



HAL
open science

Référentiel et attitude face au risque pour différents niveaux d'enjeux de loteries : une investigation théorique et expérimentale

Vincent Lenglin

► **To cite this version:**

Vincent Lenglin. Référentiel et attitude face au risque pour différents niveaux d'enjeux de loteries : une investigation théorique et expérimentale. Economies et finances. Université de Lille, 2019. Français. NNT : 2019LILUA030 . tel-03587130

HAL Id: tel-03587130

<https://theses.hal.science/tel-03587130>

Submitted on 24 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE DE LILLE

Ecole doctorale SESAM (ED n°73)

Laboratoire Lille Economie et Management (LEM) – UMR CNRS 9221

THÈSE

Pour obtenir le grade de

DOCTEUR EN SCIENCES ÉCONOMIQUES

Présentée par

Vincent LENGLIN

**RÉFÉRENTIEL ET ATTITUDE FACE AU RISQUE POUR
DIFFÉRENTS NIVEAUX D'ENJEUX DE LOTERIES**

-

Une investigation théorique et expérimentale

Sous la direction de

Jérôme FONCEL, Professeur d'Economie à l'Université de Lille

Thèse soutenue le 18 Juin 2019 devant le jury composé de :

M. Sacha BOURGEOIS-GIRONDE, Professeur d'Economie à l'Université Paris 2 Panthéon Assas, Rapporteur

Mme. Meglena JELEVA, Professeur d'Economie à l'Université Paris X Paris Nanterre, Examineur (Président du jury)

M. Jérémie JOZEFOWIEZ, Maître de conférences en Psychologie à l'Université de Lille, Examineur

M. Stefano PALMINTERI, Professeur de Psychologie à l'École Normale Supérieure de Paris, Rapporteur

Mme. Béatrice REY-FOURNIER, Professeur d'Economie à l'Université de Lyon 2, Examineur

L'université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses. Ces opinions doivent être considérées comme propres à leurs auteurs.

Remerciements

Je tiens avant tout à remercier mon directeur de thèse Jérôme Foncel pour son investissement et son soutien tout au long de cette thèse. Merci d'avoir accepté de m'encadrer et de m'avoir permis d'accomplir ce projet qui me tenait particulièrement à cœur. Ce fut un réel plaisir de travailler sous sa supervision.

Je remercie M. Darcheville, Professeur émérite en psychologie comportementale, qui fut le premier avec qui nous avons pu monter des projets en économie comportementale. La paternité de mon parcours vous revient.

Merci également M. Jozefowicz, M. Arighi et M. Champarnaud pour toutes les discussions et aides diverses que vous m'avez apportées durant toutes ces années.

Mes remerciements ne seront jamais à la hauteur du soutien et de l'aide de mon collègue et ami Laurent Ott, ingénieur de recherche au laboratoire de Scalab. Sans toi, cette thèse aurait simplement été infaisable. Merci pour ta patience, merci d'avoir partagé tes connaissances.

Merci également à mon ami Germain Lefebvre, post-doctorant au laboratoire de Neurosciences à Oxford. Comme pour Laurent, merci pour ton aide et soutien. C'est toujours un grand plaisir de travailler et discuter avec toi !

Je tiens à remercier particulièrement tous les donateurs qui ont fait des dons sur la plateforme EDUKLAB qui m'ont permis de réaliser et financer des expériences au cours de cette thèse. Bien que mes premières expériences ne fussent pas concluantes dans un premier temps, cela m'a permis de produire deux chapitres de cette thèse : le chapitre 2 et 4. Vous avez cru dans mon projet et pour cela je vous suis plus que jamais reconnaissant.

Merci au laboratoire Scalab de m'avoir accueilli dans ses murs alors que je ne faisais pas partie du laboratoire. Votre bonne humeur m'a garanti un équilibre qui m'a permis de tenir dans les moments difficiles. Un énorme merci à Alice, Jérémy, Margot, Charlotte, Audrey, Clémence, Lucie, Arthur, Mike, Francis, Laura et Sohir.

Merci à mes colocataires Alexandra et Tom !! Merci pour votre bonne humeur, votre aide à tous

niveaux, toutes ces années magnifiques passées avec vous, j'en garde un fabuleux souvenir !!

Merci aux étudiants et tous les participants d'avoir participé aux expériences et pour certains d'entre eux d'avoir travaillé sur le sujet. Et merci aux correcteurs de fautes

d'orthographe ! Lounès, Lucie, Alice et maman : merci d'avoir rendu cette thèse un peu plus lisible !

Merci à mes ami(e)s, merci de m'avoir sorti et soutenu durant toutes ces années où je n'étais pas forcément facile à vivre à certains moments ! Merci à Camille, Marc-Antoine, Cyril, Thibaut, Hubert, Bertrand, Clément, Robin, Philippe, Guillaume, Eve, Gaëtan, Pierre, Gabrielle, Lounès, Renaud, Nicolas, Louise, Baptiste, Nicolas, Jean-Charles, Yann, Guillaume.

Sandra, mes remerciements pour toi sont incalculables. Merci d'avoir été à mes côtés durant cette thèse, merci d'y avoir activement participé, d'avoir fait tout ce que tu as pu durant cette période. Mes remerciements pour toi vont bien au-delà de ces lignes.

Enfin, mes derniers remerciements inconditionnels vont à ma famille : papa, maman, cette thèse vous est dédiée. Je ne pourrai jamais vous rendre tout ce que vous m'avez donné. Votre confiance a été un trésor inestimable pour moi... Merci de m'avoir soutenu dans les bons et mauvais moments. J'espère que cette thèse, vous rendra fières. Merci à mes chères sœurs et mes beaux-frères également pour votre soutien indéfectible ! Cette expérience n'aurait pas été possible sans vous, sans votre soutien sans faille, sans votre confiance en moi.

*« Le voyage se passe de motifs. Il ne tarde pas à prouver
qu'il se suffit à lui-même. On croit qu'on va faire un
voyage, mais bientôt c'est le voyage qui vous fait, ou
vous défait ».*

Nicolas Bouvier.

A mes parents.

RÉSUMÉ

L'étude des prises de décision en environnement probabiliste constitue une part substantielle de la théorie économique et psychologique. Les modèles qui ont été proposés dans ce cadre ont considérablement évolués depuis la théorie normative de l'espérance d'utilité proposée par Von Neumann et Morgenstern (1944). Cette évolution est véritablement le fruit de la confrontation des modèles aux données expérimentales. Dès lors, les chercheurs se sont depuis évertués à renforcer le pouvoir descriptif des modèles tout en essayant de conserver une assise axiomatique convenable.

Au fur et à mesure, la théorie économique a intégré de plus en plus de composantes étudiées par la psychophysique et confère ainsi aux modèles initialement « *as if* » une réalité psychologique plus prégnante.

Dans cette perspective, la Prospect Theory a proposé un modèle de décision « référence-dépendant ». Selon cette théorie, les individus évaluent les issues impliquées dans des loteries par rapport à un référentiel au lieu d'évaluer l'utilité des richesses finales prospectives. Toutefois, la principale difficulté aujourd'hui consiste à spécifier correctement ce référentiel et d'en déduire les implications en termes de préférences face au risque.

Cette thèse propose dans un premier temps de faire l'état des lieux des différents points de référence étudiés par la littérature que ce soit dans un cadre de choix statique ou dynamique.

Nous verrons ensuite que la question du référentiel est intimement liée à la façon dont les individus jugent les issues des loteries impliquées dans un certain contexte. Dans cette perspective, une problématique fréquemment étudiée est l'influence de l'amplitude des issues sur l'attitude face au risque. Ces recherches portent plus précisément sur l'impact de la taille d'enjeux des loteries sur l'aversion au risque. Nous montrerons que le pattern comportemental qui ressort de ces expériences n'est actuellement pas capturable par les modèles existants.

Même si nous ne sommes pas en mesure de fournir une solution théorique, nous mettons en avant que la solution sera sans doute à rechercher du côté d'une spécification particulière du référentiel.

Pour finir, nous étudierons un nouveau modèle proposé par Baucells, Kontek et Lewandowski (2018, working paper) qui propose une spécification originale du référentiel. Nous verrons que pour la première fois, le référentiel est vu selon deux dimensions : une dimension séparant le domaine des pertes du domaine des gains, et une dimension relative à l'étendue des issues impliquées par des loteries d'un contexte décisionnel particulier. A travers notre dernier chapitre nous nous évertuerons à étudier quelques implications de ce modèle afin d'en éprouver les prédictions à travers une analyse expérimentale.

Table des matières

INTRODUCTION GÉNÉRALE	21
1. LES MODELES DE DECISION DANS LE RISQUE	33
INTRODUCTION	35
I. BASIQUES ET HISTORIQUE	36
1.1. ENVIRONNEMENT RISQUE, VALEUR SUBJECTIVE ET PREFERENCE	36
1.1.1. CHOIX EN ENVIRONNEMENT RISQUE	36
1.1.2. LOTERIES	36
1.1.3. VALEURS SUBJECTIVES ET PREFERENCES : NOTATION	38
1.2. LE PARADOXE DE BERNOULLI	39
1.2.1. L'ESPERANCE MATHEMATIQUE	39
1.2.2. LE PARADOXE	40
1.2.3. RESOLUTION DU PARADOXE	41
II. LA THEORIE DE L'UTILITE ESPEREE	42
2.1. AXIOMATIQUE DE L'EUT	43
2.1.1. AXIOME 1 : LES PREFERENCES SONT COMPLETES, REFLEXIVES ET TRANSITIVES	43
2.1.2. AXIOME 2 : CONTINUITÉ	44
2.1.3. AXIOME 3 : L'INDEPENDANCE	44
2.1.4. THEOREME DE L'UTILITE ESPEREE	45
2.2. LA FONCTION D'ESPERANCE D'UTILITE	45
2.3. LE TRIANGLE DE MARSCHAK-MACHINA	46
III. VIOLATIONS DE L'EUT ET GENERALISATION DU MODELE	50
3.1. LE PARADOXE D'ALLAIS	51
3.1.1. EFFET DE LA CONSEQUENCE COMMUNE : VIOLATION DE LA PROPRIETE ADDITIVE DE L'AXIOME D'INDEPENDANCE	51
3.1.2. EFFET DU RATIO COMMUN : VIOLATION DE LA PROPRIETE MULTIPLICATIVE DE L'AXIOME D'INDEPENDANCE	52
3.1.3. ANALYSE DE L'EFFET DE LA CONSEQUENCE COMMUNE AVEC LE TRIANGLE DE MARSCHAK	53
3.2. LES PREFERENCES NON-LINEAIRES DANS LES PROBABILITES	55
3.2.1. FONCTION DE PONDERATION DES PROBABILITES ET DOMINANCE STOCHASTIQUE DE PREMIER ORDRE	55
3.2.2. LA RANK-DEPENDENT UTILITY THEORY (RDUT)	57
3.2.3. LA DUAL THEORY	62
3.3. FORMALISATIONS ET PROPRIETES DES FONCTIONS DE PONDERATION DES PROBABILITES	63
3.3.1. FORMALISATION DE LA FONCTION DE PONDERATION DES PROBABILITES	66
3.4. DEPENDANCE AUX RANGS ET TRIANGLE DE MARSCHAK-MACHINA	68
4.1. L'ESPERANCE D'UTILITE ET ATTITUDE FACE AU RISQUE	74
4.2. LA RDUT ET L'ATTITUDE FACE AU RISQUE	78
V. L'APPROCHE EXPERIMENTALE	81
5.1. METHODE DE REVELATION DES PREFERENCES	82
5.1.1. LE MULTIPLE PRICE LIST (MPL)	82
5.1.2. LA METHODE DES TRADEOFFS (TO)	83
5.2. RESULTATS EXPERIMENTAUX	85

5.2.1.	LA METHODE DETERMINISTE	86
5.2.2.	L'APPROCHE STOCHASTIQUE	87
CONCLUSION DE LA PARTIE 1		89
ANNEXE		91
A. REPRESENTATION SCHEMATIQUE DE L'EUT		91
2. LES MODELES DE PREFERENCES REFERENCE-DEPENDANT		93
INTRODUCTION		95
I. LA PROSPECT THEORY		101
1.1. LES ETUDES DE KAHNEMAN ET TVERSKY		101
1.1.1.	HYPOTHESE D'INTEGRATION VS HYPOTHESE D'ISOLATION	101
1.1.2.	LE REFLECTION EFFECT	103
1.1.3.	CONCAVITE ET CONVEXITE DE LA FONCTION D'UTILITE	104
1.1.4.	L'AVERSION AUX PERTES	105
1.2. LA CUMULATIVE PROSPECT THEORY		106
1.2.1.	LA VALUE FUNCTION	106
1.2.2.	LA WEIGHTING FUNCTION	107
1.2.3.	LE MODELE DE LA CPT	108
1.2.4.	CUMULATIVE PROSPECT THEORY ET ATTITUDE FACE AU RISQUE	111
II. REFERENTIEL ET ATTITUDE FACE AU RISQUE		112
2.1. APPROCHE GRAPHIQUE		113
2.1.1.	RICHESSSE ACTUELLE ET REFERENTIEL	113
2.1.2.	PROSPECT THEORY ET TRIANGLE DE MARSCHAK-MACHINA	114
2.2. APPROCHE MATHEMATIQUE		117
2.2.1.	L'ETUDE DE DAVIES ET SATCHELL (2007)	117
2.2.2.	PRIME DE RISQUE D'UNE LOTERIE ET REFERENTIEL	122
CONCLUSION A PROPOS DE LA CUMULATIVE PROSPECT THEORY		123
III. LOCALISATION DU REFERENTIEL EN CONTEXTE DE CHOIX STATIQUE		125
3.1. LES CONTEXTES DE CHOIX STATIQUE		125
3.1.1.	DEFINITION	126
3.1.2.	EXPERIMENTATION EN CHOIX STATIQUE	126
3.1.3.	CHOIX STATIQUE ET RECHERCHE DU POINT DE REFERENCE	127
3.2. LOCALISATION DU REFERENTIEL - LES CANDIDATS		128
3.2.1.	POINT DE REFERENCE « CHOICE SPECIFIC »	128
3.2.2.	POINT DE REFERENCE « PROSPECT SPECIFIC »	130
3.3. LE(S) GAGNANT(S) EST (SONT) ...		136
3.4. LES AUTRES POINTS DE REFERENCE		136
IV. LOCALISATION DU POINT DE REFERENCE EN CONTEXTE DE CHOIX DYNAMIQUE		137
4.1. L'INFLUENCE DU PASSE DANS LES CONTEXTES DE CHOIX DYNAMIQUES		138
4.1.1.	DEFINITION	138
4.1.2.	LA (IN)CONSISTANCE DYNAMIQUE ET L'HYPOTHESE DU CONSEQUENTIALISME REMISES EN QUESTION	139
4.1.3.	L'INFLUENCE DU PASSE : LES PRIORS	141
4.2. REFERENTIEL ADAPTATIF ET PRISE EN COMPTE DU PASSE		146

4.2.1.	EXPLICATION DU PATTERN OBSERVE PAR THALER ET JOHNSON	146
4.2.2.	THEORIE DE L'UTILITE ESPEREE ET PRIORS	149
4.2.3.	EQUITY PREMIUM PUZZLE (EPP) VS DISPOSITION EFFECT (DE)	150
4.3.	LES PREUVES EXPERIMENTALES DEMONTRANT L'ACTUALISATION DU REFERENTIEL	155
4.3.1.	LA METHODE INDIRECTE DE GNEEZY (2005)	155
4.3.2.	LA METHODE DIRECTE DE BAUCCELLS ET DE ARKES	157
4.3.3.	QUELQUES MESURES COMPLEMENTAIRES	160
4.3.4.	LE GRAND PROBLEME D'UN TEST COMPORTEMENTAL	161
4.4.	MODELISATION DE L'ACTUALISATION DU REFERENTIEL	162
4.4.1.	LES MODELES RECURSIFS	162
4.4.2.	LES MODELES NON RECURSIFS	168
CONCLUSION PARTIE 2		170
<hr/>		
3.	PROBABILITY DISCOUNTING ET TAILLE D'ENJEU DES LOTERIES	172
<hr/>		
INTRODUCTION		174
<hr/>		
I.	LE MODELE DE PROBABILITY DISCOUNTING	176
<hr/>		
1.1.	LE PHENOMENE DU TEMPORAL (OU DELAY) DISCOUNTING	176
1.1.1.	DEFINITION ET ETUDE	176
1.1.2.	FORMALISATION DU TEMPORAL DISCOUNTING	178
1.2.	LE PHENOMENE DU PROBABILITY DISCOUNTING	180
1.2.1.	DEFINITION ET ETUDES	180
1.2.2.	FORMALISATION DU PROBABILITY DISCOUNTING	181
1.3.	TEMPORAL DISCOUNTING ET PROBABILITY DISCOUNTING : UN PROCESSUS SOUS-JACENT COMMUN ?	182
<hr/>		
II.	EXPLICATION DU POINT DE VUE DE L'ECONOMIE DE L'EFFET DE MAGNITUDE	185
<hr/>		
2.1.	NIVEAU D'ENJEU ET ATTITUDE FACE AU RISQUE	185
2.1.1.	L'ETUDE DE MARKOWITZ (1952)	186
2.1.2.	L'ETUDE DE HOLT ET LAURY (2002)	189
2.2.	FONCTION D'UTILITE ET AVERSION RELATIVE FACE AU RISQUE	189
2.2.1.	CADRE D'ANALYSE	190
2.2.2.	PROPRIETES DES FONCTIONS D'UTILITES ET AVERSION RELATIVE FACE AU RISQUE	193
2.2.3.	APPROXIMATION DE LA PRIME DE RISQUE RELATIVE BASE SUR L'ANALYSE DE HILTON (1988)	196
2.2.4.	ELASTICITE DES FONCTIONS D'UTILITE IRRA ET CRRA	201
<hr/>		
III.	RETOUR SUR LE MODELE DE PROBABILITY DISCOUNTING	202
<hr/>		
3.1.	RETOUR SUR LES RESULTATS DU PROBABILITY DISCOUNTING	203
3.1.1.	LA PRIME DE RISQUE RELATIVE ET L'EQUIVALENT CERTAIN NORMALISE	203
3.1.2.	LE MODELE DE PROBABILITY DISCOUNTING A L'EPREUVE DES RESULTATS	204
3.2.	SIMULATION-ESTIMATION DU MODELE MGM A TROIS PARAMETRES	206
3.2.1.	LOTeries	207
3.2.2.	SIMULATION DES EQUIVALENTS CERTAINS	208
3.2.3.	ESTIMATION	209
3.3.	RESULTATS	210
3.3.1.	CONCLUSION	212
<hr/>		
IV.	LIMITES DES MODELES ECONOMIQUES FACE AU PATTERN IRRP-CRRP	213
<hr/>		
4.1.	LE PATTERN IRRP-CRRP	213
4.2.	TENTATIVE DE CAPTURE DU PATTERN IRRP-CRRP	216

4.2.1.	LES MODELES ECONOMIQUES	216
4.2.2.	LA TENTATIVE DE MYERSON, GREEN ET MORRIS (2011)	217
4.2.3.	CRITIQUE DU MODELE A FONCTIONS NON-SEPARABLES	220
CONCLUSION PARTIE 3		222
ANNEXE		223
4	EFFET DE L'ETENDUE SUR L'ATTITUDE FACE AU RISQUE	232
INTRODUCTION		234
I.	LE RANGE EFFECT	237
1.1.	LE MODELE PSYCHOPHYSIQUE DU « RANGE EFFECT »	237
1.2.	LES ETUDES EN MARKETING	240
II.	MODELE DE BAUCCELLS, KONTEK ET LEWANDOWSKI (2018)	241
2.1.	AVANT-PROPOS	242
2.2.	LE MODELE	244
2.3.	LE FOURFOLD PATTERN	247
2.4.	RANGE-EFFECT ET ATTITUDE FACE AU RISQUE	249
2.4.1.	ANALYSE GRAPHIQUE DE L'IMPACT DE G SUR L'AVERSION AU RISQUE	251
2.4.2.	ANALYSE MATHEMATIQUE DE L'IMPACT DE G SUR L'AVERSION AU RISQUE ABSOLUE	252
III.	PROBLEMATIQUE ET PROCEDURE EXPERIMENTALE	254
3.1.	METHODOLOGIE	255
3.1.1.	GROUPES EXPERIMENTAUX	255
3.1.2.	PROCEDURE	257
3.2.	HYPOTHESES SELON LE MODELE BKL	260
3.2.1.	LE CONTEXTE DE DECISION	260
3.2.2.	HYPOTHESES	261
IV.	RESULTATS	263
4.1.	BASE DE DONNEES ET ANALYSES STATISTIQUES	263
4.1.1.	BASE DE DONNEES	263
4.1.2.	ANALYSE DES DONNEES	264
4.2.	RESULTATS DE L'EXPERIENCE	265
4.2.1.	HYPOTHESE 1	265
4.2.2.	HYPOTHESE 2	267
4.2.3.	HYPOTHESE 3	269
4.3.	ANALYSE ECONOMETRIQUE	270
4.3.1.	SPECIFICATION DU MODELE BKL	270
4.3.2.	DEPENDANCE A L'ETENDUE	271
4.3.3.	SPECIFICATION DE L'ERREUR	272
4.3.4.	ESTIMATION	272
4.3.5.	RESULTATS DE L'ESTIMATION	273
4.3.6.	OUT OF SAMPLE	278
4.3.7.	NOUVELLES PISTES DE SPECIFICATIONS DU MODELE BKL	280
V.	DISCUSSION	282

CONCLUSION	286
ANNEXE	288
A. ESTIMATION AVEC UN REFERENTIEL EGAL A L'ESPERANCE	288
B. CONSIGNES ET SCREENSHOT DE L'EXPERIENCE	292
C. MEDIANES ET BOXPLOT	293
DISCUSSION GÉNÉRALE	296
CONCLUSION GENERALE	301
BIBLIOGRAPHIE	305

ABRÉVIATIONS

Abréviations des théories

EUT	Expected Utility Theory ou théorie de l'utilité espérée
RDUT	Rank-Dependent Utility Theory ou Théorie de l'utilité dépendante des rangs
DT	Dual Theory
RSU	Range and Sign dependent Utility ou Utilité dépendante du signe et de l'étendu
PT	Prospect Theory ou Théorie des perspectives
CPT	Cumulative Prospect Theory

Abréviations sur les fonctions d'utilités

DARA	Decreasing Absolute Risk Aversion ou Aversion au risque absolue décroissante
DRRA	Decreasing Relative Risk Aversion ou Aversion au risque relative décroissante
CARA	Constant Absolute Risk Aversion ou Aversion au risque absolue constante
CRRA	Constant Relative Risk Aversion ou Aversion au risque relative constante
IARA	Increasing Absolute Risk Aversion ou Aversion au risque absolue croissante
IRRA	Increasing Relative Risk aversion ou Aversion au risque relative croissante

Autres Abréviations

RRP	Relative Risk Premium ou Prime de risque relative
ARP	Absolute risk Premium ou Prime de risque absolue
Ce	Certainty Equivalent ou équivalent certain

Notations

L	Dénote une loterie
x	Dénote une conséquence d'une loterie dégénérée ou non. Lorsque les conséquences sont dénombrables et >1 on indexera la conséquence comme suit : x_i avec $i = 1, \dots, n$
w_0	Dénote la richesse initiale d'un agent
W	Dénote la richesse finale d'un individu : $W = w_0 + x$. Lorsque les richesses finales sont dénombrables et >1 , on indexera cette richesse comme suit : $W_i = w_0 + x_i$ avec $i = 1, \dots, n$
\tilde{x}	Dénote une variable aléatoire et est assimilable à la loterie L . Lorsqu'il y a plusieurs variables aléatoires (ou loteries) on les indexera comme suit : \tilde{x}^j ou L^j avec $j = 1, \dots, m$.
\tilde{W}	Dénote une richesse finale aléatoire : $\tilde{W} = w_0 + \tilde{x}$. Lorsqu'il y a plusieurs variables aléatoires
p	Dénote la probabilité d'une conséquence. Lorsque les issues sont dénombrables on notera p_i , avec $i = 1, \dots, n$, la probabilité d'une conséquence x_i d'une loterie L .
$E(.)$	Dénote l'opérateur de l'espérance d'une variable.
R	Point de référence ou référentiel

Le tilde sur une variable dénote sa nature aléatoire.

TABLE DES FIGURES

FIGURE 1.1 REDUCTION D'UNE LOTERIE COMPOSEE. SOIT UNE LOTERIE L_C DONT LES PRIX SONT L'ACCES A DES LOTERIES L_A ET L_B AVEC LA PROBABILITE A ET $1-A$ RESPECTIVEMENT. CETTE LOTERIE EN DEUX ETAPES PEUT SE REDUIRE A UNE ETAPE PAR SIMPLE APPLICATION DU THEOREME DE BAYES (1763). L'AXIOME DE REDUCTION DES LOTERIES COMPOSEES SUPPOSE QU'UN AGENT SERA INDIFFERENT ENTRE CES DEUX LOTERIES.....	38
FIGURE 1.2 ILLUSTRATION DU TRIANGLE DE MARSCHAK-MACHINA. CES DROITES ISO-UTILITES ET ISO-ESPERANCES ONT ETE CONSTRUITES A PARTIR D'UNE LOTERIE A TROIS ISSUES : $x_1 = 0, x_2 = 500, x_3 = 1000$ ET OU $w_0 = 0$. LA FONCTION D'UTILITE EST DONNEE PAR $u(x) = \alpha x$, AVEC $\alpha = 0.8$. PAR EXEMPLE, LE POINT BLEU REPRESENTE UNE LOTERIE PRODIGUANT 1000€ AVEC LA PROBABILITE $p_3=0.3, 0$ AVEC $p_1=0.1$ ET 500 AVEC $p_2=0.6$	47
FIGURE 1.3 TRIANGLE DE MARSCHAK-MACHINA ET FONCTION D'UTILITE (ENCADRE SUPERIEUR). LES DROITES ISO-UTILITES A GAUCHE SONT ELABORES EN SUPPOSANT UNE FONCTION PUISSANCE CONVEXE. LA FIGURE AU MILIEU, A PARTIR D'UNE FONCTION D'UTILITE CONCAVE ET LA FIGURE DE DROITE A PARTIR D'UNE FONCTION D'UTILITE LINEAIRE.....	49
FIGURE 1.4 REPRESENTATION DU PARADOXE D'ALLAIS « EFFET DE LA CONSEQUENCE COMMUNE » DANS LE TRIANGLE DE MARSCHAK-MACHINA. LES POINTS A, A', B ET B' REPRESENTENT LES LOTERIES UTILISEES PAR ALLAIS DANS SES DEUX SITUATIONS.	55
FIGURE 1.5 POIDS DECISIONNELS DANS LE MODELE RDU, MONTRANT LES EFFETS DE LA DEPENDANCE DE RANG SUR LES POIDS DECISIONNELS POUR UNE DISTRIBUTION DE PROBABILITE UNIFORME DISCRETE P SUR 20 RESULTATS CLASSES i (TRIANGLE NOIRES), SELON LES CARACTERISTIQUES DE LA FONCTION DE PONDERATION DE PROBABILITE $\phi(p)$	61
FIGURE 1.6 A GAUCHE, REPRESENTATION D'UNE FONCTION DE PONDERATION DES PROBABILITES $\phi(p)$ SELON UNE SPECIFICATION DE PRELEC ET DE LA FONCTION DE PROBABILITE SANS TRANSFORMATION. A DROITE, REPRESENTATION DE LA FONCTION DERIVEE DE LA FONCTION $\phi(p)$	63
FIGURE 1.7 REPRESENTATION DE LA SUB-ADDITIVITE HAUTE ET BASSE (FIGURE TIREE DE TVERSKY ET WAKKER, 1995)	65
FIGURE 1.8 REPRESENTATION DES COURBES D'INDIFFERENCES SELON LE MODELE RDU DANS LE TRIANGLE MARSCHAK-MACHINA. ELLES SONT OBTENUES EN SUPPOSANT UNE FONCTION D'UTILITE PUISSANCE ET UNE FONCTION DE PONDERATION DES PROBABILITES EN FORME DE S INVERSE SELON LE MODELE DE PRELEC. LES POINTS A, B, A' ET B' REPRESENTENT LES LOTERIES DU PARADOXE D'ALLAIS : LA CONSEQUENCE COMMUNE	70
FIGURE 1.9 REPRESENTATION DES FONCTIONS D'UTILITE DE DEUX AGENTS A ET B , SACHANT QUE LA FONCTION D'UTILITE DE L'AGENT A EST CONSTRUITE EN PROCEDANT A UNE TRANSFORMATION CROISSANTE CONCAVE DE LA FONCTION D'UTILITE DE L'AGENT B . LA PRIME DE RISQUE DE L'AGENT A EST SUPERIEURE A CELLE DE L'AGENT B	77
FIGURE 1.10 REPRESENTATION D'UNE FONCTION D'UTILITE PUISSANCE (COURBE NOIR TRAIT PLEIN). LA FONCTION D'ESPERANCE D'UTILITE (TRAIT POINTILLE GRISE) EST CONSTRUITE SANS TRANSFORMATION DES PROBABILITES. C'EST-A-DIRE SELON UN MODELE EUT. LA FONCTION DE NON-ESPERANCE D'UTILITE (TRAITS POINTILLES NOIR) EST CONSTRUITE SELON UN MODELE RDU AVEC TRANSFORMATION DES PROBABILITES.	81
FIGURE 1.11 REPRESENTATION DES RESULTATS DE LA METHODE TRADEOFFS.....	84
FIGURE 1.12 REPRESENTATION DES LOTERIES CONSTITUEES PAR CAMERER (1989) AU SEIN DU TRIANGLE DE MARSCHAK-MACHINA.....	86
FIGURE 2.1 REPRESENTATION DE L'EXPERIENCE DE K&T TESTANT L'HYPOTHESE D'INTEGRATION ET D'ISOLATION	103
FIGURE 2.2 REPRESENTATION DE L'EXPERIENCE DE K&T TESTANT LA CONCAVITE ET LA CONVEXITE DE LA FONCTION D'UTILITE.	105
FIGURE 2.3 A GAUCHE : REPRESENTATION DE LA FONCTION D'UTILITE EN FORME DE S SELON K&T. A DROITE : REPRESENTATION DE LA FONCTION DE PONDERATION DES PROBABILITES DANS LES PERTES $\phi(-p)$ ET DANS LES GAINS $\phi + (p)$	108
FIGURE 2.4 REPRESENTATION DES DROITES D'INDIFFERENCES (EN ROUGE) DANS LE TRIANGLE DE MARSCHAK-MACHINA SELON UN MODELE CPT. LE REFERENTIEL EST EGAL A $R=1500$ ET LA RICHESSE INITIALE $w_0 = 1000$. LA LOTERIE $L \equiv x_1 = 0, p_1; x_2 = 500, p_2, x_3 = 1000, p_3$. AUTREMENT DIT, SOUS L'INFLUENCE DU REFERENTIEL L'INDIVIDU EVALUE DONC LES CONSEQUENCES SUIVANTES $Z \equiv z_1 = -500, p_1; z_2 = 0, p_2, z_3 = 500, p_3$. ON SUPPOSE UNE FONCTION D'UTILITE PUISSANCE DANS LES GAINS ET LES PERTES AVEC $\alpha = \beta = 0.88$. ON SUPPOSE EGALEMENT $\phi(p) = p$. LES DROITES EN NOIRS REPRESENTENT DES DROITES ISO-ESPERANCES. EN HAUT A DROITE ON REPRESENTE SCHEMATIQUEMENT LA LOCALISATION DES ISSUES DE LA LOTERIE PAR RAPPORT AU REFERENTIEL.....	115
FIGURE 2.5 CES FIGURES SONT ELABOREES EN CONSERVANT LES MEMES PARAMETRES ET LOTERIES QUE LA FIGURE 2.4. ON MODIFIE SIMPLEMENT LA VALEUR DU REFERENTIEL. A GAUCHE : ON SUPPOSE UN REFERENTIEL $R=0$. D'OU $Z \equiv z_1 = 1000, p_1; z_2 = 1500, p_2, z_3 = 2000, p_3$. A DROITE : ON SUPPOSE UN REFERENTIEL $R=2000$. D'OU $Z \equiv z_1 = -1000, p_1; z_2 = -500, p_2, z_3 = 0, p_3$	117
FIGURE 2.6 SIMULATION DE PRIMES DE RISQUE RELATIVES SELON UN MODELE CPT.	123
FIGURE 2.7 REPRESENTATION D'UNE SITUATION DE CHOIX DYNAMIQUE (TIREE DE MACHINA, 1989)	139
FIGURE 2.8. EN HAUT :RENDEMENTS REELS AU S&P 500 ENTRE 1889 ET 2000. EN BAS :RENDEMENTS REELS OBTENU POUR DES ACTIFS MOINS RISQUES SUR LA MEME PERIODE.....	151
FIGURE 2.9. REPRESENTATION DES DIFFERENTS PARCOURS D'ACTIFS PRESENTES AUX PARTICIPANTS DANS L'EXPERIENCE GROSSHANS ET AL, 2018).....	160

FIGURE 2.10. REPRESENTATION DE L'ADAPTATION DU POINT DE REFERENCE DANS LES PERTES ET DANS LES GAINS SELON LE MODELE DE SHI ET AL (2015).	164
FIGURE 2.11. LA FONCTION POINT-POINTILLE REPRESENTE LA FONCTION D'UTILITE PROSPECT THEORY DE LA RICHESSE AU TEMPS T+1 LORSQUE L'INDIVIDU A CONNU UN GAIN ANTERIEUREMENT. LA FONCTION POINTILLEE REPRESENTE LA FONCTION D'UTILITE PROSPECT THEORY DE LA RICHESSE AU TEMPS T+1 LORSQUE L'INDIVIDU A CONNU UNE PERTE ANTERIEUREMENT ET LA FONCTION EN TRAIN PLEIN REPRESENTE L'UTILITE DE LA RICHESSE AU TEMPS T+1 DANS LE CAS OU L'INDIVIDU N'A SUBI NI PERTE NI GAIN ANTERIEUREMENT.....	168
FIGURE 2.12 A GAUCHE REPRESENTATION DE LA FONCTION SIMILAIRE A CELLE SUPPOSE DANS LA RDU. LORSQU'ELLE A CETTE FORME ELLE IMPLIQUE LA FONCTION DE DROITE QUI REPRESENTE LA PONDERATION DES PRIX OBTENU LORS D'UNE SEQUENCE DE 100 ELEMENTS. LES PARAMETRES CHOISIS POUR CONSTRUIRE CES FONCTIONS SONT TIRES DE BAUCCELLS ET AL (2011).....	170
FIGURE 3.1 POINT D'INDIFFERENCES NORMALISE OBTENUS EN FONCTION DU DELAI D'OBTENTION DU RENFORÇATEUR. (TIRE DE TSUKAYAMA ET AL, 2010	178
FIGURE 3.2 EQUIVALENTS CERTAINS NORMALISES EN FONCTION DE LA COTE CONTRE L'OBTENTION DU RENFORÇATEUR $X1 = \$50$ (MAGNITUDE FAIBLE, TRAIT PLEIN) ET $X1 = \$5000$ (MAGNITUDE HAUTE, TRAIT POINTILLE) . (TIRE DE DAI ET AL, 2009) ..	184
FIGURE 3.3 FONCTION DE MARKOWITZ. (TIRE DE KAMMOUN, 2007)	188
FIGURE 3.4 REPRESENTATION DES PRIMES DE RISQUES SELON UN MODELE RDU ARME D'UNE FONCTION D'UTILITE PUISSANCE (A GAUCHE) OU D'UNE FONCTION D'UTILITE RACHLIN (A DROITE).....	201
FIGURE 3.5. MEDIANE DES PRIMES DE RISQUE RELATIVES OBTENUES DANS L'ETUDE DE FEHR-DUDA ET AL (2010).....	216
FIGURE 3.6 MEDIANE DES PRIMES DE RISQUES OBTENUES A PARTIR DES DONNEES DE ESTLE ET AL (2006) POUR CHAQUE NIVEAU D'ENJEU ET CHAQUE NIVEAU DE PROBABILITES. *** $p < 0.001$, ** $p < 0.01$	216
FIGURE 4.1. REPRESENTATION DU MODELE DE PARDUCCI (1965) SELON L'EQUATION (.4.1) POUR LES AGENTS A ET B EXPOSES A DIFFERENTES ETENDUES DE VOLUME D'UN SON.	239
FIGURE 4.2. REPRESENTATION DE LA FONCTION D'UTILITE NORMALISEE-TRANSFORMEE DE BAUCCELLS ET AL (2018) DANS LES PERTES ET LES GAINS.....	244
FIGURE 4.3 FONCTION D'UTILITE NORMALISEE DE BAUCCELLS ET AL (2018) DANS L'INTERVALLE $X \in [0, 500]$	250
FIGURE 4.4(EN HAUT) : REPRESENTATION DE LA FONCTION DE LA PRIME DE RISQUE ABSOLUE (GAUCHE) ET RELATIVE (DROITE) EN FONCTION DE G POUR LA LOTERIE $LX1 = 100, p1 = 0.5$; $X2 = 10, p2 = 1-p1$. (EN BAS) REPRESENTATION DE LA FONCTION DE PRELEC APPLIQUEE AUX PROBABILITES EQUIVALENTES (GAUCHE). REPRESENTATION DES FONCTIONS (EN NOIR) $U1$ ET $U2$ DEFINIES SUR L'INTERVALLE $[10, 100]$ ET $[10 500]$ RESPECTIVEMENT ET DE $P+$ EN ROUGE (DROITE).	251
FIGURE 4.5 MOYENNE DES PRIMES DES RISQUE ABSOLUES DU GROUPE 2 PHASE 2 (TRAITS PLEINS), GROUPE 1 PHASE 2 (TRAITS POINTILLES) ET GROUPE {3,4} PHASE 1 (POINTILLE) POUR LES DIFFERENTS NIVEAUX DE PROBABILITES DE GAGNER LA MEILLEURE ISSUE. LES BARRES REPRESENTENT LES ERREURS STANDARDS.....	267
FIGURE 4.6 MOYENNE DES PRIMES DES RISQUE ABSOLUES DU GROUPE 2 PHASE 1 (TRAITS POINTILLES) ET DU GROUPE 3 PHASE 2 (POINTILLES) POUR LES DIFFERENTS NIVEAUX DE PROBABILITES DE GAGNER LA MEILLEURE ISSUE. LES BARRES REPRESENTENT LES ERREURS STANDARDS.....	268
FIGURE 4.7 MOYENNE DES PRIMES DES RISQUE ABSOLUES DU GROUPE 4 PHASE 2 (TRAITS POINTILLES) ET DU GROUPE 1 PHASE 1 (POINTILLES) POUR LES DIFFERENTS NIVEAUX DE PROBABILITES DE GAGNER LA MEILLEURE ISSUE. LES BARRES REPRESENTENT LES ERREURS STANDARDS.....	269
FIGURE 4.8 REPRESENTATION DES FONCTIONS D'UTILITE DU GROUPE 1 PHASE 2 (COURBE ROUGE), DU GROUPE 2 PHASE 2 (COURBE NOIR) ET DU GROUPE {3,4} PHASE 1 (COURBE VERTE) SELON LES PARAMETRES ESTIMES PAR MLE DU MODELE BKL-PUISSANCE (HAUT) ET DU MODELE BKL-RACHLIN (BAS). A GAUCHE, LES FONCTIONS D'UTILITE TRANSFORMEE SONT REPRESENTEES EN FONCTION DE L'UTILITE NORMALISEE PAR RAPPORT A L'ETENDUE DE LA RICHESSE. A DROITE, LES FONCTIONS D'UTILITE NORMALISEES-TRANSFORMEES SONT REPRESENTEES EN FONCTION DE LA RICHESSE.	275
FIGURE 4.9.(HAUT) CONTOURS DE LA FONCTION OBJECTIF AU VOISINAGE DE SA VALEUR OPTIMALE TRACEE POUR LES PARAMETRES DU MODELE BKL-PUISSANCE: A (ABSCISSE) ET Γ (ORDONNEE) (HAUT GAUCHE), A_GROUPE1 (ABSCISSE) ET Γ _GROUPE1 (ORDONNEE) (HAUT DROITE), A_GROUPE{3,4} (ABSCISSE) ET Γ _GROUPE{3,4} (ORDONNEE) (HAUT CENTRE).	277
FIGURE 4.10 LES TRAITS PLEINS REPRESENTENT LE FIT DU MODELE BKL-PUISSANCE (HAUT) ET DU MODELE BKL-RACHLIN (BAS) AUX PRIMES DE RISQUES MOYENNES POUR LA LOTERIE LM DES GROUPE 1 (TRAITS POINTILLES ROUGES) ET 2 (TRAITS POINTILLES BLEU) EN PHASE 2. LES COURBES VERTES REPRESENTENT LES PRIMES DE RISQUES SIMULEES POUR LA LOTERIE LM A PARTIR DES PARAMETRES ESTIMES DU MODELE BKL-PUISSANCE (HAUT) ET BKL-RACHLIN (BAS) DU GROUPE 1.	280

LISTE DES TABLEAUX

TABLEAU 1.1 PROBABILITES ET PROBABILITES DECUMULEES DES ISSUES DES LOTERIES ENGAGEES DANS LE PARADOXE D'ALLAIS : EFFET DE LA CONSEQUENCE COMMUNE	66
TABLEAU 1.2 RANGS DES ISSUES ENGAGEES DANS LE PARADOXE D'ALLAIS	72
TABLEAU 2.1. LE FOURFOLD PATTERN	112
TABLEAU 2.2. EXEMPLE DE POINT DE REFERENCE POSSIBLE SELON UNE SITUATION DE CHOIX ENTRE L1 ET L2	136
TABLEAU 2.3. MEDIANE ET ECART TYPE DES DISTRIBUTION MARGINALE POSTERIEUR DES POINTS DE REFERENCES (ESTIMATION BAYESIENNE HIERARCHIQUE). TIRE DE L'ETUDE DE BAILLON ET AL (2015)	137
TABLEAU 3.1. CHOIX DE LOTERIES PROPOSE PAR MARKOWITZ (1952)	187
TABLEAU 3.2. VALEURS DES PARAMETRES AYANT ETE UTILISES POUR LA SIMULATION DES EQUIVALENTS CERTAINS ET RESULTATS DES ESTIMATIONS	211
TABLEAU 4.1 CONDITIONS EXPERIMENTALES	257
TABLEAU 4.2. CONDITIONS EXPERIMENTALES ET DETERMINATION DES TRIPLETS (=CADRE : \mathcal{U}^K)	260
TABLEAU 4.3 PRIME DE RISQUE ABSOLUE EN FONCTION DES VARIABLES GROUPE, PROBA ET DE LEURS INTERACTIONS	265
TABLEAU 4.4 ESTIMATION DES PARAMETRES DU MODELE BKL ARME D'UNE FONCTION D'UTILITE DE RACHLIN OU D'UNE FONCTION D'UTILITE PUISSANCE	273

INTRODUCTION GÉNÉRALE

La prise de décision est sans doute l'une des activités les plus régulières et les plus fondamentales de l'existence d'un organisme doté à minima d'un système nerveux. Elle conditionne nos vies, notre quotidien et notre survie. La principale difficulté pour les organismes réside dans le fait que la plupart des décisions doivent être réalisées alors que les conséquences de leurs choix sont imprévisibles. Dans certains cas, nous connaissons exactement l'ensemble des conséquences possibles qu'engendre un acte ainsi que les probabilités d'occurrences respectives mais nous ne pouvons prédire avec certitude laquelle de ces conséquences subviendra suite à l'action. Selon la classification opérée par Knight (1921), il s'agit ici d'une situation dite « risquée ». Dans d'autres cas, nous n'avons pas connaissance des probabilités d'occurrence des conséquences de nos actes, il s'agit alors d'une situation d'incertitude, et parfois même, nous ignorons toutes les conséquences possibles, on parle alors d'incertitude radicale.

