ALI M. [1977], « Probability and Utility Estimates for Racetrack Bettors », *Journal of Political Economy*, 85 (4), p. 803-815.
- ALLAIS M. [1953], « Les comportements de l'homme rationnel devant le risque : critique des postulats de l'école américaine », *Econometrica*, 21 (4), p. 503-546.
- ASCH P., MALKIEL B. et QUANDT R. [1982], « Racetrack Betting and Informed Behavior », *Journal of Financial Economics*, 10 (2), p. 187-194.
- BARSEGHYAN L., MOLINARI F., O'DONOGHUE T. et TEITELBAUM J. [2013], « The Nature of Risk Preferences: Evidence from Insurance Choices », *The American Economic Review*, 103 (6), p. 2499-2529.
- BIRD R., MCCRAE M. et BEGGS J. [1987], « Are Gamblers Really Risk Takers? », *Australian Economic Papers*, 26 (49), p. 237-253.
- CAMERER C. [2000], « Prospect Theory in the Wild: Evidence from the Field », dans D. KAHNEMAN et A. TVERSKY (dir.), *Choices, Values and Frames*, Cambridge, Cambridge University Press, p. 288-300.
- CAMERER C. et HO T. [1994], « Violations of the Betweenness Axiom and Nonlinearity in Probability », *Journal of Risk and Uncertainty*, 8 (2), p. 167-196.
- CICCHETTI C. et DUBIN J. [1994], « A Microeconomic Analysis of Risk Aversion and the Decision to Self-Insure », *Journal of Political Economy*, 102 (1), p. 169-186.
- COLEMAN L. [2002], « Risk and the Illusion of Control: Racetrack Betting in Melbourne », contribution à la 31^e Australian Conference of Economists, Adélaïde, 3 octobre.
- COLEMAN L. [2004], « New Light on the Longshot Bias », *Applied Economics*, 36 (4), p. 315-326.
- GANDAR J., ZUBER R. et JOHNSON R. [2001], « Searching for the Favourite-Longshot Bias Down Under: An Examination of the New Zealand Pari-Mutuel Betting Market », *Applied Economics*, 33 (13), p. 1621-1629.
- GONZALEZ R. et WU G. [1999], « On the Shape of the Probability Weighting Function », *Cognitive Psychology*, 38 (1), p. 129-166.
- GRIFFITH R. [1949], « Odds Adjustments by American Horse-Race Bettors », *The American Journal of Psychology*, 62 (2), p. 290-294.
- JULLIEN B. et SALANIÉ B. [2000], « Estimating Preferences under Risk: The Case of Racetrack Bettors », *Journal of Political Economy*, 108 (3), p. 503-530.
- KAHNEMAN D. et TVERSKY A. [1979], « Prospect Theory: An Analysis of Decision under Risk », *Econometrica*, 47 (2), p. 263-292.
- KLIGER D. et LEVY O. [2009], « Theories of Choice Under Risk : Insights from Financial Markets », *Journal of Economic Behavior and Organization*, 71 (2), p. 330-346.
- KÖSZEGI B. et RABIN M. [2006], « A Model of Reference-Dependent Preferences », *Quarterly Journal of Economics*, 121 (4), p. 1133-1165.
- KÖSZEGI B. et RABIN M. [2007], « Reference-Dependent Risk Attitudes », *The American Economic Review*, 97 (4), p. 1047-1073.
- KÖSZEGI B. et RABIN M. [2009], « Reference-Dependent Consumption Plans », *The American Economic Review*, 99 (3), p. 909-936.
- LATTIMORE P., BAKER J. et WITTE A. [1992], « The Influence of Probability in Risky Choices: A Parametric Examination », *Journal of Economic Behavior and Organization*, 17 (3), p. 377-400.
- LEVITT S. et LIST J. [2008], « Homo Economicus Evolves », *Science*, 319 (5865), p. 909-910.
- LIST J. [2003], « Does Market Experience Eliminate Market Anomalies? », *Quarterly Journal of Economics*, 118 (1), p. 41-71.
- LIST J. [2004], « Neoclassical Theory versus Prospect Theory: Evidence from the Market-Place », *Econometrica*, 72 (2), p. 615-625.
- MCGLOTHLIN W. [1956], « Stability of Choices among Uncertain Alternatives », *The American Journal of Psychology*, 69 (4), p. 604-615.