Dans cette optique, certains domaines, dont les plus prolifiques sont sans doute l'économie et la psychologie, ont cherché à capturer et à prédire les comportements décisionnels dans le risque à travers des formalisations mathématiques. Ces modèles sont alors évalués sur au moins deux critères. Un critère normatif et un critère descriptif. Le critère normatif dicte, si l'on accepte un certain cadre axiomatique, la manière dont devraient se comporter les individus et il permet en de raisonner en termes de rationalité et d'optimalité. Le critère descriptif relève plutôt de la capacité d'un modèle à expliquer les comportements réellement observés, c'est-à-dire son succès à faire correspondre ses prédictions avec la réalité des phénomènes observables.

Bien qu'à ce titre les différentes formalisations ne soient pas toutes aussi performantes les unes que les autres, elles reposent néanmoins toutes sur le même principe fondamental : la maximisation de la « valeur ». En effet, l'ensemble de ces théories supposent que les individus assignent une « valeur subjective » à chaque option disponible dans un environnement de choix avant de choisir celle dont la valeur est la plus importante. Ces options, que l'on peut conceptualiser comme des variables aléatoires, peuvent engendrer différentes conséquences dont on ne sait pas à l'avance laquelle surviendra suite à la décision.

Cette valeur dépend naturellement 1) des caractéristiques objectives de chacune des options : la nature des conséquences (tangibles vs non tangibles, nourriture, argent, eau...), les quantités qu'il est possible d'obtenir ou de perdre, les probabilités d'occurrence objective ou subjective de chaque conséquence etc... 2) des caractéristiques inhérentes à l'individu lui-même : sa personnalité, son vécu, ses réussites, ses échecs, son état de déprivation... Dans l'ensemble de ces modèles, il est alors supposé que l'individu choisisse l'option à laquelle est associée la plus grande valeur. Autrement dit, il est suggéré que la décision de cet agent est conduite par la maximisation de quelque chose ; il reste toutefois à savoir quoi et comment. Quoi qu'il en soit, un bon modèle devra au minimum être en mesure de rendre compte le mieux possible de l'ordre des préférences exprimé par des agents entre différentes options risquées.

Quant à savoir ce que l'individu semble maximiser, les philosophes du courant utilitariste comme Bentham ou encore John Stuart Mill répondront que les agents cherchent à maximiser leur plaisir ou à minimiser leur déplaisir. Bernoulli dans sa résolution du paradoxe de Saint Pétersbourg ne propose pas autre chose. Selon lui, les individus ne s'appliquent pas à maximiser leur profit, auquel cas le critère de décision s'appuierait uniquement sur celui de l'espérance mathématique de chaque alternative et à choisir l'option ayant la plus grande espérance. Selon Bernoulli, les individus maximisent une fonction d'espérance d'utilité (qu'on pourrait traduire de façon approximative par le terme de « satisfaction ») qu'ils pensent pouvoir dériver des

issues des options proposées dans l'environnement. Autrement dit, les individus maximisent une espérance de satisfaction tirée des issues financières impliqués dans une situation de choix. C'est sur la base des travaux menés par Bernoulli que Von Neumann & Morgenstern (1944) ont alors proposé un fondement axiomatique à la théorie de la décision dans le risque et plus spécifiquement au modèle d'espérance d'utilité. Les auteurs ont construit une théorie rationnelle de la décision : la théorie de l'utilité espérée (EUT) qui, par ce biais, offre des prescriptions normatives substantielles puisqu'elle permet d'établir ce que *devrait* faire un agent étant donné sa fonction d'utilité. Selon ce modèle, il est donc supposé que les préférences des individus soient fondées sur l'objectif de maximiser leur utilité.

Bien que ce modèle soit encore très largement utilisé pour analyser les décisions en environnement risqué, la capacité de celui-ci à correspondre aux comportements réellement observés a sérieusement été mise à mal. En effet, cette théorie est en mesure de répondre à la question de ce que *devrait faire* un individu dans un cadre axiomatique précis, mais ne semble pas pouvoir expliquer ce que *fait* réellement un agent. Il faudra attendre une vingtaine d'année après la publication phare de VnM (*Theory of games and economic behavior*, 1944) pour que le potentiel descriptif de cette théorie (c'est-à-dire sa capacité à décrire les comportements effectivement réalisés) soit finalement ébranlé par les premières investigations expérimentales menées par Maurice Allais (1953), Ellsberg (1961) puis par Kahneman et Tversky (1979, 1992). C'est alors que l'on vit apparaître un certain nombre de modèles alternatifs qui visèrent à pallier les insuffisances descriptives de l'EUT. On verra émerger parmi les théories les plus importantes : la Rank Dependent Utility (Quiggin, 1982), la Regret Theory (Loomes et Sugden, 1982) ou encore la Prospect Theory (PT) (Kahneman et Tversky, 1979). La PT, par exemple, ne présume pas de la rationalité des individus mais va au contraire introduire des éléments non-

normatifs, des biais heuristiques induits notamment par les capacités de traitement limitées des individus.

L'accumulation de ces preuves a finalement ouvert la voie à un nouveau champs d'investigation : l'économie expérimentale qui explore, entre-autres, les causes et implications des violations de la rationalité (Ho et al., 2006). La conjugaison de l'économie et de la psychologie bat désormais son plein, ce qui finalement correspond à l'esprit du courant utilitariste. Toutefois, l'avènement du programme néoclassique et la primeur qu'il accordait au raisonnement hypothético-déductif plutôt qu'inductif avait séparé les deux domaines. Aujourd'hui, suite aux progrès de la psychologie expérimentale, la psychologie et l'économie se retrouvent sur un terrain commun dont l'un des objectifs est de produire des modèles descriptifs plus performants.

Dès lors, la recherche expérimentale en économie s'est principalement concentrée sur deux pans. Le premier consistait à tester la validité axiomatique des différents modèles de décision. Très souvent, il a été constaté que certains axiomes comme celui d'indépendance, ou de transitivité, n'étaient pas toujours exhibés par les agents testés. Dès lors, d'autres modèles furent élaborés au prix du relâchement, ou de l'affaiblissement de certains axiomes. C'est dans ce cadre que fut modifié la théorie standard d'espérance d'utilité avec l'intégration de nouveaux éléments comme la fonction de pondération des probabilités, les poids décisionnels ou encore le point de référence. Le second pan, s'est davantage focalisé sur la capacité descriptive des modèles à expliquer les comportements des agents, en intégrant des critères statistiques comme le BIC (Bayesian Information Criterion) ou l'AIC (Akaike Information Criterion) pour juger et comparer les performances des modèles. Dans ce type d'étude, l'adéquation entre l'axiomatique et les comportements n'est pas évaluée ; on cherchera plutôt à récupérer des données relatives

à la propension à prendre des risques de la part des individus, puis à comparer les modèles qui parviennent à s'approcher au mieux de ces données observées.

Les deux approches sont toutefois complémentaires et nécessaires à l'élaboration de modèles cohérents normativement, et suffisamment performants pour prédire et capturer les comportements des agents en environnement probabiliste.

Dans cette espace interdisciplinaire, la discussion entre économistes et psychologues n'est toutefois pas toujours évidente. Une grande partie des psychologues fut dans un premier temps relativement réfractaire aux exigences axiomatiques et aux raisonnements hypothético-déductifs des économistes. Ce qui se conçoit facilement dans la mesure où la culture des psychologues est plutôt tournée vers les raisonnements inductifs. Dès lors, une première approche des psychologues fut de rechercher l'ensemble des biais cognitifs humains qui ne sont pas pris en compte par la théorie standard en économie et qui prouverait, selon eux, l'invalidité de cette théorie. Aujourd'hui, l'interdisciplinarité entre ces champs semble avoir trouvé des voies davantage exploitables :

- 1- On observe dans ce cadre, l'intégration d'éléments cognitifs dans les modèles économiques, ce qui a permis de dérigidifier les modèles économiques. Cette voie est sans doute celle qui fut, et reste encore, la plus prolifique. En particulier, l'élément cognitif fondamental qui fut incorporé à la théorie économique est ce que l'on appelle le « point de référence » ou le « référentiel ». Ce construit hypothétique fut étudié en premier lieu par un psychologue de la « Gestalt » : Helson (1947) et fut ensuite intégré à la théorie économique par Markowitz (1952) puis par Kahneman et Tversky (1979). Ce concept permet simplement de prendre en compte le fait que les individus perçoivent de manière relative les informations (ou stimuli) de l'environnement. Le référentiel constitue donc un « point de comparaison » qui sert à évaluer/juger la valeur d'un stimulus.

L'intégration du point de référence, a profondément modifié la théorisation de la prise de décision en univers risqué et a permis d'accroître la capacité descriptive des modèles. Toutefois, nous verrons que ce référentiel est également le talon d'Achille de ces nouvelles théories puisque la spécification de ce référentiel reste encore un véritable challenge aujourd'hui. Sans spécification, ces modèles décrivent bien les comportements, mais ne les prédisent pas. Les performances de ces modèles sont donc principalement *ad hoc*.

- 2- Une autre voie très en vogue actuellement consiste à rechercher une « réalité cognitive » aux modèles « *as if* » de l'économie. En effet, en économie personne ne prétend que les individus calculent réellement des espérances d'utilité plus ou moins complexes consciemment ou non. Dans cette optique, on se préoccupe davantage de la performance du modèle à capturer les comportements plutôt que de sa réalité cognitive en termes de traitement de l'information. D'une certaine manière, la conception économique des modèles correspond à ce que disait le statisticien George Box : « tous les modèles sont faux, mais certains sont utiles » (1976). L'idée n'est donc pas de construire des modèles qui reposent sur des éléments réalistes de la psychologie ou de la cognition des organismes, mais d'établir des modèles prédictifs, normatifs et prescriptifs. La psychologie ainsi que les neurosciences se sont cependant emparée de cette question, avec pour objectif de procurer une assise cognitive (et neuronale) aux modèles abstraits de l'économie. La recherche conduite dans ce cadre consiste dans un premier temps à fournir une interprétation cognitive des éléments mathématiques impliqués dans les modèles, puis d'en tester les implications sur un plan de « traitement de l'information ». Dès lors, les résultats obtenus dans ces expériences ne visent pas à valider une théorie plutôt qu'une autre, mais plutôt à appuyer certaines interprétations des modèles plutôt que d'autres (Pachur et al., 2018). A titre d'illustration, certains chercheurs cherchent en ce moment à estimer des fonctions d'utilité

à partir des ressources attentionnelles mesurées par « eye tracking » allouées par un individu entre différentes options risquées (Pachur et, 2018). D'autres encore cherchent à estimer les fonctions d'utilité à partir de temps de réaction (Echeinique et Sato, 2017).

3- Enfin, une troisième voie que nous avons suivi dans notre partie 3 vise à confronter les prédictions de modèles provenant de la psychologie et de l'économie. En particulier, nous avons étudié le modèle de Probability discounting développé par Myerson, Green et Morris (2011). En étudiant en profondeur ce modèle à travers une perspective économiste, nous avons mis en évidence l'inadéquation des prédictions du modèle avec les patterns décisionnels communément observés en environnement probabiliste. Toutefois, nous avons également montré (comme d'autres économistes d'ailleurs) que ce pattern décisionnel révélé n'est pas accommodable par les théories économiques non plus. Cette inadéquation devrait pousser la recherche à trouver de nouveaux modèles afin de capturer ce pattern qui semble robuste, puisque répliqué maintes fois. Une fois de plus, la spécification d'un certain référentiel est une piste envisageable, mais pour le moment aucun chercheur n'est parvenu à proposer un modèle satisfaisant.

RESUME DES PARTIES DE LA THESE

PARTIE 1.

Cette partie est consacrée à la présentation de la théorie de l'espérance d'utilité. Dans un premier temps, nous présentons brièvement l'émergence de cette théorie puis nous en fournissons une description formelle. Il est important de présenter clairement cette théorie dans la mesure où les modèles alternatifs que nous aborderons sont des généralisations de celle-ci. Dans un second temps, nous présenterons le paradoxe d'Allais qui est véritablement le point de départ qui a permis l'émergence de nouveaux modèles. Dans ce cadre, nous nous intéresserons

au modèle de la Dual Theory de Yaari (1987) et surtout au modèle de la Rank-Dependent Utility de Quiggin (1982). Nous verrons que la théorie de l'espérance d'utilité explique les différentes attitudes face au risque en supposant l'existence d'une fonction d'utilité définie sur la richesse certaine, dont la forme et les propriétés déterminent la prise de décision dans le risque. La Dual Theory ainsi que la Rank-Dependent Utility Theory, intègrent pour la première un nouveau déterminant de l'attitude face au risque en plus de la fonction d'utilité : la fonction de pondération des probabilités décumulées.

Nous préciserons alors dans une troisième partie, l'influence de ces fonctions sur l'attitude face au risque. Puis nous aborderons les principaux résultats expérimentaux qui ont éprouvé la validité axiomatique de ces théories ainsi que leurs performances descriptives.

PARTIE 2.

Cette partie est entièrement consacré à la théorie concurrente la plus connue de l'espérance d'utilité ou de la Rank-Dependent Utility : la Prospect Theory de Kahneman et Tversky (1979). Cette partie nous fournira l'occasion de présenter un large spectre de la littérature portant sur la notion de référentiel. Nous montrerons dans un premier temps en quoi cette théorie permet d'expliquer certains patterns comportementaux que les théories alternatives ne sont pas capables d'accommoder. Nous entrerons alors dans le détail du formalisme de la seconde version de la Prospect Theory : la Cumulative Prospect Theory (CPT).

Nous focaliserons notre troisième et quatrième point sur le point de référence dont nous ferons une revue de littérature. Le référentiel est l'élément déterminant qui sépare cette théorie des autres alternatives théoriques. Pour la première fois, il est suggéré que les individus n'évaluent pas l'utilité des différentes positions de richesses finales atteignables dans une situation risquée (il s'agit de l'hypothèse d'intégration selon laquelle les individus intègrent les richesses impliquées dans les loteries à leur richesse initiale). Selon la CPT, les individus dérivent plutôt

l'utilité de variations de richesse entre une position finale et un point de référence (hypothèse d'isolation). Ce point de référence permet alors pour un individu de séparer ce qu'il considère être de l'ordre des pertes de ce qu'il considère être de l'ordre des gains. En observant que les individus n'adoptent pas la même attitude face au risque dans ces deux régions (pertes versus gains), les auteurs ont alors inféré l'existence d'une fonction d'utilité en deux parties : une fonction concave dans les gains et une fonction convexe dans les pertes.

Nous mettrons alors en évidence que la principale difficulté dans cette théorie réside dans la spécification exacte de ce point de référence. Nous verrons entre-autres que sans spécification, cette théorie jouit d'un trop grand degré de liberté. Dès lors, les explications des patterns comportementaux apportés par ce modèle sont très souvent *ad hoc*.

PARTIE 3.

Dans cette partie nous nous focalisons sur le modèle de Probability Discounting proposé par Myerson et al. (2011). Dans ce contexte, les chercheurs du courant comportementaliste, étudient le phénomène de dévaluation subjective d'un renforçateur probabiliste. Il s'agit du phénomène selon lequel la valeur subjective d'une récompense diminue conjointement avec sa probabilité d'obtention. Traduit dans des termes économiques, il s'agit d'étudier comment évolue l'attitude face au risque lorsque la probabilité d'occurrence d'un renforçateur diminue. Dans ce cadre, les comportementalistes se sont focalisés sur l'effet de la magnitude du renforçateur probabiliste. Ils notent que plus l'amplitude augmente, plus la dévaluation subjective relative augmente.

Ce phénomène, également étudié par des économistes connus comme Markowitz (1952) ou encore Fehr-Duda et al (2010), est appelé : l'effet de l'enjeu, ou « stake size effect » et correspond au fait que les individus deviennent de plus en plus averses au risque à mesure que l'enjeu d'une loterie augmente.

En cherchant à estimer les paramètres libres de leur modèle, les comportementalistes observent alors qu'un des paramètres n'est pas stable à travers les différents niveaux d'enjeu. Plus exactement, à mesure que l'amplitude du renforçateur probabiliste augmente, un paramètre du modèle semble également augmenter. En général, un paramètre instable est le signe d'un problème d'indétermination ou de spécification du modèle. Nous verrons ici qu'il s'agit en fait des deux. L'indétermination est liée au fait que les loteries engagées dans la méthodologie des comportementalistes produisent un modèle dont les paramètres libres ne peuvent être identifiés séparément. Le problème de spécification est lié à la nature de la fonction d'utilité supposée par ce courant. En particulier, ils supposent (implicitement) une fonction d'utilité puissance dont la caractéristique est d'être isoélastique. Nous verrons alors que cette propriété est directement en contradiction avec les données expérimentales.

Pour finir, nous montrerons que le pattern comportemental observé lorsque l'enjeu d'une loterie augmente n'est pas aussi simpliste que prévu. Plus spécifiquement, nous verrons que l'effet de l'enjeu sur le degré d'aversion au risque relatif semble prépondérant lorsque la probabilité de gagner est faible, mais disparaît à mesure que la probabilité de gagner augmente. Ce résultat est un véritable challenge y compris pour la théorie économique, puisqu'aucun modèle jusqu'à présent n'est en mesure de capturer ce résultat. Dans ce cadre nous explorerons de façon purement théorique une piste envisageable tirée de la littérature comportementaliste.

PARTIE 4.

Dans cette dernière partie, nous nous intéresserons à un nouveau modèle proposé par Baucells, Kontek et Lewandowski (2018) qui intègre un effet connu des psychologues : l'effet de l'étendu (ou range effect). Cet effet correspond à la tendance des individus à évaluer (ou juger) des stimuli en fonction de l'étendu des valeurs de ce stimuli auxquels ces individus ont été exposés. En fait, l'étendue des valeurs d'un stimulus permet de juger subjectivement ce qui est de l'ordre

du « petit » et du « grand ». A titre d'illustration par exemple, le prix d'un billet d'avion peut être perçu comme élevé ou faible en fonction de la gamme de prix d'autres billets présentés en concurrence. La localisation du prix d'un billet par rapport à cette étendue détermine alors son jugement en termes de « cher » « moyennement cher » ou « pas cher ». Ce dernier modèle peut être vu comme une extension de la Prospect Theory qui intègre cette fois non pas un point de référence, mais trois au sein du modèle.

En intégrant cette notion d'étendue dans un modèle de décision dans le risque, les auteurs prodiguent une nouvelle formalisation de l'effet de l'enjeu sur l'attitude face au risque. En effet, nous avons évoqué lors de la partie III que les individus sont plutôt chercheurs de risque lorsque l'enjeu est faible et plutôt averses au risque lorsque l'enjeu est important. Markowitz (1952) avait dans ce cadre proposé une fonction d'utilité convexe pour des petits gains et concave pour de plus grands gains permettant d'accommoder ce pattern comportemental. Toutefois, Markowitz ne spécifie pas ce qui est de l'ordre du « petit » ou du « grand », rendant impossible une déclaration fiable du point d'inflexion de la fonction d'utilité. Avec l'intégration de l'étendu dans l'évaluation subjective des issues monétaires de loterie, Baucells et al. (2018) fournissent une spécification intuitive à la notion de « grandeur ». En particulier une valeur monétaire sera jugée grande ou petite en fonction de sa localisation au sein de l'étendu des valeurs monétaires engagées dans une situation de choix.

Nous verrons alors que le modèle de Baucells et collaborateurs doit toutefois être précisé. En effet, selon les auteurs, l'étendue des valeurs monétaires correspond au montant maximum et minimum qu'il est possible d'obtenir au sein d'un contexte décisionnel. Toutefois, les auteurs ne précisent pas ce qu'ils entendent par « contexte décisionnel ». S'agit-il des loteries en présence au moment du choix, ou bien ce contexte intègre-t-il également des loteries antérieurement présentées mais qui n'ont plus cours au moment du choix ? A travers cette

question, nous nous demanderons d'où peuvent être tirées les valeurs qui sont considérées comme minimum et maximum par les individus.

En dérivant certaines prédictions du modèle de Baucells et al., nous implémentons une expérience qui nous a permis d'apporter certaines réponses. En particulier, nous montrerons que la préexposition à des loteries à certains niveaux d'enjeux influence la prise de décision subséquente face à d'autres loteries à niveaux d'enjeu différents. Ce résultat semble potentiellement indiquer que les valeurs maximum et minimum pour juger de l'étendu des valeurs monétaires peuvent être tirées de loteries qui ne sont pas forcément en présence au moment du choix, mais aussi de loteries antérieures.

1. LES MODELES DE DECISION DANS LE RISQUE

INTRODUCTION

Dans cette section nous présentons la théorie normative de la prise de décision dans le risque : la théorie de l'espérance d'utilité (noté EUT). Après avoir défini les fondamentaux à savoir le cadre historique mais aussi conceptuel de notre objet d'étude, nous présenterons l'axiomatisation de l'EUT et le modèle de décision qui en découle. Nous nous focaliserons ensuite sur les remises en question de cette théorie qui mena au relâchement de l'axiome d'indépendance. Sur la bases de ces critiques, nous présenterons la Dual Theory ainsi que la Rank-Dependent Utility Theory (RDUT) qui constitue encore à ce jour l'une des principales alternatives à l'EUT. Pour finir, nous présenterons différentes expériences qui furent principalement menées dans les années 1990 afin de départager les différents modèles vis-à-vis de leurs capacités descriptives et leurs réalités axiomatiques.

I. BASIQUES ET HISTORIQUE

Cette section est consacrée à la présentation du cheminement conceptuel qui a permis l'émergence de la théorie de l'utilité espérée telle qu'elle fut formulée rigoureusement par Von Neumann et Morgenstern (1944). Dans un premier temps nous exposerons un bref rappel concernant ce qu'est une situation de choix risquée, ainsi que la relation entre modèle et préférence. Dans un second temps nous présenterons les deux modèles fondamentaux antérieurs à celui de l'EUT avant de présenter ce dernier ainsi que son axiomatique.

1.1. Environnement risqué, valeur subjective et préférence

1.1.1. Choix en environnement risqué

Lorsqu'un agent doit prendre une décision, il est implicitement supposé que l'individu à face à lui différentes alternatives qui mènent, chacune potentiellement, à différentes conséquences possibles mutuellement exclusives. Si les conséquences de chaque alternative sont connues, c'est-à-dire que si l'individu sait exactement avant de prendre sa décision quelle conséquence surviendra à lui suite de n'importe quel de ses choix, alors il s'agit d'un environnement où l'information est parfaite et donc où toute incertitude est levée. Le risque, quant à lui, n'est toutefois pas une situation d'ignorance totale sur les conséquences des choix (Cayatte, 2009). L'individu peut connaître avec précision les différentes conséquences possibles de chaque choix disponible, seulement, il ne peut prévoir laquelle de ces conséquences surviendra à la suite de son choix. Le résultat d'une décision est donc ici l'effet combiné d'une alternative choisie et de la réalisation de l'aléa.

1.1.2. Loteries

En économie, les décisions en situation risquée impliquent des décisions d'investissements financiers, de portefeuilles ou encore d'épargne qui enrichissent ou appauvrissent le décideur par rapport à sa richesse initiale. Ces situations rappellent en fait la structure d'une loterie. C'est pourquoi, les diverses situations risquées sont simplifiées en les représentant par des loteries où les probabilités d'occurrence de chaque état de la nature sont connues par l'individu et où la plupart du temps les conséquences sont des montants monétaires¹. Une loterie noté $L^j \in \mathcal{L}$ avec \mathcal{L} représentant un ensemble fini de loteries où $j = 1, \dots, m$, peut être décrite comme une variable aléatoire $\tilde{x}^j: \Omega \rightarrow \mathbb{R}$ où toutes les conséquences monétaires mutuellement exclusives prise par \tilde{x}^j sont notées : $\tilde{x}^j(\Omega) = \{x_1^j, \dots, x_i^j, \dots, x_n^j\}$. La distribution des probabilités sur \tilde{x}^j assigne à chaque conséquence x_i^j la probabilité p_i^j avec $0 \leq p_i^j \leq 1$ et $\sum_i^n p_i^j = 1$. Par exemple, la loterie $L^j \equiv (x_1^j, p_1^j; x_2^j, 1 - p_1^j)$, produit la conséquence x_1^j avec la probabilité p_1^j ou bien la conséquence x_2^j avec la probabilité complémentaire. L'ensemble des loteries présentes dans un contexte de décision, noté \mathbb{C} , peut être représenté ainsi : $\mathbb{C} \equiv \mathcal{L}(\tilde{x}^1, \dots, \tilde{x}^j, \dots, \tilde{x}^m)$. Pour l'instant nous définirons le contexte de décision comme l'ensemble des loteries en présence au moment du choix. C'est-à-dire l'ensemble des alternatives sur lequel l'individu peut opérer un choix. Enfin, lorsqu'une variable aléatoire ne contient qu'une seule conséquence possible on parle d'une loterie dégénérée.

On peut également considérer le cas de loterie composée : il s'agit de loterie dont certaines conséquences sont elles-mêmes des loteries. C'est-à-dire qu'une (ou plusieurs) conséquence possible d'une loterie est de gagner le droit de participer à une autre loterie. Prenons le cas

¹ On pourrait penser que supposer des agents qu'ils aient une connaissance parfaite de la distribution des probabilités objectives sur les issues soit une approche très restrictive. Une approche alternative et plus générale proposé par Savage (1972) consiste à établir que les probabilités, comme les préférences, sont des éléments subjectifs et propre à chaque individu. Il n'est donc pas obligatoire de supposer que les individus aient une information parfaite sur les probabilités. Toutefois, dans la mesure où ce relâchement ne produit pas de différence majeure, nous en resterons à la présentation des théories de la décision dans le cas où les probabilités sont objectives (c'est-à-dire en univers risqué et non incertain).

d'une loterie composé L_c qui permet d'accéder à la loterie L_a avec la probabilité α ou à la loterie L_b avec la probabilité $1 - \alpha$. On peut facilement montrer que la loterie composée L_c se

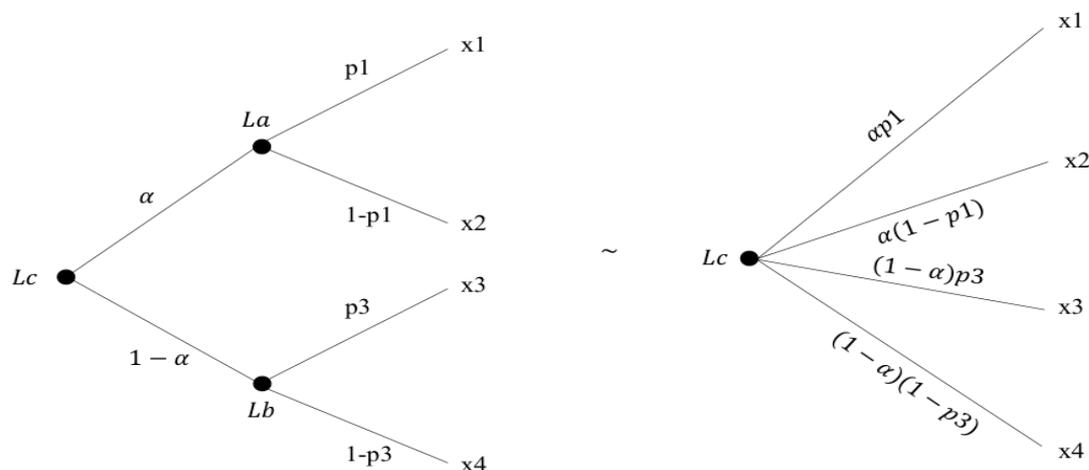


Figure 1.1 Réduction d'une loterie composée. Soit une loterie L_c dont les prix sont l'accès à des loteries L_a et L_b avec la probabilité α et $1 - \alpha$ respectivement. Cette loterie en deux étapes peut se réduire à une étape par simple application du théorème de Bayes (1763). L'axiome de réduction des loteries composées suppose qu'un agent sera indifférent entre ces deux loteries.

réduit à une loterie simple car elle possède le même système de probabilité que la combinaison convexe suivante : $L_c = \alpha L_a + (1 - \alpha)L_b$.

1.1.3. Valeurs subjectives et préférences : notation

Les préférences des individus sont définies sur les conséquences des décisions. En effet, les choix sont en fait des moyens de passer d'un état initial (ou d'une richesse initiale) à un état final (ou à une richesse finale). Par corolaire, les préférences sur les décisions sont donc des préférences sur les richesses finales qu'elles permettent d'atteindre (Cayatte, 2009). Lorsque les individus choisissent parmi plusieurs alternatives, ils expriment et révèlent ainsi leurs préférences sur ces dernières. On utilise alors la notation conventionnelle suivante : Si l'individu a une préférence faible pour une loterie A par rapport à une loterie B , on note $A \succcurlyeq B$, ou une préférence stricte : $A \succ B$ ou s'il est indifférent entre les 2 loteries dans ce cas on note : $A \sim B$.

Supposons un ensemble \mathcal{L} de loteries. Les préférences de l'agent sont représentées par la relation \succsim (préféré ou indifférent à) supposée complète, réflexive et transitive. L'objectif d'un modèle de décision est alors de proposer une fonction de valeur $V: \mathcal{L} \rightarrow \mathbb{R}$ qui représente l'ordre des préférences de l'individu sur les loteries. Ainsi, $\forall L, L' \in \mathcal{L}, L \succsim L' \Leftrightarrow V(L) \geq V(L')$ (Gayant, 1995). Autrement dit, le classement des valeurs numériques associées à chaque loterie donnée par le modèle, doit correspondre au classement des préférences sur les loteries exhibées par l'agent. Il s'agit donc ici d'une mesure ordinale des préférences.

1.2. Le paradoxe de Bernoulli

1.2.1. L'espérance mathématique

D'un point de vue économique, il pourrait sembler naturel de supposer que les agents choisissent systématiquement l'option qui prodigue la plus grande espérance de richesse, c'est-à-dire choisir l'option dont la somme des produits des probabilités à leurs issues donne le plus grand résultat (on peut parler également de richesse moyenne). A richesse initiale donnée w_0 et face à une loterie $L(x_i, p_i)_{i=1, \dots, n}$ l'espérance mathématique de richesse finale $E(\tilde{W})$, où $\tilde{W} = w_0 + \tilde{x}$, s'écrit alors :

$$E(\tilde{W}) = \sum_{i=1}^n p_i (w_0 + x_i) \quad (1.1)$$

Lorsqu'un individu est confronté à plusieurs loteries $L^j \in \mathcal{L}$, avec $j = 1, \dots, n$, un choix de loterie basé sur un critère de maximisation de l'espérance est noté : $\tilde{W}^* \in \operatorname{argmax}\{E(\tilde{W}^j)\}$.

Toutefois, cette règle de maximisation, appelée critère de Pascal (ou critère de Laplace quand la distribution des probabilités suit une loi uniforme), ne correspond pas aux comportements observés dans la réalité, et semble même contraire à l'intuition dans certain cas. En effet, ce modèle implique que les individus feraient tous les mêmes choix face à n'importe quelle

situation de choix risqué. Or il est évident que les attitudes face au risque sont hétérogènes et dépendent d'éléments subjectifs propre à chaque agent.

Une expérience de pensée, appelé le paradoxe de Saint Pétersbourg exposé pour la première fois par Nicolas Bernoulli (1713) puis résolu par son cousin Daniel Bernoulli (1738) a permis d'en révéler les limites.

1.2.2. Le paradoxe

On envisage un jeu à 2 joueurs jouant à pile ou face. Le joueur A joue et relance systématiquement tant que face apparaît. Le jeu s'arrête quand la pièce tombe sur pile. Le nombre n de fois que le joueur devra lancer est donc aléatoire. A chaque lancer, le joueur A a une chance sur deux d'obtenir face. On peut établir facilement par dénombrement que la probabilité d'obtenir 2 fois face de suite est égale à $(1/2)^2$, 3 fois de suite : $(1/2)^3$... on obtient en définitive la loi de distribution sur le nombre de lancer égale à $\left(\frac{1}{2}\right)^n$. Le gain du joueur A est établi à 2^n que le joueur B doit lui procurer. La question est alors la suivante : quelle somme, b , le joueur A devrait-il être prêt à miser pour participer à ce jeu si celui-ci utilise le critère d'espérance mathématique de gain comme critère de décision ? Etudions les alternatives : S'il ne joue pas, la richesse finale de l'individu sera égale à sa richesse initiale w_0 . S'il joue, sa richesse finale de l'individu sera $W = w_0 - b + 2^n$. Pour résoudre ce problème, il faut simplement calculer l'espérance du jeu et la somme trouvée sera la mise que l'individu A sera prêt à investir. Cette espérance est alors égale à $E(2^n) = \frac{1}{2} \times 2 + \left(\frac{1}{2}\right)^2 \times 2^2 + \dots + \left(\frac{1}{2}\right)^n \times 2^n = 1 + 1 + \dots + 1 + \dots = +\infty$. Chaque terme de la somme est égal à 1, produisant ainsi une espérance de gain infinie. Dès lors, un individu ayant recours à l'espérance comme critère de décision devrait être disposé à investir toute sa fortune pour jouer à ce jeu. Or il est clair que personne n'est prêt à prendre un tel risque. En effet, ne serait-ce qu'obtenir 10 fois de

suite le côté face est un évènement pour le moins peu probable car il équivaut à une probabilité de $1/1024 < 0.01$.

1.2.3. Résolution du paradoxe

Pour résoudre ce paradoxe, Daniel Bernoulli considéra d'emblée que le problème inhérent au critère d'espérance mathématique est qu'aucune caractéristique des personnes n'est pris en compte, seuls comptent les termes du pari. Or, pour lui : la détermination de la valeur d'un objet ne doit pas être basée sur ses avantages², mais seulement sur l'utilité qu'il procure. Les avantages de l'objet dépendent seulement de lui-même et sont les mêmes pour tout le monde ; l'utilité, par contre, dépend des caractéristiques propres de la personne qui fait l'évaluation [...]. Daniel Bernoulli suggéra donc que les individus ne maximisent pas l'espérance mathématique de gain mais plutôt leur espérance de satisfaction ou d'utilité dérivé de la richesse prospective. Autrement dit, les agents évaluent subjectivement une option risquée sur la base des moyennes des utilités qu'ils espèrent tirer des résultats monétaires de cette option plutôt que sur la moyenne objective de la richesse que cette dernière prodigue. Dès lors, ce n'est pas le gain en lui-même qui intéresse l'individu mais l'utilité que celle-ci est susceptible de procurer. Formellement, un agent doté d'une richesse initiale w_0 faisant face à une loterie $L(x_i, p_i)_{i=1, \dots, n}$ évalue subjectivement la richesse finale aléatoire par la fonction d'évaluation suivante :

$$EUT(\tilde{W}) \equiv E(u(\tilde{W})) = \sum_{i=1}^n p_i u(w_0 + x_i) \quad (1.2)$$

Ici, le critère de décision consiste à maximiser l'espérance d'utilité de la richesse finale : $\tilde{W}^* \in \operatorname{argmax}\{EUT(\tilde{W}^j)\}$. Bernoulli impose une restriction fondamentale vis-à-vis de la fonction

² C'est-à-dire, sur ce que rapporte objectivement cet objet en terme de richesse

d'utilité u . Il suppose que l'utilité résultant de tout petit accroissement de la richesse sera inversement proportionnelle à la quantité de biens antérieurement possédés. Autrement dit, l'accroissement de l'utilité décélère à mesure qu'un individu est de plus en plus riche, c'est ce que l'on appelle : « la décroissance marginale de l'utilité ». Cette hypothèse détermine dès lors la forme de fonction d'utilité qui doit présenter les caractéristiques d'être croissante ($u' > 0$) et concave ($u'' < 0$) avec l'accroissement de la richesse. Bernoulli proposa alors la fonction $u(x) = \log(x)$. Avec ces propriétés, le paradoxe de St. Pétersbourg est résolu : on obtient un calcul d'espérance d'utilité fini car l'utilité augmente beaucoup moins vite que la richesse elle-même³. Dans cette configuration, un agent n'a pas intérêt à risquer tout son patrimoine pour finalement récolter une utilité supplémentaire qui s'avère négligeable. Ainsi, contrairement aux prédictions établies sur la base de la maximisation de l'espérance, les individus sont prêts à renoncer à des montants monétaires lorsque ces dernières sont risquées. Selon Bernoulli, cette aversion pour le hasard est en fait une conséquence de la concavité de la fonction d'utilité.

II. LA THEORIE DE L'UTILITE ESPEREE

Cette théorie fut proposée par Von Neumann et Morgenstern dans « Theory of games and economic behavior » paru en 1944. Dans cet ouvrage, les auteurs proposent un fondement axiomatique à la prise de décision dans le risque, où : 1. les probabilités sur les conséquences sont connues, 2. Les objets sur lesquels portent les choix sont des loteries (ou de manière équivalente sur des distributions de probabilités sur les conséquences). Comme Bernoulli, les auteurs montrent que les agents cherchent à maximiser l'espérance d'utilité qu'ils anticipent pouvoir dériver des richesses prospectives induites par les loteries. Toutefois, à la différence de Bernoulli, en procédant à l'établissement d'un fondement axiomatique aux fonctions d'utilité

³ En supposant, $w_0 = 0$, $U = \lim_{n \rightarrow +\infty} \sum_{n=1}^N \frac{1}{2^n} \ln(2^n) = \ln(2) \lim_{n \rightarrow +\infty} \sum_{n=1}^N \frac{n}{2^n} \approx \ln(4)$

espérée, les auteurs expriment l'idée selon laquelle si l'on accepte ces axiomes, on doit nécessairement admettre qu'il existe une fonction u , représentant les préférences d'un individu sur les richesses (Cayatte, 2009). En ce sens, la démarche des auteurs consiste alors à fournir un cadre hypothético-déductif à la prise de décision dans le risque (Laine, 2014) plutôt que de justifier l'existence d'une telle fonction sur l'intuition ou sur l'observation empirique des comportements décisionnels face au risque.

2.1. Axiomatique de l'EUT

La relation de préférence qu'un agent exprime sur l'ensemble des loteries L doit satisfaire un certain nombre d'axiomes pour que la fonction d'utilité qui représente ses préférences existe. C'est une façon détournée ici d'établir que ces axiomes doivent être respectés pour considérer l'agent comme rationnel et ainsi construire une théorie du choix dans l'incertain cohérente.

2.1.1. Axiome 1 : Les préférences sont complètes, réflexives et transitives

Les préférences sont complètes (préordre complet) signifie que les agents sont en mesure de comparer les loteries entres-elles et donc d'en préférer une par rapport à l'autre ou bien d'être indifférent entre les deux. Dès lors, on estime que la relation de préférence \succsim est complète car toutes les préférences sur loteries peuvent être représentées par cette relation. Supposons deux loteries L_a et L_b appartenant à l'ensemble des loteries \mathcal{L} :

$$\forall L_a, L_b \in \mathcal{L}, L_a \succsim L_b \text{ ou } L_b \succsim L_a$$

La relation d'indifférence entre L_a et L_b est englobée dans la relation \succsim car :

$$\forall L_a, L_b \in \mathcal{L}, L_a \sim L_b \Leftrightarrow L_a \succsim L_b \text{ et } L_b \succsim L_a$$

Les préférences sont réflexives signifie qu'une loterie L quelconque est au moins préféré à elle-même :

$$\forall L \in \mathcal{L}, L \succcurlyeq L$$

Les préférences sont transitives correspond à une hypothèse de rationalité parfaite :

$$\forall L_a, L_b, L_c \in \mathcal{L}, \text{ si } L_a \succcurlyeq L_b \text{ et } L_b \succcurlyeq L_c \implies L_a \succcurlyeq L_c$$

2.1.2. Axiome 2 : Continuité

La relation de préférence \succcurlyeq sur l'ensemble \mathcal{L} des loteries vérifie : $\forall L_a, L_b, L_c \in \mathcal{L}$, si $L_a \succcurlyeq L_b \succcurlyeq L_c$ alors il existe une probabilité $\alpha \in [0,1]$ tel que :

$$L_b \sim \alpha L_a + (1 - \alpha)L_c.$$

α doit être vu comme la probabilité de donner accès à la loterie A et $(1 - \alpha)$ donnant accès à la loterie B. Alors, dit grossièrement, si L_a est préféré à L_b qui est préféré à L_c , il existe tout de même une certaine probabilité α pour que l'agent soit indifférent entre L_b et une composition de L_a et L_c .

2.1.3. Axiome 3 : L'Indépendance

Cet axiome signifie que si deux loteries L_a et L_b sont mélangées de façon identique avec une troisième loterie L_c alors la relation de préférence entre les deux loteries initiales n'est pas modifiée. La relation de préférence \succcurlyeq sur l'ensemble \mathcal{L} des loteries est telle que : $\forall L_a, L_b, L_c \in \mathcal{L}$ et pour tout $\alpha \in [0, 1]$,

$$L_a \succcurlyeq L_b \Leftrightarrow \alpha L_a + (1 - \alpha)L_c \succcurlyeq \alpha L_b + (1 - \alpha)L_c$$

L'axiome d'indépendance est absolument fondamental car il assure que les probabilités objectives sont indépendantes des préférences et implique donc la linéarité de la représentation des préférences par rapport aux probabilités.

2.1.4. Théorème de l'utilité espérée

Supposons que la relation de préférence \succcurlyeq définie sur l'ensemble des loteries satisfasse les axiomes 1 à 3. Dans ce cas, cette relation peut être représentée par une fonction d'évaluation linéaire dans les probabilités d'occurrence des événements $E(u(\tilde{W}))$. Autrement dit, il existe une fonction d'utilité u (définie à une transformation croissante affine près) définie sur les conséquences est à valeur dans \mathbb{R} tel que pour toute loterie $L_a(W_i^a, p_i^a)_{i=1,\dots,n}$ et $L_b(W_i^b, p_i^b)_{i=1,\dots,n}$ alors l'équivalence suivante est respectée :

$$\forall L_a, L_b \in \mathcal{L}, L_a \succcurlyeq L_b \Leftrightarrow E(u(\tilde{W}^a)) \geq E(u(\tilde{W}^b))$$

2.2. La fonction d'espérance d'utilité

Selon cette théorie, la valeur subjective d'une loterie est égale à l'espérance mathématique des utilités des gains monétaires de cette loterie. La fonction d'évaluation $E(u(\tilde{W}))$, est donnée par l'Equation (1.2) et porte sur les richesses finales prospectives induites par l'intégration des issues probabilistes des loteries à la richesse initiale de l'individu. Dans cette logique dite d'hypothèse d'intégration des issues à la richesse initiale, un individu devrait alors accepter un pari à la position initiale w_0 si et seulement si $u(w_0) < u(w_0 + x_1, p_1 + \dots + w_0 + x_n, p_n)$.

Par ailleurs, une propriété importante des fonctions d'utilité espérée est qu'elles sont définies à une transformation affine strictement croissante près, tout en conservant la propriété de

représenter les mêmes préférences. Une transformation strictement croissante de l'espérance d'utilité pourrait potentiellement préserver l'ordre des préférences comme par exemple la fonction carrée $\left[E \left(u(\tilde{W}) \right) \right]^2$, mais dans ce cas il ne s'agit plus d'une espérance. La fonction d'utilité élémentaire u , est elle-même définie à une transformation affine strictement croissante, de sorte que $E \left(u(\tilde{W}) \right)$ représente les mêmes préférences que $E \left(g \left(u(\tilde{W}) \right) \right)$ pourvu que la fonction g soit affine et strictement croissante. Ces précisions sont importantes car en substance cela indique que l'utilité peut être mesurée sur une échelle dont l'origine et l'unité peuvent être modifiées. Autrement dit, il s'agit d'une mesure cardinale de l'utilité où quelque-soit la transformation croissante et affine de u , le rapport des utilités induites par deux richesses sera malgré tout conservé et par conséquent l'ordre des préférences également (Cayatte, 2009).

Pour finir, on notera que le critère de maximisation basé sur l'espérance mathématique n'est pas rejeté par le critère basé sur celui de l'espérance d'utilité, mais s'en trouve être un cas spécifique lorsque la fonction d'utilité est linéaire : $u(x) = x$. Toutefois, la fonction d'évaluation $E \left(u(\tilde{W}) \right)$ conserve une propriété clé de la fonction d'espérance mathématique : la linéarité dans les probabilités.

2.3. Le triangle de Marschak-Machina

Le triangle de Marschak (1950) et Machina (1982) est une façon commode de représenter l'ensemble des loteries possibles à trois issues financières données et ordonnées $x_1 < x_2 < x_3$, $p_i \in [0, 1]$ pour $i \in \{1, 2, 3\}$ et $\sum_{i=1}^3 p_i = 1$. La probabilité p_1 du plus petit montant financier x_1 est portée par l'abscisse et la probabilité p_3 du meilleur résultat financier x_3 est portée par les ordonnées. Puisque tous les côtés du triangle sont égaux à l'unité, la probabilité $p_2 = (1 - p_1 - p_3)$ est égale à la longueur du segment entre l'hypoténuse et le point d'intersection entre p_1 et p_3 (Munier, 1995). Dans le cas où une loterie n'admettrait aucune chance d'obtenir x_1 ,

alors la probabilité de gagner x_1 sera forcément accolée à l'axe des ordonnées. Si $p_2 = 0$, la loterie sera sur le segment de l'hypoténuse, si $p_3 = 0$, alors elle sera placée sur l'axe des abscisses, enfin les loteries à l'intérieur du triangle sont des loteries où chaque probabilité est inférieure à 1 mais où bien entendu leur somme donne 1.

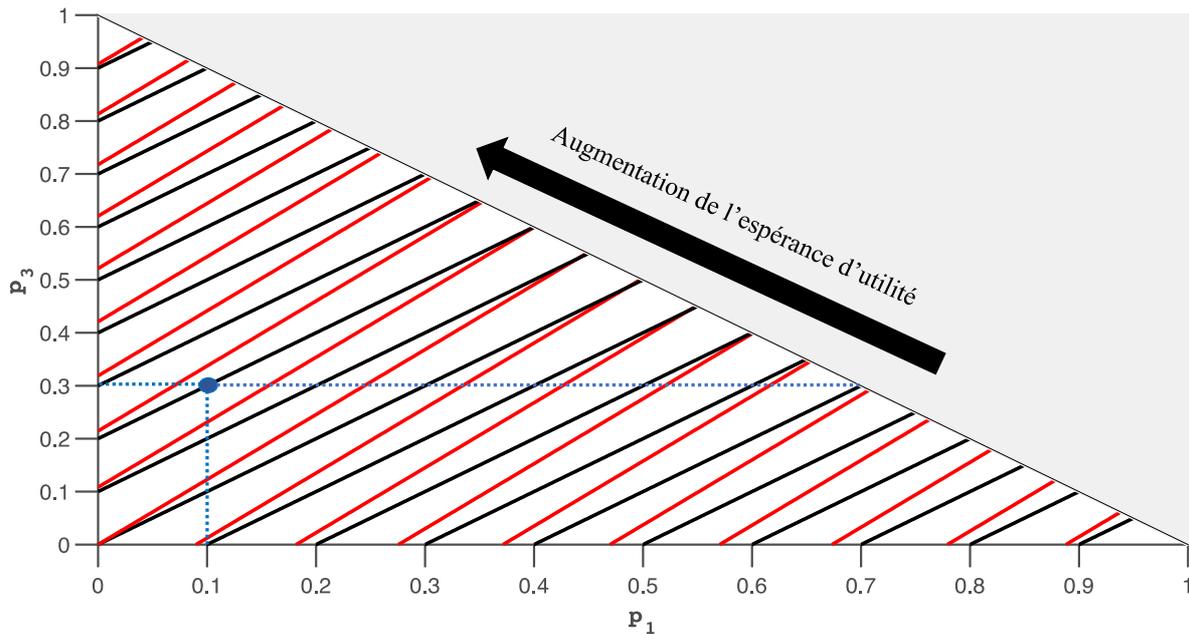


Figure 1.2 Illustration du triangle de Marschak-Machina. Ces droites iso-utilités et iso-espérances ont été construites à partir d'une loterie à trois issues : $x_1 = 0, x_2 = 500, x_3 = 1000$ et où $w_0 = 0$. La fonction d'utilité est donnée par $u(x) = x^\alpha$, avec $\alpha = 0.8$. Par exemple, le point bleu représente une loterie prodiguant 1000€ avec la probabilité $p_3=0.3$, 0 avec $p_1=0.1$ et 500 avec $p_2=0.6$

Les loteries situées sur une droite en trait noir de la Figure 1.2 donnent toutes les loteries ayant la même espérance de gain pour différentes combinaisons de p_1, p_2 et p_3 (« mean preserving spread »). Cependant, le déplacement « nord-est » sur ces droites « iso-espérance » représentent des loteries de plus en plus risquées car la variance s'accroît (puisque p_1 et p_3 augmente et p_2 diminue) tout en préservant la même espérance, on parle aussi d'augmentation « pure » du risque (Rothschild & Stiglitz, 1970). Par ailleurs, tout déplacement dans la direction « nord-ouest » du triangle mène à des loteries (ou des distributions) stochastiquement dominantes au

premier ordre⁴ car la probabilité du meilleur résultat augmente tandis que la probabilité du pire résultat diminue.

Il est possible à partir de ces informations de représenter graphiquement l'hypothèse d'espérance d'utilité en traçant les droites d'indifférences et d'espérances (Abdellaoui & Munier, 2001). Une droite d'indifférence représente l'ensemble des loteries qui procurent le même niveau d'utilité (iso-utilité), et une droite d'espérance représente toutes les loteries qui procurent une espérance de richesse équivalente (iso-espérance). Pour obtenir une droite d'indifférence ou d'espérance il suffit de fixer $E(u(\tilde{x})) = K$ et $E(\tilde{x}) = K$ respectivement. La question que l'on se pose alors est celle-ci : pour un accroissement de la probabilité (p_1) de la moins bonne issue, de combien doit augmenter la probabilité (p_3) de la meilleure issue pour maintenir la même utilité ? Etant donné que notre espace probabiliste est représenté par p_1 et p_3 , le terme p_2 devra s'écrire $1 - p_3 - p_1$. Ainsi, l'espérance d'utilité s'écrit $E(u(\tilde{x})) = p_1u(x_1) + (1 - p_1 - p_3)u(x_2) + p_3u(x_3)$. En dérivant l'expression par rapport à p_1 et p_3 et en maintenant l'expression égale à 0 pour préserver un niveau d'utilité constant, on obtient le taux marginal de substitution probabilistique suivant : $\frac{dp_3}{dp_1} = \frac{u(x_2) - u(x_1)}{u(x_3) - u(x_2)}$ qui représente la pente de la tangente de la fonction d'utilité en tout points⁵.

⁴ Soit F et G deux fonctions de répartition, F domine stochastiquement au premier ordre G si pour $\forall x \in X, F(x) \leq G(x)$ avec une inégalité stricte pour au moins un x .

⁵ Plus explicitement, on obtiendra une carte des droites d'indifférences ainsi :

$$\begin{aligned} K &= p_1u(x_1) + (1 - p_1 - p_3)u(x_2) + p_3u(x_3) \\ \Leftrightarrow K &= -p_1(u(x_2) - u(x_1)) + u(x_2) + p_3(u(x_3) - u(x_2)) = K \\ \Leftrightarrow p_3 &= \frac{K - u(x_2)}{u(x_3) - u(x_2)} + p_1 \frac{u(x_2) - u(x_1)}{u(x_3) - u(x_2)} \end{aligned}$$

Avec K une constante. Pour les droites d'iso-espérance, la fonction u disparaît car l'utilité est supposée linéaire avec la richesse $u(x) = x$.

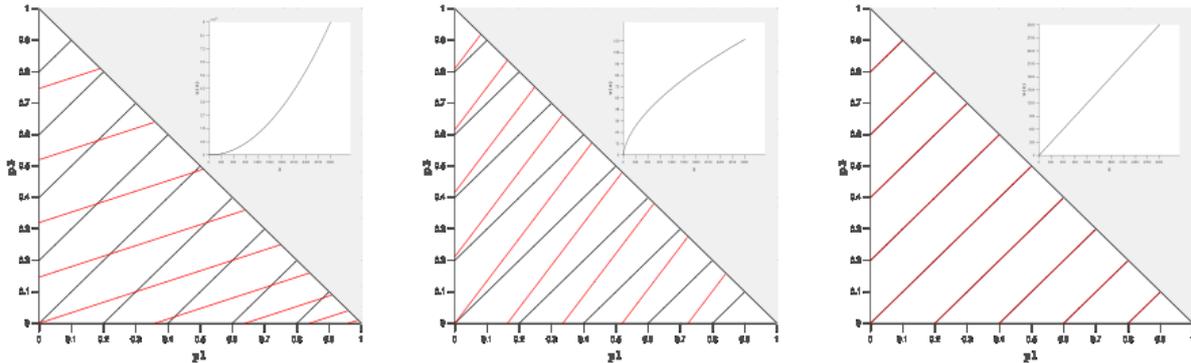


Figure 1.3 Triangle de Marschak-Machina et fonction d'utilité (encadré supérieur). Les droites iso-utilités à gauche sont élaborés en supposant une fonction puissance convexe. La figure au milieu, à partir d'une fonction d'utilité concave et la figure de droite à partir d'une fonction d'utilité linéaire.

Dans la mesure où cette pente est une constante (à x_1 , x_2 et x_3 donnée) car elle est indépendante de p , les droites d'indifférences sont donc linéaires et parallèles dans le modèle d'espérance d'utilité⁶. Dès lors, sous l'hypothèse d'utilité espérée, si deux substitutions de probabilités, à niveau d'utilité différents, ont le même impact sur ces niveaux d'utilités, alors ces substitutions sont de même amplitudes (Abdellaoui & Munier, 2001). Ce qui va nous intéresser ici plus particulièrement est la comparaison des pentes des droites d'indifférences et celle des droites d'espérances pour répondre à notre question et ce que l'on peut en déduire. On montre facilement que si u est fonction croissante concave, la pente de la droite d'indifférence est plus importante que la pente de la droite d'espérance : $\frac{u(x_2)-u(x_1)}{u(x_3)-u(x_2)} > \frac{x_2-x_1}{x_3-x_2}$. Autrement dit, pour maintenir la même utilité, la compensation que l'on doit appliquer à la probabilité (p_3) de la meilleure issue lorsque la probabilité du plus mauvais résultat (p_1) augmente, est plus importante que la compensation que l'on doit appliquer à p_3 pour maintenir la même espérance. Cette analyse en termes de taux de substitution probabilistique nous renseigne donc sur les

⁶ Plus exactement, la linéarité et le parallélisme des droites d'indifférences découle directement de l'axiome d'indépendance qui présume l'indifférence aux informations non pertinentes : La linéarité des droites d'indifférences est induite par $La \sim Lb \Leftrightarrow La \sim \alpha La + (1-\alpha)Lb \sim Lb$ avec $\alpha \in [0, 1]$. Le parallélisme des droites d'indifférences est induit par $La \sim Lb \Leftrightarrow \alpha La + (1-\alpha)Lc \sim \alpha Lb + (1-\alpha)Lc$. Comportementalement, la propriété de linéarité dans les probabilités selon l'EUT peut donc être vue comme une restriction des préférences sur des loteries composées (Machina, 1987).

préférences de l'individu face au risque : plus la pente des droites d'indifférences est supérieure à la pente des droites d'espérances, plus l'individu est averse au risque. En substance, cela nous indique que pour un agent averse au risque, l'augmentation pure du risque doit être compensée par une augmentation de l'espérance pour préserver le même niveau d'utilité.

III. VIOLATIONS DE L'EUT ET GENERALISATION DU MODELE

Une fois la théorie de l'Utilité Espérée énoncée sous sa forme axiomatique par VnM, il faudra attendre quelques années pour que les premières voix dissidentes s'élèvent (Allais, 1953; Ellsberg, 1961; Kahneman & Tversky, 1979). La première critique qui produisit un retentissement considérable dans la communauté des économistes fut adressée par Maurice Allais en 1953 avec son paradoxe éponyme. A partir de là, une profusion de contre-exemples expérimentaux et empiriques ont démontré, ou semble avoir démontré, l'inaptitude de l'EUT à rendre compte des comportements décisionnels réellement opérés par les agents. C'est aussi pendant cette période que foisonnèrent des modèles alternatifs à l'EUT : certains en sont des raffinements élaborés au prix du relâchement de l'axiome d'indépendance, d'autres s'en départissent plus radicalement jusqu'à rejeter l'idée même de rationalité telle que définie par l'EUT.

En définitive, les chercheurs contemporains en économie expérimentale doivent désormais choisir très finement les loteries qu'ils décideront de soumettre aux participants car la plupart de ces modèles ne divergent dans leurs prédictions que pour des situations de choix bien précises (ce qui d'une certaine manière est logique dans la mesure où ces modèles se veulent les plus généraux possible) (Cavagnaro, Gonzalez, Myung, & Pitt, 2013). D'ailleurs, la théorie de l'utilité espérée étant elle-même en capacité de capturer une grande partie des décisions humaines en environnement risqué, certains tenants de cette théorie reprochent aux dissidents

de confectionner des situations de choix qui en favorisent la violation (Binmore & Shaked, 2007). Vraisemblablement, on trouvera toujours d'autres exemples, où des manières spécifiques de présenter les choix (ou les conséquences des choix) qui favoriseront l'occurrence de résultats captés soit par l'EUT soit par un autre modèle. Quoiqu'il en soit, nous ne pouvons ignorer l'ensemble des travaux qui révéleront les défauts de la théorie standard en microéconomie de l'incertitude, et nous allons dans cette section en exposer les principaux.

Dans un premier temps nous analyserons le paradoxe d'Allais qui mena au relâchement de l'axiome d'indépendance, puis dans un second temps nous étudierons les paradoxes ayant conduit à de nouvelles formalisations de la fonction d'espérance utilité.

3.1. Le paradoxe d'Allais

3.1.1. Effet de la conséquence commune : violation de la propriété additive de l'axiome d'indépendance

Situation 1 :

Option A : 100% de chance de gagner 100 millions (ou écrit autrement : 11% de chance de gagner 100 millions et 89% de chance de gagner 100 millions)

Option B : 10% de chance de gagner 500 millions ; 89% de chance de gagner 100 millions ; 1% de ne rien gagner.

Situation 2 :

Option A' : 11% de chance de gagner 100 millions ; 89% de chance de ne rien gagner.

Option B' : 10% de chance de gagner 500 millions ; 90% de chance de ne rien gagner.

Les options A et B ont en commun la conséquence : « 89% de chance de gagner 100 millions ». Dans les options A' et B', cette conséquence est tout simplement supprimée. L'axiome d'indépendance nous indique que si $A > B$ ET si A et B sont mélangées de façon identique avec une loterie C, alors la relation de préférence initiale reste identique. Dans la situation de Allais, les loteries A' et B' se déduisent des loteries A et B dans la mesure où la conséquence « 89% de chance de gagner 100M » a été supprimée. Mathématiquement l'inconsistance est particulièrement visible. On observe que :

$$A > B \Leftrightarrow 1u(100M) > 0.89u(100M) + 0.01u(0) + 0.1u(500M)$$

$$A' < B' \Leftrightarrow 0.11u(100M) < 0.9u(0) + 0.1u(500M)$$

On peut réécrire la dernière équation comme suit :

$$1u(100M) - 0.89u(100M) < 0.01u(0) + 0.1u(500M)$$

$$1u(100M) < 0.89u(100M) + 0.01u(0) + 0.1u(500M)$$

On remarque une contradiction puisque la dernière équation déduite des préférences entre A' et B' implique une préférence contraire à celle observé entre A et B. Ce résultat équivaut à un renversement des préférences et remet donc en question la propriété additive de l'axiome d'indépendance (Camerer & Ho, 1994; Kammoun, 2007).

3.1.2. Effet du ratio commun : violation de la propriété multiplicative de l'axiome d'indépendance

Situation 1 :

Option A : 100% de chance de gagner 100 millions.

Option B : 98% de chance de gagner 500 millions ; 2% de chance de ne rien gagner.

Situation 2 :

Option A' : 1% de chance de gagner 100 millions ; 99% de chance de ne rien gagner.

Option B' : 0.98% de chance de gagner 500 millions ; 99.02% de chance de ne rien gagner.

Dans ces situations, les individus préfèrent A à B et B' à A' . A nouveau cela est inconsistant avec l'axiome d'indépendance et sa propriété cette fois multiplicative. En effet, on remarque que les loteries A' et B' sont obtenues en mixant les loteries A et B avec une autre loterie que nous appellerons $Z(0\text{€}, 1)$. En effet, $A' = [0.01(A) 0.99(Z)]$ et $B' = [0.01(B) 0.99(Z)]$. En fait, les probabilités de la situation 1 ont simplement été divisées par le même facteur : $1/100$, par conséquent le rapport des probabilités de gagner est conservé : $p(500)/p(100) = 0.98$ dans les deux situations. Dès lors, étant donné le traitement linéaire des probabilités, cela ne devrait pas affecter les préférences de l'individu selon l'EUT. En effet, l'axiome d'indépendance implique que pour deux probabilités p et q on obtient alors l'égalité suivante $p/q = rp/rq$ avec $0 < r < 1$. Cet effet a été expliqué par ce que l'on appelle « l'effet de certitude », c'est-à-dire par une préférence disproportionnée des agents pour la certitude par rapport à l'incertitude.

3.1.3. Analyse de l'effet de la conséquence commune avec le triangle de Marschak

Dans cette section nous représentons le paradoxe d'Allais au sein du triangle de Marschak-Machina. Nous voyons dans la Figure 1.4 que les quatre loteries forment un parallélogramme (rond noir sur la Figure 1.4), où p_1 représente la probabilité d'obtenir $x_1 = 0$, p_2 la probabilité d'obtenir $x_2 = 100$ millions et p_3 la probabilité d'obtenir $x_3 = 500$ millions. Les droites bleues représentent les droites d'espérance mathématique égales. On observe que les loteries B et B' donnent un plus grand gain espéré que les loteries A et A' respectivement. Les droites rouges sont les droites d'indifférences. Un individu qui préfère A à B est averse au risque, les droites d'indifférence sont donc plus pentues que les droites iso-espérées. Ainsi, on voit que l'utilité procurée par A est plus grande que celle procurée par B . Les droites d'indifférence étant linéaires et parallèles, l'utilité prodiguée par A' devrait alors être plus importante que celle de B' , d'où le paradoxe puisque les individus se contredisent en préférant B' à A' .

Depuis le travail de M Allais, d'autres études ont montré des paradoxes similaires (Ellsberg, 1961; Kahneman & Tversky, 1979). A la lumière de ces paradoxes, d'autres modèles ont été élaborés pour capturer ces préférences renversées. Toutefois, avant d'en faire l'état des lieux, il est important de réhabiliter les intentions véritables de Allais. A l'origine, M Allais développe son paradoxe dans l'optique d'illustrer son intuition selon laquelle l'axiome d'indépendance est « incompatible avec la préférence pour la sécurité au voisinage de la certitude » (Allais, 2008, p.4). Cependant, il ajoute que « loin de la certitude », les individus agissent comme indiqué par l'EUT. Cette remarque corrobore l'observation de Camerer (1992) selon laquelle, l'essentiel des violations de l'EUT est observable aux extrémités du triangle et non à l'intérieur puisque qu'aucune des issues n'est sûr ou proche de la certitude. Dans cette optique, de nombreuses études montrent en effet que les agents semblent traiter différemment les probabilités proches des bornes du triangle et les probabilités intermédiaires (Abdellaoui & Munier, 1998; Camerer & Ho, 1994; Harless & Camerer, 1994). Ainsi, les paradoxes d'Allais sont moins une violation générale de l'EUT qu'une violation locale qui apparaît lorsque certaines issues sont proches

de la certitude (Andreoni & Sprenger, 2010). Par conséquent, selon Andreoni & Sprenger (2010) “if the violation is isolated very close to certainty, it may prove useful to represent it as a violation of continuity. Individuals may exhibit discontinuous preferences over certain and uncertain outcomes” (Andreoni & Sprenger, 2010, p4).

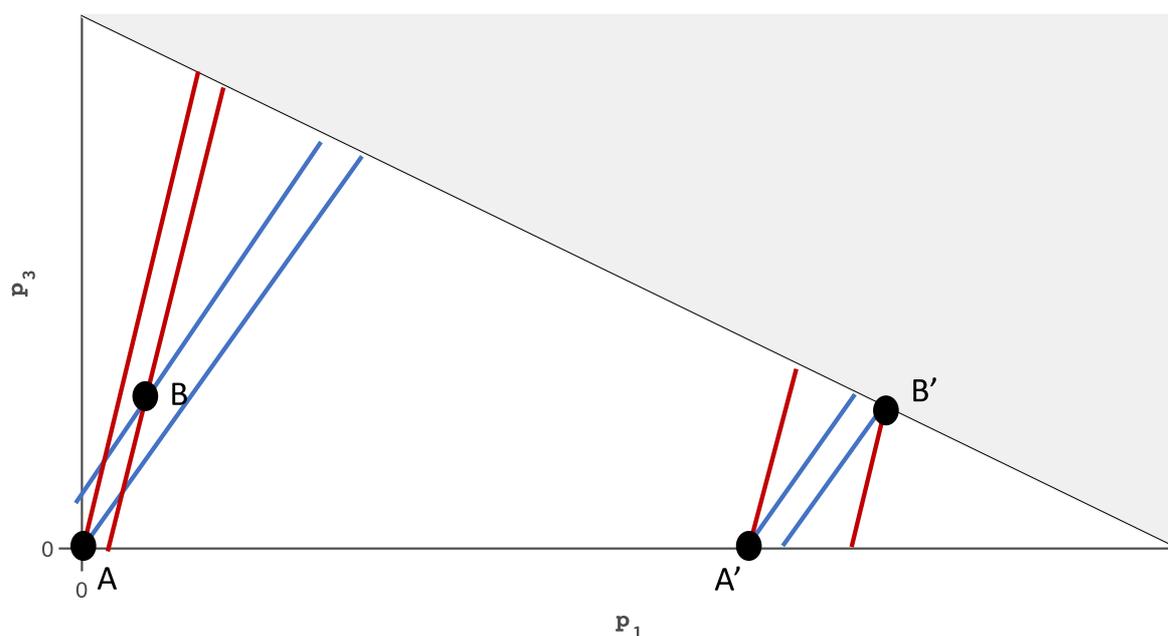


Figure 1.4 Représentation du paradoxe d'Allais « effet de la conséquence commune » dans le triangle de Marschak-Machina. Les points A, A', B et B' représentent les loteries utilisées par Allais dans ses deux situations.

3.2. Les préférences non-linéaires dans les probabilités

3.2.1. Fonction de pondération des probabilités et dominance stochastique de premier ordre

Une façon de résoudre ces paradoxes consiste à relâcher, voire à abandonner complètement l'axiome d'indépendance, en supposant l'existence d'une fonction de pondération des probabilités non-linéaires représentant le fait que les individus altèrent subjectivement les probabilités. D'une certaine manière, si l'on accepte l'idée d'une transformation subjective de la richesse, rien ne nous empêche d'appliquer cette conception aux probabilités. Toutefois, cette voie fut dans un premier temps moins prolifique car la fonction de transformation fut appliquée

directement aux probabilités des issues indépendamment des issues elles-mêmes ce qui mène à des prédictions normatives indésirables (Willinger, 1990). On peut en effet facilement montrer qu'une telle application conduit à la violation de la dominance stochastique (qui est la version probabiliste de l'axiome de non-satiété en univers certain : plus = mieux). De telles fonctions de pondération furent proposées par de nombreux auteurs comme Wu & Gonzalez (1999), Karmakar (1979) puis Kahneman & Tversky (1979) dans leur première version de la prospect theory. Ces formalisations consistent simplement en un calcul d'espérance dans lesquelles les probabilités et les paiements sont pondérés par des fonctions différentes. Dès lors, un individu doté d'une richesse initiale w_0 faisant face à une loterie $L(x_i, p_i)_{i=1, \dots, n}$ évalue subjectivement la richesse finale aléatoire comme suit :

$$U(\tilde{W}) = \sum_{i=1}^n \varphi(p_i) u(W_i) \quad (1.3)$$

Où $\varphi(p_i)$ désigne la fonction de pondération des probabilités définie dans l'intervalle $[0, 1]$. Bien que l'introduction d'une fonction altérant les probabilités permette une nette amélioration descriptive du modèle, la validité normative est toutefois douteuse car le modèle présenté par l'Equation (1.3) le principe de dominance stochastique de premier ordre⁷ et donc l'axiome de non-satiété. Pour le démontrer, prenons une loterie à n issues équiprobables prodiguant $x + \varepsilon_i$, avec $i = 1, \dots, n$. Supposons une fonction de pondération des probabilités convexe telle que $\varphi(p) = \varphi\left(\frac{1}{n}\right) < \frac{1}{n}$. On peut montrer qu'en prenant $\varepsilon > 0$ mais suffisamment petit, on trouvera

⁷ La loterie A, de fonction de répartition Fx , domine stochastiquement à l'ordre 1 la loterie B, de fonction de répartition Gx , si et seulement si $\forall x, Fx \leq Gx$, l'inégalité étant stricte pour au moins une valeur de x de probabilité non nulle.

qu'on peut obtenir : $u(x) > \sum_{i=1}^n \varphi\left(\frac{1}{n}\right) u(x + \varepsilon_i)$. Dans cette situation, l'individu préfère un gain certain à un jeu qui lui propose pourtant le gain certain + un supplément ε_i . Autrement dit, même si l'individu est certain de gagner davantage que le gain certain en choisissant la loterie, le modèle prédit qu'il peut préférer le gain certain. Cela est dû au fait que nous avons choisi une fonction de pondération qui sous-pondère les probabilités et un supplément ε suffisamment petit.

3.2.2. La Rank-Dependent Utility Theory (RDUT)

Dans cette section nous allons exposer le modèle d'utilité dépendant des rangs (RDU). Ce modèle présente l'avantage, contrairement aux modèles précédents, de respecter le principe de dominance stochastique. Elle fut énoncée en premier lieu par Quiggin (1982) qu'il appela « la théorie de l'utilité anticipée » bien que la paternité de l'idée puisse être attribuée à Allais. L'existence d'une fonction de pondération non linéaire est maintenue mais ne porte pas sur les probabilités prises chacune isolément. Cette fonction s'applique désormais aux probabilités décumulées. Dès lors, le poids associé à une conséquence dépend non seulement de sa probabilité mais également de son rang (sa place) par rapport aux autres conséquences (Diecidue & Wakker, 2001).

Présentons maintenant le modèle de décision RDU. A richesse initiale w_0 , supposons sans perte de généralité que les issues x_i d'une loterie L soient ordonnées, on obtient alors les richesses finales suivantes : $W_1 < W_2 < \dots < W_n$. Ces conséquences sont attachées aux probabilités p_1, p_2, \dots, p_n respectivement et où $\sum_{i=1}^n p_i = 1$. Dans le modèle RDU, l'utilité d'une richesse finale aléatoire dépend d'une fonction d'utilité $u(\cdot)$ et d'une fonction de pondération des probabilités $\varphi(\cdot)$. La fonction φ est strictement croissante avec $\varphi(0) = 0$ et $\varphi(1) = 1$. La valeur subjective d'une loterie à richesse initiale donnée, selon le modèle RDU s'écrit alors :

$$\begin{aligned}
RDU(\tilde{W}) &\equiv E_{\Phi} \left(u(\tilde{W}) \right) \\
&= \varphi(p_n)u(W_n) + \sum_{k=1}^{n-1} \left[\varphi \left(\sum_{j=0}^k p_{n-j} \right) - \varphi \left(\sum_{j=0}^{k-1} p_{n-j} \right) \right] u(W_{n-k}) \quad (1.4)
\end{aligned}$$

Ce modèle est souvent écrit selon une expression réduite :

$$E_{\Phi} \left(u(\tilde{W}) \right) = \sum_{i=1}^n \Phi_i u(W_i) \quad (1.5)$$

Où $\Phi_i = \varphi \left(\sum_{j=0}^{i=k} p_{n-j} \right) - \varphi \left(\sum_{j=0}^{i=k-1} p_{n-j} \right)$ représente ce que l'on appelle le poids décisionnel d'une conséquence x_i . De plus, la somme des poids décisionnels est égale à 1. Comme nous l'avons évoqué la fonction $\varphi(\cdot)$ porte désormais non pas sur les probabilités individuelles mais sur les probabilités décumulées. Pour le comprendre, notons $F_{\tilde{W}}(W)$ la fonction de répartition de la richesse finale aléatoire \tilde{W} . Cette fonction indique pour chaque issue la probabilité d'obtenir un résultat inférieur ou égal à l'issue considérée W_i . A l'inverse, dans le modèle RDU, la fonction φ s'applique à l'ensemble de probabilités des issues aux moins aussi bénéfique, $\varphi \left(\sum_{j=0}^{i=k} p_{n-j} \right)$, et strictement plus bénéfique, $\varphi \left(\sum_{j=0}^{i=k-1} p_{n-j} \right)$, que l'issue considérée W_i . On considère donc la fonction décumulative $G_{\tilde{W}}(W) = 1 - F_{\tilde{W}}(W)$. Dès lors, le poids Φ_i d'une richesse W_i est défini comme la différence entre le poids d'obtenir n'importe quel gain aussi bénéfique que W_i et le poids d'obtenir n'importe quel gain strictement plus bénéfique que W_i . En définitive, le poids décisionnel d'une issue dans une loterie dépend de différents éléments.

1. Le rang de l'issue.
2. La probabilité de l'issue ainsi que celle des autres issues (autrement dit la distribution des probabilités sur les conséquences).
3. La fonction de pondération des probabilités.

La forme de la fonction $\varphi(p)$ capture l'attitude des individus face aux probabilités. En l'occurrence, on parlera d'optimisme ou de pessimisme lorsque la fonction φ est concave ou convexe respectivement. La Figure 1.5 illustre l'impact de la forme de φ sur les poids décisionnels des conséquences d'une loterie comprenant 20 conséquences équiprobables.

Pour le comprendre considérons un individu pessimiste confronté à une loterie produisant la conséquence w_1 avec la probabilité p_1 ou w_2 avec la probabilité p_2 , où $w_1 > w_2$. Le poids décisionnel de w_2 est donné par $\Phi_2 = \varphi(p_1 + p_2) - \varphi(p_2)$.

Si φ est convexe (individu pessimiste) alors $\varphi(p_1 + p_2) - \varphi(p_2) < \varphi(p_1 + p_2') - \varphi(p_2')$ avec $p_2' \geq p_2 \geq 0$. Autrement dit, une augmentation de probabilité de la plus mauvaise conséquence (mauvaise nouvelle) produit une augmentation du poids décisionnel de ce dernier (et réciproquement une diminution du poids de la meilleure conséquence).

Si φ est concave (individu optimiste) alors $\varphi(p_1 + p_2) - \varphi(p_2) > \varphi(p_1 + p_2') - \varphi(p_2')$ avec $p_2' \geq p_2 \geq 0$. Autrement dit, une augmentation de probabilité de la plus mauvaise conséquence (mauvaise nouvelle) produit une diminution du poids décisionnel de ce dernier (et réciproquement une augmentation du poids de la meilleure conséquence).

En définitive une fonction φ concave signifie que l'agent accorde un poids plus important aux meilleures conséquences, alors qu'une fonction convexe signifie que l'agent accorde plus d'importance aux pires conséquences.

Ce concept de poids décisionnel est souvent mal compris car il est souvent interprété à tort comme une distorsion subjective des probabilités individuelles (Fehr-Duda & Epper, 2012). Diecidue et Wakker (2001) proposent une intuition psychologique derrière la notion de poids décisionnel. Il s'agit plutôt de l'importance relative qu'accorde l'individu à une issue possible. Une manière un peu grossière de représenter cette idée peut être évoquée de la façon suivante. L'attractivité d'une loterie dépend évidemment des issues et de leurs probabilités respectives. Les poids décisionnels représentent la participation relative de chacune de ces conséquences à

l'attractivité de la loterie. Par exemple, un poids décisionnel de 0.3 signifie que la conséquence participe à 30% de l'attractivité de la loterie.

Cette importance relative dépend du rang de l'issue, de sa probabilité ainsi que de la forme de la fonction φ . Pour comprendre l'impact du rang, supposons un individu pessimiste confronté à une loterie à 3 issues équiprobables : 10, 20 et 30€. Le poids décisionnel de 10€ sera surreprésenté par rapport aux autres issues étant donné qu'elle est la pire conséquence. Si on suppose maintenant une loterie à 3 issues mais où 10€ est la meilleure issue, dans ce cas son poids sera sous-représenté par rapport aux autres.

Il est donc important de noter que les poids décisionnels dépendent des rangs et non des amplitudes de gains possibles conformément à l'axiome d'indépendance comonotonique (voir section Dépendance aux rangs et Triangle de Marschak-Machina de ce chapitre).

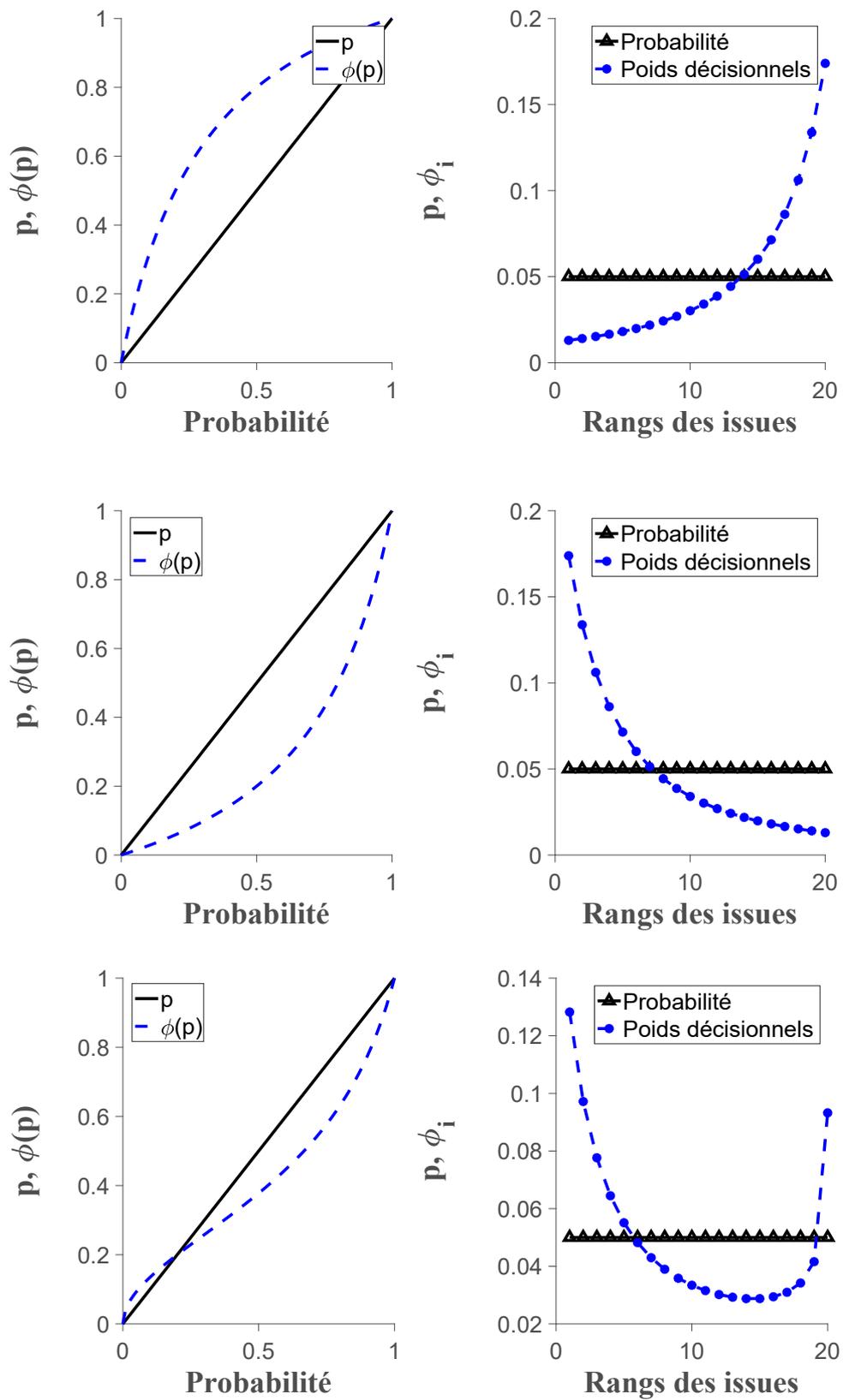


Figure 1.5 Poids décisionnels dans le modèle RDU, montrant les effets de la dépendance de rang sur les poids décisionnels pour une distribution de probabilité uniforme discrète p sur 20 résultats classés i (triangle noires), selon les caractéristiques de la fonction de pondération de probabilité $\phi(p)$.

3.2.3. La Dual Theory

La « dual theory » introduite par Yaari (1987) est un cas particulier de la théorie de l'utilité dépendante des rangs bien que Yaari développa indépendamment son modèle de celui de Quiggin (1982). L'auteur maintient l'idée d'une transformation non linéaire des probabilités cumulées mais suppose une fonction d'utilité linéaire : $u(x) = x$, de sorte qu'en reprenant l'Equation (1.5), on obtient :

$$DT(\tilde{W}) = \sum_{i=1}^n \Phi_i W_i \quad (1.6)$$

Ce modèle étant un cas particulier de la RDU, on pourrait estimer en ce sens qu'il n'a que peu d'intérêt. Toutefois son originalité provient de sa dualité avec l'EUT. La dualité exprime l'idée selon laquelle une même chose peut provenir de deux processus différents. Ainsi, alors que l'EUT suppose que l'attitude face au risque est une conséquence d'une transformation de la richesse en utilité non linéaire de la richesse, la dual theory établit au contraire la linéarité dans l'utilité dérivée des paiements mais la non-linéarité dans les probabilités. Dès lors l'attitude face au risque est une conséquence d'une transformation des probabilités. La théorie standard traduit la linéarité de $u(\cdot)$ comme la neutralité face au risque, alors que dans la théorie de Yaari cette linéarité exprime simplement l'idée que l'utilité marginale est constante⁸.

⁸ On généralise également cette remarque dans le modèle RDU. En effet dans ce modèle, on distingue l'attitude face au risque de la fonction d'utilité alors que dans la théorie standard ces deux éléments sont confondus. En effet dans l'EUT, une fonction d'utilité croissante concave (utilité marginale décroissante) implique l'aversion au risque ; croissante convexe (utilité marginale croissante) = riscophilie ; croissante linéaire (utilité marginale constante) = neutralité. Dans le modèle RDU, une fonction u croissante concave représente uniquement la décroissance marginale de l'utilité avec la richesse mais n'implique pas forcément l'aversion au risque puisque cette dernière dépend également de l'attitude face aux probabilités. Ainsi, dans cette théorie, il est tout à fait envisageable d'avoir une utilité marginale décroissante et avoir de la préférence pour le risque si la fonction de pondération des probabilités est suffisamment concave (Quiggin, 1982).

3.3. Formalisations et propriétés des fonctions de pondération des probabilités

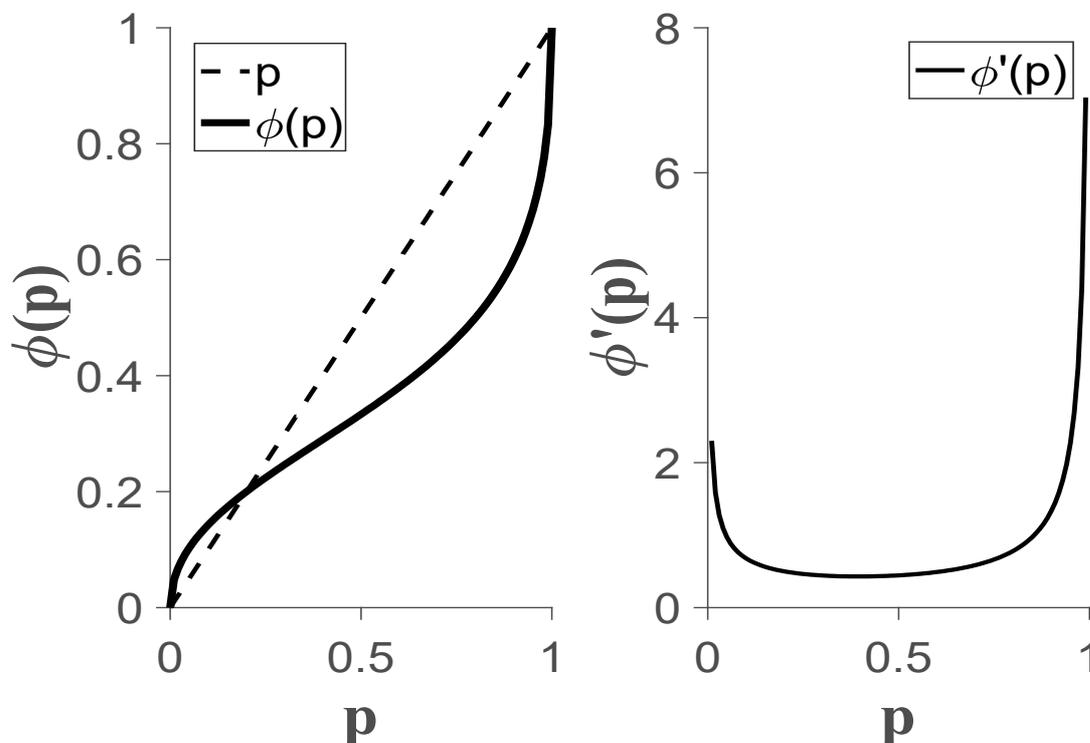


Figure 1.6 A gauche, représentation d'une fonction de pondération des probabilités $\phi(p)$ selon une spécification de Prelec et de la fonction de probabilité sans transformation. A droite, représentation de la fonction dérivée de la fonction $\phi(p)$.