- NEUMANN J. VON et MORGENSTERN O. [1947], *Theory of Games and Economic Behavior*, Princeton, Princeton University Press.
- OTTAVIANI M. et SØRENSEN P. [2008], « The Favorite-Longshot Bias: An Overview of the Main Explanations », dans D. HAUSCH et W. T. ZIEMBA (dir.), *Handbook of Sports and Lottery Markets*, Amsterdam, North Holland, p. 83-101.
- POPE D. et SCHWEITZER M. [2011], « Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition and High Stakes », *The American Economic Review*, 101 (1), p. 129-157.
- POST T., VAN DEN ASSEM M., BALTUSSEN G. et THALER R. [2008], « Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show », *The American Economic Review*, 98 (1), p. 38-71.
- PRELEC D. [1998], « The Probability Weighting Function », *Econometrica*, 66 (3), p. 497-527.
- QUIGGIN J. [1982], « A Theory of Anticipated Utility », *Journal of Economic Behavior and Organization*, 3 (4), p. 323-343.
- SAUER R. [1998], « The Economics of Wagering Markets », *Journal of Economic Literature*, 36 (4), p. 2021-2064.
- SNOWBERG E. et WOLFERS J. [2010], « Explaining the Favorite-Long Shot Bias: Is It Risk-Love or Misperceptions? », *Journal of Political Economy*, 118 (4), p. 726-746.
- SNYDER W. [1978], « Horse Racing: Testing the Efficient Markets Model », *The Journal of Finance*, 33 (4), p. 1109-1118.
- TVERSKY A. et KAHNEMAN D. [1992], « Advances in Prospect Theory: Cumulative Representation of Uncertainty », *Journal of Risk and Uncertainty*, 5, p. 297-323.
- WAKKER P., THALER R. et TVERSKY A. [1997], « Probabilistic Insurance », *Journal of Risk and Uncertainty*, 15, p. 7-28.
- WEITZMAN M. [1965], « Utility Analysis and Group Behavior: An Empirical Study », *Journal of Political Economy*, 73 (1), p. 18-26.
- WILLIAMS L. V. et PATON D. [1997], « Why Is There a Favourite-Longshot Bias in British Racetrack Betting Markets », *The Economic Journal*, 107 (440), p. 150-158.
- WU G. et GONZALEZ R. [1996], « Curvature of the Probability Weighting Function », *Management Science*, 42 (12), p. 1676-1690.

ANNEXES

I. VALEUR GLOBALE D'UN PARI DANS CHAQUE MODÈLE

Le parieur représentatif parie a dans chaque course et est doté d'une richesse initiale M . Le choix d'un cheval dépend uniquement des probabilités de gagner des chevaux et des cotes finales écrites $((O_1, p_1), (O_2, p_2), \dots, (O_N, p_N))$ pour les N chevaux d'une course.

Modèle d'utilité espérée

La richesse finale du parieur après un pari sur le cheval i peut s'élever soit à $M - a$ si le cheval perd la course, soit à $M + aO_i$ si le cheval termine premier. Le premier se produit avec la probabilité $1 - p_i$ et le second avec la probabilité p_i . La valeur globale d'un pari sur le cheval i est donc $W(p_i, a, M, O_i) = p_i * u(M + aO_i) + (1 - p_i) * u(M - a)$, où u est une fonction d'utilité continue et strictement croissante avec $u(0) = 0$.

Modèle d'utilité dépendante du rang

Dans le modèle d'utilité dépendante du rang (Quiggin [1982]), la valeur globale d'une loterie avec deux résultats possibles est égale à l'utilité dérivée du pire résultat, que le décideur est sûr d'obtenir, plus l'augmentation possible d'utilité liée à l'obtention du meilleur résultat, pondérée par la probabilité pondérée d'obtenir le meilleur résultat.

La valeur globale d'un pari sur le cheval i est donc $W(p_i, a, M, O_i) = \varphi(p_i) * u(M + aO_i) + (1 - \varphi(p_i)) * u(M - a)$, où φ est la fonction de pondération des probabilités, continue et strictement croissante de $[0, 1]$ à $[0, 1]$ et satisfaisant $\varphi(0) = 0$ et $\varphi(1) = 1$.

Modèle de la théorie des perspectives cumulées

Le modèle de la théorie des perspectives cumulées développé par Tversky et Kahneman [1992] s'écarte du modèle d'utilité dépendante du rang, car les résultats sont perçus comme des gains et des pertes par rapport à un point de référence. De plus, la fonction d'utilité diffère pour les gains et pour les pertes. Elle est généralement concave pour les gains et convexe pour les pertes et plus pentue pour les pertes que pour les gains pour refléter l'aversion aux pertes.

La fonction de pondération des probabilités surpondère les probabilités faibles et sous-pondère les probabilités moyennes et élevées. Elle est en forme de S inversé, ce qui signifie concave puis convexe. Plus elle est courbée, plus la sensibilité à des faibles variations des probabilités près des extrêmes de l'échelle des probabilités est élevée. Cette propriété est appelée sensibilité décroissante. Le point où la fonction coupe la diagonale se situe à un niveau de probabilité d'environ 1/3.

Un parieur gagnant obtient aO_i . Cela arrive avec la probabilité p_i . Les pertes s'élèvent à a et se produisent avec la probabilité $1 - p_i$. La fonction de pondération des probabilités des gains est écrite φ^+ et celle des pertes φ^- . Les deux fonctions sont strictement continues et croissantes à partir de l'intervalle unitaire à l'intervalle unitaire et satisfont $\varphi^+(0) = \varphi^-(0) = 0$ et $\varphi^+(1) = \varphi^-(1) = 1$. La valeur globale d'un pari sur le cheval i est $W(p_i, a, O_i) = \varphi^+(p_i) * u(aO_i) + \varphi^-(1 - p_i) * u(-a)$.

Notons que l'utilité est la même pour les gains et les pertes. La raison en est que a n'est pas observé, donc je n'ai pas assez de données pour identifier u dans le domaine des pertes. L'aversion aux pertes n'est donc pas modélisée ici. Notons également que le point de référence est le *statu quo*, qui correspond à ne pas parier. Deux aspects clés de la théorie des perspectives cumulatives sont donc modélisés ici : la dépendance au point de référence et la pondération différente des gains et des pertes.

II. CALCUL DE P1 POUR ESTIMER LES PARAMÈTRES DU MODÈLE

Modèle d'utilité espérée

Soit a la somme mise par le parieur représentatif, M sa richesse initiale, O_i la cote associée au cheval i , p_i la probabilité de gagner du cheval i et w une constante. Je définis aussi u , une fonction d'utilité continue et strictement croissante avec $u(0) = 0$.

À l'équilibre,

$$\forall i \in 1, \dots, N, p_i * u(M + aO_i) + (1 - p_i) * u(M - a) = w.$$

Donc :

$$p_i = \frac{w - u(M - a)}{u(M + aO_i) - (M - a)}. \quad (6)$$

Comme $\sum_{i=1}^N p_i = 1$,

$$w = u(M - a) + \frac{1}{\sum_{j=1}^N \frac{1}{u(M + aO_j) - u(M - a)}}. \quad (7)$$

Combiner les équations 6 et 7 permet de trouver p_i :

$$p_i = \frac{1}{u(M + aO_i) - u(M - a)} * \frac{1}{\sum_{j=1}^N \frac{1}{u(M + aO_j) - u(M - a)}}. \quad (8)$$

Étant donné l'équation 8 et la forme de u ,

$$p_i = \frac{1}{e^{a\theta} - e^{-a\theta O_i}} * \frac{1}{\sum_{j=1}^N \frac{1}{e^{a\theta} - e^{-a\theta O_j}}}. \quad (9)$$

Notons que a , qui n'est pas observé dans les données, ne peut pas être séparé de θ .

Modèle d'utilité dépendante du rang

Soit a la somme mise par le parieur représentatif, M sa richesse initiale, O_i la cote associée au cheval i , p_i la probabilité de gagner du cheval i et w une constante. Je définis aussi u , une fonction d'utilité continue et strictement croissante avec $u(0) = 0$ et φ une fonction de pondération des probabilités, continue et strictement croissante de $[0,1]$ à $[0,1]$ et satisfaisant $\varphi(0) = 0$ et $\varphi(1) = 1$.

À l'équilibre,

$$\forall i \in 1, \dots, N, \varphi(p_i) * u(M + aO_i) + (1 - \varphi(p_i)) * u(M - a) = w.$$

Donc :

$$\varphi(p_i) = \frac{w - u(M - a)}{u(M + aO_i) - (M - a)}. \quad (10)$$

En notant Ψ la réciproque de φ (qui existe puisque φ est strictement croissante) :

$$p_i = \Psi \left(\frac{w - u(M - a)}{u(M + aO_i) - u(M - a)} \right). \quad (11)$$

Comme $\sum_{i=1}^N p_i = 1$:

$$\sum_{j=1}^N \Psi \left(\frac{w - u(M - a)}{u(M + aO_j) - u(M - a)} \right) = 1. \quad (12)$$

Résoudre cette équation, ce qui ne peut être fait analytiquement, permet d'obtenir w . Remplacer w dans l'équation 11 permet de trouver p_i .