3.3.1. Propriétés

Une littérature vaste (comprenant le paradoxe d'Allais), montre que les individus sont particulièrement sensibles aux variations de probabilités dans les extrêmes mais sont par contre peu sensibles aux variations de probabilités intermédiaires (Quiggin, 1982; Tversky & Wakker, 1995; Weber, 1994). Une augmentation de probabilité de gagner de 0.1 à un impact important sur l'accroissement de l'utilité de la loterie lorsque l'on passe de très probable (90% de chance de gagner) à la certitude de gagner (100%) ou lorsque l'on passe de l'impossibilité (0%) à l'envisageable (10%). En revanche, lorsqu'on considère cette même variation sur un intervalle intermédiaire (passer de 40% de chance de gagner à 50% de chance de gagner), l'accroissement

de l'utilité de la loterie est plus faible. Par conséquent, la forme de cette fonction adopte celle d'un S inversé : concave puis convexe (avec p en abscisse et $\varphi(p)$ en ordonnée).

On peut interpréter cette forme comme étant le résultat d'une loi psychophysique établit par Weber & Fechner (1860) : la décroissance marginale de la sensibilité à mesure que l'intensité d'un stimulus augmente. En particulier, dans une situation probabiliste, on peut établir que le plus haut degré d'incertitude est au point 0.5 ce qui correspond à l'intensité maximale du stimulus « incertitude », alors que l'intensité minimale se situe aux deux pôles $p = 0$ (impossibilité) et $p = 1$ (certitude). Dès lors, à mesure que l'on s'éloigne de ces deux points, la sensibilité aux variations de probabilités diminue jusqu'à un certain point p (au niveau du point d'inflexion). De ces considérations on tire plusieurs propriétés de la fonction de pondération Prelec (1998):

Régressive : $\varphi(p) > p$ quand p est proche de 0 et $\varphi(p) < p$ quand p proche de 1.

Sous-additivité (voir Figure 1.7):

basse : Si on considère un intervalle de probabilité proche de 0 : $\varphi(q) - \varphi(0) \geq \varphi(p + q) - \varphi(p)$ quel que soit $p + q \leq 1 - \varepsilon$, où $\varepsilon > 0$. Cette propriété signe la concavité de la fonction de pondération des probabilités quand p est petit. Elle fait référence à l'effet de « possibilité », qui traduit le sentiment des individus qu'un évènement rare reste possible. Ainsi, passer d'un gain impossible à 5% de chance de l'obtenir a plus d'impact que de passer de 30% à 35% de chance de l'obtenir.

haute : Si on considère un intervalle de probabilité proche de 1 : pour une constante $\varepsilon' \geq 0$, $\varphi(1) - \varphi(1 - q) \geq \varphi(p + q) - \varphi(p)$ quel que soit $p \geq \varepsilon'$. Cette propriété signe la nature convexe de la fonction de pondération des probabilités lorsque p est élevé. Elle fait généralement référence à l'effet de certitude qui traduit le sentiment des individus que la différence entre un gain probable et un gain certain est beaucoup plus

important qu'une différence comparable dans l'intervalle des probabilités intermédiaires (Kahneman & Tversky, 1979).

La sous-additivité haute explique l'effet de la conséquence commune du paradoxe d'Allais (voir Tversky & Wakker, 1995). Pour le comprendre, il nous faut représenter la situation de ce paradoxe en considérant les fonctions de répartition décumulatives de chaque loterie, c'est-à-dire la probabilité d'obtenir un résultat égal ou supérieur à l'issue considérée. Dans le Tableau 1.1, nous représentons la distribution de probabilité sur les conséquences de chaque loterie A, B et A' et B'. Nous répertorions également les distributions de probabilités décumulées sur les conséquences pour chaque loterie.

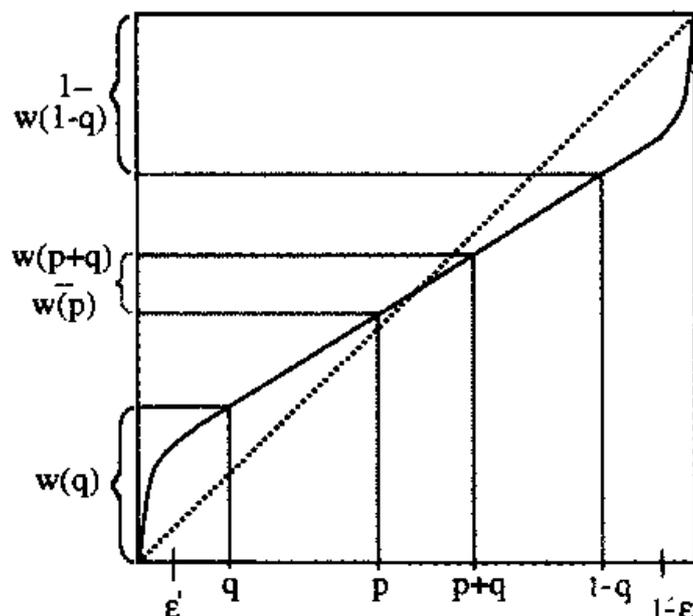


Figure 1.7 Représentation de la sub-additivité haute et basse (figure tirée de Tversky et Wakker, 1995)

On constate qu'en termes de probabilités décumulées, les situations sont identiques sauf concernant l'issue x_2 . Concernant la situation 1 et 2, la probabilité décumulée de x_2 diminue de 0.1 entre les loteries A et B et entre les loteries A' et B'. Etant donné que les préférences s'inversent entre ces deux situations ($A > B$ et $A' < B'$), l'impact de cette diminution n'est donc pas la même. Vraisemblablement, la perte de la certitude de gagner dans la situation 1 produit

une plus forte diminution de la valeur subjective que lorsque cette même diminution est appliquée sur une issue incertaine. Mathématiquement : $\varphi(1) - \varphi(0.99) > \varphi(0.11) - \varphi(0.1)$.

Tableau 1.1 Probabilités et probabilités décumulées des issues des loteries engagées dans le paradoxe d'Allais : Effet de la conséquence commune

Issues	Probabilités			Probabilités décumulées		
	$x_1 = 0$	$x_2 = 100M$	$x_3 = 500M$	$x_1 = 0$	$x_2 = 100M$	$x_3 = 500M$
A	$p_1 = 0$	$p_2 = 1$	$p_3 = 0$	$P_1 = 1$	$P_2 = 1$	$P_3 = 0$
B	$p_1 = 0.01$	$p_2 = 0.89$	$p_3 = 0.1$	$P_1 = 1$	$P_2 = 0.99$	$P_3 = 0.1$
A'	$p_1 = 0.11$	$p_2 = 0.11$	$p_3 = 0$	$P_1 = 1$	$P_2 = 0.11$	$P_3 = 0$
B'	$p_1 = 0.9$	$p_2 = 0$	$p_3 = 0.1$	$P_1 = 1$	$P_2 = 0.1$	$P_3 = 0.1$

Sous-proportionnalité : explique l'effet du ratio commun

$$\frac{\varphi(p)}{\varphi(q)} > \frac{\varphi(rp)}{\varphi(rq)}, 0 < r < 1, q < p \leq 1$$

Dans l'effet du ratio commun, nous avons pour la première situation la relation de préférence suivante – $A > B \Leftrightarrow \varphi(1)u(100M) > \varphi(0.98)u(500M)$. En réarrangeant les termes on obtient $\frac{\varphi(1)}{\varphi(0.98)} > \frac{u(500M)}{u(100M)}$. Pour la seconde situation, les probabilités de gagner sont divisées par 100. Nous avons alors la relation de préférence suivante – $A' < B' \Leftrightarrow \varphi(0.01)u(100M) < \varphi(0.098)u(500M)$. En réarrangeant les termes on obtient : $\frac{\varphi(0.01)}{\varphi(0.098)} < \frac{u(500M)}{u(100M)}$. Par conséquent, conformément à la propriété sous-proportionnelle des fonctions de pondération des probabilités, on observe que $\frac{\varphi(1)}{\varphi(0.98)} > \frac{\varphi(0.01)}{\varphi(0.098)}$.

3.3.1. Formalisation de la fonction de pondération des probabilités

De nombreuses formalisations de la fonction de pondération des probabilités ont été proposées. On peut distinguer deux types de formalisations : Celles à un paramètre libre et ceux à 2 paramètres libres. Tversky & Kahneman (1995) ont proposé le modèle suivant :

$$\varphi(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \quad (1.7)$$

Où γ contrôle la courbure de la fonction. Si γ est trop petit la fonction n'est pas strictement croissante en fonction de p . Si $\gamma = 1$ le modèle se réduit à l'EUT car $\varphi(p)$ devient linéaire.

Plus γ est petit plus la forme S inversé est prononcée.

Parmi les formalisations à deux paramètres libres on trouve parmi les plus connues : le modèle de Prelec (1998) et les modèles de la famille hyperbolique proposé par Wu et Gonzalez (1998), Lattimore, Baker, & Witte (1992) et Myerson, Green et Morris (2011). Le modèle de Prelec (1998) s'écrit :

$$\varphi(p) = \exp(-\delta(-\ln p)^\gamma) \quad (1.8)$$

Le modèle hyperbolique de Wu & Gonzalez (1998) :

$$\varphi(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma} \quad (1.9)$$

Le modèle hyperboloïde de Myerson, Green et Morris⁹ (2011)

⁹ Le modèle de Myerson et al (2011) est toutefois plus difficile à manier dans la mesure où h et s sont corrélés. C'est véritablement la conjugaison des 2 paramètres qui donne la forme en S inversé à la fonction alors que dans le modèle de Wu et Gonzalez (ou encore de Prelec), le paramètre γ est uniquement consacré à la courbure de la fonction et le paramètre δ à l'élévation (c'est-à-dire le point d'intersection entre la fonction $\varphi(p)$ et la fonction p).

$$\varphi(p) = \left(\frac{p}{p + h(1-p)} \right)^s \quad (1.10)$$

Les modèles à un paramètre libre capturent très bien la courbure convexe et concave de la fonction de pondération. En revanche, ils ne nous disent rien sur la sur et sous-pondération par rapport aux probabilités objectives, c'est-à-dire par rapport à la fonction identité à 45°. En effet, des agents peuvent tout à fait exhiber des préférences non linéaires de type S shape inversé, tout en sur-pondérant ou sous-pondérant l'ensemble des probabilités. Dans ces cas extrêmes, la fonction $\varphi(p)$ ne croise pas la fonction p et peut se trouver soit complètement au-dessus (sur-pondération complète) soit complètement en dessous (sous pondération complète) de la fonction p . Le point d'intersection entre la fonction $\varphi(p)$ et p exprime ce que l'on appelle l'élévation de la fonction et s'interprète comme la tendance générale de l'individu à surpondérer ou sous pondérer les probabilités. Cette propriété est capturée par le paramètre β chez Prelec et δ chez Wu & Gonzalez. La courbure des fonctions de pondération des probabilités à deux paramètres libres est quant à elle exprimée par le paramètre γ .

3.4. Dépendance aux rangs et Triangle de Marschak-Machina

La fonction des courbes d'indifférence selon un modèle RDU sont obtenues ainsi : On pose K une constante représentant un niveau fixe d'utilité. On suppose une loterie à trois issues $x_3 > x_2 > x_1$ associés aux probabilités p_3 , p_2 et p_1 respectivement. La valeur subjective de la loterie L est donnée par :

$$\begin{aligned} RDEU(L) &= K \\ &= \varphi(p_3)u(x_3) + [\varphi(p_3 + p_2) - \varphi(p_3)]u(x_2) \\ &\quad + [\varphi(p_3 + p_2 + p_1) - \varphi(p_3 + p_2)]u(x_1) \end{aligned}$$

On sait que $p_3 + p_2 + p_1 = 1$ et que $\varphi(1) = 1$. De plus, comme nous devons exprimer les fonctions de courbes d'indifférence dans un plan (O, p_3, p_1) on exprimera p_2 tel que $p_2 = 1 - p_3 - p_1$. On réécrit alors l'expression ainsi :

$$RDU(L) = K = \varphi(p_3)u(x_3) + [\varphi(1 - p_1) - \varphi(p_3)]u(x_2) + [1 - \varphi(1 - p_1)]u(x_1) \quad (1.11)$$

On isole ensuite $\varphi(p_3)$ et on obtient :

$$\varphi(p_3) = \frac{K - u(x_1) - \varphi(1 - p_1)(u(x_2) - u(x_1))}{u(x_3) - u(x_2)}$$

Il est alors intéressant d'étudier le taux marginal de substitution probabilistique induit par un modèle RDU (voir Camerer, 1989, p.77). C'est-à-dire, de combien doit augmenter p_3 lorsque p_1 augmente pour maintenir une utilité (ou une valeur subjective) constante ? Pour cela on procède classiquement. On dérive l'expression (1.11) par rapport à p_3 , $RDU(L)_{p_3}' = \frac{\partial RDU(L)}{\partial p_3}$.

Ainsi, $RDU(L)_{p_3}' \cdot dp_3$ représente l'augmentation marginale de RDU lorsque p_3 augmente. Puis

on dérive cette même expression par rapport à p_1 , $RDU(L)_{p_1}' = \frac{\partial RDU(L)}{\partial p_1}$. $RDU(L)_{p_1}' \cdot dp_1$

représente l'augmentation marginale de RDU lorsque p_1 augmente. En différenciant

l'expression (1.11) par rapport à p_3 et p_1 on obtient alors $dRDU(L) = RDU(L)_{p_3}' \cdot dp_3 +$

$RDU(L)_{p_1}' \cdot dp_1$ Pour maintenir $RDU(L)$ constant on pose $dRDU(L) = 0$. Ce qui implique :

$\frac{dp_3}{dp_1} = -\frac{RDU(L)_{p_1}'}{RDU(L)_{p_3}'}$. Ou plus explicitement, le taux marginal probabilistique donne :

$$\frac{dp_3}{dp_1} = \frac{\varphi'(1-p_1)[u(x_2) - u(x_1)]}{\varphi'(p_3)[u(x_3) - u(x_2)]}$$

Ce qui fournit la pente des courbes d'indifférences. On comprend alors que les courbes d'indifférences ne sont pas linéaires si on suppose des fonctions de pondération des probabilités non-linéaires spécifiés selon l'un des modèles que nous avons présenté dans la section 3.3.2 de cette partie.

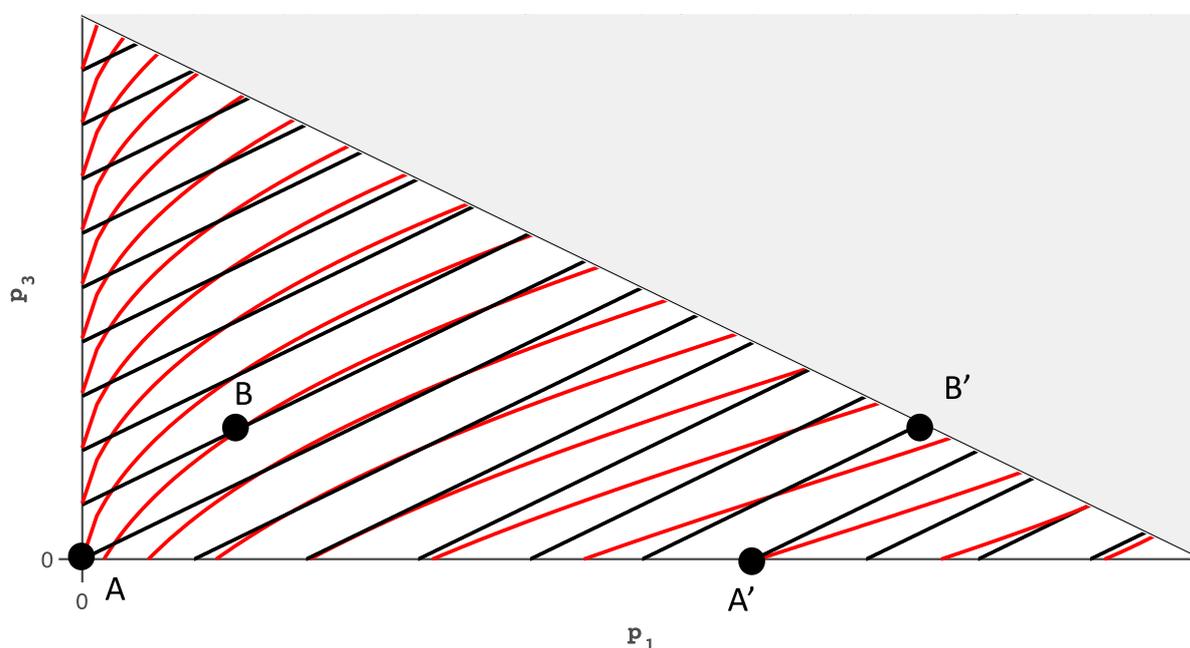


Figure 1.8 Représentation des courbes d'indifférences selon le modèle RDU dans le triangle Marschak-Machina. Elles sont obtenues en supposant une fonction d'utilité puissance et une fonction de pondération des probabilités en forme de S inversé selon le modèle de Prelec. Les Points A, B, A' et B' représentent les loteries du paradoxe d'Allais : la conséquence commune

Analysons les courbes d'indifférences obtenues avec le modèle RDU (courbes rouges dans la Figure 1.8). Ces dernières sont élaborées en supposant une fonction d'utilité puissance concave et en supposant une fonction de pondération des probabilités en forme de S inversé obtenu à partir de la spécification de Prelec (1998). On remarque que lorsque la probabilité (\$p_1\$) de la pire conséquences (\$x_1\$) est petite, une importante augmentation de la probabilité \$p_3\$ de la meilleure conséquence (\$x_3\$) est nécessaire pour compenser une petite augmentation de \$p_1\$

(aversion au risque). A l'inverse, lorsque p_1 est élevé, les individus sont prêts à accepter une petite augmentation de p_3 pour compenser l'augmentation de p_1 . Cette compensation est d'ailleurs plus petite que la compensation induite par un modèle d'espérance mathématique de gain (courbes noirs), marquant ici une attitude de recherche de risque. Ainsi, le modèle RDU est en mesure de prédire des changements d'attitude face au risque en fonction des probabilités et des rangs des issues engagées.

Nous montrons par ailleurs que le modèle RDU est en mesure d'accommoder le paradoxe d'Allais (1953). Comme le montre la Figure 1.8, la courbe d'indifférence passant par le point A est supérieure à la courbe d'indifférence passant par le point B marquant la préférence du sujet pour la loterie A. A l'inverse, on remarque que la courbe d'indifférence passant par le point A' est inférieure à la courbe d'indifférence passant par le point B' révélant ainsi une préférence du sujet pour la loterie B', conformément aux résultats obtenus par Allais (1953).

Puisque la RDUT est en mesure d'accommoder le paradoxe d'Allais, on en déduit que ce modèle ne respecte pas l'axiome d'indépendance énoncé dans le cadre de l'EUT. En fait, cet axiome est relâché pour être énoncé sous une forme plus faible appelé « axiome d'indépendance comonotique ». Deux variables aléatoires X et Y qui induisent le même classement des conséquences (c.-à-d. qui génèrent les mêmes nombres $1 : n, \dots, n : n$), sont appelées comonotoniques. Nous avons vu dans la section 3.2.2 de cette partie que le poids décisionnel d'une conséquence dépend du rang de cette conséquence, c'est-à-dire de son classement par rapport aux autres issues. Si ces variables aléatoires ont un résultat en commun tel que $x_i = y_i$ et que l'on applique une transformation identique à ces deux résultats $x_i' = y_i'$ sans modifier l'ordre de classement des résultats, les poids décisionnels restent identiques et donc la préférence entre ces deux loteries ne devrait pas être affecté. C'est ce que l'on appelle l'indépendance comonotique. A contrario, si la modification des deux résultats implique une

modification du classement des issues, les poids décisionnels sont alors modifiés et l'ordre des préférences peut potentiellement changer (Wakker, Erev, & Weber, 1994).

C'est exactement ce qu'il se passe dans le cadre du paradoxe d'Allais et de l'effet de la conséquence commune. Dans le Tableau 1.2 on remarque que le retrait de la conséquence « 89% de chance de gagner 100M » a modifié les rangs des issues ce qui peut induire un changement de préférence.

Tableau 1.2 Rangs des issues engagées dans le paradoxe d'Allais

RANGS			
LOTERIE	1	2	3
A	100		
B	0	100	500

RANGS			
LOTERIE	1	2	3
A'	0	100	
B'	0	500	

IV. L'ATTITUDE FACE AU RISQUE DANS LES DIFFERENTS MODELES DE DECISION

Il existe plusieurs manières de définir l'attitude face au risque. Les deux principales sont : l'aversion forte pour le risque (Rothschild & Stiglitz, 1970) et l'aversion faible pour le risque (Arrow, 1965; Pratt, 1964) L'aversion forte pour le risque, telle que définie par Rothschild et Stiglitz (1970) établit qu'un agent a de l'aversion forte pour le risque s'il préfère la variable aléatoire X à n'importe quelle variable aléatoire Y déduite de X dont l'étalement de la fonction de densité est plus important que celui de la fonction de densité de X tout en préservant la même espérance. On parle alors d'augmentation pure du risque à moyenne constante, ou de « mean preserving spread » (MPS).

La distinction entre aversion faible et forte pour le risque est importante. En effet, comme nous allons le voir, dans le cadre de L'EUT, la concavité de la fonction d'utilité implique naturellement l'aversion faible pour le risque (par simple application des inégalités de Jensen). Autrement dit, la diminution marginale de l'utilité est une condition nécessaire et suffisante à l'aversion au risque. Avec l'avènement des nouveaux modèles tels que la RDUT, la concavité de la fonction d'utilité n'implique plus forcément l'aversion faible pour le risque. C'est pourquoi une définition plus générale de l'aversion au risque fut introduite par Rotschild et Stiglitz (1970).

L'aversion faible pour le risque réfère au fait qu'un individu préfère à n'importe quelle richesse finale aléatoire \tilde{W} , son espérance $E(\tilde{W})$, tel que $E(\tilde{W}) > \tilde{W}$. De manière générale, l'attitude face au risque est définie en termes d'équivalent certain, $Ce(\tilde{W})$, par rapport à l'espérance d'une richesse finale non-dégénérée. L'équivalent certain est le montant d'argent qui a la même utilité (la même valeur subjective) qu'une loterie. Il s'agit donc d'un point d'indifférence entre une quantité d'argent certaine et une richesse finale non-dégénérée, de sorte que $u(Ce(\tilde{W})) \sim E(u(\tilde{W}))$. On obtient l'équivalent certain d'une loterie à richesse initiale donnée, en calculant l'inverse de l'utilité appliqué à la valeur subjective de la richesse finale prospective :

$$Ce(\tilde{W}) = u^{-1} \left[E(u(\tilde{W})) \right] \quad (1.12)$$

Ici, $U(\tilde{W})$ réfère à n'importe quel modèle de décision : RDU, EUT, DT ou même cumulative prospect theory que nous développons dans la partie 2. On notera que l'équivalent certain de la loterie elle-même est simplement $Ce(L) = Ce(\tilde{W}) - w_0$. La différence entre l'espérance de la richesse finale aléatoire et son équivalent certain donne la prime de risque absolue de la loterie : $\pi_a(\tilde{W}) = E(\tilde{W}) - Ce(\tilde{W}) \Leftrightarrow \pi_a(L) = (w_0 + E(L)) - (Ce(L) + w_0) = E(L) -$

$Ce(L)$. Dès lors, il est inutile de préciser si nous parlons de la prime absolue de la loterie ou de la prime de risque absolue de la richesse finale aléatoire puisqu'elles sont égales. Nous désignerons donc la prime de risque simplement par la notation suivante : π_a . La prime de risque absolue est donc la quantité d'argent que l'individu est prêt à payer pour éviter le risque. On peut désormais adopter la classification suivante :

Si $Ce(\tilde{W}) < E(\tilde{W})$ ou si $\pi_a > 0$, l'agent est averse au risque

Si $Ce(\tilde{W}) > E(\tilde{W})$ ou si $\pi_a < 0$, l'agent est chercheur de risque

Si $Ce(\tilde{W}) = E(\tilde{W})$ ou si $\pi_a = 0$, l'agent est neutre au risque

La prime de risque constitue donc une mesure très utile pour quantifier et comparer les degrés d'aversion au risque d'individus à loterie et richesse initiale donnée. Nous verrons dans la partie suivante qu'il est possible de séparer les déterminants de la prime de risque, à savoir, séparer les contenus objectifs de la situation des éléments subjectifs propres à chaque agent.

4.1. L'espérance d'utilité et attitude face au risque

La théorie de l'espérance d'utilité suppose l'existence d'une fonction d'utilité monotone croissante puis soit concave, soit convexe, soit linéaire. Par ailleurs aucune transformation des probabilités n'est présente dans ce modèle. La forme de la fonction d'utilité détermine donc à elle seule l'attitude face au risque et le degré d'aversion (ou de recherche) de risque. Le recours aux inégalités de Jensen permet de montrer qu'une fonction d'utilité concave (convexe) implique toujours l'aversion au (la recherche de) risque face à n'importe quelle loterie non-dégénérée. Autrement dit, la diminution marginale de l'utilité ($u' > 0$ et $u'' < 0$) implique l'aversion au risque.

A richesse initiale et loterie donnée, nous avons vu que la prime de risque constitue une quantification de l'aversion au risque d'un individu. Dans ce cadre, les analyses de Arrow (1965) et de Pratt (1964) fournissent une approximation de la prime de risque pour un petit risque. L'intérêt fondamental de cette approximation est qu'elle permet de décomposer la prime de risque en séparant les éléments objectifs de la situation (richesse initiale, caractéristiques de la loterie) des éléments subjectifs (fonction d'utilité) qui déterminent le niveau de la prime de risque d'un agent.

Le but est de trouver une approximation de la prime de risque à partir de l'expression suivante :

$$u(E(\tilde{W}) - \pi_a) = E(u(\tilde{W}))$$

Cette expression représente l'égalité entre l'utilité de l'équivalent certain ($E(\tilde{W}) - \pi_a = Ce$) et l'espérance d'utilité de la loterie. Les auteurs effectuent alors un développement limité des deux termes pour un petit risque. Par petit risque on entend qu'un individu doté d'une richesse initiale w_0 est confronté à une loterie composée d'issues aléatoires x_i dont l'espérance est nulle, $E(\tilde{x}) = 0$ et la variance infinitésimalement petite. Dès lors l'approximation s'effectue au voisinage de la richesse initiale w_0 ou de façon équivalente au voisinage de l'espérance de richesse finale¹⁰ $E(\tilde{W})$.

Concernant le terme de gauche, on effectue maintenant un développement limité d'ordre 1 du terme de gauche au voisinage de $E(\tilde{W})$:

$$u(E(\tilde{W}) - \pi_a) \approx u(E(\tilde{W})) - \pi_a u'(E(\tilde{W}))$$

Concernant le terme de droite, on effectue un développement limité d'ordre 2 autour du point $E(\tilde{W})$ on obtient :

¹⁰ En effet, on sait que $E(\tilde{W}) = E(w_0 + \tilde{x})$. Or puisque l'espérance d'une somme est égale à la somme des espérances on a $E(\tilde{W}) = E(w_0) + E(\tilde{x})$. Enfin, étant donné que $E(\tilde{x}) = 0$ alors $E(\tilde{W}) = w_0$

$$E(u(\tilde{W})) \approx u(E(\tilde{W})) + E(\tilde{W} - E(\tilde{W}))u'(E(\tilde{W})) + \frac{1}{2}E(\tilde{W} - E(\tilde{W}))^2 u''(E(\tilde{W}))$$

Or puisque :

$$E(\tilde{W} - E(\tilde{W})) = 0 \text{ et } E(\tilde{W} - E(\tilde{W}))^2 = \sigma^2(\tilde{W})$$

On a alors :

$$E(u(\tilde{W})) \approx \frac{1}{2}\sigma^2(\tilde{W})u''(E(\tilde{W}))$$

En égalisant ces deux développements on obtient une approximation de la prime de risque :

$$\pi_a \approx \frac{1}{2}\sigma^2(\tilde{W}) \left[-\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} \right] \quad (1.13)$$

On remarque d'après l'Equation (1.13) que la prime de risque absolue dépend de la moitié de la variance de la richesse finale aléatoire, $\frac{1}{2}\sigma^2(\tilde{W})$, (élément objectif de la situation) et du degré d'aversion au risque local de l'agent, dénoté Γ_a , $\Gamma_a = -\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))}$, appelé coefficient d'Arrow-Pratt¹¹ (élément subjectif de la situation). Ainsi, le degré d'aversion au risque absolue dépend du degré de concavité de la fonction d'utilité en un point (ici l'espérance de richesse finale) normalisé par sa dérivée première. En particulier, plus la fonction d'utilité est concave plus le degré d'aversion au risque est important. Lorsque $\Gamma_a > 0$ ($\Gamma_a < 0$) l'individu est averse au (chercheur de) risque.

¹¹ Il s'agit finalement d'une mesure de l'élasticité de l'utilité marginale.

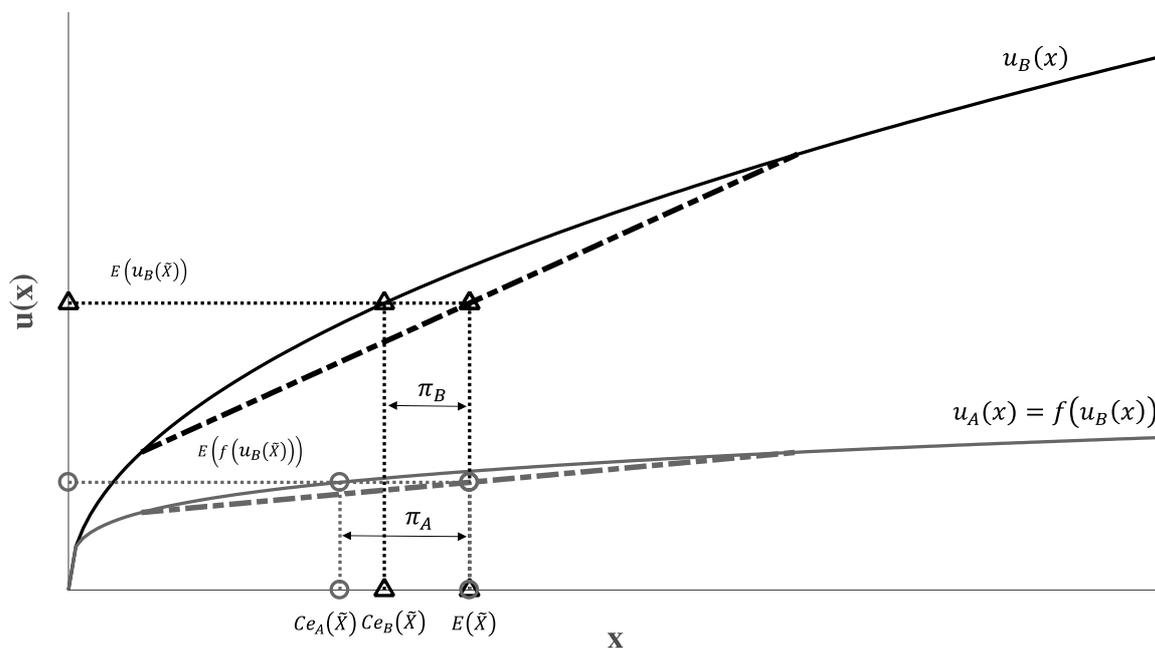


Figure 1.9 Représentation des fonctions d'utilité de deux agents A et B, sachant que la fonction d'utilité de l'agent A est construite en procédant à une transformation croissante concave de la fonction d'utilité de l'agent B. La prime de risque de l'agent A est supérieure à celle de l'agent B.

Ce coefficient est spécifique à chaque individu puisqu'il dépend de la fonction d'utilité propre à l'agent. Dès lors si deux individus A et B possèdent la même richesse initiale et sont confrontés à une loterie, la différence de prime de risque entre les deux ne pourra s'expliquer que par leur degré d'aversion au risque différent.

On peut alors énoncer le théorème de Pratt (1964) : Soit deux individus A et B dont les préférences sont représentées respectivement par les fonctions d'espérances d'utilité :

$E(u_A(\tilde{W}))$ et $E(u_B(\tilde{W}))$, les trois propriétés suivantes sont équivalentes :

u_A est une transformation croissante concave de u_B telle que : $\exists f, f' > 0, f'' < 0 : u_A(x) = f(u_B(x))$ (voir Figure 1.9).

La prime de risque absolue de A associée à la richesse finale aléatoire \tilde{W} est supérieure à la prime de risque absolue de B associée à la même richesse pour tout petit risque \tilde{x} , tel que $\tilde{W} = w_0 + \tilde{x}$. D'où $\pi_A(\tilde{x}) > \pi_B(\tilde{x})$.

En n'importe quel point de x , le coefficient d'Arrow-Pratt de A est supérieur à celui de B.

$$\forall x, \frac{u''_A(x)}{u'_A(x)} > \frac{u''_B(x)}{u'_B(x)}$$

4.2. La RDUT et l'attitude face au risque

Nous venons d'exposer dans la section précédente que concavité de la fonction de la fonction d'utilité était synonyme d'aversion pour le risque dans le cadre de l'EUT. Chez Yaari au contraire, la fonction d'utilité est linéaire et ne peut donc pas capturer les différentes attitudes face au risque. Dans ce cadre c'est la fonction de pondération des probabilités, $\varphi(p)$, qui est en mesure d'accommoder l'attitude face au risque des individus. En particulier, Yaari montre qu'une fonction $\varphi(p)$ convexe (concave) implique l'aversion au (la recherche de) risque.

Dans la RDUT la caractérisation de l'attitude face au risque est plus compliquée car elle est déterminée par deux fonctions : la fonction d'utilité et la fonction de pondération des probabilités décumulées. Par exemple, en considérant l'aversion faible pour le risque, un individu peut être averse au risque même si la fonction d'utilité est convexe, à la condition que la fonction de pondération des probabilités soit « suffisamment » convexe. Nous cernons ici un indice de possibles problèmes d'indétermination où un même comportement (par exemple un équivalent certain révélé d'un agent) peut être capturé par différentes paramétrisations du modèle (Kirby, 2011). Nous exposerons plus en détail ce problème dans la partie 3.

L'idéal serait d'être en mesure de séparer les effets de ces deux fonctions sur l'attitude face au risque. Cela fut entrepris par Wakker (1994) selon une approche axiomatique mais également par Hilton (1988) selon une approche dans la continuité directe des analyses de Arrow (1964) et Pratt (1965).

L'approche de Hilton (1988)

Pour comprendre l'analyse de Hilton, nous devons dans un premier temps étudier les déterminants de l'attitude face au risque lorsqu'un agent est confronté à une loterie ordinaire

composée d'un vecteur de conséquences possibles attachées à une distribution de probabilités. L'évaluation d'une telle loterie est donnée par l'Equation (1.5). Nous avons vu que l'attitude face au risque était la résultante de la fonction d'utilité et de la distorsion des probabilités décumulées transformée en poids décisionnel. Dès lors, la prime de risque absolue est en fait conditionnée par l'aversion au risque elle-même ainsi que par l'attitude de l'individu face aux probabilités.

Afin de séparer l'impact de ces deux attitudes, Hilton suppose que la prime de risque « global », noté π_a , est composée d'une prime de poids décisionnel (« decisional weight premium », noté π_Φ , et d'une prime de risque standard dans le sens d'Arrow-Pratt, noté π_θ , mais calculé en fonction des probabilités transformées en poids décisionnels. Formellement, la prime de poids décisionnel est la différence entre l'espérance objective de la loterie et l'espérance transformée, notée $E_\Phi(\tilde{W})$, calculée à partir des poids décisionnels tels que :

$$\pi_\Phi = \sum_{i=1}^n p_i W_i - \sum_{i=1}^n \Phi_i W_i = E(\tilde{W}) - E_\Phi(\tilde{W}) \quad (1.14)$$

Or, puisque selon Hilton $\pi_a = \pi_\Phi + \pi_\theta$, alors la prime de risque standard est bien calculée en fonction des probabilités transformées, tel que :

$$\pi_\theta = \pi_a + \sum_{i=1}^n \Phi_i W_i - \sum_{i=1}^n p_i W_i = \pi_a - \pi_\Phi \quad (1.15)$$

La Figure 1.10 donne une illustration graphique de cette prime de risque composite.

Présentons maintenant l'approximation de la prime de risque dans un modèle RDU selon Hilton (1988). Supposons une variable aléatoire de richesse finale \tilde{W} associée à la distribution de

probabilité P , une espérance de 0 et une variance infinitésimalement petite. Le but est de trouver une approximation locale de la relation : $u(E(\tilde{W}) - \pi_a) = E_\Phi(u(\tilde{W}))$ qui correspond à l'égalité entre la valeur subjective de la loterie donnée par le modèle RDU et l'utilité de l'équivalent certain (qui est égal à l'utilité de la différence entre l'espérance et la prime de risque).

En procédant aux développements limités des deux termes au voisinage de $E(\tilde{W})$ (au premier ordre pour le terme de gauche puis au second ordre pour le terme de droite) et isolant la prime de risque, on obtient l'approximation suivante :

$$\pi_a \approx \frac{1}{2} \sigma_\Phi^2(\tilde{W}) \left[-\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} \right] + [E(\tilde{W}) - E_\Phi(\tilde{W})] \quad (1.16)$$

Le premier terme de l'Equation (1.16) est similaire à l'approximation local de Arrow-Pratt au point $E(\tilde{W})$ présente dans l'Equation (1.13). La différence est que l'approximation est calculée ici en fonction de la distribution des probabilités transformée. Le terme $\left[-\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} \right]$ est la traditionnelle fonction du coefficient d'Arrow-Pratt. Enfin le second terme, $[E(\tilde{W}) - E_\Phi(\tilde{W})]$, est, ce que Hilton appelle, la prime du poids décisionnel.

On remarque donc très clairement que l'attitude générale de l'individu face au risque (en l'occurrence évaluée par la prime de risque) est la combinaison, ici une somme, de l'effet de l'aversion au risque de l'individu et de l'effet de son optimisme/pessimisme. Il est donc tout à fait envisageable qu'un individu se révèle averse au risque même si son degré d'aversion au risque est négatif (chercheur de risque) $\Gamma_a < 0$. Pour cela, il faut qu'il soit suffisamment pessimiste de sorte que $E(\tilde{W}) - E_\Phi(\tilde{W}) > 0$.

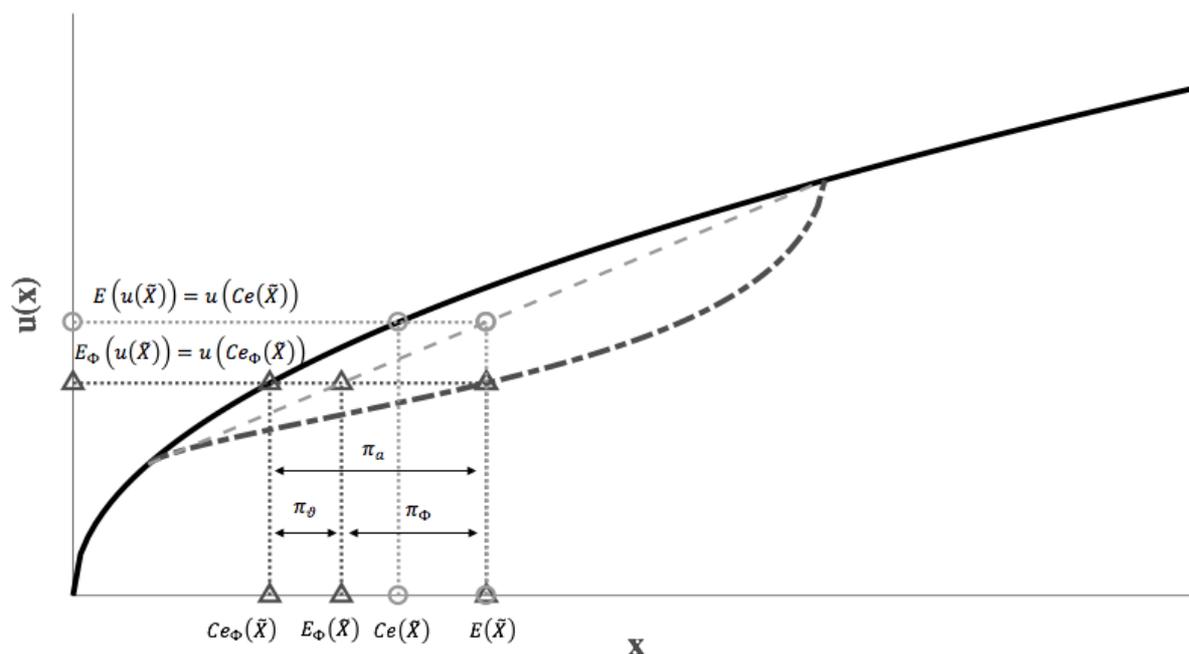


Figure 1.10 Représentation d'une fonction d'utilité puissance (courbe noir trait plein). La fonction d'espérance d'utilité (trait pointillé grise) est construite sans transformation des probabilités. C'est-à-dire selon un modèle EUT. La fonction de non-espérance d'utilité (traits pointillés noir) est construite selon un modèle RDU avec transformation des probabilités.

V. L'approche expérimentale

La révélation de différents paradoxes, comme celui du paradoxe d'Allais, ont poussé la recherche à développer de nouveaux modèles capables de les accommoder. Dès lors, de nombreuses études ont cherché à confronter les théories entre-elles vis-à-vis de leurs pouvoirs descriptifs. Dans ce chapitre nous n'avons présenté que l'EUT, la DT et la RDUT. Dans la section suivante nous présenterons la Cumulative Prospect theory. Toutefois, de nombreux autres modèles furent proposés comme la « Regret Theory » (Loomes & Sugden, 1982, 1987), la « Disappointment aversion theory » (Gul, 1991) ou encore la « Weighted Utility » (Karmarkar, 1978, 1979). Nous ne présentons pas les théories Disappointment aversion theory, ni la Weighted Utility, ni la Regret Theory ici, mais d'excellentes descriptions de ces modèles sont fournies dans « The foundations of Behavioral Economic Analysis » (Dhmi, 2016).

Après avoir passé en revue quelques méthodes de révélations des préférences célèbres, nous reporterons certains résultats importants qui ont tenté de départager les modèles.

5.1. Méthode de révélation des préférences

Les méthodes utilisées pour révéler l'attitude face au risque d'agents ne produisent pas les mêmes résultats. Autrement dit, la méthode influence l'objet qu'il est censé mesurer (Charness, Gneezy, & Imas, 2013; Crosetto & Filippin, 2013). Par exemple, en moyenne, l'aversion au risque observé par la méthode de l'équivalent certain est moins importante que lorsqu'elle est mesurée par la méthode du « probability equivalent » (Hershey & Schoemaker, 1985). Ici nous présentons le Multiple Price List et le tradeoff method¹².

5.1.1. Le Multiple Price List (MPL)

Le MPL est sans doute la méthode la plus utilisée dans la littérature économique. Il en existe différentes versions mais les plus connues sont celle proposées par Holt et Laury (2002) ainsi que la méthode de l'équivalent certain. Dans ces deux méthodes, l'individu est confronté à un choix entre deux alternatives. Chez Holt et Laury, le participant doit choisir entre deux loteries non-dégénérées, impliquant chacune deux issues possibles. L'une des loterie, appelée $R \equiv (r_1, p_1, r_2, 1 - p_1)$ est plus risquée que l'autre, appelée $S \equiv (s_1, p_1, s_2, 1 - p_1)$ de sorte que $r_1 > s_1 > s_2 > r_2$. Les auteurs font varier la probabilité p_1 de S et R dans l'intervalle $[0, 1]$. Dès lors, l'espérance de S est supérieure à celle de R quand p_1 est petit (≤ 0.4) et cette relation s'inverse quand $p_1 > 0.4$. Les auteurs recherchent le point de shift, c'est-à-dire le niveau de probabilité pour lequel l'individu préfère R à S. Si le point de shift est supérieur à 0.5 on dira que l'individu est averse au risque. Holt et Laury (2002) ont notamment montré que plus l'enjeu

¹² D'autres méthodes existent comme la méthode d'investissement de Gneezy et Potters (1997), le BART (Lejuez et al., 2002), le Bomb Risk Elicitation Task (Crosetto & Filippin, 2013)

des issues était important, plus le point de shift se produisait pour des valeurs de p_1 élevées (révélant une aversion au risque dite IRRA sur laquelle nous reviendrons dans la partie 3).

Une autre méthode du MPL est la méthode de l'équivalent certain. Il s'agit ici de présenter une loterie non dégénérée L que l'on confronte à une liste de montants financiers certains¹³ Y compris entre le minimum de la loterie et son maximum, $Y \in [\min(L), \max(L)]$ (voir par exemple Fehr-Duda et al, 2010). Pour chaque montant certain, l'individu doit signaler s'il préfère la loterie pour laquelle les issues et leurs probabilités sont fixées, ou le montant certain. Cette méthode permet de révéler directement l'équivalent certain d'une loterie, tel que $L(x_i, p_i)_{i=1, \dots, n} \sim Ce(L) \Leftrightarrow U(L) = u(Ce(L))$. Pour cela, il suffit de trouver le point de shift entre la préférence pour les montants certains et la loterie. L'équivalent certain est alors donné par la moyenne entre le dernier montant certain préféré à la loterie et le premier montant certain pour lequel la loterie est préférée. L'une des principales critiques portées contre la méthode de l'équivalent certain est que les résultats dépendent fortement des probabilités choisies. McCord et Neufville (1986) montrent que les fonctions d'utilités générées à partir des équivalents certains sont différentes en fonction des probabilités choisies de la loterie. Or, l'ensemble des modèles en économie supposent l'indépendance des fonctions d'utilité par rapport aux probabilités (ou aux fonctions de pondération des probabilités). Il s'agit du principe de séparabilité sur lequel nous revenons dans la partie 3. La méthode des tradeoffs a, entre-autre, été établie pour contrecarrer ce problème.

5.1.2. La méthode des tradeoffs (TO)

¹³ Il existe plusieurs méthodes d'élicitation de l'équivalent certain comme par exemple la méthode itérative (Estle, Green et Myerson, 2006).

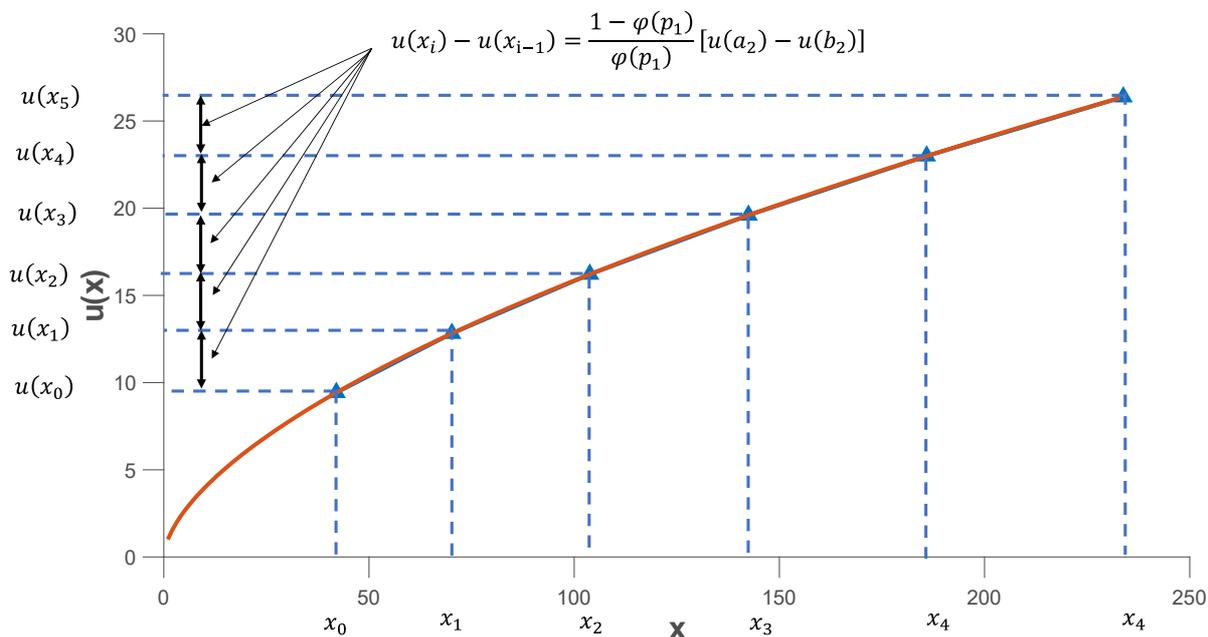


Figure 1.11 Représentation des résultats de la méthode tradeoffs

La méthode des tradeoffs a été proposée par Wakker et Deneffe (1996). Elle permet notamment de construire les fonctions d'utilité de façon non-paramétrique. La méthode consiste à déterminer l'indifférence entre deux loteries non-dégénérées. Une loterie $A \equiv (x_0, p_1; a_2, 1 - p_1)$ est fixée. Une seconde loterie $B \equiv (x_1, p_1; b_2, 1 - p_1)$ est établie où b_2 est fixé. Avec $x_1 > x_0 > a_2 > b_2$. On demande alors à l'individu pour quelle valeur de x_1 il est indifférent entre A et B. Au point d'indifférence, on a par définition une égalité des valeurs subjectives associées à A et B, de sorte que $U(A) = U(B)$. On va alors procéder par itérations successives pour construire la fonction d'utilité. Pour le comprendre, supposons un modèle de décision par exemple RDU. On sait que si l'agent est indifférent entre A et B alors :

$$\varphi(p_1)u(x_0) + (1 - \varphi(p_1))u(a_2) = \varphi(p_1)u(x_1) + (1 - \varphi(p_1))u(b_2)$$

Le montant x_1 est alors égal à $u(x_1) = \frac{1 - \varphi(p_1)}{\varphi(p_1)} [u(a_2) - u(b_2)] + u(x_0)$. Pour construire la fonction d'utilité on va maintenir constante la distance entre l'utilité de x_0 et celle de x_1 , de

sorte que $u(x_1) - u(x_0) = \frac{1-\varphi(p_1)}{\varphi(p_1)} [u(a_2) - u(b_2)]$. Le terme de droite est bien constant car l'ensemble des informations contenues dans cette expression est fixé. Ainsi, dès que l'on révèle x_1 , à la prochaine itération on remplace x_0 par la valeur de x_1 révélé et on va alors rechercher x_2 tel que $A \equiv (x_1, p_1; a_2, 1 - p_1)$ et $B \equiv (x_2, p_1; b_2, 1 - p_1)$. La Figure 1.11 illustre ce procédé et son intérêt pour construire de façon non-paramétrique la fonction d'utilité (on procède par interpolation linéaire pour relier les points, ou bien en supposant une forme paramétrique de la fonction d'utilité).

Notons par ailleurs que sur la base des résultats donnés par la méthode TO on pourra estimer la fonction de pondération des probabilités (Abdellaoui, 2000).

5.2. Résultats expérimentaux

De nombreuses expériences furent conduites dans le but de tester la robustesse et les capacités descriptives des différents modèles. Deux approches furent adoptées. L'approche déterministe consiste à compter le nombre de comportements (de décisions) en accord avec les prédictions d'une théorie. Le modèle qui explique la plus grande fréquence de choix est considéré comme le meilleur. La seconde approche, dite stochastique, consiste à estimer les paramètres libres des fonctions impliquées dans un modèle de décision puis de comparer les capacités descriptives de ces modèles (c'est-à-dire leurs capacités à capturer les observations empiriques) à l'aide du BIC¹⁴ ou de l'AIC¹⁵. Dans cette méthode, tout un débat, encore ouvert aujourd'hui, porte sur la spécification de l'erreur (Hey & Orme, 1994).

¹⁴ BIC : Bayesian Information Criterion, où $BIC = -2 \ln(L) + k \ln(N)$, avec L la vraisemblance du modèle estimée, N le nombre d'observations dans l'échantillon et k le nombre de paramètres libres du modèle.

¹⁵ AIC : Akaike information criterion, où $AIC = 2k - 2 \ln(L)$, avec k est le nombre de paramètres libres à estimer et L le maximum de la fonction de vraisemblance du modèle. Le BIC comme l'AIC permet de pénaliser les modèles en fonction du nombre de paramètres libres qu'ils intègrent afin de satisfaire le principe de parcimonie.

5.2.1. La méthode déterministe

Table 2. Gamble pairs presented to subjects

Pair no.	Less risky gamble			More risky gamble		
	P_L	P_M	P_H	P_L	P_M	P_H
1	0	.2	.8	.1	0	.9
2	0	.6	.4	.1	.4	.5
3	0	.6	.4	.2	.2	.6
4	.1	.4	.5	.3	0	.7
5	0	1.0	0	.1	.8	.1
6	0	1.0	0	.2	.6	.2
7	.3	.4	.3	.5	0	.5
8	.4	.2	.4	.5	0	.5
9	.4	.6	0	.5	.4	.1
10	.4	.6	0	.6	.2	.2
11	.5	.4	.1	.7	0	.3
12	.8	.2	0	.9	0	.1
13	.2	.2	.6	.3	0	.7
14	.6	.2	.2	.7	0	.3

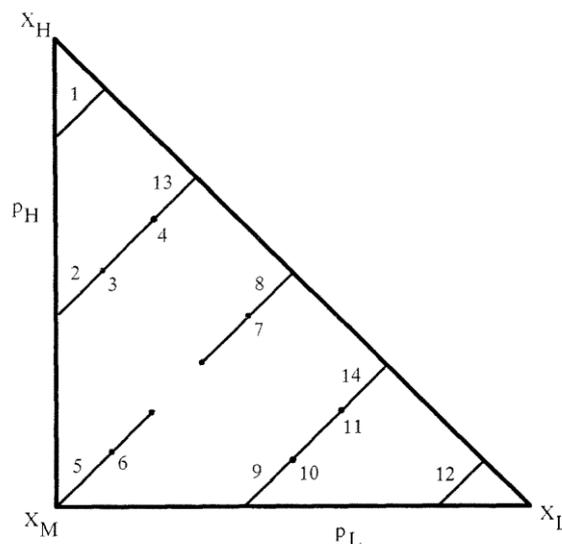


Fig. 8. Gamble pairs presented to subjects.

Figure 1.12 Représentation des loteries constituées par Camerer (1989) au sein du triangle de Marschak-Machina

L'approche déterministe consiste dans un premier temps à définir des loteries au sein du triangle de Marschak-Machina. Pour construire les courbes d'indifférence au sein du triangle, deux méthodes sont possibles. La première consiste à demander directement aux agents les loteries pour lesquels ils sont indifférents à une loterie donnée. Cette technique, bien que permettant de figurer simplement les résultats dans le triangle est en fait assez difficile à mettre en place (Kontek, 2015). L'autre technique, consiste à présenter des situations de choix entre deux loteries pour lesquelles l'agent devra simplement exprimer sa préférence. La plupart du temps, ces loteries fournissent la même espérance mais sont de variances différentes (mean preserving spread approach). La Figure 1.12 est une illustration provenant de Camerer (1989) de cette seconde solution. Cette seconde solution ne permet pas de former des courbes d'indifférence. Tout ce que l'expérimentateur peut faire est de tester les hypothèses qu'induisent les différents modèles vis-à-vis des courbes d'indifférence dans le triangle dans différentes régions du triangle (Kontek, 2015).

L'étude de Camerer (1989) a montré la supériorité du modèle RDU par rapport au modèle EU avec toutefois un certain nombre de nuances. Au milieu du triangle, les comportements sont en accord avec l'EUT (droites iso-espérances parallèles). Aux limites du triangle, par contre le modèle RDU capture mieux les décisions que le modèle EU (voir aussi Harless & Camerer, 1994 ; Kontek, 2015). De plus, il faut noter que le modèle RDU ne capture pas bien les résultats dans le domaine des pertes (nous verrons pourquoi dans le prochain chapitre) et lorsque l'enjeu des loteries est faible.

5.2.2. L'approche stochastique

Aujourd'hui la recherche privilégie ce genre d'approche. Concernant la spécification de l'erreur, c'est-à-dire la structure stochastique qui sous-tend les observations empiriques, deux approches sont possibles. Soit on considère que les erreurs proviennent essentiellement de l'inattention de l'agent (hypothèse de la main tremblante), soit on considère qu'il ne connaît pas lui-même précisément ses préférences.

Dans le premier cas, la solution la plus naturelle consiste à ajouter aux modèles déterministes de décision un terme d'erreur, en l'occurrence un bruit blanc, normalement distribué centré sur 0 et qui peut être estimé grâce à la méthode du maximum de vraisemblance. Le second cas, consiste à implanter une structure stochastique directement sur les fonctions d'utilité ou/ les fonctions de pondération des probabilités de chaque individu (Loomes & Sugden, 1995). On parle alors de Random Preference Model que Loomes et Sugden résument ainsi : « The random preference approach proposes that, for each individual, there is a set of alternative preference relations; facing any particular decision problem, the individual acts on one of these preference relations, selected at random" (Loomes, Moffatt, & Sugden, 2002, p.104). Il existe depuis peu une nouvelle voie concernant l'explication stochastique des choix. Jusqu'à maintenant, les modèles de décisions sont toujours déterministes par nature. Les choix sont déterminés par un

ordre de préférence ou par un état de préférence existant juste avant le choix (Busemeyer & Rieskamp, 2014). Les variations stochastiques sont alors le produit soit d'erreur dans l'exécution des préférences (approche de l'erreur constante – main tremblante), soit d'erreurs dans la formation des préférences (l'erreur intervient au moment du calcul de la valeur subjective des alternatives, appelé Fechner model), soit au niveau des préférences elles-mêmes (Loomes et al., 2002; Loomes & Sugden, 1995). Quoiqu'il arrive, ces différentes structures stochastiques sont résolues avant la réalisation du choix. La nouvelle voie que nous évoquions au-dessus consiste à supposer que le choix est probabiliste par nature car il n'y a pas d'ordre de préférences déterminé avant le choix. Ou plutôt, il peut exister simultanément différents ordres de préférence en état « superposé » au moment du choix. Cette nouvelle voie empreinte les outils de la mécanique quantique pour résoudre de nombreux paradoxes et inconsistances de choix en économie de la décision (Busemeyer, Wang, & Townsend, 2006; V. I. Yukalov & Sornette, 2009; V. Yukalov & Sornette, 2017; Vyacheslav I. Yukalov, Yukalova, & Sornette, 2018). On l'appelle la théorie quantique de la décision.

On pourra également différencier les études qui estiment les paramètres par « Pooling » (Harless & Camerer, 1994) des études qui procèdent à l'estimation des paramètres des modèles de décision pour chaque individu (Hey & Orme, 1994). Chacune de ces analyses a ses avantages et inconvénients. La première technique consiste à faire « comme si » nous n'avions qu'un seul sujet : l'agent représentatif. On travaille ici sur des données agrégées. Cette procédure est avantageuse lorsque les données sont peu nombreuses et particulièrement bruitées, puisque cela permet de réduire le bruit (Nilsson, Rieskamp, & Wagenmakers, 2011). Toutefois, la principale critique vis-à-vis de la première approche est qu'elle ne permet pas de prendre en compte l'hétérogénéité des attitudes face au risque des individus. Pousser le raisonnement à l'extrême, cette approche pourrait même conduire à de mauvaises conclusions. Par exemple si dans un

échantillon, la moitié des individus sont averses au risque et l'autre moitié chercheur de risque, on conclura que l'agent représentatif est neutre face au risque.

La seconde technique qui consiste à estimer les paramètres libres pour chaque individu est avantageuse pour appréhender l'hétérogénéité d'un échantillon. Toutefois, l'estimation risque d'être moins précise dans la mesure où peu de données par individu sont en général disponibles. Aujourd'hui des approches Bayésiennes Hierarchiques sont de plus en plus utilisées car elles permettent un compromis entre ces deux techniques (Nilsson et al, 2011).

Les différentes études qui ont adopté l'approche stochastique ont également conclu à la supériorité de la RDUT sur l'EUT (Carbone & Hey, 1994; Harless & Camerer, 1994; Hey & Orme, 1994).

Pour finir on notera toutefois que même si la RDUT semble montrer de meilleures capacités descriptives que l'EUT, l'axiome de comonotonicité est violé comme le montre l'expérience menée par Wakker, Erev et Weber (1994).

CONCLUSION DE LA PARTIE 1

Nous avons présenté dans cette partie 1 la théorie standard de l'utilité espérée basé sur l'axiomatique proposé par Von Neumann et Morgenstern (1944). Après avoir évoqué ses limites dans sa capacité à accommoder les résultats empiriques, notamment à travers le paradoxe d'Allais, nous avons vu que la Rank-dependent utility theory était en mesure de combler cette lacune. La RDU est elle-même une généralisation de l'EUT au prix du relâchement de l'axiome d'indépendance, remplacé par l'axiome de comonotonicité.

Nous verrons lors du second chapitre que la RDUT n'est toutefois pas en mesure d'expliquer les renversements de préférences des agents lorsque ces derniers sont confrontés à des changements de cadrage (framing effect). Comme nous l'exposerons, la Cumulative Prospect Theory développé par Kahneman et Tversky apporte une solution en supposant une fonction

d'utilité en deux parties portant non pas sur les richesses finales prospectives (hypothèse d'intégration) mais sur les variations de richesses par rapport à un référentiel.

ANNEXE

A. Représentation schématique de l'EUT

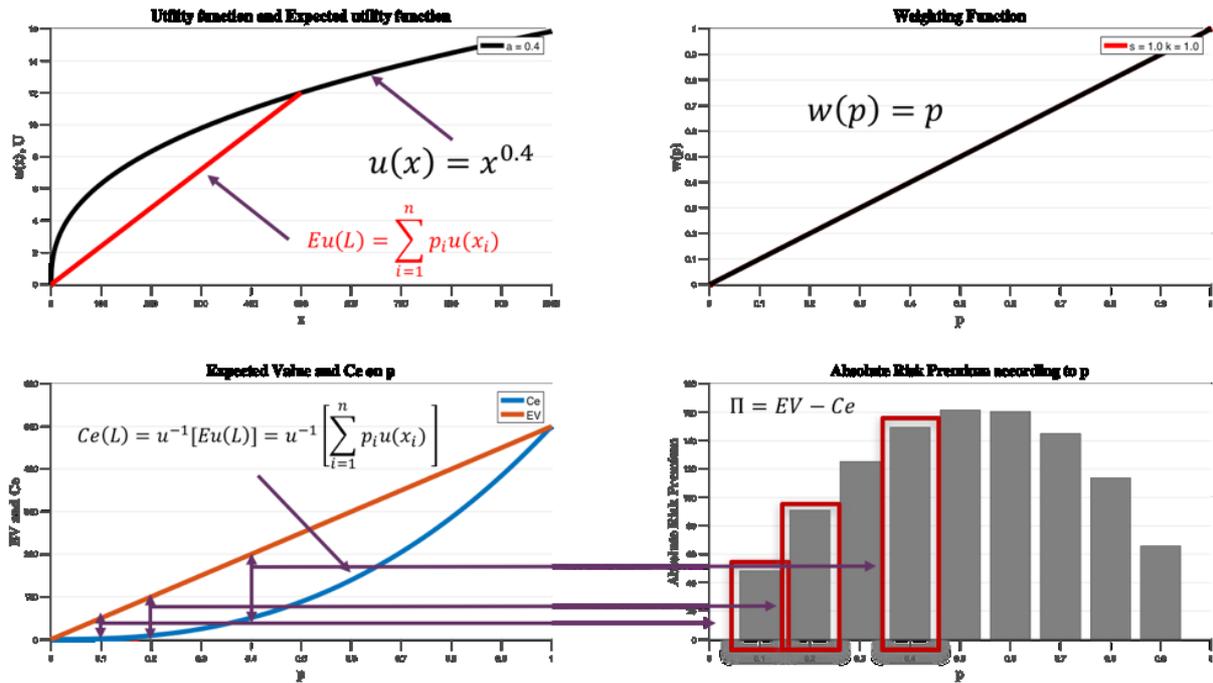


Figure A.1. Représentation résumant le modèle d'espérance d'utilité. On suppose une loterie à deux issues possibles. En haut à gauche : la courbe en noir représente la fonction d'utilité puissance : $u(x) = x^\alpha$ et la droite en rouge représente la fonction d'espérance d'utilité. En haut à droite : la droite en noir représente la fonction des probabilités. En bas à gauche : La droite en rouge représente la fonction d'espérance

B. Fonction de pondération des probabilités

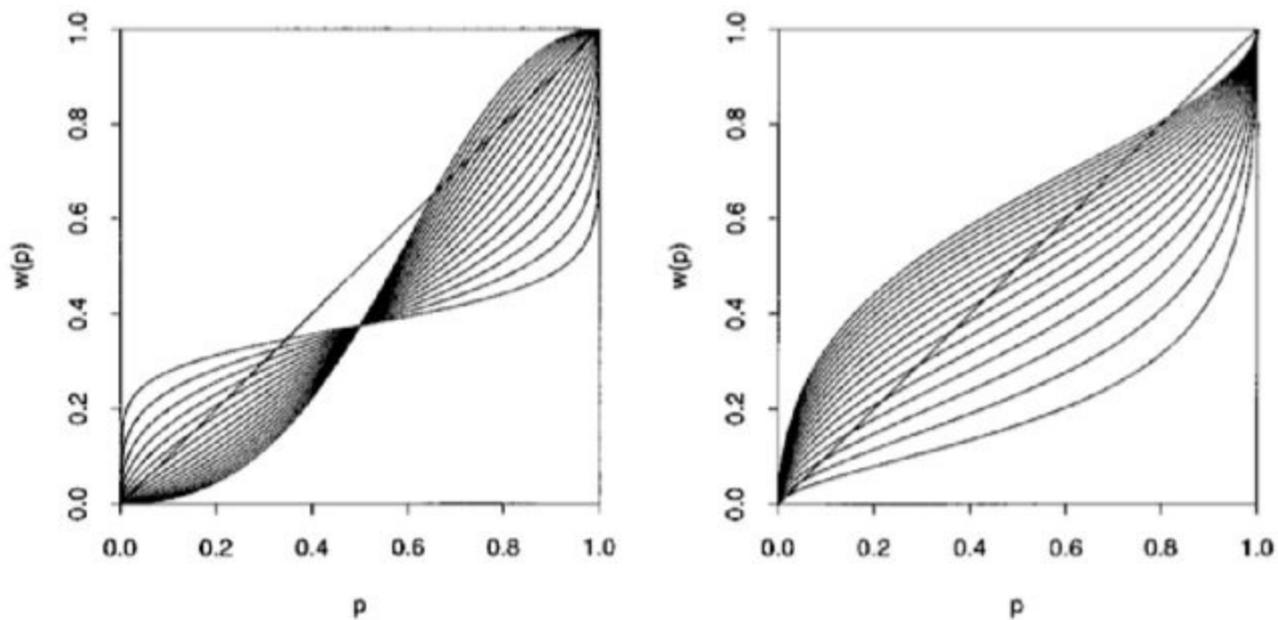


Figure A.2. Représentation de la fonction de pondération des probabilités de Wu et Gonzalez (1999). Effet de γ sur la courbure et δ sur l'élévation de la fonction de pondération de probabilité. La figure de gauche fixe $\delta = 0,6$ et varie $0,2 < \gamma < 1,8$. La figure de droite fixe $\gamma = 0,6$ et varie $0,2 < \delta < 1,8$. Tiré de Gonzalez et Wu (1999).

2. LES MODELES DE PREFERENCES REFERENCE- DEPENDANT

Introduction

Comme nous l'avons exposé dans la section précédente, la théorie de l'utilité espérée a dominé l'analyse de la prise de décision dans le risque durant des décennies. Elle a été largement acceptée par la communauté scientifique comme le modèle normatif du choix rationnel (Kahneman & Tversky, 1979). Selon l'EUT, seule la forme de la fonction d'utilité détermine l'attitude face au risque (la linéarité implique la neutralité face au risque, la concavité implique l'aversion au risque et la convexité implique la recherche de risque). Une autre approche normative, la dual theory, formalisée par Yaari (1987) établit que l'attitude face au risque dépend de la déformation des probabilités cumulées par les agents. Cette déformation est induite par une fonction de pondération des probabilités, la weighting function, qui, lorsqu'elle est convexe implique l'aversion au risque et, lorsqu'elle concave, implique le goût du risque. Les recherches empiriques ont démontré depuis que ces deux modèles manquaient de précision pour décrire l'attitude face au risque. Aujourd'hui, une grande part de la communauté scientifique semble admettre que l'attitude face au risque dépend à la fois de la sensibilité aux issues financières (capturée par la fonction d'utilité) et de la déformation des probabilités cumulées (capturée par la weighting function) comme le propose la théorie de l'utilité espérée dépendante des rangs (Abdellaoui, l'Haridon, & Zank, 2010).

Toutefois, les résultats expérimentaux introduits par Kahneman et Tversky (noté K&T) dans leur article séminal de 1979, ont profondément bousculé la recherche en économie de l'incertitude. Leurs expériences ont en particulier montré que les choix des individus dépendent de ce qu'ils appellent un « point de référence » ou dit plus trivialement d'un « point de comparaison ». Dans cette perspective, K&T ont observé qu'un agent pouvait révéler des attitudes face au risque opposés pour des situations pourtant strictement identiques en termes de conséquences finales prospectives.

L'exemple sans doute le plus connu concerne l'expérience de la « maladie » (Tversky & Kahneman, 1981) : Les expérimentateurs ont présenté à deux groupes de participants différents un scénario spécifique. Pour le premier groupe, on demandait aux participants d'imaginer qu'ils étaient médecin et qu'il faisait face à une maladie faisant des ravages dans leur ville. Deux vaccins étaient disponibles. Le premier sauve 200 personnes avec certitude et le deuxième en sauve 600 avec une chance sur trois ou aucune avec deux chances sur trois. Les participants ont préféré la première option. Pour le second groupe, la même épidémie est présentée puis on leur propose le choix suivant entre deux vaccins : Avec le premier vaccin, 400 personnes mourront avec certitude. Avec le second, il y a une chance sur 3 que personnes ne meurent et deux chances sur trois que 600 personnes meurent. Cette fois les participants ont majoritairement choisit l'option risquée. Pourtant, en valeur absolue, le nombre de survivants et de décès est identique dans les deux situations. Dès lors, les auteurs concluent que la manière de cadrer une même situation influence les comportements des individus car ils ne traitent pas de la même manière les pertes et les gains. Or la notion de perte et de gain est relatif à un référentiel qui, s'il est manipulé, modifie la perception de ce qui relève des pertes et des gains dans une situation et donc modifie le comportement face au risque des individus.

L'étude de ce référentiel sur le comportement a donné lieu à la révélation de différents effets désormais connues même du grand public comme : l'effet de dotation (Kahneman, Knetsch, & Thaler, 1991), l'effet d'ancrage (Tversky & Kahneman, 1974), l'effet de cadrage (Kahneman & Tversky, 1979) ou encore l'effet de disposition (Weber & Camerer, 1998) et trouve des applications variées dans des domaines tels que la psychologie, le marketing, le management, et la finance. Ces expériences ont alors mené à l'introduction d'un nouvel élément non-normatif dans les modèles de décisions : le point de référence. Ce dernier est défini comme étant un point subjectivement neutre divisant l'échelle de perception de l'intensité d'un stimulus en 2 parties :

une partie « négative » (en dessous du point de référence) une partie « positive » (au-dessus du point de référence).

Dans ce chapitre, après avoir défini plus exactement le point de référence et son impact sur l'attitude face au risque, nous verrons que sa spécification est encore aujourd'hui un mystère substantiel. Les interrogations portent sur deux points. La première concerne sa localisation. En effet dans une situation de choix risqué, le référentiel pourrait être par exemple, la richesse actuelle de l'individu, l'espérance mathématique, l'équivalent certain etc. La seconde problématique concerne son actualisation dans des contextes de choix dynamique. En particulier, est-ce que le référentiel s'adapte parfaitement à chaque nouvelle situation ou non ? La recherche se focalise de plus en plus sur ces questions car sans spécification précise du référentiel, on concède alors un degré de liberté trop important aux modèles « Référence-dépendant ». Autrement dit, ces modèles sont en mesure de capturer un grand nombre de phénomènes *a posteriori* car « il suffit » de sélectionner le référentiel permettant de capturer les résultats obtenus. Cette perspective *ad hoc* n'est bien évidemment pas satisfaisante pour une science qui souhaite prédire les phénomènes. En effet la multiplicité des référentiels potentiels implique que ces modèles peuvent produire *ex ante* des prédictions contradictoires en fonction du référentiel choisi *a priori*.

On pourrait penser que la recherche dans ce domaine consiste donc à éliminer progressivement les points de références potentielles pour finalement ne sélectionner que celui ou ceux qui semblent à l'œuvre dans les prises de décision. Pourtant ce n'est pas le cas. De nombreuses recherches consistent à étudier un point de référence en observant si les résultats comportementaux correspondent aux prédictions établies conditionnellement à ce point de référence. Le problème réside dans le fait que ces études ne s'assurent pas que d'autres points de références puissent expliquer leurs résultats¹⁶. Dès lors, la littérature sur le sujet ne permet

¹⁶ Voir l'étude de Baillon, Bleichrodt et Spinu (2015) qui montrent que les résultats obtenus par Abeler, Falk, Goette et Huffman (2007) sont consistants avec leur spécification du point de référence (ici il s'agit

pas de faire émerger des référentiels « vainqueurs ». On assiste plutôt à une surenchère de théories sans qu'elles soient véritablement comparées entre elles.

Toutefois, nous comprendrons dans ce chapitre que l'entreprise qui consisterait à comparer toutes ces différentes spécifications du référentiel est particulièrement ardue. En l'occurrence nous n'évoquerons qu'une seule étude ayant tenté d'accomplir cela : l'étude de Baillon et al (2017) qui n'est d'ailleurs toujours pas publiée depuis sa première version en working paper qui date de 2015.

Toutefois, l'aspect intuitif derrière la notion de référentiel lui accorde un attrait important. Avant de procéder à son analyse, nous prodiguons deux exemples qui permettent d'en capturer la nature.

Prenons l'exemple d'un phénomène d'hystérèse dont nous avons tous fait l'expérience un jour : éteindre le chauffage quand nous avons suffisamment chaud et allumer le chauffage lorsque nous avons froid. On pourrait supposer à priori que nous l'allumons ou l'éteignons lorsque la température ambiante est en-dessous ou au-dessus d'un certain seuil critique. Ce seuil pourrait correspondre à notre point de référence. Au-dessus de ce seuil, j'éteins le chauffage car je considère qu'il fait trop chaud. En dessous, je juge la température comme étant trop froide alors j'allume le chauffage. Pourtant, en fonction de l'ordre dans lequel nous vivons les variations de température (chaud vers froid ou froid vers chaud) nous n'allons pas allumer ou éteindre le chauffage à un même niveau de température. Il semble donc dans cet exemple que le ressenti subjectif d'un même stimulus (ici la température) n'est pas perçu de la même manière par un même individu en fonction de son histoire d'exposition aux différents degrés de ce même stimulus. Par exemple, la température 20° peut être perçue comme froide si l'individu a eu

d'une spécification tirée de Koszegi et Rabin, 2006) mais sont également cohérent d'après une spécification qui n'est pas envisagé par les auteurs de l'étude : la règle MaxMin. Ces référentiels seront abordés dans ce chapitre.

l'habitude de vivre sous 40°, ou bien chaude s'il a eu l'habitude de vivre sous -10°. Nous voyons dans cet exemple que nos choix sont sensibles à notre histoire. Plus exactement, on peut donc supposer l'existence d'un point de référence qui varie d'une certaine façon et qui en fonction de sa localisation nous fait percevoir un même stimulus de façon différente.

Dans le domaine de l'économie, le paradoxe d'Easterlin (2005) est également une excellente illustration de l'impact du référentiel. Ce dernier observe qu'une hausse du PIB ne conduit pas forcément à une hausse du niveau de bien-être ressenti : le paradoxe ici étant que « plus » ne signifie pas « mieux ». Selon lui, le niveau de bien être d'un individu est en fait lié relativement à celle de la richesse des membres les plus riches de la société dans laquelle il vit. Dès lors, les individus les plus fortunés sont les plus heureux non pas du fait de leur richesse, mais plutôt parce que leurs richesses s'accroissent économiquement plus vite que le reste de la population. Dans la même perspective, cela signifie qu'un individu d'une société dont l'augmentation de la fortune se situe dans la moyenne ne se déclarera pas plus heureux. C'est donc le revenu relatif (c'est-à-dire en comparaison avec le revenu moyen du reste de la société), et non sa progression brute, qu'il convient de prendre en compte. C'est bien cette relativité qui conduit à ce paradoxe.

A priori, il ne s'agit donc pas d'une incroyable découverte étant donné sa relative évidence dans notre vie quotidienne. D'ailleurs, précisons que la paternité de l'idée peut en être attribuée entre autres au psychologue Helson qui la formula 20 ans plus tôt que K&T, en 1947 sous le nom de « niveau d'adaptation » ou encore à l'économiste Markowitz en 1952 sous le nom de « customary wealth ». En revanche les difficultés résident dans la capture de ce point de référence, les facteurs d'influence de sa constitution, son éventuelle évolution au cours du temps ou encore dans ses implications sur les choix en univers risqué.

Ce chapitre est consacré au cadre conceptuel des modèles de préférences référence-dépendant. Dans un premier temps nous présenterons le paradoxe de K&T qui a permis la mise en évidence de ce référentiel, puis nous exposerons plus en détail la théorie des auteurs qui en découla : la Prospect Theory (PT). Dans un second temps, nous étudierons les implications sur la décision en univers risqué d'un point de référence. Dans un troisième temps nous reporterons les différents « candidats » au poste de référentiel dans les contextes de choix statiques. Enfin nous terminerons par une analyse approfondie des référentiels dans les contextes de choix dynamique et nous aborderons entre autres les tentatives de captures de l'actualisation du référentiel sur plusieurs périodes.

I. LA PROSPECT THEORY

Dans cette section nous exposerons dans un premier temps les expériences de K&T qui ont permis de révéler la potentielle existence d'un référentiel. Sur la base de ces expériences nous présenterons en détail la fonction d'utilité qu'ils proposèrent : la value function. Enfin, nous reviendrons sur le traitement non linéaire des probabilités mais cette fois vue par la prospect theory. Pour finir, nous traiterons la seconde version de la Prospect Theory qui intègre la « value function », la fonction de pondération des probabilités et un référentiel.

1.1. Les études de Kahneman et Tversky

1.1.1. Hypothèse d'intégration vs Hypothèse d'isolation

Nous avons vu au cours du premier chapitre que l'évaluation des loteries portait sur l'espérance d'utilité des richesses finales prospectives. Cette hypothèse, dite d'intégration, stipule que les individus incrémentent les issues des loteries à leurs richesses initiales : $W_i = w_0 + x_i$. Dès lors selon cette hypothèse, la loterie $(x_1, p_1; \dots; x_n, p_n)$ est acceptable à la position de richesse actuelle w_a , si et seulement si $u(w_0 + x_1, p_1; \dots; w_0 + x_n, p_n) > u(w_0)$. K&T ont remis en question cette hypothèse à travers le paradoxe suivant :

Kahneman et Tversky (1979) exposèrent deux situations fictives à un ensemble de participants (Figure 2.1). Dans la première situation, K&T proposèrent la consigne suivante au premier groupe de sujets : « En plus de ce que vous avez déjà, ici 1000€, choisissez entre la loterie $A \equiv (x_1 = 1000€, p = 0.5; x_2 = 0€, 1 - p = 0.5)$ et la loterie $B \equiv (x_3 = 500€; p = 1)$ ». Dans la seconde situation, la consigne donnée au second groupe était la même sauf que les individus possédaient 2000€ initialement au lieu des 1000€ de la première situation. Puis on leur demandait de choisir entre $A' \equiv (x'_1 = -1000€, p = 0.5; x'_2 = 0€, 1 - p = 0.5)$ et $B' \equiv (x'_3 = -500€; p = 1)$.

En se référant à la Figure 2.1 on remarque facilement ici que les situation 1 et 2 sont identiques : la situation 2 étant obtenue à partir de la situation 1 en ajoutant 1000€ à la richesse initiale et en soustrayant 1000€ à l'ensemble des issues des loteries. Les deux situations sont donc strictement identiques en termes de richesse finale $A \equiv (2000, 0.5 ; 1000, 0.5) \equiv A'$ et $B \equiv (1500, 1) \equiv B'$. Dès lors, dans la mesure où les richesses finales étaient identiques et selon l'hypothèse d'intégration, les individus choisissant A (B) dans la situation 1 sont riscophiles (riscophobes) devraient choisir A' (B') dans la situation 2. Ici les groupes étaient indépendants, donc l'hypothèse consiste à supposer que si les deux groupes ne sont pas significativement différents initialement en termes d'attitude face au risque, alors nous ne devrions pas observer de différence de choix entre les deux groupes. Pourtant les résultats montrent que 84% des sujets du groupe 1 choisissent B (aversion au risque) et 70% du groupe 2 choisissent A' (recherche de risque).

La différence entre ces deux situations se situe au niveau du cadrage de la situation (framing effect), ce qui semble avoir déterminé les réponses des agents. En effet, bien que les richesses finales prospectives soient identiques, la première situation est présentée sous l'angle de gains alors que la seconde situation est présentée sous l'angle des pertes.

Il ressort de cette expérience deux éléments : 1. Il semble que les individus traitent différemment les pertes et les gains, et c'est de ce traitement différent que l'on peut déduire que les individus n'intègrent pas les issues des loteries à la richesse initiale. 2. Il semble plutôt que les individus isolent les résultats des loteries de leur richesse actuelle (appelé hypothèse d'isolation) : ici, donc, les issues sont évaluées par rapport au point de référence w_a ou w'_a .

La première observation est plus spécifiquement corroborée par les expériences portant sur ce que K&T appellent le « reflection effect » comme nous le verrons dans la section suivante. La seconde observation fait dire à K&T que les individus n'évaluent non pas l'utilité d'une position finale procurée par une alternative (pas d'intégration des issues à la richesse initiale), mais

plutôt l'utilité dérivée de la variation de richesse par rapport à un point de référence. Nous reviendrons plus précisément sur cette assertion dans la section 2 de ce chapitre.

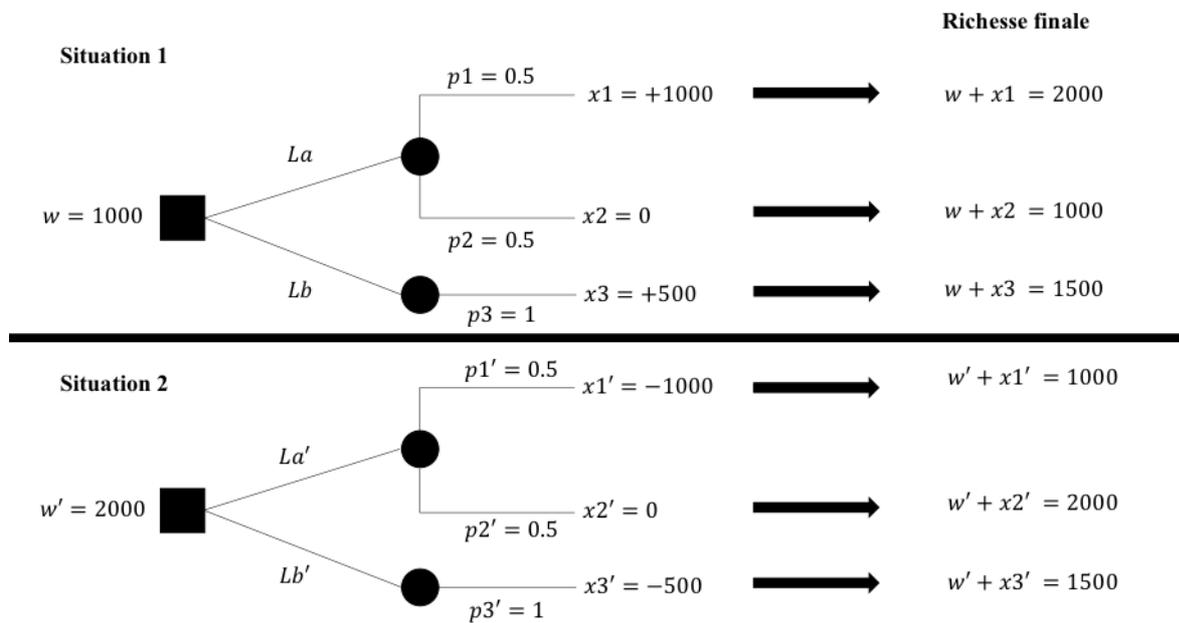


Figure 2.1 Représentation de l'expérience de K&T testant l'hypothèse d'intégration et d'isolation

1.1.2. Le Reflection Effect

Pour mettre en évidence le traitement différent appliqué par les individus entre les gains et les pertes, K&T exposèrent des sujets à des variantes du « common ratio effect » de Allais dans lesquelles les issues des loteries étaient soit entièrement positives (gains) soit toutes négatives (pertes). Il ressort de cette expérience que les individus se montraient averse au risque dans le domaine des gains lorsque la probabilité de gagner était élevée et chercheur de risque lorsque la probabilité de gagner était faible, ce qui est en accord avec le paradoxe d'Allais. En revanche, en situation de perte, l'attitude face au risque s'inversait lorsque l'on considère le domaine des pertes : lorsque la probabilité de perdre était élevée les individus se montrèrent chercheur de risque alors que lorsque la probabilité de perdre était faible les individus se montraient averse au risque. Ces résultats concernant l'attitude face au risque des individus sont regroupés sous le terme de « fourfold pattern ».