Théorie des perspectives cumulées

Soit a le montant misé par le parieur représentatif, O_i la cote associée au cheval i et p_i la probabilité de gagner du cheval i . Je définis aussi u , une fonction d'utilité continue et strictement croissante avec $u(0) = 0$ et j'écris φ^+ la fonction de pondération des probabilités des gains et φ^- celle de pertes. Les deux fonctions de pondération des probabilités sont strictement continues et croissantes à partir de l'intervalle unitaire en lui-même et satisfont $\varphi^+(0) = \varphi^-(0) = 0$ et $\varphi^+(1) = \varphi^-(1) = 1$.

À l'équilibre,

$$\forall i \in 1, \dots, N, \varphi^+(p_i) * u(aO_i) + \varphi^-(1 - p_i) * u(-a) = w. \quad (13)$$

Je combine le fait que $\sum_{i=1}^N p_i = 1$ et l'équation 13 pour trouver p_i . Cela ne peut être fait analytiquement, donc p_i est obtenu numériquement.

III. DÉTAILS SUR LES FONCTIONS DE PONDÉRATION DES PROBABILITÉS PAR MODÈLE

Modèle d'utilité dépendante du rang

Fonction de pondération des probabilités puissance

La fonction de pondération des probabilités a la forme suivante :

$$\varphi(p) = p^\alpha,$$

où $\alpha \geq 0$. Si la probabilité pondérée de gagner le pari est inférieure à la probabilité ($\alpha > 1$), les parieurs sous-estiment la valeur globale du pari, ils sont donc pessimistes. Le modèle d'utilité espérée est équivalent à ce modèle lorsque $\alpha = 1$.

Fonction de pondération des probabilités Cicchetti et Dubin

La fonction introduite par Cicchetti et Dubin [1994] est :

$$\frac{\varphi(p)}{1 - \varphi(p)} = \left(\frac{p}{1 - p} \right)^{a_1} * \left(\frac{p_0}{1 - p_0} \right)^{1 - a_1}.$$

$\varphi(p)$ croise la diagonale en p_0 . a_1 est positif. Si $a_1 < 1$, la fonction est convexe, puis concave. Inversement, si $a_1 > 1$, elle est d'abord concave, puis convexe. Plus a_1 est proche de 1, plus la fonction est proche de la diagonale, ce qui signifie relativement peu de sensibilité aux petites variations de probabilité près de l'extrême de l'échelle des probabilités et une haute sensibilité loin des extrêmes de l'échelle des probabilités. Lorsque $a_1 = 1$, nous revenons au modèle d'utilité espérée et p_0 n'est pas identifié.

La fonction est strictement croissante, son inverse est :

$$\Psi(p) = \frac{\left(\frac{p}{(1-p)*A} \right)^{1/a_1}}{1 + \left(\frac{p}{(1-p)*A} \right)^{1/a_1}} \quad \text{avec } A = \left(\frac{p_0}{(1-p_0)} \right)^{1-a_1}.$$

Fonction de pondération des probabilités Lattimore, Baker et Witte

La fonction proposée par Lattimore, Baker et Witte [1992] est :

$$\varphi(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma},$$

où δ et γ sont strictement positifs. δ contrôle principalement l'élévation de la fonction. Il capte l'ampleur du pessimisme ou de l'optimisme. γ contrôle principalement la courbure (c'est-à-dire la sensibilité aux changements de probabilité). Lorsque $\delta=1$ et $\gamma=1$, nous revenons au modèle d'utilité espérée. Cette fonction est strictement croissante, son inverse Ψ est :

$$\Psi(p) = \frac{\left(\frac{p}{\delta(1-p)} \right)^{1/\gamma}}{1 + \left(\frac{p}{\delta(1-p)} \right)^{1/\gamma}}.$$

Fonction de pondération des probabilités Prelec

La fonction de pondération des probabilités proposée par Prelec [1998] est :

$$\varphi(p) = e^{-\beta(-\ln p)^\alpha},$$

où $0 < \alpha < 1$ et $\beta > 0$. Si $\beta=1$ et $\alpha=1$, nous revenons au modèle d'utilité espérée. Il est équivalent à la fonction puissance pour $\alpha=1$. α représente la sensibilité aux probabilités : plus il est petit, plus la fonction est incurvée. $\beta < 1$ montre l'optimisme, $\beta > 1$ le pessimisme. β contrôle donc l'élévation de la fonction.