Cette expérience est particulièrement instructive car elle démontre, comme le paradoxe d'Allais, le traitement non linéaire des probabilités opéré par les agents. Cependant, elle montre également que le traitement des pertes et des gains n'est pas identique. En fait, pour expliquer ce renversement des préférences, K&T stipule que la fonction d'utilité serait concave dans les gains et convexe dans les pertes. C'est donc la combinaison des fonctions de pondérations des probabilités en S inversé et de la fonction d'utilité en S qui explique le fourfold pattern.

1.1.3. Concavité et Convexité de la fonction d'utilité

Pour justifier de la concavité de la fonction d'utilité dans les gains et de sa convexité dans les pertes, K&T font appel à un argument provenant de la psychophysique et de sa loi la plus célèbre : la loi de Weber-Fechner (1860). Cette loi stipule que l'intensité des réactions physiologiques augmentent de moins en moins rapidement (mais reste croissante) avec l'accroissement d'intensité d'un stimulus : « la sensation varie comme le logarithme de l'excitation », d'où la loi : $I = k \log(s)$, avec I l'intensité de la sensation, k une constante et s l'intensité du stimulus. Dès lors, une implication de cette loi est qu'il est plus aisé de discriminer une petite variation d'intensité Δs lorsque l'intensité initiale n'est pas trop importante que de discriminer cette même variation lorsque l'intensité initiale est forte. Ce principe appliqué à la richesse revient au principe énoncé à l'origine, intuitivement, par Bernoulli : l'accroissement d'utilité obtenue à partir d'un gain de 100€ lorsque l'on passe d'une richesse de 100€ à 200€ est plus grand que lorsque l'on passe d'une richesse de 100100€ à 100200€. La même considération peut être effectuée avec les pertes : une perte de 100€ est plus douloureuse lorsque l'on passe de -100 à -200€ alors que cette même perte est moins douloureuse lorsque l'on passe d'une richesse de -100100€ à -100200€. Ainsi, les individus sont de moins en moins sensibles aux variations de richesse à mesure que les pertes ou les gains augmentent. Nous reviendrons sur les implications sur l'attitude face au risque de cette fonction d'utilité dans la section II de ce chapitre. Pour tester cette idée, K&T présentèrent les 2 situations suivantes :

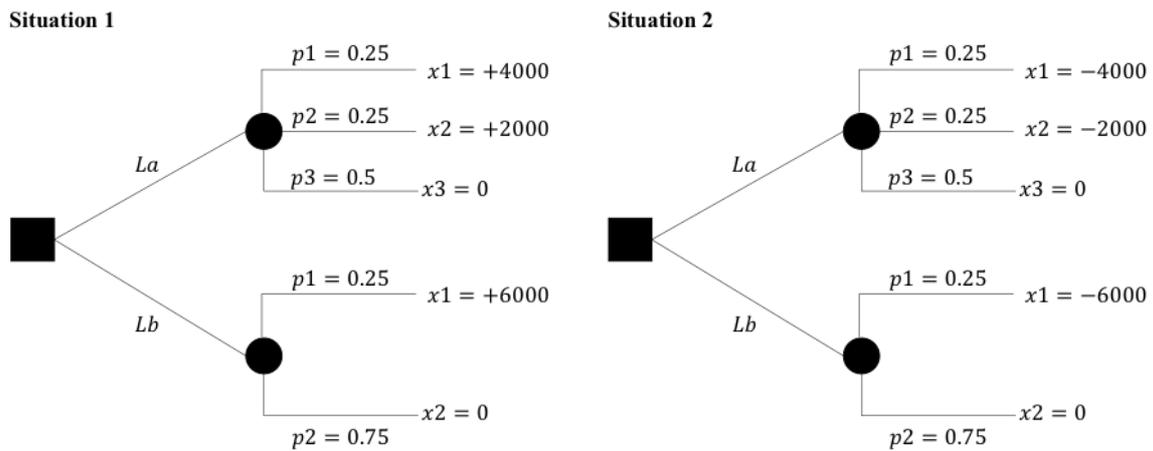


Figure 2.2 Représentation de l'expérience de K&T testant la concavité et la convexité de la fonction d'utilité.

Dans cette expérience, 82% des sujets préférèrent A dans la situation 1 alors qu'ils préférèrent la loterie B dans la situation 2.

Une fonctionnelle appropriée doit donc rendre compte de ces préférences. En supposant que les individus pondèrent les probabilités selon la fonction $w(p)$ et la richesse selon une fonction $v(x)$, alors cette fonctionnelle doit produire :

$$w(0.25)u(6000) < w(0.25)[v(4000) + v(2000)],$$

$$w(0.25)u(-6000) > w(0.25)[v(-4000) + v(-2000)]$$

Ainsi, $v(6000) < v(4000) + v(2000)$ ce qui correspond à l'implication du fonction concave et $v(-6000) > v(-4000) + v(-2000)$, ce qui correspond à l'implication d'une fonction convexe.

1.1.4. L'aversion aux pertes

L'aversion aux pertes fait référence au traitement asymétrique des pertes et des gains par rapport à un point de référence. K&T observent en effet qu'une perte à plus de poids qu'un gain de même amplitude. Autrement dit, le déplaisir lié à une perte est plus grand que le plaisir dérivé d'un gain de même amplitude. Les auteurs observent ce phénomène à travers des loteries

mixtes, c'est-à-dire des loteries impliquant des pertes et des gains (par opposition aux loteries dites « régulières » impliquant soit uniquement des pertes soit uniquement des gains, avec dans les deux cas 0 compris). Selon K&T, cette tendance à surpondérer les pertes par rapport au gain permet d'expliquer pourquoi les individus rejettent des loteries équitables même lorsque le montant des issues est faible ou modeste. Formellement, la fonction d'utilité devrait exhiber les propriétés suivantes : $\forall x > y > 0, v(y) + v(-y) > v(x) + v(-x)$ et $v(-y) - v(-x) > v(x) - v(y)$. Comme on peut le voir sur la Figure 2.3, l'aversion aux pertes est représentée par un kink dans la région des pertes et proche du point de référence. La pente de la fonction $v(x)$ est plus pentue, plus raide, dans les pertes que dans les gains.

1.2. La Cumulative Prospect Theory

1.2.1. La value function

La fonction d'utilité de type Prospect Theory (ou « value function ») proposée par Kahneman et Tversky est 1. Définie sur les déviations de richesses par rapport à un point de référence R ; 2. Concave dans les gains et convexe dans les pertes ; 3. Plus pentu dans les pertes que dans les gains. La fonction d'utilité proposée par les auteurs s'écrit selon une spécification en deux morceaux :

$$v(x) = \begin{cases} x^\alpha & \text{si } x \geq 0 \\ -\lambda(-x)^\beta & \text{si } x < 0 \end{cases} \quad (2.1)$$

Avec x représentant le montant financier d'une issue engagée dans une loterie. $\alpha < 1$ et $\beta < 1$ sont des paramètres libres et définissent la forme la fonction d'utilité dans les pertes (convexe) et dans les gains (concave), $\lambda > 1$ est le facteur d'aversion aux pertes.

Le fait que x soit l'argument de la fonction d'utilité doit faire ici l'objet d'une précision importante. Nous avons précisé que selon la Prospect Theory, les individus dérivent l'utilité

d'une variation de richesse. Mais nous n'avons pas spécifié les éléments constituant cette variation. En fait, une spécification possible consiste à considérer que l'individu dérive l'utilité d'une variation entre la richesse finale, $(w_0 + x)$, et le point de référence. On notera v_+ la partie de la fonction d'utilité associé aux gains et v_- la partie de la fonction d'utilité associé aux pertes. D'où :

$$v(z) = \begin{cases} v_+(z) & \text{si } z \geq 0 \\ -\lambda v_-(|z|) & \text{si } z < 0 \end{cases} \quad (2.2)$$

Où z représente une variation de richesse entre la richesse finale et le point de référence, tel que, $z = (w_0 + x) - R = W - R$. On notera que si la richesse initiale est égale au point de référence, $w_0 = R$, alors z est égal aux montants des issues financières de la loterie : $z = x$. Autrement dit, dans ce cas, l'individu ne dérive l'utilité qu'à partir des montants financiers engagés dans les loteries en présences. Nous voyons ici que l'argument x est donc simplement un cas particulier de la Prospect Theory. Par ailleurs, au niveau du point de référence où sa valeur est par définition 0, on spécifie que $v'_-(0) = v'_+(0) = v'(0)$.

Intuitivement, on comprend selon la spécification donnée à la fonction d'utilité que les individus sont davantage sensibles à des variations de richesses proches du point de référence qu'à des variations « loin » du référentiel, ce qui est une implication directe de la concavité dans les gains et de la convexité dans les pertes. Autrement dit la diminution marginale de l'utilité (ou de la désutilité) est appliqué au domaine des gains comme au domaine des pertes.

1.2.2. La Weighting function

La première version de la Prospect Theory appliquait la fonction de pondération aux probabilités et non aux probabilités cumulées. Les critiques portant sur le non-respect de la dominance stochastique d'ordre 1 induit par cette formalisation, poussa K&T à adopter une

seconde version de la Prospect Theory appelé la Cumulative Prospect Theory intégrant le formalisme Quiggin (1982) dans lequel les probabilités cumulées sont transformées en poids décisionnels. Toutefois, dans la logique du traitement différent opéré par les agents entre les pertes et les gains (reflection effect), K&T montrent l'existence de deux fonctions de pondération : une pour les pertes, notée $\phi^-(p)$ et une pour les gains, notée $\phi^+(p)$ (voir Figure 2.3). Leurs formes sont similaires mais diffèrent dans leurs courbures et dans leurs élévations. Les résultats montrent que la courbure de la fonction est moins importante dans les pertes que dans les gains et que l'élévation est plus élevée pour les pertes que pour les gains.

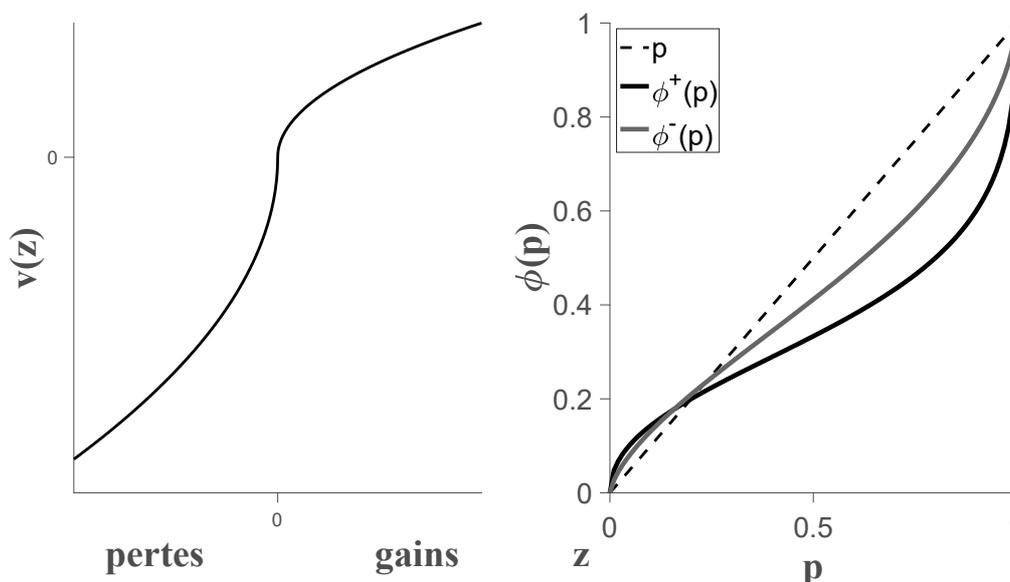


Figure 2.3 A gauche : représentation de la fonction d'utilité en forme de S selon K&T. A droite : représentation de la fonction de pondération des probabilités dans les pertes $\phi^-(p)$ et dans les gains $\phi^+(p)$.

1.2.3. Le modèle de la CPT

Supposons que les issues financières d'une loterie L soient ordonnées : $x_{-m} < \dots < x_{-1} < x_0 < x_1 < \dots < x_n$, lesquelles sont attachées aux probabilités p_{-m}, \dots, p_n respectivement. Dans le modèle CPT, l'utilité d'une loterie dépend d'une fonction d'utilité définie sur les pertes et gains et d'une fonction de pondération des probabilités. Toutefois la fonction de pondération des probabilités est différente dans les pertes et dans les gains, que l'on notera ϕ^+ et ϕ^-

respectivement. Ces fonctions sont strictement croissantes sur l'intervalle $[0,1]$ avec $\varphi^\pm(0) = 0$ et $\varphi^\pm(1) = 1$. L'évaluation d'une loterie L dont les probabilités des résultats sont connues s'écrit en deux parties, l'une pour les gains et l'autre pour les pertes :

$$CPT(L) = CPT(L^-) + CPT(L^+) \quad (2.3)$$

Avec L^+ désigne l'ensemble des issues financières de la loterie L supérieures ou égales au point de référence (gains) et L^- désigne l'ensemble des issues financières inférieures au point de référence (pertes). Le premier terme de l'Equation (2.3) évalue la désutilité des pertes et le second terme l'utilité des gains puis ces deux valeurs sont additionnées. Sous CPT, la fonctionnelle d'une loterie s'écrit :

$$CPT(L) = \sum_{i=-m}^{-1} \Phi_i^- v(z_i) + \sum_{i=0}^n \Phi_i^+ v(z_i) \quad (2.4)$$

Avec $v(z)$, la fonction d'utilité décrite par l'Equation (2.1). Mais bien entendu la fonction puissance n'est qu'une formalisation parmi d'autres. Les poids décisionnels dans les gains sont notés Φ_i^+ et Φ_i^- pour le domaine des pertes.

On spécifie $\Phi_n^+ = \varphi_+(p_n)$ et $\Phi_{-m}^- = \varphi_-(p_{-m})$. Dans le domaine des gains, Φ_i^+ est défini par $\Phi_i^+ = \varphi_+(p_i + \dots + p_n) - \varphi_+(p_{i+1} + \dots + p_n)$ pour $0 \leq i \leq n - 1$. Dans le domaine des pertes, Φ_i^- est défini par, $\Phi_i^- = \varphi_-(p_{-m} + \dots + p_i) - \varphi_-(p_{-m} + \dots + p_{i-1})$ pour $1 - m \leq i < 0$. On peut également reformuler le modèle (2.4) ainsi :

$$CPT(L^+) = \varphi_+(p_n)v_+(z_n) + \sum_{k=1}^{n-1} \left[\varphi_+ \left(\sum_{j=0}^k p_{n-j} \right) - \varphi_+ \left(\sum_{j=0}^{k-1} p_{n-j} \right) \right] v_+(z_{n-k}) \quad (2.5)$$

$$CPT(L^-) = \varphi_-(p_{-m})v_-(z_{-m}) + \sum_{k=1}^m \left[\varphi_- \left(\sum_{j=0}^k p_{-(m-j)} \right) - \varphi_- \left(\sum_{j=0}^{k-1} p_{-(m-j)} \right) \right] v_-(z_{-(m-k)}) \quad (2.6)$$

Avec $\Phi_i^+ = \varphi_+(\sum_{j=0}^{i=k} p_{n-j}) - \varphi_+(\sum_{j=0}^{i=k-1} p_{n-j})$ et $\Phi_i^- = \varphi_-(\sum_{j=0}^{i=k} p_{-(m-j)}) - \varphi_-(\sum_{j=0}^{i=k-1} p_{-(m-j)})$.

Passons à l'interprétation de ce modèle. Considérons tout d'abord le domaine des gains, Equation (2.5). Comme nous pouvons le constater, la fonction de pondération des probabilités ne s'applique pas aux probabilités individuelles comme c'est le cas dans la première version de la Prospect Theory, mais plutôt aux probabilités décumulées. Pour le comprendre, notons F_{L^+} la fonction de réparation de la loterie L . Cette fonction indique pour chaque éléments z la probabilité d'obtenir un résultat inférieur ou égal à l'élément considéré z_i . A l'inverse, dans le modèle CPT (comme dans le modèle RDUT), la fonction φ_+ s'applique à l'ensemble de probabilités des issues aux moins aussi bénéfiques, $\varphi_+(\sum_{j=0}^{i=k} p_{n-j})$, et strictement plus bénéfique, $\varphi_+(\sum_{j=0}^{i=k-1} p_{n-j})$, que l'élément considéré z_i . On considère donc la fonction décumulative $G_L = 1 - F_{L^+}$. Dès lors, le poids Φ_i^+ d'un gain z_i est défini comme la différence entre le poids d'obtenir n'importe quel gain aussi bénéfique que z_i et le poids d'obtenir n'importe quel gain strictement plus bénéfique que z_i . Autrement dit, le poids d'un gain dans une loterie dépend de différents éléments. 1. Le rang de l'issue. 2. La probabilité de l'issue ainsi que celle des autres issues (autrement dit la distribution des probabilités sur les

conséquences). 3. S'il s'agit d'une perte ou d'un gain. 4. La fonction de pondération des probabilités.

L'une des différences fondamentales avec la RDUT est le traitement des pertes (Equation (2.6) où la fonction de pondération opère cette fois sur les probabilités cumulées. En effet, le poids Φ_i^- d'une perte z_i est défini comme la différence entre le poids d'obtenir une perte au moins aussi importante, $\varphi_-(\sum_{j=0}^{i=k} p_{-(m-j)})$, et le poids d'obtenir une perte strictement plus importante, $\varphi_-(\sum_{j=0}^{i=k-1} p_{-(m-j)})$. En ce sens, la fonction de pondération porte cette fois sur la fonction cumulative des probabilités F_L^- . Dès lors, le poids d'une perte dépend des poids des autres pertes plus conséquentes qu'elle-même.

En conclusion, l'impact de chaque conséquence d'une loterie est pondéré par un poids et ce poids dépend du rang de la conséquence. Ainsi une même conséquence z_i , n'a pas le même poids selon son rang au sein d'une loterie. En particulier, avec une fonction de pondération des probabilités S-shape inversé, si z est une des conséquences extrêmes (le plus petit résultat ou le plus grand résultat d'une loterie) d'une loterie engageant $n > 3$ résultats et à distribution de probabilité uniforme alors son poids sera surpondéré alors que si z_i est une des conséquence intermédiaire son poids sera sous-pondérées.

1.2.4. Cumulative Prospect Theory et Attitude face au risque

La CPT contrairement aux autres modèles présentés est la seule à pouvoir capturer un pattern expérimentale connue aujourd'hui sous le nom de « fourfold pattern » (voir Tableau 2.1) en combinant l'impact de la weighting function de forme S-shape inversé avec une fonction d'utilité convexe dans les pertes et concave dans les gains¹⁷.

¹⁷ On notera toutefois qu'un très récent modèle développé par Baucells, Kontek et Lewandowski (2018, working paper) que nous présentons au chapitre 4, est également en mesure de capturer le fourfold pattern.

Tableau 2.1. Le fourfold pattern

	Gains	Pertes
Grande Probabilité	<p>Aversion au risque</p> <p>Sous-pondère grande probabilité de gagner. $\varphi(p_{high})$ convexe + $v_+(z)$ concave.</p>	<p>Recherche de risque</p> <p>Sous-pondère la grande probabilité de perdre. $\varphi(p_{high})$ convexe + $v_-(-z)$ convexe. La convexité de $v_-(-z)$ contrecarre la convexité de $\varphi(p)$.</p>
Petite Probabilité	<p>Recherche de risque</p> <p>Surpondère la petite probabilité de gain. $\varphi(p_{small})$ concave + $v_+(z)$ concave. La concavité de $\varphi(p)$ contrecarre¹⁸ la concavité de $v_+(z)$</p>	<p>Aversion au risque</p> <p>Surpondère la petite probabilité de perte. $\varphi(p_{small})$ concave + $v_-(-z)$ convexe. La convexité de $v_-(-z)$ contrecarre la concavité de $\varphi(p)$</p>

II. REFERENTIEL ET ATTITUDE FACE AU RISQUE

L'analyse de l'attitude face au risque dans un modèle CPT est particulièrement difficile dans la mesure où cette dernière dépend : de la courbure de la fonction d'utilité dans les gains et dans les pertes, de l'aversion aux pertes ainsi que de la distorsion des poids décisionnels. Nous présentons dans un premier une première approche graphique qui fournit l'intuition générale de l'influence du point de référence sur l'attitude face au risque. Dans un second temps, nous adoptons une approche mathématique dans la lignée de l'analyse d'Arrow (1965) et Pratt (1964). Nous reprenons en fait ici l'analyse exposée par Davies et Satchell (2007) qui consiste à fournir une approximation de la prime de risque par l'intermédiaire d'une expansion d'une suite de Taylor. Cette approximation permet de décomposer la prime de risque et d'en analyser

¹⁸ On entend par « contrecarre » l'idée que la fonction $w(p)$ est suffisamment concave (convexe) pour générer de la recherche (de l'aversion pour le) de risque malgré une fonction d'utilité concave (convexe).

les déterminants. Toutefois, pour simplifier l'analyse nous n'étudierons pas l'impact de la déformation des probabilités sur l'attitude face au risque. En outre, nous supposons $\varphi(p) = p$.

Pour finir, nous donnerons une représentation graphique de l'effet du référentiel sur la prime de risque absolue.

2.1. Approche graphique

2.1.1. Richesse actuelle et référentiel

La formalisation de la value function proposée par K&T telle qu'elle est souvent formalisée dans les articles (voir Equation (2.1)) est trompeuse. En effet, l'argument " x " de la fonction d'utilité est souvent confondu avec les issues des loteries. Or, nous avons vu précédemment que cette assertion n'est vraie que si l'on suppose que le point de référence est égal à la richesse actuelle de l'individu. Ainsi, l'argument de la fonction d'utilité " x " représente bien une déviation de la richesse finale par rapport à un référentiel mais dans la mesure où ce référentiel peut être différent de la richesse actuelle, l'argument « x » peut être différent des issues monétaires engagés dans les loteries.

K&T ont toutefois argumenté en faveur d'un point de référence égal à la richesse actuelle. Ils appellent ce référentiel : le statu-quo et le choisissent pour deux raisons : 1) Par commodité pour simplifier les équations et 2) Car ils estiment que la plupart du temps c'est probablement ce référentiel qui est utilisé par les individus comme semble le démontrer leur paradoxe présenté dans ce chapitre à la section 1.1.1.

Les auteurs sont toutefois conscients, comme Markowitz au demeurant, que le point de référence peut dévier de la richesse initiale. « [...] reference point was taken to be the status quo, or one's current assets. Although this is probably true for most choice problems, there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo » (Kahneman & Tversky, 1979, p. 286). Depuis, de nombreux

chercheurs se sont attachés à découvrir et étudier les différents candidats potentiels au poste de « référentiel ». Un champ d'étude diversifié s'est alors constitué autour de cette question et envisage le point de référence sous différentes perspectives, qui peuvent plus ou moins coexister simultanément.

On trouve aujourd'hui comme référentiels envisagés : les normes sociales (Gomez et al, 2009), la comparaison sociale (Fiegenbaum & Thomas, 1988; Kameda & Davis, 1990; Lehner, 2000), les gains et pertes antécédentes (Arkes, Hirshleifer, Jiang, & Lim, 2008, 2010; Baucells, Weber, & Welfens, 2011; Post, Van den Assem, Baltussen, & Thaler, 2008), un pic de richesse antécédent (Gneezy, 2005), le niveau d'aspiration (Camerer, Babcock, Loewenstein, & Thaler, 1997; Gerhard, Hoffmann, & Post, 2015; Payne, Laughunn, & Crum, 1980), un référentiel stochastique basé sur les anticipations rationnelles des agents (Koszegi & Rabin, 2006, 2007), l'équivalent certain (Gul, 1991), le meilleur gain de la loterie qui n'a pas été choisi (Loomes & Sugden, 1982, 1987)...

Avant de d'analyser plus en détail les différents candidats au point de référence, nous étudierons tout d'abord l'effet de différentes localisations du référentiel par rapport aux issues finales sur l'attitude face au risque.

2.1.2. Prospect Theory et Triangle de Marschak-Machina

On peut dégager trois grands cas de figures concernant l'impact de la position du référentiel par rapport aux richesses finales prospectives sur l'attitude face au risque. 1. R est situé entre les richesses finales prospectives ; 2. R est supérieur à l'ensemble des issues ; 3. R est inférieur à l'ensemble des issues.

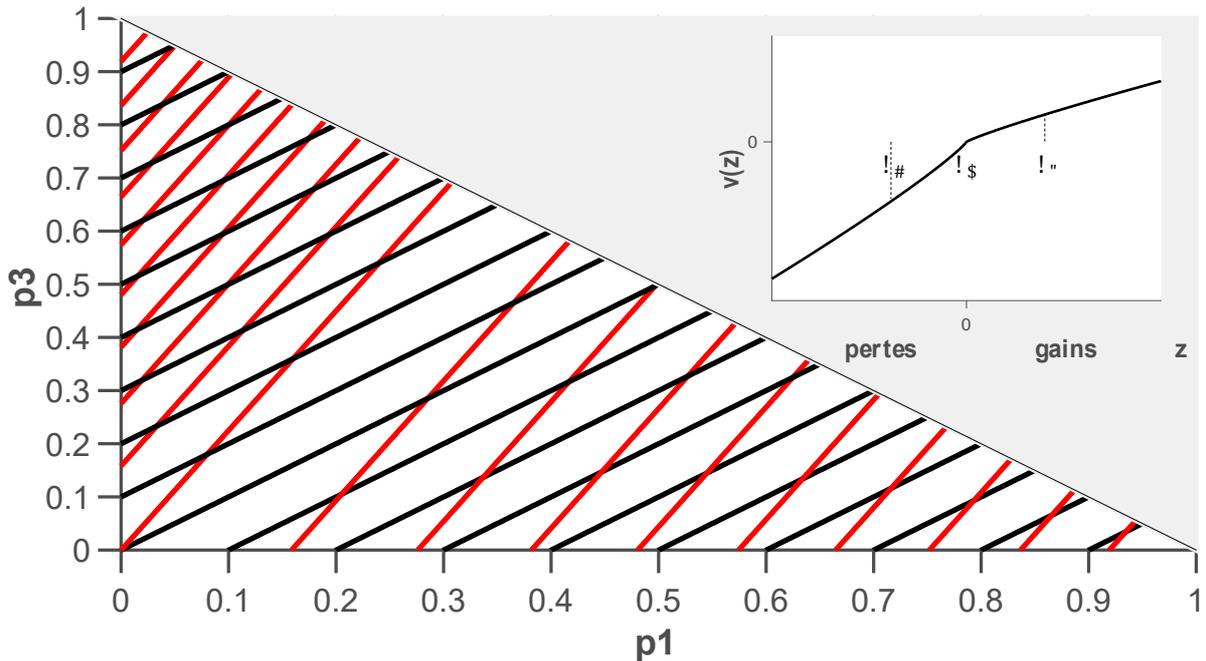


Figure 2.4 Représentation des droites d'indifférences (en rouge) dans le triangle de Marschak-Machina selon un modèle CPT. Le référentiel est égal à $R=1500$ et la richesse initiale $w_0 = 1000$. La loterie $L \equiv (x_1 = 0, p_1; x_2 = 500, p_2, x_3 = 1000, p_3)$. Autrement dit, sous l'influence du référentiel l'individu évalue donc les conséquences suivantes $Z \equiv (z_1 = -500, p_1; z_2 = 0, p_2, z_3 = 500, p_3)$. On suppose une fonction d'utilité puissance dans les gains et les pertes avec $\alpha = \beta = 0.88$. On suppose également $\varphi(p) = p$. Les droites en noirs représentent des droites iso-espérances. En haut à droite on représente schématiquement la localisation des issues de la loterie par rapport au référentiel.

Cas 1. $W_1 < W_2 < \dots < W_k < R < W_{k+1} < W_{k+2} < \dots < W_n$

Ce cas de figure correspond à une loterie dite « mixte ». Une loterie mixte est une loterie qui engage à la fois des pertes et des gains, à la différence d'une loterie dite « régulière » dans laquelle les issues sont soit uniquement des pertes (0 non-compris) soit uniquement des gains (0 compris). Dans une situation où le référentiel peut être différent de la richesse actuelle, w_0 , il est tout à fait envisageable que des loteries n'engageant objectivement que des gains (donc régulières) puissent malgré tout être perçues par l'agent comme des pertes et réciproquement. Dès lors, dans le cas où R est positionné entre les issues finales, une partie des richesses finales prospectives sont perçues comme des pertes et les autres comme des gains rendant la loterie mixte alors qu'elle peut être objectivement régulière. Les prédictions de la CPT établissent

qu'un agent est particulièrement averse au risque dans le cas de loterie mixte à cause de l'aversion aux pertes. Si on analyse la fonction d'utilité, on remarque facilement que la région la plus concave de la fonction est située autour du point de référence à cause de l'aversion aux pertes comme l'illustre la Figure 2.4 générant une aversion pour le risque importante. Dit autrement, dans la mesure où la désutilité marginale liée à une perte est renforcée par le paramètre d'aversion aux pertes, elle est donc bien plus importante que l'utilité marginale liée à un gain.

Cas 2. $W_1 < W_2 < \dots < W_n < R$ ou bien **cas 3.** $R < W_1 < W_2 < \dots < W_n$

Dans ces situations l'ensemble des issues d'une loterie sont soit perçues exclusivement comme des pertes (cas 2) soit exclusivement comme des gains (cas 3). Nous avons déjà vu que l'individu devrait se montrer riscophile ou riscophobe respectivement. En revanche, dans la section des gains, l'aversion au risque n'est pas renforcée par l'aversion aux pertes. Dès lors l'individu devrait se montrer moins averse au risque que dans la situation d'une loterie mixte vu juste au-dessus.

De plus, l'hypothèse de diminution marginale de la sensibilité aux pertes ou aux gains introduit une prédiction supplémentaire. Dans la mesure où la désutilité (l'utilité) marginale d'une perte (gain) décroît, la fonction devient donc de plus en plus linéaire. Par conséquent, pour une loterie donnée, à mesure que le point de référence devient plus grand que la richesse actuelle, l'individu devrait se montrer de moins en moins riscophile et tendre vers la neutralité. De même, à mesure que le référentiel devient plus petit que la richesse actuelle, l'individu devrait se montrer de moins en moins averse au risque et tendre vers la neutralité.

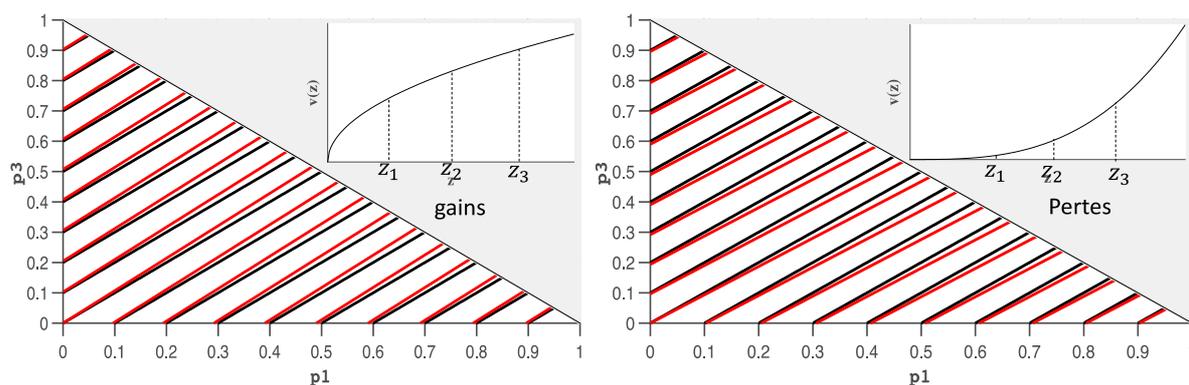


Figure 2.5 Ces figures sont élaborées en conservant les mêmes paramètres et loteries que la figure 2.4. On modifie simplement la valeur du référentiel. A gauche : On suppose un référentiel $R=0$. D'où $Z \equiv (z_1 = 1000, p_1; z_2 = 1500, p_2, z_3 = 2000, p_3)$. A droite : on suppose un référentiel $R=2000$. D'où $Z \equiv (z_1 = -1000, p_1; z_2 = -500, p_2, z_3 = 0, p_3)$.

2.2. Approche mathématique

2.2.1. L'étude de Davies et Satchell (2007)

Comme nous l'avons évoqué dans la partie 1, la prime de risque constitue un bon indicateur du degré d'aversion au risque d'une individu. Les approches d'Arrow et Pratt puis ensuite de Hilton ont permis de révéler les déterminants de la prime de risque absolue. Ici, nous reprenons l'analyse opérée par Davies & Satchell (2007) qui consiste à approximer la prime de risque par l'intermédiaire de développement limité sous différentes conditions concernant la localisation du point de référence par rapport aux richesses finales. Cette investigation permet d'étudier formellement l'impact du référentiel sur la prime de risque.

Comme Arrow (1965) et Pratt (1964), Davies & Satchell (2007) restreignent leur analyse à des « petites loteries », c'est-à-dire d'espérance 0 et de variance infinitésimalement petite. Rappelons que nous supposons également aucune distorsion des probabilités mais le lecteur intéressé pourra se référer à l'étude de Davies et al qui prend en compte la déformation des probabilités.

Formellement, nous considérons un individu doté d'une richesse initiale w_0 et confronté à une loterie composée d'issues aléatoires x_i dont l'espérance est nulle, $E(\tilde{x}) = 0$ et la variance infinitésimalement petite (négligeable).

Nous effectuons dans un premier temps une approximation au voisinage d'un point de référence égal à la richesse initiale w_0 . Nous avons vu dans la section 1.2.1 de ce chapitre que lorsque le point de référence est égal à la richesse initiale, la fonction d'utilité porte alors sur les issues de la loterie elle-même. Dès lors, l'approximation s'effectue au voisinage de l'espérance de la loterie, c'est-à-dire ici 0. On dégagera alors deux cas : le cas d'une prime de risque négative (recherche de risque) et le cas d'une prime de risque positive (aversion au risque).

Point de référence = Richesse initiale

Cas 1 : recherche de risque : $E(\tilde{W}) - \pi_a < 0$, $\pi_a < 0$ et donc $-\pi_a > 0$; On notera q la probabilité de perdre, $q = \sum_{i=1}^n (p_i | W_i < R)$ et donc $(1 - q)$ la probabilité d'obtenir un gain. Cette probabilité q permet de partager en deux la contribution des gains et la contribution des pertes dans la formation de la valeur subjective d'une loterie. Ce partage permet en outre d'éviter un problème de singularité au niveau du point de référence induit par la spécification de la fonction d'utilité dans la Prospect Theory (cette dernière n'est pas dérivable au point du référentiel).

De plus, étant donné que nous traitons le cas où $w_0 = R$, alors l'argument de la fonction d'utilité z est égal au montant des issues financières x , tel que $v(z) = v(x)$ (voir section 1.2.1 de ce chapitre). On cherche alors une approximation de la prime à partir de l'égalité suivante

$$v_+(E(\tilde{x}) - \pi_a) = q\lambda E[v_-(\tilde{x})|\tilde{x} < 0] + (1 - q)E[v_+(\tilde{x})|\tilde{x} > 0] \quad (2.7)$$

En exerçant une expansion d'une suite de Taylor au premier ordre du terme de gauche au voisinage du point de référence R , on obtient :

$$v_+(E(\tilde{x}) - \pi_a) \approx v_+(E(\tilde{x})) + \pi_a v'_+(E(\tilde{x}))$$

Ou bien, puisque $E(\tilde{x}) = 0$, alors on peut réécrire l'expression ci-dessus ainsi :

$$v_+(-\pi_a) \approx v_+(0) + \pi_a v'_+(0)$$

Une approximation de deuxième ordre au voisinage du point de référence du terme de droite est :

$$\begin{aligned} & q\lambda E[v_-(\tilde{x})|\tilde{x} < 0] + (1 - q)E[v_+(\tilde{x})|\tilde{x} > 0] \\ &= q\lambda E \left[v_-(0) + v'_-(0)(\tilde{x}|\tilde{x} < 0) + \frac{1}{2} v''_-(0)(\tilde{x}^2|\tilde{x} < 0) \right] \\ &+ (1 - q)E \left[v_+(0) + v'_+(0)(\tilde{x}|\tilde{x} > 0) + \frac{1}{2} v''_+(0)(\tilde{x}^2|\tilde{x} > 0) \right] \end{aligned}$$

Pour alléger l'écriture on adopte les notations suivantes. La variance conditionnelle aux pertes est notée $\sigma_-^2 \equiv E[(\tilde{x}^2 - \tilde{x})|\tilde{x} < 0]$ et la variance conditionnelle aux gains est notée $\sigma_+^2 \equiv E[(\tilde{x}^2 - \tilde{x})|\tilde{x} > 0]$. La moyenne conditionnelle aux pertes est notée $\mu_- \equiv E[\tilde{x}|\tilde{x} < 0]$ et l'espérance conditionnelle aux gains est notée $\mu_+ \equiv E[\tilde{x}|\tilde{x} > 0]$. De plus, puisque $v(0) = 0$ alors :

$$-\pi_a v'_+(0) \approx q\lambda \left(v'_-(0)\mu_- + \frac{1}{2} v''_-(0)\sigma_-^2 \right) + (1 - q) \left(v'_+(0)\mu_+ + \frac{1}{2} v''_+(0)\sigma_+^2 \right)$$

Puisque l'espérance de \tilde{x} est nulle et puisque $v'_+(0) = v'_-(0) = v'(0)$, on obtient l'approximation simplifiée de la prime de risque (pour un chercheur de risque) suivante :

$$\pi_a \approx -\frac{q v''_-(0)\sigma_-^2 + (1 - q)v''_+(0)\sigma_+^2}{2v'(0)} - q\mu_-(\lambda - 1) \quad (2.8)$$

Cas 2 : Aversion au risque : $E(\tilde{W}) - \pi_a < 0$, $\pi_a > 0$. L'expansion d'une suite de Taylor au voisinage du point de référence du terme de gauche donne :

$$v_-(-\pi_a) \approx v_-(0) - \pi_a \lambda v'_-(0)$$

L'approximation au second ordre du terme de droite est identique au cas 2. Ainsi on obtient l'approximation de la prime de risque suivante :

$$\pi_a \approx \frac{1}{\lambda} \left[-\frac{qv''_-(0)\sigma_-^2 + (1-q)v''_+(0)\sigma_+^2}{2v'(0)} - q\mu_-(\lambda - 1) \right] \quad (2.9)$$

Analyse de la décomposition de la prime de risque

Nous sommes désormais en mesure d'examiner les différentes composantes de la prime de risque ou plus généralement de l'attitude face au risque lorsque $R = w_0$ et que les issues financières d'une loterie sont réparties de part et d'autre du point de référence. Nous avons vu dans le modèle EUT que la prime de risque dépendait du produit entre la variance et le degré de concavité normalisé. Concernant le modèle CPT, le premier terme de l'Equation (2.8),

$-\frac{qv''_-(0)\sigma_-^2 + (1-q)v''_+(0)\sigma_+^2}{2v'(0)}$, indique que la prime de risque dépend de deux produits : (1) Le produit entre la variance conditionnelle aux pertes et le degré de concavité/convexité dans les pertes, (2) Le produit entre la variance dans les gains et le degré de concavité/convexité dans les gains, (3) ces deux produits sont pondérés par la masse des probabilités de pertes et de gains.

On déduit alors les résultats suivants :

Ce terme est nécessairement positif (négatif) si la fonction d'utilité est concave (convexe) dans les pertes et les gains. Autrement dit, l'individu est nécessairement averse au risque (chercheur de risque) dans ce cas.

Empiriquement nous avons vu que la fonction d'utilité semble concave dans les gains et convexe dans les pertes. D'après l'Equation (2.8), il est possible que ce terme soit positif (donc

aversion pour le risque) si la convexité dans les pertes est compensée par une suffisamment grande concavité dans les gains¹⁹, tel que $-qv''_-(0)\sigma_-^2 > (1-q)v''_+(0)\sigma_+^2$.

Le second terme, $q\mu_-(\lambda - 1)$, n'est pas présent dans la décomposition d'Arrow-Pratt. Il marque ici l'influence de l'aversion aux pertes. Nous remarquons dans les Equation (2.8) et 2.9 que l'aversion aux pertes renforce toujours l'aversion au risque tant que $\lambda > 1$ (c'est-à-dire lorsque l'individu est averse aux pertes). Pour le voir il suffit de dériver partiellement les deux expressions par rapport à λ .

On obtient pour l'Equation (2.8) (lorsque $\pi_a < 0$) : $\frac{\partial \pi_a}{\partial \lambda} = -\frac{qv''_-(0)\sigma_-^2}{2v'(0)} - q\mu_-$. Nous pouvons alors signer l'expression. Si la fonction dans les pertes est concave alors le premier terme est positif. Puisque $\mu_- < 0$ la fonction dérivée de la prime de risque par rapport à λ , quand la prime de risque est négative, est donc croissante. Autrement dit, l'individu est de moins en moins chercheur de risque puisque le degré d'aversion au risque dans les pertes est renforcé par le second terme. Si la fonction d'utilité est convexe dans les pertes, le premier terme est alors négatif. Dans ce cas, tant que $-\frac{v''_-(0)\sigma_-^2}{2v'(0)} > \mu_-$ la fonction dérivée de la prime de risque reste croissante.

Concernant l'équation (2.9) (lorsque $\pi_a > 0$), $\frac{\partial \pi_a}{\partial \lambda} = -\frac{qv''_-(0)\sigma_-^2}{2v'(0)} - q\mu_- + \frac{\lambda^2 \pi_a \ln \lambda}{q}$. Le dernier terme est positif tant que $\lambda > 1$. Lorsque la fonction d'utilité est concave dans les pertes, la fonction dérivée reste croissante tant que $\lambda > 1$. Si la fonction est convexe, la fonction dérivée de la prime de risque par rapport à λ reste croissante tant que $-\frac{v''_-(0)\sigma_-^2}{2v'(0)} > \mu_- - \frac{\lambda^2 \pi_a \ln \lambda}{q}$.

Point de référence = Richesse initiale

Nous traitons maintenant le cas où le point de référence est différent de la richesse initiale et en dehors de l'intervalle des richesses finales. Cette section ne nécessite pas de développement

¹⁹ Ou bien à l'inverse, ce terme peut être positif si la convexité dans les gains est compensée par une suffisamment grande concavité dans les pertes.

approfondi. On notera simplement que dans la situation où toutes les richesses finales sont supérieures au référentiel, toutes les conséquences sont perçues comme des gains. Dans la situation où toutes les richesses finales sont inférieures au référentiel, toutes ces conséquences sont perçues comme des pertes. Dès lors, l'aversion aux pertes n'a aucun impact sur l'évaluation des conséquences. Dans ces cas, seule la concavité ou la convexité des fonctions déterminent l'aversion au risque et par conséquent l'analyse d'Arrow-Pratt s'applique à nouveau. La seule différence avec Arrow-Pratt est que la prime de risque sera évaluée soit à partir de v_- soit de v_+ en fonction de la position du référentiel par rapport aux issues finales.

2.2.2. Prime de risque d'une loterie et Référentiel

Ici, pour une loterie donnée, nous représentons l'évolution de la prime de risque en fonction du référentiel. Dans cet exemple, l'individu possède une richesse actuelle de $w_0 = 1000\text{€}$ et fait face à une loterie $L(500, 0.5; 0, 0.5)$. En termes de richesse finale, on a alors $W(1500, 0.5; 1000, 0.5)$. Pour construire la fonction de la prime de risque, nous supposons une fonction d'utilité puissance et aucune transformation des probabilités. $u(W) = \begin{cases} (W - R)^{0.88} & W \geq R \\ -2.25(R - W)^{0.88} & W < R \end{cases}$, et $\varphi(p) = p$. Les paramètres de la fonction d'utilité sont tirés des études menées par Kahneman et Tversky (1979). La fonction d'équivalent certain est alors donnée par l'expression suivante :

$$Ce(\tilde{W}) = v^{-1}[CPT(\tilde{W})] = \begin{cases} R + [CPT(\tilde{W})]^{\frac{1}{\alpha}} & \text{si } CPT(\tilde{W}) \geq 0 \\ R - \left[-\frac{CPT(\tilde{W})}{\lambda} \right]^{\frac{1}{\beta}} & \text{si } CPT(\tilde{W}) < 0 \end{cases} \quad (2.10)$$

Et les primes de risque relatives sont simplement calculées en faisant varier le référentiel dans l'intervalle $[0, 2000]$. La prime de risque relative est obtenue en divisant la prime de risque

absolue, $\pi_a = E(\tilde{W}) - Ce(\tilde{W})$, par l'espérance de richesse finale : $\pi_r = \pi_a/E(\tilde{W})$. A partir de ces données, nous représentons l'évolution de la prime de risque dans la Figure 2.6.

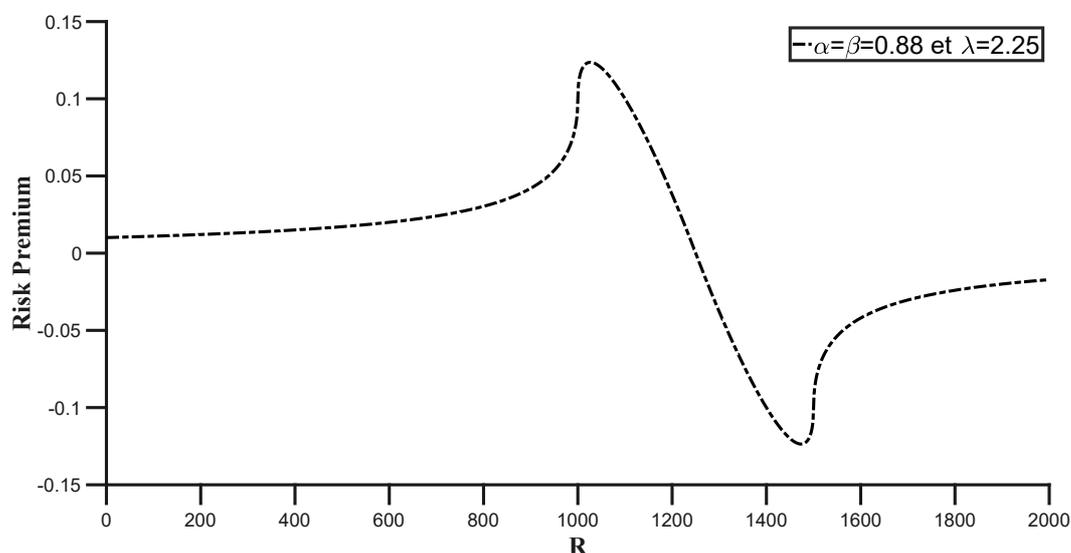


Figure 2.6 Simulation de primes de risque relatives selon un modèle CPT.

Nous observons que lorsque le référentiel est proche de l'espérance et que donc les richesses finales sont réparties autour du point de référence, l'aversion au risque est importante à cause de l'aversion aux pertes. Par ailleurs, on remarque que lorsque le point de référence est égal à la richesse finale la plus élevée ($R = W = 1500$), la prime de risque est alors à son minimum. Ce résultat révèle que les issues financières sont alors évaluées sur la section la plus convexe de la fonction d'utilité à cause de l'impact de l'aversion aux pertes. Enfin, on remarque qu'à mesure que le référentiel s'écarte de l'intervalle des issues finales prospectives, la riscophobie (riscophilie) diminue dans les gains (pertes) et l'individu tend vers la neutralité vis-à-vis du risque.

Conclusion à propos de la Cumulative Prospect Theory

Nous avons vu dans cette section qu'en fonction du point de référence une même loterie peut paraître favorable ou inacceptable. Par conséquent, un individu averse au risque face à une loterie peut se révéler riscophile pour cette même loterie si son point de référence est modifié.

Cela s'explique par le fait qu'en fonction de la valeur du point de référence, les issues d'une loterie sont situées soit dans la région des pertes soit dans la région des gains soit dans les deux régions. De manière générale, le point de référence pose la question fondamentale : « Qu'est-ce qu'une perte et qu'est-ce qu'un gain ? » (Gneezy, 2005).

En définitive nous avons montré les éléments suivants : 1. A mesure que R est éloigné des richesses finales prospectives, la section de la fonction d'utilité sur laquelle les richesses finales prospectives sont évaluées est de plus en plus linéaire conduisant à la neutralité face au risque. 2. A contrario, si les issues finales prospectives sont réparties autour du point de référence, la section de la fonction d'utilité sur laquelle \tilde{W} est évaluée est fortement concave à cause de la présence du kink. Cela traduit l'idée selon laquelle l'aversion au risque est conduite par le fait que le gain d'un euro est moins plaisant que la perte d'un euro est douloureuse (c.a.d aversion aux pertes). Or cette différence ($|v'(-z)| - |v'(z)| > 0$) est maximisée au niveau du point de référence. 3. Dans le cas où le référentiel est moins élevé que les richesses finales prospectives alors ces issues sont perçues comme des gains. Dans la région des gains, la différence entre l'utilité marginale d'un gain supplémentaire et la désutilité marginale d'une perte supplémentaire est moins importante que lorsque les issues financières sont réparties autour du point de référence (puisque le kink n'est pas présent dans la section concave de la fonction). Dès lors, l'individu sera moins averse au risque dans cette situation. Si le référentiel est plus élevé que les richesses finales prospectives alors le bénéfice de gagner un euro supplémentaire est plus importante que la désutilité marginale liée à une perte d'un euro supplémentaire générant ainsi la recherche de risque (en effet la pente de la fonction dans les pertes est plus élevée que la pente de la fonction dans les gains).

Dans la prochaine section nous allons désormais étudier les différents points de référence qui ont été envisagés dans la littérature. Nous différencions deux types d'études : celles qui recherchent le point de référence dans un environnement de décision statique et celles qui

recherchent le point de référence et son évolution dans un environnement de décision dynamique. Nous verrons entre autres que les environnements dynamiques soulèvent une question supplémentaire à propos du référentiel qui ne peut être appréhendé dans des situations de choix statiques : le point de référence s'adapte-t-il, et si oui, à quoi et comment ? Avant de débiter l'analyse, précisons néanmoins un élément : notre étude va se concentrer sur certains types de référentiels tel que : « un niveau de richesse » ou encore « un niveau de prix d'actif » ou « un rendement ». Il est implicitement supposé que la nature du point de référence dépend en fait du contexte : si on présente une situation impliquant des loteries, on supposera que le référentiel sera un niveau de richesse. Si on présente une situation montrant l'évolution d'un prix, on supposera que le référentiel est un prix.

III. LOCALISATION DU REFERENTIEL EN CONTEXTE DE CHOIX STATIQUE

Le problème fondamental de la PT et des autres théories référence-dépendant est qu'elles ne spécifient pas comment le point de référence est constitué (Baillon, Bleichrodt, & Spinu, 2015). Même si K&T choisissent la richesse initiale comme point de référence les auteurs sont conscients que le référentiel peut différer de la richesse initiale. Markowitz (1952), avait lui-même déjà soulevé cette question à propos de ce qu'il appelait le « customary wealth » qui faisait office de point de référence dans sa théorie : « it would be convenient if I had a formula from which customary wealth could be calculated when this was not equal to present wealth. But I do not have such rule and formula ». Comme le font remarquer Baillon & al, (2015), une telle ambiguïté est pour le moins regrettable dans la mesure où un degré de liberté trop important est alors laissé aux modèles dans leurs prédictions, si bien qu'il est en définitive particulièrement difficile de tester rigoureusement ces théories.

3.1. Les contextes de choix statique

3.1.1. Définition

Un environnement incertain implique un choix « statique » quand la décision de l'individu peut être prise avant toute résolution de l'aléa d'une loterie (ou d'un niveau d'une loterie dans le cas de loterie composée). « In other words, « nature » does not make any moves until the decision maker has irrevocably made all of his own moves » (Machina, 1989, p.1632). Jusqu'à présent dans cette thèse nous n'avons envisagé que des situations de choix statique, des représentations telle que la Figure 2.1. ou la Figure 2.2. en donne des bons exemples. Dans ces figures, les carrés (ou nœuds de décision) des arbres représentent les moments de choix. Ces choix mènent à une des loteries choisies représenté par les ronds. Chaque branche des ronds représentant les différents « mouvements » possible de la nature. Dans un contexte de choix entre des loteries composés, comme la Figure 2.1, on remarque bien que le premier mouvement est opéré par l'individu et qu'ensuite plus aucun choix n'est possible pour ce dernier. Seule la ou les réalisations de l'aléa se produiront successivement.

3.1.2. Expérimentation en choix statique

Les choix statiques sont très utilisés dans les expérimentations. Elles ont l'avantage d'être particulièrement simples à comprendre pour le participant. En revanche, il est clair que ne présenter qu'un seul choix à un individu limite la possibilité de tester des hypothèses. L'idéal consiste donc à présenter un ensemble de choix aux participants de façon séquentiel. Mais dans ce cas peut-on encore parler de choix statique ? Le problème étant qu'en fonction des choix et de la réalisation de l'aléa, la richesse de l'individu évolue après chaque essai, et nous rentrons dès lors dans une situation de choix dynamique. Pour contrecarrer ce problème, les chercheurs ont deux solutions : Soit utiliser des gains et pertes hypothétiques tout en informant les sujets que les situations de choix successives sont bien indépendantes entre-elles. Soit, ne récompenser qu'un seul choix effectué au cours d'une session (voir par exemple Holt & Laury, 2002). Avec la dernière méthode, l'individu est amené à considérer chaque choix comme un

« one shot » puisque chaque choix a autant de probabilité d'être réellement joué et payé à la fin de l'expérience. De la sorte, les chercheurs parviennent à supprimer l'aspect dynamique dans un processus de choix séquentiel. L'avantage avec cette méthode est qu'elle permet de motiver les individus à répondre aux questions selon leurs vrais préférences. Il a en effet été constaté maintes fois que les gains hypothétiques et réels ne produisaient pas forcément les mêmes comportements.

3.1.3. Choix statique et recherche du point de référence

Les études portant sur le point de référence en environnement statique doivent impérativement présenter de nombreuses loteries aux individus dans le but d'être en mesure de distinguer les référentiels opérants dans une situation de choix. Seulement, la plupart du temps, les auteurs choisissent en amont de leur expérience, un référentiel qui leur paraît crédible puis, sur la base des prédictions qu'induit cette spécification, le testent afin de vérifier sa véracité. Le problème avec cette méthode est que les auteurs ne vérifient pas si d'autres référentiels pourraient aussi bien expliquer leurs hypothèses. Il existe néanmoins une étude qui nous paraît particulièrement pertinente et fut proposée récemment par Baillon et al (2015). Cette expérience oppose huit points de références envisagés de façon récurrente par différents chercheurs. Pour parvenir à sélectionner les référentiels qui semblent être utilisés par les sujets, les auteurs ont alors constitué plus de 70 situations de choix différentes auxquelles les sujets devaient répondre. Systématiquement, les individus avaient le choix entre 2 loteries composées chacune de deux à quatre issues. Pour chaque situation de choix, les différents points de référence prédisaient quelle loterie devrait être choisie. Bien évidemment, différents référentiels peuvent prédire le même choix. C'est pourquoi Baillon et al (2015) ont élaboré un grand nombre de loterie, afin de procéder par exclusion progressive. Nous exposons ici cette étude car elle nous permet de passer en revue les différents points de références qui furent examinés dans la littérature et de déterminer dans un second temps ceux qui semblent efficaces.

3.2. Localisation du référentiel - les candidats

Les règles de détermination du point de référence peuvent être établies selon 3 dimensions (Baillon & al, 2015) : 1) Le point de référence est-il « choice specific » ou « prospect specific ». Dans le cas d'un point de référence « choice specific » une règle unique est utilisée et appliquée à toutes les loteries présentes dans l'ensemble des choix possibles. Dans le cas d'un point de référence « prospect specific » un point de référence unique est constitué pour chaque loterie séparément. Ces référentiels peuvent être la valeur espérée de chaque loterie (Bell, 1985), l'équivalent certain (Gul, 1991) ou bien les issues des loteries elles-mêmes (Koszegi & Rabin, 2007). 2) Le point de référence est-il déterministe ou aléatoire. 3) Les probabilités sont-elles prises en compte pour déterminer le point de référence ou bien seules les issues financières peuvent-elles servir de point de référence ?

3.2.1. Point de référence « choice specific »

Selon cette spécification du point de référence, chaque issue des différentes loteries présentes dans une situation de choix sont comparées à un seul point de référence. Nous passons ici en revue les différents référentiels envisagés dans la littérature.

Le statut quo

Nous avons déjà évoqué ce cas de figure notamment avec le paradoxe de K&T. Nous n'en dirons pas davantage à ce sujet. Par ailleurs, de nombreuses études semblent confirmer l'intuition de K&T (voir par exemple Samuelson & Zeckhauser, 1988). Toutefois, il nous apparaît important ici de préciser un point. Lorsque nous parlons de du statut quo, nous évoquons la richesse actuelle de l'individu. Autrement dit, sa richesse à un moment t . Il ne faut pas confondre avec le statut quo avec la richesse (ou la dotation) initiale de l'individu. Il est facile de les confondre car la plupart du temps, les décisions sont prises en « one shot » : c'est-à-dire dans une situation non dynamique où, dès lors, la richesse initiale de l'individu n'évolue

pas en fonction de ses choix et de la réalisation de l'aléa. Dans ce cas de figure, le statut quo est égal à la richesse actuelle qui est elle-même égale à la richesse initiale.

De nombreux exemples appuient cette spécification du point de référence. L'effet de dotation par exemple révélé par la célèbre expérience du « mug » menée par Kahneman, Knetsch, & Thaler (1990) en est une excellente illustration²⁰ en situation de certitude. Dans cette expérience, les expérimentateurs fournissaient un mug uniquement à la moitié des participants et pas aux autres. Les auteurs ont ensuite demandé aux sujets possédant un mug, leur prix de réserve, c'est à dire à quel prix ils seraient prêts à le céder (willingness to accept, WTA), et ont parallèlement demandé aux sujets n'ayant pas de mug, à quel prix ils seraient prêts à l'acheter (willingness to pay, WTP). Les auteurs observent un écart significatif entre le WTP et le WTA²¹. Plus précisément, les possesseurs de mug étaient prêts à vendre plus cher que les non possesseurs de mug étaient prêts à l'acheter ($WTA > WTP$). Rabin (1997) suppose que les individus ont tendances à donner une valeur plus importante à ce qu'ils possèdent déjà qu'à ce qu'ils ne possèdent pas. Pour expliquer cet effet, il faut considérer 2 éléments inscrits dans l'ADN de la prospect theory : un point de référence et l'aversion aux pertes. Ici, on comprend bien que les possesseurs de mug ont bien comme référence « avoir un mug ». Dès lors, l'état « ne pas avoir de mug » est une perte. Pour les individus n'ayant pas de mug, leur référentiel est 0. Par conséquent, l'état « avoir un mug » est un gain. Puisqu'on sait que l'impact d'une perte est plus importante que celui d'un gain (aversion à la dépossession ou à la perte), les individus possédant un mug souhaitent le vendre plus cher pour contrebalancer la désutilité liée à la perte du mug. Les « non possesseurs » de mug n'ont pas à contrebalancer cette souffrance dans la mesure où le référentiel est 0 (ils ne peuvent pas perdre moins que 0 mug) d'où l'écart entre WTP et WTA.

²⁰ Voir Hanemann (1991), pour une explication néoclassique sans avoir recours à la prospect theory pour expliquer cet effet.

²¹ Le prix de réserve médian des vendeurs était 2.5 fois plus élevé que celui des acheteurs.

La règle MaxMin :

Face à un ensemble de loteries, l'individu prête une attention particulière à toutes les issues minimums de chaque loterie. Il sélectionne ensuite la plus grande des issues minimums comme référentiel (Birnbaum & Schmidt, 2010). Cette règle est associée à ce que l'on appelle le niveau de sécurité (« security level »). En particulier, un individu ayant des préférences « MaxMin » est un individu qui choisit la loterie assurant le meilleur des pires résultats possibles.

La règle MinMax :

Face à un ensemble de loteries, l'individu prête une attention particulière à toutes les issues maximales de chaque loterie. Il sélectionne ensuite la plus petite des issues maximales comme référentiel (Yang, 2000).

La règle X at max P :

Le point de référence est l'issue ayant la plus grande probabilité d'occurrence. On précisera que si certaines issues ont les mêmes probabilités, l'individu choisira la plus petite comme référentiel.

3.2.2. Point de référence « prospect specific »

Köszegi & Rabin (2007) spécifient une fonction d'utilité différentes de celle que nous avons décrite précédemment. Selon eux l'utilité en situation de certitude est dérivée d'un niveau absolu de consommation c (fixe) et d'un niveau relatif représentant le sentiment de perte et de gain par rapport à un point de référence r , d'où :

$$u(c|r) = u(c) + \mu[u(c) - u(r)] \quad (2.11)$$

Avec u une fonction d'utilité de type VnM et μ une fonction d'utilité définie sur les pertes et les gains par rapport au point de référence comme déclaré par la prospect theory²².

$$\mu(y) = \begin{cases} \eta y & \text{si } y > 0 \\ \eta \lambda y & \text{si } y < 0 \end{cases} \quad \eta > 0, \lambda > 1 \quad (2.12)$$

Pour les petits stacks et conformément à A3', $\mu(y)$ est supposé linéaire avec $\eta = 1$ et $\lambda > 1$.

En contexte d'incertitude évidemment c est stochastique, nous dénommerons alors l'équivalent stochastique de c par sa distribution F , d'où $c \sim F$. Ainsi quand c est aléatoire et r fixe, l'espérance d'utilité s'obtient par :

$$U(F|r) = \int [u(c) + \mu[u(c) - u(r)]] dF(c) \quad (2.13)$$

Si le point de référence est lui-même stochastique alors la distribution de r sera appelée G , $r \sim G$, d'où :

$$U(F|G) = \iint [u(c) + \mu[u(c) - u(r)]] dG(r) dF(c) \quad (2.14)$$

Cette formulation stipule que les agents évaluent l'utilité d'une action en se basant sur la distribution F définie sur les conséquences de cette action par rapport à la distribution de G .

²² Les auteurs présument les propriétés suivantes, A0. $\mu(x)$ est continue pour tout x , deux fois dérivables pour $x \neq 0$ et $\mu(0) = 0$. A1. $\mu(x)$ est strictement croissante conformément à l'axiome de monotonie. A2. Si $y > x > 0$, donc $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$, ce qui représente l'aversion aux pertes pour les grands enjeux. A3. $\mu''(x) \leq 0$ pour $x > 0$ et $\mu''(x) \geq 0$ pour $x < 0$, ce qui signifie que plus=mieux mais seulement à la marge. Il capture une autre propriété importante qui est la diminution marginale de la sensibilité : les agents sont plus sensibles à une variation proche du point de référence que lorsque cette même variation à lieu « loin » du référentiel. Toutefois, pour les petit stacks : A3'. $\forall x \neq 0, \mu'' = 0$. A4. $(\mu'_-(0))/(\mu'_+(0)) \equiv \gamma > 1$, où $\mu'_+(0) \equiv \lim_{x \rightarrow 0} \mu'_+(|x|)$ et $\mu'_-(0) \equiv \lim_{x \rightarrow 0} \mu'_-(-|x|)$, ce qui représente l'aversion aux pertes pour les petits enjeux.

Les questions qui se posent alors sont : sur quoi est définie la distribution G et pourquoi penser que le point de référence puisse être stochastique ?

Selon les auteurs, le point de référence capture les croyances de l'individu à propos des futures potentielles conséquences d'un acte. Dès lors, le point de référence ne peut évidemment être fixe puisque par définition, plusieurs conséquences sont possibles au moment d'un choix risqué. Cette conception du référentiel est assez proche de celle développée par Bell (1985), Gul (1991) ou encore Loomes et Sugden (1986) dans leurs théories de l'aversion à la déception. En effet dans ces dernières, les issues d'une loterie sont chacune comparées à une loterie de référence. Sauf que selon ces théories, contrairement à K&R, la loterie de référence est résumée à sa forme dégénérée : l'espérance de la loterie (Bell, 1985; Loomes & Sugden, 1986) ou l'équivalent certain (Gul, 1991). La valeur d'une loterie est alors donnée par :

$$U(L|\bar{r}) = \sum_{i=1}^n p_i v(c_i|\bar{r}) \quad (2.15)$$

Avec v la fonction d'utilité de type Prospect Theory et \bar{r} l'équivalent certain ou l'espérance de la loterie.

En revanche, dans la théorie de K&R chaque issue est comparé aux autres issues de la loterie. En fait, chaque élément de la distribution des issues possibles peut potentiellement servir de point de référence (Goette, Harms, & Sprenger, 2014). La sélection des issues qui serviront de référence dépend des anticipations rationnelles opérées par les agents. Nous pouvons maintenant écrire l'Equation (2.15) de manière discrète :

$$U(F|G) = \sum_{i=1}^n p_i u(c_i) + \sum_{i=1}^n p_i \left[\sum_{j=1}^n \varphi_j \mu(u(c_i) - u(r_j)) \right] \quad (2.16)$$

Selon cette formulation, le niveau de consommation c_1 par exemple sera comparé à toutes les réalisations possibles de r ²³. Il est alors envisageable que l'agent anticipe un sentiment de gain pour l'une des comparaisons et au contraire un sentiment de perte, de déception, pour une autre comparaison. Le même processus est opéré pour tous les autres niveaux de consommations.

La spécification de r , dépend des récentes anticipations rationnelles formées dans le passé par l'agent. Sur cette base, les auteurs définissent deux²⁴ solutions d'équilibres en fonction des anticipations rationnelles des individus. Dans la première spécification appelée « Choice Acclimating Personal Equilibrium » (CPE), le point de référence de chaque loterie est la loterie elle-même (autrement dit, chaque issue d'une loterie est un point de référence de la loterie) : $U(F|F)$. La seconde spécification est appelée « Choice Unacclimating Personal Equilibrium » (UPE). Ici r correspond à la loterie préférée de l'agent : par exemple si l'individu préfère F' à F , alors les espérances d'utilité des deux loteries seront calculées ainsi : $U(F|F')$ et $U(F'|F')$. En considérant ces concepts, on comprend que tout l'esprit du modèle de K&R réside dans l'anticipation des résultats de l'individu qui déterminent alors la manière de percevoir les issues et donc en définitive, conduisent l'attitude face au risque. Par exemple, si l'agent s'attend à une perte de $x\text{€}$, cette perte sera perçue comme neutre et non comme une véritable perte. Dans cette même optique, le « biais du statut quo » observé dans les expériences provient selon K&R du fait que les agents s'attendent à maintenir au moins leur statut quo. Ainsi, dans ces théories

²³ Exemple : $L1(c_1, p_1, c_2, p_2)$ où $c_1 < c_2$ et $\sum_{i=1}^2 p_i = 1$, et avec $w_0 \geq 0$ compris dans c . La distribution des probabilités sur c est notée F et le point de référence r a la distribution G , tel que $r = r_1$ avec la probabilité φ_1 et $r = r_2$ avec la probabilité φ_2 et $\sum_{j=1}^2 \varphi_j = 1$. $U(F|G) = \sum_{i=1}^2 p_i u(c_i) + p_1 [\sum_{j=1}^2 \varphi_j \mu(u(c_1) - u(r_j))] + p_2 [\sum_{j=1}^2 \varphi_j \mu(u(c_2) - u(r_j))]$.

²⁴ En fait K&R en précisent trois mais nous ne parlerons pas du Personal Equilibrium qui est une restriction de l'UPE : la loterie préférée et qui fait office de point de référence est la loterie qui donne la plus grande espérance d'utilité *ex ante*

basées sur l'anticipations des résultats, une perte non attendue est plus douloureuse qu'une perte attendue, tout comme un gain inattendu est perçue comme plus bénéfique qu'un gain attendu.

Prenons un exemple : Un agent ayant une richesse de w a le choix entre F' : s'assurer contre un risque pour un prix de $x_2 = 55\text{€}$ ou bien F : prendre le risque de perdre $x_1 = 100\text{€}$ ou ne rien perdre avec respectivement $p_1 = 0.5$ et $1 - p_1 = 0.5$ de chance : $F(w - x_1, p_1; w, (1 - p_1))$ vs $F'(w - x_2, 1)$, avec $x_1 < x_2 < 0$. En condition UPE où l'individu préfère par exemple maintenir son statut quo w . On sait qu'à cause de la fonction convexe dans les pertes, l'individu est riscophile et devrait donc préférer la loterie plutôt que la certitude de perdre x_2 . Imaginons une condition UPE où, avant de faire son choix, l'individu préfère l'assurance. L'individu ne considère plus x_2 comme une perte. Par contre, dans le cas où malgré sa préférence il envisagerait de ne pas s'assurer, alors il s'exposerait à une perte de $x_1 - x_2$ (si la perte de la loterie se réalise) ou bénéficierait d'un gain de x_2 (si la perte ne se réalise pas). Or à cause de l'aversion aux pertes, l'individu décidera donc de s'assurer étant donné son anticipation initiale. On applique alors la formulation suivante : $U(F|F')$ vs $U(F'|F')$. On supposera, comme le spécifie A3' : la linéarité de la fonction d'utilité pour les stack moyen avec toutefois la présence de l'aversion aux pertes.

Cas 1 UPE : préférence pour le statu quo : Trivial $\Rightarrow U(F|w) > U(F'|w)$

Cas 2 UPE : préférence pour l'assurance : $U(F|F') \text{ vs } U(F'|F')$

$$\begin{aligned} U(F|F') &= p_1(w - x_1) + (1 - p_1)w + p_1[\mu[(w - x_1) - (w - x_2)]] \\ &\quad + (1 - p_1)[\mu[w - (w - x_2)]] \\ &= p_1[w - x_1 - \lambda\eta(x_1 - x_2)] + (1 - p_1)[\eta(w + x_2)] \end{aligned}$$

$$U(F|F') = \frac{1}{2}[(w - 100 - \lambda\eta 45)] + \frac{1}{2}[\eta(w + 55)]$$

$$U(F'|F') = w - x_2 + [0]$$

$$\Rightarrow U(F|F') < U(F'|F')$$

Cas 3 CPE : $U(F|F)$ vs $U(F'|F')$

$$\begin{aligned} U(F|F) &= p_1(w - x_1) + (1 - p_1)w \\ &\quad + p_1[\varphi_1\mu[(w - x_1) - (w - x_1)] + (1 - \varphi_1)\mu[(w - x_1) - w]] \\ &\quad + (1 - p_1)[\varphi_1\mu[w - (w - x_1)] + (1 - \varphi_1)\mu[w - w]] \\ &= p_1[(w - x_1) + \varphi_1\lambda\eta(-x_1)] + (1 - p_1)[w + (1 - \varphi_1)\eta(x_1)] \end{aligned}$$

$$U(F|F) = \frac{1}{2} \left[(w - 100) + \frac{1}{2} \eta\lambda(-100) \right] + \frac{1}{2} \left[w + \frac{1}{2} \eta(100) \right]$$

$$\Rightarrow U(F|F) < U(F'|F')$$

Résumé des spécifications de référentiel répertoriées par Baillon et al (2015)

Nous donnons ici un exemple de situation de choix entre une loterie L1 et L2 auquel les individus ont été exposé durant l'expérience. Nous reportons les différents de références possible selon les spécifications que nous avons évoquées jusqu'à maintenant. Il est intéressant de noter dans ce Tableau 2.2 que différentes spécifications du point de référence peuvent mener au même point de référence, comme c'est le cas de la spécification MaxMin et X at max P.

Exemple : Choice Specific et Prospect specific

Le statu quo est égal à 0, et les référentiels selon le CPE de K&R sont les x et y eux-mêmes :

$$U(X|X) \cong U(Y|Y).$$

Tableau 2.2. Exemple de point de référence possible selon une situation de choix entre L_1 et L_2

L1						L2						Point de référence				
x_1	x_2	x_3	p_1	p_2	p_3	y_1	y_2	y_3	p_1	p_2	p_3	MaxMin	MinMax	X at max p	E V X	E V Y
15	22	40	0.	0.	0.	3	9	34	0.	0.	0.	159	346	159	215	240.1
9	1	8	7	1	2	4	7	6	1	3	6					

3.3. Le(s) gagnant(s) est (sont) ...

Les tentatives d'identification du point de référence parmi plusieurs candidats possibles ont été peu nombreuses. Cette rareté s'explique par le fait que la plupart du temps, les chercheurs font une hypothèse à priori sur ce que peut être le point de référence, puis testent les implications théoriques de ce référentiel à travers une expérience pour valider ou non cette spécification. Toutefois, comme nous l'avons signalé, des points de références différents peuvent toutefois conduire aux mêmes prédictions rendant les conclusions confuses. L'étude de Baillon & al (2017) est sans doute l'étude la plus aboutie qui ait tenté d'identifier le ou les meilleurs candidats au poste de référentiel. Après avoir récolté les données d'une centaine de participants à travers 70 loteries, les auteurs ont traité leurs données selon une modélisation bayésienne hiérarchique. Ils estiment que 30% de leur échantillon utilisaient le statu quo comme référentiel et 30% des sujets utilisaient plutôt la spécification MaxMin (voir Tableau 2.3). On peut noter que selon leurs résultats 20% semblent suivre une spécification à la Koszegi & Rabin malgré la critique que nous avons soulevée à savoir l'importante charge cognitive que suppose ce modèle.

3.4. Les autres points de référence

Nous ne présenterons pas les autres points de références tels que les normes sociales (Gómez-Mejía, Haynes, Núñez-Nickel, Jacobson, & Moyano-Fuentes, 2007; Kahneman, 1992), la

comparaison sociale (Fafchamps, Kebede, & Zizzo, 2015; Fehr & Schmidt, 1999; Lehner, 2000). Le lecteur intéressé par ces problématiques pourra en revanche se tourner vers la thèse de Stommel (2012) qui passe en revue ces types de référentiels.

Tableau 2.3. Médiane et écart type des distribution marginale posterieur des points de références (estimation Bayésienne Hierarchique). Tiré de l'étude de Baillon et al (2015)

	Median	SD
Status Quo	0.30	0.06
MaxMin	0.30	0.06
MinMax	0.10	0.04
X atMax P	0.01	0.02
ExpectedValue	0.06	0.04
Prospect Itself	0.20	0.06

IV. LOCALISATION DU POINT DE REFERENCE EN CONTEXTE DE CHOIX DYNAMIQUE

Nous avons vu dans la précédente section les implications qu'induit un point de référence sur la perception des issues et sur la perception du risque. De nombreuses études révèlent que la formation du référentiel est elle-même influencée par le contexte (Kahneman & Tversky, 1979; Slovic, 2000). Le terme de « contexte » est toutefois très général et il convient d'en spécifier les différents aspects (Cohen, Etner, & Jeleva, 2008). a) Le contexte peut correspondre à l'expérience passée comme par exemple l'évolution des rendements d'actifs dans les marchés financiers (Hirshleifer & Teoh, 2003) ou dans la demande d'assurance (Kunreuther, 1996). b) Le contexte peut également correspondre à certaines caractéristiques sociales présentes chez les pairs. Par exemple, l'individu peut se comparer à ses collègues et plus spécifiquement à leurs performances ou à leur richesses (Fafchamps et al, 2015). c) La manière de présenter les issues financières influence également les décisions comme nous l'avons vu avec le framing effect de K&T.

Dans cette section nous nous focalisons sur l'influence des expériences passées sur les décisions subséquentes. Les expériences passées correspondent à un ou des évènements ayant eu lieu avant le moment d'une décision à un moment t . Ces évènements peuvent être des réalisations d'aléas financiers passées provenant d'investissements antérieures de l'agent. Il peut également s'agir de réalisations passées indépendantes de l'individu comme par exemple l'occurrence de catastrophe naturel ou autre ... Pour analyser ce type d'influence, il convient de considérer les environnements de prise de décision dynamiques. Dans un tel contexte, les réalisations financières passées sont appelées par certains auteurs « les priors » (Weber & Zuchel, 2005). La question qui se pose alors est de localiser le point de référence dans une chaîne d'évènement et potentiellement d'en traquer l'évolution. Le point de référence est-il égal à la richesse initiale ? actuelle ? 0 ? la moyenne des richesses obtenues (Andersen, Harrison, & Rutström, 2006)? L'argument du référentiel est-il forcément une richesse antérieure ou bien peut-il correspondre par exemple à des aspirations de richesses antérieures²⁵ ?

4.1. L'influence du passé dans les contextes de choix dynamiques

4.1.1. Définition

Dans cette section nous considérons une situation de choix dynamique. Edwards (1962) spécifie 4 propriétés indispensables (Haag, 1989) :

- Des séries de décisions sont requises pour atteindre un but.
- Les décisions ne sont pas indépendantes entre elles car les issues des actions présentes à un moment t dépendent des actions précédentes et de leurs réalisations antérieures.

²⁵ Par exemple Buffat et Senn (2015) montrent que des individus ayant antérieurement gagné une certaine richesse pourtant très incertaine prenaient ensuite davantage de risque que des individus ayant antérieurement gagné la même richesse avec une probabilité élevée. Les auteurs montrent par ce biais que l'aspiration antérieure de richesse peut influencer la prise de risque subséquente.

- L'état de l'environnement présent change au cours du temps et dépend des résolutions aléatoires des actions antérieures ou/et d'évènements en dehors du control de l'individu.
- Les décisions sont effectuées en temps réel

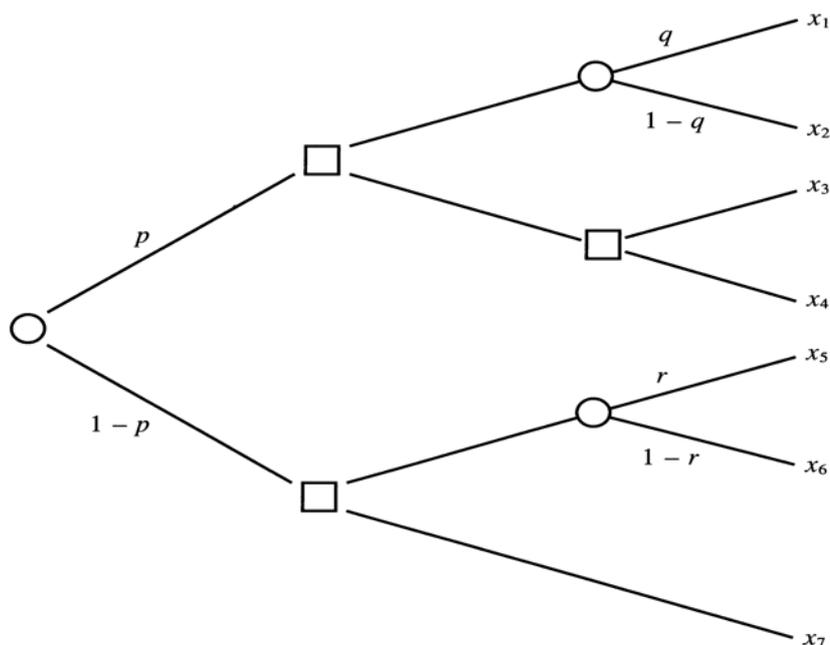


Figure 2.7 Représentation d'une situation de choix dynamique (tirée de Machina, 1989)

Une situation de choix est dite « dynamique » si des choix sont produit après résolution de l'aléa. « [...] a dynamic choice situation will include at least some choices that the individual can postpone until after nature has made at least some of her moves” (Machina, 1989, p.1632). Dans l'arbre de décision présenté dans la Figure 2.7 nous représentons un exemple de situation dynamique impliquant des choix. Le premier élément (le premier rond) signale que l'individu aura à faire un choix (représenté par les carré) après réalisation du premier aléa.

4.1.2. La (in)consistance dynamique et l'hypothèse du conséquentialisme remises en question

Les questions essentielles qui se posent dans un environnement dynamique sont celles-ci : comment sont-évaluées initialement les actions (stratégies) possibles alors que différents nœuds de décision se succéderont et qu'aucun mouvement de la nature n'a encore eu lieu ? Comment

établir une stratégie alors que l'on ne sait pas exactement à l'avance quelles seront les futures décisions à prendre ? Est-ce l'individu devrait changer sa stratégie initiale au milieu de l'arbre de décisions en fonction des réalisations aléatoires de la nature ?

Pour répondre à ces questions en conservant l'hypothèse de rationalité des agents dans le cadre normatif de l'EUT, nous avons besoin de trois ingrédients : la consistance dynamique, le conséquentialisme et l'axiome de réduction des loteries composées (Karni & Schmeidler, 1991; Volij, 1994). Ces trois éléments permettent en fait d'assurer le respect de l'axiome d'indépendance dans un contexte de choix dynamique alors que cet axiome est une propriété des choix en one-shot (à un niveau). Karni et al (1991), montrent qu'un contexte dynamique peut être réduit à sa version réduite en un niveau : 1) Le conséquentialisme suppose que l'agent est seulement intéressé par les résultats terminaux des différentes décisions. Dès lors, l'évaluation d'une stratégie est indépendante de l'histoire du processus de décision qui mène à ces résultats finaux (Hammond, 1988; Machina, 1989). 2) L'axiome de réduction des loteries composées stipule que des loteries à plusieurs niveaux peuvent être réduites à un seul niveau par simple application de calcul de dénombrement que nous avons évoqué dans la partie 1 (voir Figure 1.1).

Nous voyons donc ici que l'individu évalue en amont de l'arbre de décision, l'utilité des distributions finales provenant des différentes stratégies. Une telle analyse, supprime l'aspect temporel des décisions séquentielles, et cela est très important puisque par conséquent aucune dévaluation temporelle appliquée aux issues finales n'est alors supposée ici (« atemporal sequential decision », Karni et al, 1991). Au lieu de cela, chaque issue finale est évaluée conditionnellement à sa probabilité finale d'occurrence. Ce type d'évaluation n'est finalement qu'un raisonnement vers l'arrière (backward induction, Barkan & Busemeyer, 2003) : l'individu évalue les issues finales puis évalue les différents chemins qui mènent à ces issues, lui permettant de savoir à l'avance quoi choisir à chaque nœuds de décision quelles que soient

les réalisations de la nature. C'est en ce sens que l'on parle de consistance dynamique : l'individu ne devrait pas réévaluer ses plans au milieu de l'arbre de décision, mais même s'il le faisait, cela ne changerait rien à sa stratégie initiale puisque la même fonctionnelle de préférence serait appliquée aux distributions restantes²⁶ (Machina, 1989). En d'autres mots, si un investisseur ne prend en compte que les distributions terminales, et se trouve par conséquent indifférent vis-à-vis du processus menant à ces différentes richesses terminales, alors une fonction d'utilité classique définie sur les résultats finaux suffit à capturer les préférences des individus.

Cependant, de nombreuses études montrent que les individus réévaluent leurs plans régulièrement et les modifient faisant dès lors preuve d'inconsistance dynamique. Les agents semblent donc sensibles au processus d'atteinte d'une certaine richesse, conditionnant alors leurs prises de risque subséquente (Barkan & Busemeyer, 1999, 2003). Ce que cette assertion veut dire est que lorsque l'individu atteint une certaine richesse, l'utilité de cette dernière est dépendante du parcours qui a permis son atteinte, signifiant donc une dépendance de l'utilité au passé. Or si cette dépendance s'avère vraie, un individu qui se pense plus riche qu'avant ne prendra pas le même risque qu'un individu qui se pense plus pauvre, et ce même si les deux individus sont supposés identiques et que leur richesse terminales l'est également (Khaw, Glimcher, & Louie, 2017). Le fait de percevoir sa propre richesse actuelle dans la région des pertes ou des gains semble dépendre du parcours de richesse antérieures, et c'est en ce point que l'application d'un modèle dynamique référence-dépendant semble tout indiqué.

4.1.3. L'influence du passé : les priors

Jusqu'à présent nous avons considéré les déterminants de la prise de décision dans le risque (ou de la perception du risque) selon le couple : { *perception des probabilités*,

²⁶ Machina montre dans son article de 1989, que la théorie RDU ne respecte pas la consistance dynamique à cause de la non linéarité des préférences dans les probabilités.

perception des issues financières }. Désormais, nous y ajoutons la composante { *perception des évènements passés* }. Ici nous nous inscrivons dans un champ d'étude qui gagne de plus en plus d'ampleur dans la littérature et qui est souvent résumé sous le terme « History Preference Dependent Theory » ou encore « History Dependent Risk Attitude » (Dillenberger & Rozen, 2015; Tserenjigmid, 2015). Une représentation intertemporelle générale des comportements de choix d'un individu doit alors inclure dans l'idéal les termes qui représentent toutes les dimensions d'un comportement intertemporel (Haag, 1989) :

$$\begin{aligned}
 & \text{(choix actuel)} \\
 & = f \left\{ \begin{array}{l} \text{(niveau actuel, passé et futurs des variables exogènes),} \\ \text{(effets de l'histoire passé pertinente),} \\ \text{(cumul de l'expérience la plus récente dans un état),} \\ \text{(prendre en compte l'hétérogénéité), (Persistence de l'habitude)} \end{array} \right\}
 \end{aligned}$$

a. Impact sur la perception des probabilités

Dans un contexte de choix dynamique, l'influence du passé peut être très prégnante sur le comportement décisionnel subséquent. La demande d'assurance est par exemple particulièrement sensible à ces évènements antérieurs malgré l'indépendance de l'occurrence des pertes à travers les périodes successives (McClelland, Schulze, & Coursey, 1993; Papon, 2004). Ici, on observe deux types d'effets sur la demande d'assurance ultérieure à l'occurrence d'une perte (ou d'une catastrophe ou d'un dommage). Soit à cause du biais de disponibilité révélé par K&T, les individus finissent par surestimer la probabilité d'un prochain désastre, ce qui conduit à une augmentation de la demande d'assurance. Soit sous l'effet du « gambler's fallacy », les individus considèrent que la probabilité qu'un tel évènement se reproduise à nouveau est très faible et la sous-estiment (Ayton & Fischer, 2004), conduisant à une diminution de la demande d'assurance. Sous EUT, de tels comportements ne peuvent être expliqués étant donné le caractère indépendant des évènements. Dans ce modèle, une perte

mène simplement à la diminution de la richesse qui peut alors modifier l'aversion au risque en altérant la perception des gains et des pertes mais n'est pas censé altérer la perception des probabilités. Dès lors, si l'on veut capturer l'influence des événements passés sur la décision actuelle, il faut relâcher l'axiome d'indépendance au contexte (Cohen et al., 2008)

b. Impact sur la perception des richesses finales prospectives

Une telle relax peut produire des comportements déviants largement des prédictions normatives de l'EUT, en particulier si l'on considère une fonction d'utilité de type PT. Il est possible dans ce cadre, qu'une même richesse finale prospective soit perçues dans le domaine des pertes ou dans le domaine des gains selon le parcours de richesses antérieure de l'individu. Or, selon la PT, nous avons vu que le traitement différent appliqué aux gains et aux pertes conduit à des attitudes face au risque radicalement différentes. La difficulté réside en fait dans l'application de la fonction d'utilité de type PT à un contexte dynamique inter-temporel alors qu'à l'origine la PT fut développé dans des contextes statiques en « one shot » (Barberis, Huang, & Santos, 2001). Selon K&T, l'application aux contextes dynamiques devait attendre afin d'accumuler suffisamment de preuves et de régularités à propos de la façon dont les individus traitent les séquences de pertes et de gains (Tversky & Kahneman, 1981). Depuis, quelques chercheurs se sont penchés sur ce problème et ont révélé le pattern suivant (Baucells et al., 2011; Chen & Rao, 2002; Cohen et al., 2008; Franken, Georgieva, & Muris, 2006; Post et al., 2008; Thaler & Johnson, 1990) :

L' « House Money Effect » (HME) : Après un gain, l'aversion au risque diminue.