La fonction est strictement croissante, je note son inverse Ψ .

$$\Psi(p) = \exp \left[- \left(- \frac{\ln p}{\beta} \right)^{1/\alpha} \right].$$

Fonction de pondération des probabilités Tversky et Kahneman

La fonction de pondération des probabilités de Tversky et Kahneman [1992] est :

$$\varphi(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}}.$$

Lorsque $\gamma=1$, nous revenons au modèle d'utilité espérée.

Théorie des perspectives cumulées

Fonctions de pondération des probabilités puissance

Les fonctions de pondération des probabilités ont les formes suivantes :

$$\varphi^+(p) = p^\alpha, \varphi^-(p) = p^\beta.$$

Lorsque $\alpha=1$ et $\beta=1$, nous revenons au modèle d'utilité espérée, sauf que l'utilité s'applique à des gains et des pertes plutôt qu'à la richesse finale. Comme M s'annule dans le modèle d'utilité espérée, le modèle d'utilité espérée est imbriqué dans ce modèle.

Fonctions de pondération des probabilités Cicchetti et Dubin

On suppose que les fonctions de pondération des probabilités ont les formes suivantes :

$$\frac{\varphi^+(p)}{1-\varphi^+(p)} = \left(\frac{p}{1-p} \right)^{a_1} * \left(\frac{p_0}{1-p_0} \right)^{1-a_1},$$

$$\frac{\varphi^-(p)}{1-\varphi^-(p)} = \left(\frac{p}{1-p} \right)^{a'_1} * \left(\frac{p'_0}{1-p'_0} \right)^{1-a'_1}.$$

Ces fonctions croisent la diagonale en p_0 et p'_0 . a_1 et a'_1 sont positifs. Lorsque $a_1 = a'_1 = 1$, nous retournons au modèle d'utilité espérée. $a_1 = a'_1$ et $p_0 + p'_0 = 1$ est équivalent à l'utilité dépendante du rang avec la fonction de pondération des probabilités de Cicchetti et Dubin. Les fonctions de pondération des gains et des pertes sont identiques lorsque $a_1 = a'_1$ et $p_0 = p'_0$.

Fonctions de pondération des probabilités Lattimore, Baker et Witte

On suppose que les fonctions de pondération des probabilités ont les formes suivantes :

$$\varphi^+(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma},$$

$$\varphi^-(p) = \frac{\delta' p^{\gamma'}}{\delta' p^{\gamma'} + (1-p)^{\gamma'}},$$

où δ, γ, δ' et γ' sont strictement positifs. Quand $\delta = \gamma = \delta' = \gamma' = 1$, nous revenons au modèle d'utilité espérée. Les fonctions de pondération des gains et des pertes sont identiques lorsque $\delta = \delta'$ et $\gamma = \gamma'$.

Fonctions de pondération des probabilités Prelec

On suppose que les fonctions de pondération des probabilités ont les formes suivantes :

$$\varphi^-(p) = e^{-\beta(-\ln p)^\alpha},$$

$$\varphi^-(p) = e^{-\beta'(-\ln p)^{\alpha'}},$$

où $0 < \alpha < 1$, $\beta > 0$, $0 < \alpha' < 1$ et $\beta' > 0$. Lorsque $\beta = \alpha = \beta' = \alpha' = 1$, nous revenons au modèle d'utilité espérée. Ce modèle équivaut à la fonction puissance lorsque $\alpha = \alpha' = 1$. Les fonctions de pondération des gains et des pertes sont identiques lorsque $\alpha = \alpha'$ et $\beta = \beta'$.

Fonctions de pondération des probabilités Tversky et Kahneman

On suppose que les fonctions de pondération des probabilités ont les formes suivantes :

$$\varphi^+(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}},$$

$$\varphi^-(p) = \frac{p^{\gamma'}}{[p^{\gamma'} + (1-p)^{\gamma'}]^{1/\gamma'}}.$$

γ et γ' sont positifs. Si $\gamma = \gamma' = 1$, nous revenons au modèle d'utilité espérée. Les fonctions de pondération des gains et des pertes sont identiques lorsque $\gamma = \gamma'$.