Le « Break Even Effect » (BEE) : Après une perte, l'aversion au risque diminue.

Le « Make Peace with Loss » (MPwL) : Après une perte, l'aversion au risque n'est pas altérée.

Expérience de Thaler et Johnson (1990) (1)

HOUSE MONEY EFFECT

Problème 1 : (en deux étapes)

Vous venez de gagner \$15. Maintenant choisissez entre :

$$L_1(\$4.5, 0.5 ; -\$5.5, 0.5) [81\%]$$

$$L_2(\$0, 1)$$

Problème 1' : (équivalent au problème 1 en termes de richesses finales mais en une étape)

Choisissez entre :

$$(A') L_1'(\$19.5, 0.5 ; \$10.5, 0.5) [40\%]$$

$$(B') L_2'(\$15, 1)$$

Les résultats montrent que malgré des richesses finales identiques, le bonus de \$30 du problème 1 qui fait office de « prior gain » semble mener les individus à prendre davantage de risque.

Expérience de Thaler et Johnson (1990) (2)

BREAK EVEN EFFECT

Problème 2 : (en deux étapes)

Vous avez perdu \$2.25. Maintenant choisissez entre :

$$L_1(\$2.25, 0.5 ; -\$2.25, 0.5) [72\%]$$

$$L_2(\$0, 1)$$

Problème 2' : (en une étape)

(A') $L_1'(\$4.5, 0.5 ; \$0, 0.5)$ [78%]

(B') $L_2'(\$ - 2.25, 1)$

Expérience de Thaler et Johnson (1990) (3)

MAKE PEACE WITH LOSSES

Problème 3 : (en deux étapes)

Vous avez perdu \$7.5. Maintenant choisissez entre :

$L_1'(\$2.25, 0.5 ; -\$2.25, 0.5)$ [42%]

$L_2'(\$0, 1)$

Problème 3' : (en une étape)

Choisissez entre :

(A') $L_1'(-\$5.25, 0.5 ; -\$9.75, 0.5)$ [70%]

(B') $L_2'(\$ - 7.5, 1)$

Les problèmes 2 et 3 illustrent la nature ambivalente d'un « prior loss » sur l'attitude face au risque subséquente. Le problème 2 montre qu'après une perte les individus prennent moins de risque. Le problème 3 montre qu'après une perte les agents prennent davantage de risque. La différence entre les 2 problèmes est que le problème 3 donne l'opportunité de couvrir la perte initiale, stimulant ainsi la prise de risque.

La perte semble pouvoir provoquer deux types de comportements : soit l'individu renonce à prendre davantage de risque, soit l'individu au contraire, augmente sa propension à prendre des risques. Thaler et Johnson, remarquent que cet effet différencié, provenant pourtant de la même cause, est fonction des issues monétaires de la future loterie. Si les issues permettent de combler

la perte vécue, alors les individus auront tendances à miser. A l'inverse, si la loterie ne permet pas une telle couverture, alors les individus renoncent à parier. Ce pattern est néanmoins assez difficile à capturer, c'est pourquoi certains chercheurs comme Barberis & al (2001) ou Arkes & al (2010) ne parviennent à modéliser soit que le BEE soit que le MPwL mais jamais les deux en même temps.

4.2. Référentiel adaptatif et prise en compte du passé

4.2.1. Explication du pattern observé par Thaler et Johnson

En ce qui concerne le caractère potentiellement adaptatif du référentiel, il y a deux cas possibles : 1. Un référentiel parfaitement adaptatif et donc sans mémoire du passé 2. Un référentiel pas ou partiellement adaptatif et donc avec mémoire du passé. Commençons par définir une distinction plus précise de ces deux cas de figure.

Un référentiel parfaitement adaptatif est un référentiel qui s'adapte à un élément de la situation actuelle de l'individu. Autrement dit, n'importe quel point de référence envisagés dans la section 3 peut être envisagé comme référentiel du contexte dynamique à condition d'y intégrer les gains et pertes vécus dans la situation. Un référentiel parfaitement non-adaptatif est un référentiel au contraire toujours égal à un élément de la situation initiale vécu par l'individu.

Point de référence parfaitement adaptatif

Illustrons formellement le cas d'un point de référence parfaitement adaptatif et le cas d'un référentiel non-adaptatif. Pour cela nous devons considérer un jeu sur plusieurs périodes. Ici nous nous limiterons à deux périodes, t_1 et t_2 . Nous présumons que l'individu possède une richesse initiale w_0 et qu'il est confronté sur les deux périodes à l'évaluation d'une même loterie L_1 . Nous pouvons considérer n'importe quelle spécification du point de référence, cela ne modifiera pas le raisonnement. Nous maintiendrons donc la formulation générale R_0 pour

désigner le référentiel à la période 1. Comme l'indique l'équation 2.2, l'évaluation d'une issue de la loterie à la période 1 devrait s'opérer ainsi :

$$u(z_{t_1}) = \begin{cases} u^+ \left((w_0 + x_{t_1}) - R_0 \right) & \text{si } (w_0 + x_{t_1}) \geq R_0 \\ u^- \left(R_0 - (w_0 + x_{t_1}) \right) & \text{si } (w_0 + x_{t_1}) < R_0 \end{cases}$$

Ce qui nous intéresse c'est désormais l'évaluation des issues à la période 2 d'un référentiel parfaitement adaptatif. A la période 2, la richesse de l'individu est désormais de $w_0 + x_{t_1}$ avec x_{t_1} la réalisation de l'aléa à la période 1. Comme le référentiel est parfaitement adaptatif, il évolue similairement à la richesse actuelle de l'individu. Dès lors le référentiel à la période 2 est de $R_0 + x_{t_1}$. L'évaluation d'une issue de la loterie à la période 2 est donc :

$$u(z_{t_2}) = \begin{cases} u^+ \left((w_0 + x_{t_1} + x_{t_2}) - (R_0 + x_{t_1}) \right) & \text{si } (w_0 + x_{t_1} + x_{t_2}) \geq R_0 + x_{t_1} \\ u^- \left((R_0 + x_{t_1}) - (w_0 + x_{t_1} + x_{t_2}) \right) & \text{si } (w_0 + x_{t_1} + x_{t_2}) < R_0 + x_{t_1} \end{cases}$$

Dès lors l'évaluation d'une issue à la période t_2 devrait être identique à l'évaluation opérée durant la période t_1 puisqu'en simplifiant l'expression de la période t_2 on retrouve la même expression que la période t_1 . En effet, puisque $x_{t_1} = x_{t_2}$ alors $u(z_{t_1}) = u(z_{t_2})$.

Par corollaire, l'attitude face au risque n'évolue pas quelle que soit l'augmentation ou la diminution de richesse de l'individu lorsque le référentiel est parfaitement adaptatif. Cette prédiction peut être en accord avec l'effet « Make Peace with Loss » puisqu'un référentiel parfaitement adaptatif implique que l'agent ne prendra pas davantage de risque après une perte. Toutefois, il n'est pas capable d'accommoder le pattern HME et BEE qui suppose une augmentation de la prise de risque après une perte ou après un gain.

Point de référence parfaitement non-adaptatif

On considère maintenant le cas d'un référentiel non-adaptatif. L'analyse est plus délicate car l'impact d'un gain ou d'une perte sur l'attitude face au risque subséquente dépend de nombreux éléments comme la localisation du référentiel par rapport aux richesses finales prospectives, les issues engagées dans une situation risquée et bien évidemment des caractéristiques de la fonction d'utilité et de la fonction de pondération des probabilités. Nous présenterons ici l'analyse qui est habituellement rapportée dans les publications scientifiques sur le sujet (Eil & Lien, 2014; Post et al., 2008).

Tout d'abord on maintiendra l'idée d'une fonction concave dans les gains et convexe dans les pertes. La fonction $\varphi(p)$ et sa forme n'ont ici pas d'importances concernant les conclusions auxquelles nous allons aboutir. Supposons l'existence d'une loterie L engageant un certain nombre d'issues positives (gains) et négatives (pertes). Fixons un point de référence quelconque et une richesse actuelle quelconque.

Analysons dans un premier temps l'impact de gains lorsque l'ensemble (ou une grande part) des richesses finales sont situées dans la région des gains. La fonction d'utilité est concave et implique donc l'aversion pour le risque. Toutefois, à mesure que l'individu gagne, sa richesse actuelle augmente et par conséquent les richesses finales prospectives sont évaluées sur une section de plus en plus linéaire (de moins en moins concave) de la fonction d'utilité. En termes d'analyse marginale, nous pourrions dire que le rapport entre l'utilité marginale d'un gain et la désutilité marginale d'une perte diminue. Dès lors, à mesure que l'agent gagne, il devient de moins en moins averse au risque et tend vers la neutralité. Cette analyse peut aisément s'accorder avec l'HME.

Considérons maintenant l'impact d'une perte lorsque l'ensemble (ou une grande part) des richesses finales sont situées dans la région des pertes. La fonction d'utilité est convexe et implique donc la recherche de risque. Toutefois, à mesure que l'individu perd, sa richesse

actuelle diminue et par conséquent les richesses finales prospectives sont évaluées sur une section de plus en plus linéaire (de moins en moins convexe) de la fonction d'utilité. A nouveau, le rapport entre l'utilité marginale d'un gain et la désutilité marginale d'une perte diminue. Dès lors, à mesure que l'agent perd, il devient de moins en moins chercheur de risque et tend vers la neutralité. Cette analyse peut aisément s'accorder avec le MPwL.

Ici nous nuancions le propos. On peut également capturer le HME en supposant que les issues de la loterie se situent dans la région des pertes. Dans ce cas, à mesure que l'individu gagne, les richesses finales sont évaluées sur une section de plus en plus convexe de la fonction d'utilité générant ainsi une recherche pour le risque de plus en plus forte. De même, on peut accommoder le MPwL, en supposant que les issues sont situées dans la région des gains. Dans ce cas, à mesure que l'individu perd, les richesses finales sont évaluées dans une section de plus en plus concave, générant une aversion pour le risque de plus en plus prononcée.

Jusqu'à présent nous n'avons donné aucune explication vis-à-vis du BEE. L'analyse classique consiste à considérer le passage de la fonction concave à la fonction convexe. Nous avons évoqué dans ce chapitre section II que l'aversion au risque était maximisée lorsque les issues finales sont réparties autour du point de référence. En revanche, si l'individu continue de perdre et que le référentiel ne s'adapte pas, les issues finales finiront par être évalués dans la section la plus convexe de la fonction d'utilité. Dès lors, à cause du kink dans la région des pertes à proximité du référentiel, l'utilité marginale d'un gain sera bien plus élevée que l'utilité marginale d'une perte induisant donc la recherche de risque après une perte.

4.2.2. Théorie de l'Utilité Espérée et priors

Il n'est pas complètement juste de considérer que l'EUT ne prend pas en compte les priors et qu'elle n'est donc pas capable d'expliquer les comportements révélés par Thaler et Johnson. Sous EUT, les gains et pertes antérieures affectent la prise de décision subséquente car ils modifient la richesse actuelle, ce qui à son tour, détermine l'aversion au risque de l'agent selon

les propriétés des fonctions d'utilités considérés (Weber & Zuchel, 2005). Une diminution de l'aversion au risque après un gain est cohérent avec l'EUT si la spécification de la fonction d'utilité implique une diminution de l'aversion absolue au risque avec l'augmentation de la richesse provoqué par le gain antécédent (fonction Decreasing Absolute Risk Aversion = DARA). De même, l'aversion au risque peut diminuer après une perte si on spécifie une fonction d'utilité exhibant la propriété selon laquelle l'aversion absolue augmente avec la richesse (fonction Increasing Absolute Risk Aversion = IARA).

Toutefois, ces fonctions d'utilités ne sont pas supposées changer en fonction du contexte. Les préférences d'un individu ne peuvent donc pas être capturées à la fois par une fonction DARA et IARA selon ce qui nous arrange pour expliquer son comportement. C'est donc en ce sens que l'EUT ne peut rendre compte des résultats de Thaler et Johnson (1990).

4.2.3. Equity Premium Puzzle (EPP) vs Disposition Effect (DE)

Ici nous verrons comment le recours à la spécification d'un référentiel parfaitement adaptatif ($w_t = R$) permet d'expliquer l'EPP, puis comment un référentiel parfaitement non adaptatif permet d'expliquer le DE ($w_0 = R$). Ces deux exemples, interrogent sur le degré de liberté accordé aux modèles de préférence référence-dépendant qui, sans spécification rigoureuse du référentiel, peuvent capturer de nombreux phénomènes de façon purement ad-hoc, mais prodiguent des prédictions *ex ante* contradictoires.

a. Equity premium puzzle

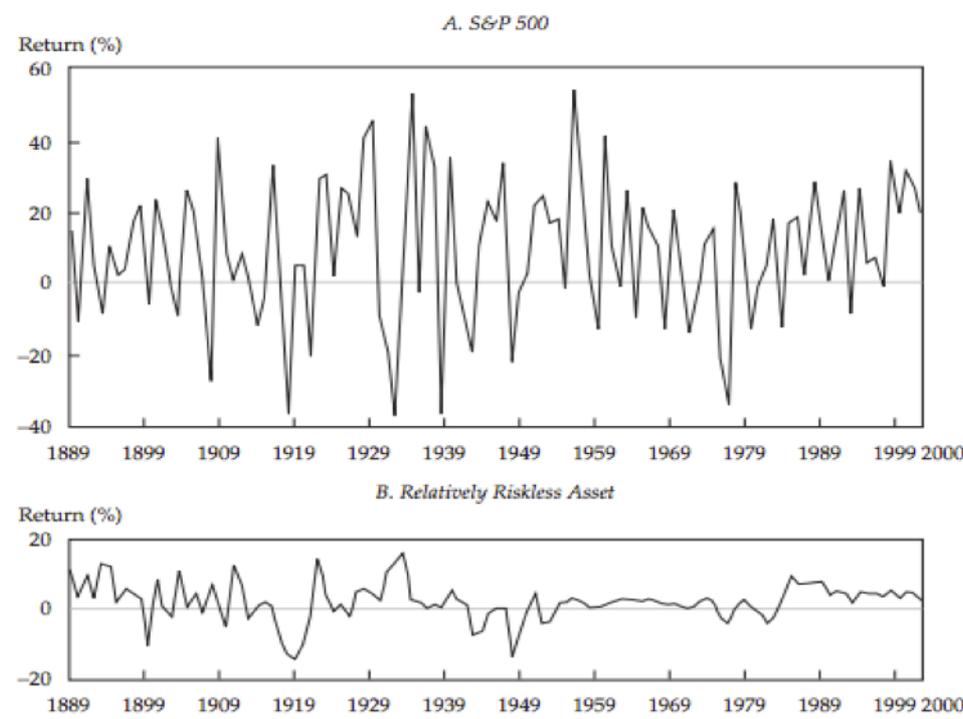


Figure 2.8. En haut : rendements réels au S&P 500 entre 1889 et 2000. En bas : rendements réels obtenu pour des actifs moins risqués sur la même période.

L'EPP a été mis en évidence par Mehra & Prescott (1985). Ces derniers montrent que le rendement des actifs financiers durant le dernier siècle était beaucoup plus élevé que le rendement des obligations d'Etat²⁷. On pourrait considérer qu'étant donné la variance plus élevée du rendement des actions par rapport à la variance des obligations, les agents réclament naturellement une protection (une prime de risque calculée comme la différence entre le rendement moyen des actions par rapport au rendement moyen des obligations) importante pour supporter le risque additionnel. Pour comprendre au moins grossièrement le problème, il faut se demander quel niveau d'aversion au risque doit-on supposer pour que le rendement des

²⁷ Aux Etats-Unis, le rendement moyen des actions entre 1889 et 2000 était de 7.9% quand le rendement moyen des obligations était de 1% pour la même période. La différence entre ces deux rendements, 6.9 points, représente la prime de risque (données actualisées par Mehra, 2003).

actions soit si élevé par rapport à celui des obligations ? Sheffrin & Statman (1985) montre qu'il faut supposer un niveau d'aversion au risque relatif absurdement élevé pour expliquer cette prime de risque. Avec un niveau d'aversion au risque convenable on estime que le rendement ajusté au risque²⁸ des actions devrait être supérieur d'un point par rapport à celui des obligations, or il est beaucoup plus élevé dans la réalité, d'où le puzzle²⁹ (Mehra, 2003).

Une tentative d'explication célèbre et qui s'intègre parfaitement dans notre analyse du point de référence, a été proposée par Bernatzi & Thaler (1995) à travers le concept de « Myopic Loss Aversion ». Cette résolution suppose deux ingrédients : l'aversion aux pertes et une fréquence d'évaluation de la performance du portefeuille élevée. L'ensemble des richesses prospectives sont évaluées par rapport à la richesse obtenue lors de la dernière évaluation du portefeuille. Autrement dit, le point de référence est parfaitement actualisé à chaque fois que l'individu examine les performances de son portefeuille. Nous avons vu précédemment qu'un référentiel parfaitement adaptatif impliquait que les issues finales seraient toujours évaluées de la même façon. Or, le fait que le référentiel s'adapte parfaitement annule par conséquent l'effet de richesse de l'individu. Rappelons que dans l'EUT, si l'on suppose une fonction d'utilité DARA l'aversion au risque absolue diminue avec la richesse. Ici, puisque le référentiel s'adapte parfaitement, l'individu est toujours autant averse au risque quelle que soit l'évolution de sa richesse.

²⁸ Le risque en finance n'est pas mesuré par la simple variance d'un rendement d'une action. En fait il est mesuré, entre-autre, par le paramètre $\beta = \frac{cov(r_a, r_b)}{var(r_b)}$, avec r_a le rendement du portefeuille, r_b le rendement de l'indice du marché. Ce paramètre mesure la volatilité de l'action par rapport à la volatilité du marché. Autrement dit, combien de fois un titre est-il plus ou moins variable que le marché. Il s'agit bien d'un indicateur du risque normalisé, car par exemple si le marché est à la baisse, le cours du titre peut baisser par exemple deux fois plus si $\beta = 2$. Un rendement ajusté au risque établit donc le rendement par unité de risque du titre : $r_a^{adjusted} = \frac{r_a}{\beta}$.

Pour capturer le phénomène d'Equity Premium puzzle, il faut de plus supposer un degré d'aversion au risque très élevé, ce qui peut être accommodé en supposant que les individus sont averses aux pertes (nous avons vu que l'aversion aux pertes renforce l'aversion au risque).

Prenons un individu pauvre, $w_0 = 0$, confronté à une loterie impliquant des montants tels que $L(200, 0.5 ; -50, 0.5)$. Il rejette la loterie à cause de son aversion aux pertes et préfère un équivalent certain de 50€ par exemple. Imaginons que l'individu soit maintenant plus riche de 1000 euros. Normalement, l'équivalent certain devrait augmenter car son aversion diminue avec la richesse. En effet, si le point de référence ne s'adapte pas et reste à 0, les issues sont alors intégrées à la richesse actuelle de l'individu : $L(1200, 0.5 ; 950, 0.5)$. Étant donné que le coefficient d'aversion aux pertes ne s'applique pas ici, l'individu devrait considérablement être moins averse au risque, la concavité de la fonction étant moins importante dans la section considéré ici que dans la région proche du point de référence. Cependant, si l'agent adapte parfaitement son point de référence à sa richesse actuelle, alors l'évaluation de la loterie porte à nouveau sur les issues $+200 - 50$ équiprobable, provoquant à nouveau une forte aversion aux pertes et donc un équivalent certain identique au moment où l'individu était pauvre.

L'explication du puzzle implique donc de supposer que les agents sont averses aux pertes et initialisent leurs référentiels au niveau de richesse obtenu lors du dernier examen du portefeuille. Cette réinitialisation rend effective l'application de l'aversion aux pertes pour n'importe quel niveau de richesse actuelle de l'individu et donc une forte aversion aux risques même pour des montants modérés voire petits, poussant les individus à demander une forte protection quand le risque augmente (Barberis & Huang, 2008).

b. Le disposition effect

Sherfrin et Statman (1985) décrivent la tendance selon laquelle les agents tendent à maintenir dans leurs portefeuilles des actifs perdants par rapport à leurs prix d'achat initial, alors qu'ils

vendent plus rapidement les actifs gagnants. Ce phénomène est appelé l'effet de disposition. Formellement, le disposition effect est visible quand la Proportion de Gain Réalisées (PGR) est supérieure à la Proportion de Pertes Réalisées (PLR) :

$$PGR = \frac{\text{Gains réalisés}}{\text{Gain réalisés} + \text{Gain potentiels}} > \frac{\text{perte réalisées}}{\text{Perte réalisées} + \text{Perte potentielles}} = PLR$$

Les gains réalisés sont les gains obtenus en ayant vendu les actifs gagnant (où le prix de vente P_v est supérieur au prix d'achat P_a), et les gains potentiels sont les actifs qui ne sont pas encore vendus mais dont le prix actuel est supérieur au prix d'achat. Concernant les pertes réalisées, $P_v < P_a$, et les pertes potentielles, P_p , décrivent les actifs dont le prix actuel est inférieur aux prix d'achat mais ne sont pas encore vendu (Odean, 1998).

Comment expliquer ce phénomène ? Selon Bernatzi & Thaler (1985) le point de référence s'adapte parfaitement au statut quo de la dernière évaluation du portefeuille. Si l'on applique la même spécification on constate facilement que les individus devraient vendre aussi rapidement les actifs gagnants que les actifs perdants : puisque les mêmes issues prospectives sont évaluées de part et d'autre du point de référence, et donc dans la région la plus concave de la fonction d'utilité. A l'inverse si le point de référence ne s'adapte pas et reste égale au prix d'achat P_a que constate-t-on ?

Cas 1 : Si $R = P_a > P_{actuelle} \rightarrow u'(+P_{actuelle}) > u'(-P_{actuelle})$ (section convexe)

Cas 2 : Si $R = P_a < P_{actuelle} \rightarrow u'(+P_{actuelle}) < u'(-P_{actuelle})$ (section concave)

Dans le cas 1, l'utilité marginale d'un gain est plus grande que la désutilité marginale d'une perte quand l'individu juge sa richesse actuelle dans la région des pertes, grâce au kink proche du point de référence. Ici un gain (i.e. une hausse du prix) devient une opportunité pour l'agent

de couvrir sa perte et donc de supprimer la « douleur », procurant alors à ce gain potentiel une utilité plus conséquente qu'une perte de même ampleur. L'investisseur choisit alors de maintenir son actif perdant dans l'espoir de recouvrer son ancienne perte, se comportant comme un riscophile (break even effect).

Dans le cas 2, l'utilité d'un gain potentiel est plus petite que la désutilité d'une perte potentielle et décide donc de vendre rapidement pour se protéger d'une perte éventuelle. Toutefois sa propension à vendre rapidement devrait ralentir à mesure qu'il est riche puisque l'écart entre utilité marginale et désutilité marginale diminue (house money effect).

Nous voyons à travers ces deux exemples qu'il faut supposer deux systèmes d'adaptation différent du référentiel pour accommoder ces phénomènes. Donc, la PT est capable de capturer les phénomènes de façon ad hoc. En revanche, aucun des référentiels n'est en mesure de produire les bonnes prédictions pour capturer l'ensemble de ces phénomènes.

4.3. Les preuves expérimentales démontrant l'actualisation du référentiel

Depuis une quinzaine d'année, certains chercheurs ont tenté de traquer l'adaptation du référentiel à travers des méthodes que l'on peut qualifier de « directes » ou d'« indirectes ». La méthode directe consiste à révéler le point de référence des individus en leurs demandant de le spécifier eux-même. La méthode indirecte consiste, quant à elle, en une suite de choix permettant ensuite d'inférer la valeur du point de référence. Dans cette section nous passons en revue les expériences qui ont tenté de capturer la manière dont les agents traitent des séquences d'évolution de richesse dans des contextes dynamiques.

4.3.1. La méthode indirecte de Gneezy (2005)

Dans cette expérience des participants sont exposés à l'évolution du prix d'un actif sur plusieurs périodes. L'ensemble des sujets devaient décider à chaque nouvelle période s'ils souhaitaient vendre ou maintenir leur actif. Les sujets étaient tous exposés aux mêmes séquences de prix.

L'idée de Gneezy consiste à vérifier ou infirmer certaines hypothèses en fonction du comportement de vente de l'individu.

Si l'individu vend à un certain prix alors qu'il avait décidé antérieurement de maintenir son actif à ce même prix, alors cela est inconsistant avec l'hypothèse de consistance dynamique de l'EUT. Gneezy établit deux autres hypothèses à propos cette fois du point de référence. Selon lui, le référentiel peut être égal au prix initial ou à un pic de richesse obtenu antérieurement. Pour le vérifier, si le pic de richesse est le référentiel, il devrait observer que les individus ne vendent pas tant que le prix de l'actif est en dessous d'un ancien pic de prix. Si c'est le prix initial qui est le référentiel, il devrait observer que les individus ne vendent pas tant que le prix est en dessous du prix initial (conformément au disposition effect). On constate que la première hypothèse est incluse dans la seconde dans la mesure où une vente est cohérente avec les deux hypothèses lorsque le prix actuel est au-dessus du prix initial et au-dessus du pic du prix. Cependant, la réciproque n'est pas vraie, il existe en effet un intervalle entre le pic du prix et le prix initial dans lequel si les individus ne vendent pas alors on peut conclure que le pic de richesse est effectivement le référentiel (tout en s'assurant que les individus par contre vendent bien à un prix égal ou supérieur au pic de prix antérieur).

Gneezy semble conclure que le pic du prix fait office de référentiel, pourtant l'analyse de ses résultats sont tout aussi cohérents avec un référentiel égal au prix initial. Pour départager les deux prédictions, Gneezy utilise une méthodologie proposée par Selten (1991). Cette méthodologie consiste à mesurer l'ensemble des prédictions possible d'une théorie « area theory », dénotée a , puis de comparer le nombre de prédictions possibles au nombre de prédictions correctes, c'est-à-dire réellement observées dans une expérience « hit rate : r ». La combinaison des deux facteurs donne une idée de la précision et de la performance d'une théorie, appelé le « succès prédictif : m » :

$$m = r - a$$

Une théorie qui peut prédire n'importe quel ensemble de résultats est également une théorie qui manque de précision et la rend donc inutile. Gneezy observe à travers des simulations opérées après l'expérience, que la zone théorique est très élevée lorsque l'on considère le prix initial comme point de référence alors que cette zone est beaucoup plus faible lorsqu'on envisage un pic de prix antérieure comme référentiel. Le pourcentage de prédictions correctes étant similaire selon ces deux hypothèses, mais la zone prédictive étant plus faible pour l'hypothèse du *pic* de richesse alors on peut conclure que l'hypothèse du pic de prix est plus précise que l'hypothèse du prix initial.

4.3.2. La méthode directe de Baucells et de Arkes

Les mesures de Baucells et al (2011) et de Arkes et al (2008, 2010) sont assez similaires puisque les méthodes ont la particularité de passer par un questionnaire. Toutefois, leurs mesures diffèrent sur un point : Arkes et al s'interrogent sur l'adaptation du point de référence après la hausse (gain) ou la baisse (perte) du prix d'un actif. Dès lors la question posée par les auteurs aux participants ne porte pas sur la valeur du référentiel en tant que tel, mais plutôt sur l'adaptation du référentiel. Baucells et al (2011) s'interrogent quant à eux directement sur la valeur du référentiel suite à l'exposition d'une séquence de prix d'un actif.

a. La méthode de Arkes et al (2008, 2010)

Arkes et al (2008, 2010) développent un contexte expérimental en lien avec la finance. Il est dans un premier temps demandé aux participants d'imaginer qu'ils ont acheté il y a deux mois un actif au prix de P_0 , puis que celui-ci a finalement augmenté (ou diminué) le mois suivant pour arriver à un prix p_1 . Il est ensuite demandé aux sujets : « A quel prix voudriez-vous vendre cet actif de sorte à être aussi heureux (triste) que la première fois qu'il a augmenté (diminué) ? ». La question est assez alambiquée mais s'explique par la logique suivante : si le prix de l'actif

est p_1 et que le précédent référentiel est R_0 alors la différence entre p_1 et R_0 devrait être la même qu'entre le nouveau prix p^* annoncé par le participant et le nouveau point de référence R^* . Dès lors si :

$$p_1 - R_0 = p^* - R^* \Rightarrow \Delta R = R^* - R_0 = p^* - p_1$$

Avec cette égalité les auteurs sont en mesure d'estimer l'adaptation du référentiel ΔR . Les auteurs font l'observation que l'adaptation du référentiel après un gain ou après une perte serait asymétrique³⁰ : l'adaptation du point de référence est significativement plus grande après un gain qu'après une perte. Si le prix initial est de 30€ et qu'il augmente le mois suivant de 6€ pour atteindre 36€ les sujets répondaient en moyenne qu'un prix de 40.24€ produirait le même gain d'utilité que la première augmentation du prix. Autrement dit un gain de +4.24€. Etant donné que la fonction d'utilité est supposée inchangée quel que soit la richesse par les auteurs, on peut établir que le point de référence s'est adapté de 4.24 (40.24€ - 36€) après un gain de 6€ (36€ - 30€). Après une perte d'amplitude équivalente, les participants déclarent qu'un prix de 21.49€ serait aussi douloureux que la première perte. Autrement dit après une perte de 6€ le point de référence s'adapte de 2.51€.

Ces résultats vont clairement dans le sens de l'hypothèse de Thaler à propos du mental accounting : Après un gain, le point de référence s'adapte à la nouvelle richesse si bien que pour obtenir la même augmentation d'utilité il suffit d'obtenir le même gain que précédemment. A l'inverse si le référentiel ne s'adaptait pas il faudrait un gain plus important pour générer la même augmentation d'utilité. En ce qui concerne les pertes, le point de référence ne s'adapte pas (ou peu). Dans ce cas, cela implique qu'une future perte est moins douloureuse que la

³⁰ Après avoir contrôlé notamment le fait que l'adaptation du référentiel n'était pas dû à une attente de l'individu vis-à-vis du futur mouvement du prix de l'actif comme l'aurait pourtant spécifié Koszegi & Rabin (voir la méthode employée par Arkes et al, 2008, p73-74).

première perte de même amplitude, ce qui est une implication de l'hypothèse d'intégration des pertes selon Thaler.

b. La méthode de Baucells et al (2011)

La méthode de Baucells et al (2011) est encore plus explicite que la précédente. Dans celle-ci, les individus sont exposés séquentiellement à l'évolution sur n période du prix $y_i, i = 1, \dots, n$ d'un actif. A la fin de cette séquence, il est demandé aux participants : « à quel prix voudriez-vous vendre cet actif de sorte à n'être ni heureux ni triste de le vendre ? ». Cette question renvoie directement à la définition même du point de référence qui est un niveau de richesse (ou un prix) pour lequel l'individu ne se sent ni en perte ni en gain, c'est-à-dire un état neutre. Les auteurs ont alors élaboré 60 séquences différentes avec un nombre de période pouvant aller de 3 à 10. Le nombre important de séquences élaborées est justifié par le fait que les auteurs cherchent à discerner ce qui, dans une chaîne d'évènements, influence le futur point de référence. Ils identifient au préalable cinq points de références possibles : le prix initial, le prix final, le prix moyen, le plus haut prix vécu (pic du prix) et le plus bas prix de la séquence. De plus, ils ajoutent quelques séquences dans l'optique de mesurer si l'ordre d'apparition des prix (le pattern d'évolution du prix) à une influence sur le futur point de référence. Notons deux suppositions que les auteurs précisent : 1) Ils espèrent que le point de référence sera bien un niveau de prix et non autre chose dans la mesure où ils présentent aux sujets des prix d'actifs (Baucells et al, 2011). 2) Les auteurs supposent que seule les prix passés influencent la formation du point de référence au temps t .

Les auteurs montrent que le point de référence au moment t est fonction de l'ensemble des anciens prix avec des poids plus importants pour le prix initial et le prix final et une sous-pondération des prix intermédiaires. Par ailleurs, ils montrent également qu'il ne semble pas

exister d'effet d'ordre, leur permettant de rejeter une modélisation de l'actualisation du point de référence sous forme récursive.

4.3.3. Quelques Mesures complémentaires

Aujourd'hui des chercheurs emploient ces méthodologies ou des variantes afin d'affiner les recherches portant sur le traitement et le jugement des individus des séquences d'évènements passés et de leurs influences sur les futurs comportements (Barkan & Busemeyer, 1999; Chen & Rao, 2002; Cohen et al., 2008; Nolte & Schneider, 2018). Les mesures portent notamment sur la satisfaction d'obtenir une certaine richesse (ou un certain prix d'actif) à la fin d'une séquence d'évolution de cette richesse. Il a notamment été observé que des agents ayant vécu le parcours de prix d'un actif ascendant puis descendant « up-down » sont moins satisfaits que les agents ayant vécu le parcours « down-up », en ayant pris soins de maintenir égal la prix initiale et finale entre les groupes (Grosshans & Zeisberger, 2018) (voir Figure 2.9).

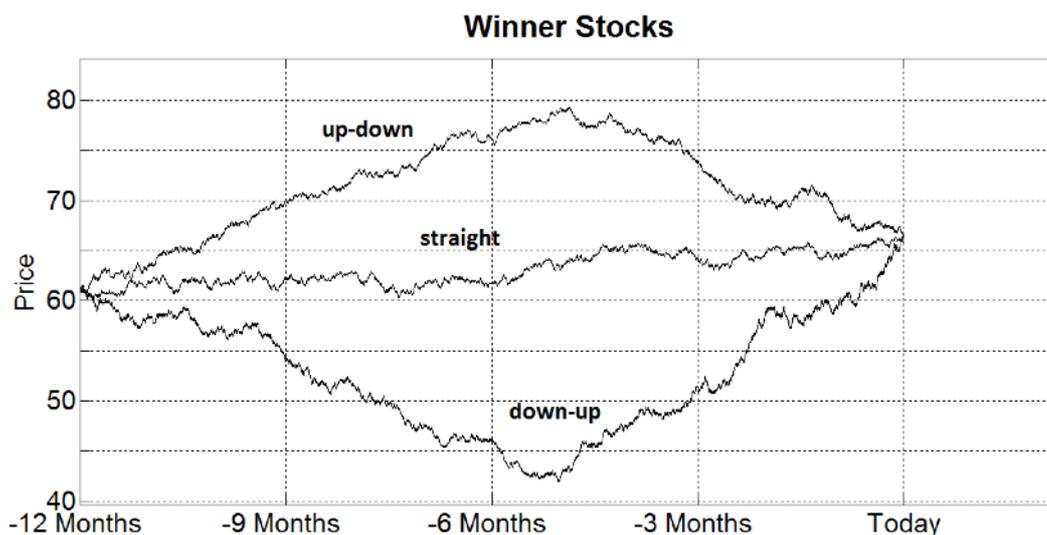


Figure 2.9. Représentation des différents parcours d'actifs présentés aux participants dans l'expérience Grosshans et al, 2018).

Les auteurs observent également que la propension à vendre l'actif à la fin des périodes est plus élevée pour le groupe down-up que le groupe up-down. Ce résultat est intéressant à analyser lorsqu'on se réfère à ce que nous avons pu évoquer au cours du chapitre à propos du point de référence. Une manière de capturer ad hoc ce résultat est de supposer que le point de référence s'adapte au pic de richesse comme l'évoque Gneezy ($R=80$ pour le groupe « up-down ») puis s'adapte peu lorsque le prix diminue. Dès lors à la fin de la période, le prix de l'actif est considéré dans le domaine des pertes pour le groupe « up-down ». Concernant le groupe « down-up », le prix est devenu plus grand que le prix initial à la fin des périodes. Avant cela le prix était toujours inférieur. Dès lors le prix final est considéré dans le domaine des gains. Or, nous avons vu avec le disposition effect que les individus avaient tendances à vendre plus rapidement les actifs considérés comme gagnant (groupe « down-up ») que les actifs considérés comme perdant (groupe « up-down »).

4.3.4. Le grand problème d'un test comportemental

Supposons la situation suivante : on cherche à tester l'impact d'une chaîne d'évènement (un parcours de richesse sur n périodes) sur la prise de risque subséquente. Supposons le parcours de richesse de l'agent A : 100 500 (donc un gain de 400€ entre la période 1 et 2) et le parcours de l'agent B : 900 500 (donc une perte de 400 entre la période 1 et 2). Si le point de référence est égal à la richesse initiale, les deux agents devraient se montrer moins averses au risque au coup suivant : L'agent A à cause de l'house money effect et l'agent B sous l'effet potentiel du Break Even Effect. Il est donc envisageable que les deux individus fassent le même choix subséquent mais pour des raisons différentes. Si au contraire le point de référence s'adapte parfaitement, les deux agents feront le même choix aussi. Par conséquent, discriminer entre les points de références devient donc assez problématique si l'on n'observe pas de différence entre les individus aux parcours différents...

La Figure 2.6 portant sur la prime de risque en fonction du point de référence est révélatrice de problème. Comme nous pouvons le voir, la même prime de risque peut être obtenues pour différentes valeurs de référentiel, marquant ici un problème d'indétermination du modèle si on considère le point de référence comme paramètre libre.

4.4. Modélisation de l'actualisation du référentiel

Dans cette section, nous présentons différentes modélisations du processus adaptatif du référentiel. Nous pouvons dégager de cette littérature deux grandes familles : Les modèles récursifs (Barberis et al., 2001; Shi, Cui, Yao, & Li, 2015) et les modèles intégrant une fonction de pondération des événements (prix ou richesses) passés (Baucells et al., 2011; Post et al., 2008). Nous référençons ici trois modèles qui ont peu trouver un certain succès dans la littérature : le modèle de Shi et al (2015), le modèle de Baucells et al (2011) et le modèle de Barberis et al (2001).

4.4.1. Les modèles récursifs

a. Modèle de Shi et al (2015)

Le modèle de Shi et al (2015) se base sur les travaux opérés par Arkes et al (2008, 2010) dans lesquels ces derniers révèlent que le référentiel s'adapte de façon asymétrique. En particulier, (1) après un gain le point de référence augmente et s'adapte partiellement à la nouvelle richesse de l'individu, (2) Après une perte, le point de référence diminue et s'adapte partiellement à la nouvelle richesse, (3) l'adaptation du référentiel est plus importante après un gain qu'après une perte. Sur le base de ce pattern, Shi et al (2015) propose le modèle récursif d'adaptation du référentiel W_{t+1}^{REF} suivant :

$$W_{t+1}^{REF} = \begin{cases} W_t^{REF} + \alpha_g [W_t - W_t^{REF}] & \text{si } W_t \geq W_t^{REF} \text{ (prior gain)} \\ W_t^{REF} + \alpha_l [W_t - W_t^{REF}] & \text{si } W_t < W_t^{REF} \text{ (prior perte)} \end{cases}$$

Où W_t représente la richesse de l'individu au temps t et W_t^{REF} représente la richesse de référence au temps t . Le paramètre $\alpha_g \in [0, 1]$ est le coefficient d'adaptation après un gain et $\alpha_l \in [0, 1]$ le coefficient d'adaptation après une perte. Une adaptation partielle du référentiel est représentée par le fait que α_g et α_l sont inférieures à 1 et l'asymétrie du référentiel implique $\alpha_g > \alpha_l$.

Le point de référence au temps $t + 1$ est formé sur la base de deux éléments : a) le point de référence au temps t et b) sur le domaine de richesse considéré, c'est-à-dire que si c'est-à-dire si $W_t > W_t^{REF}$ nous sommes dans le domaine des gains et si $W_t < W_t^{REF}$ nous sommes dans le domaine des pertes. Tiré de l'article de Shi et al (2015) nous représentons graphiquement l'adaptation asymétrique du point de référence à la richesse $t+1$ en fonction du domaine de richesse considéré (perte ou gain) (voir Figure 2.10).

Une des implications de ce modèle est qu'en fonction du parcours de richesse de l'individu, son attitude face au risque à un moment t pour une richesse finale aléatoire donnée, ne sera pas la même. Prenons un exemple pour comprendre intuitivement une implication de ce modèle : Supposons deux individus A et B, identiques en termes de préférences (même fonction d'utilité, même fonction de pondération des probabilités), ayant la même richesse initiale et finale mais connaissant deux parcours de richesses différents. Supposons trois périodes d'évolution de la richesse :

$$A : [W_{t1} : 500\$; W_{t2} : 1000\$, W_{t3} : 500\$]$$

$$B : [W_{t1} : 500\$; W_{t2} : 0\$, W_{t3} : 500\$]$$

Si l'asymétrie d'adaptation du référentiel est effective, on comprend que le point de référence à la dernière période ne devrait pas être le même pour les deux individus. Dès lors, si ces deux agents sont alors confrontés à une même loterie à la période t_4 , ces derniers ne devraient pas exprimer les mêmes attitudes face au risque, ou le même degré d'aversion au risque. En outre,

si l'adaptation du référentiel est plus élevée dans les gains que pour les pertes, on devrait observer que le référentiel de l'agent A est plus élevé que celui de l'agent B. L'idée sous-jacente est qu'un individu « s'habitue » plus vite à un gain qu'à une perte. Il aura tendance à s'approprier plus rapidement une richesse gagnante qu'une richesse perdante.

Toutefois, une telle dépendance au parcours passé de richesse rend très complexe les prédictions et formalisations à propos de l'évolution de l'attitude face au risque sur plusieurs périodes.

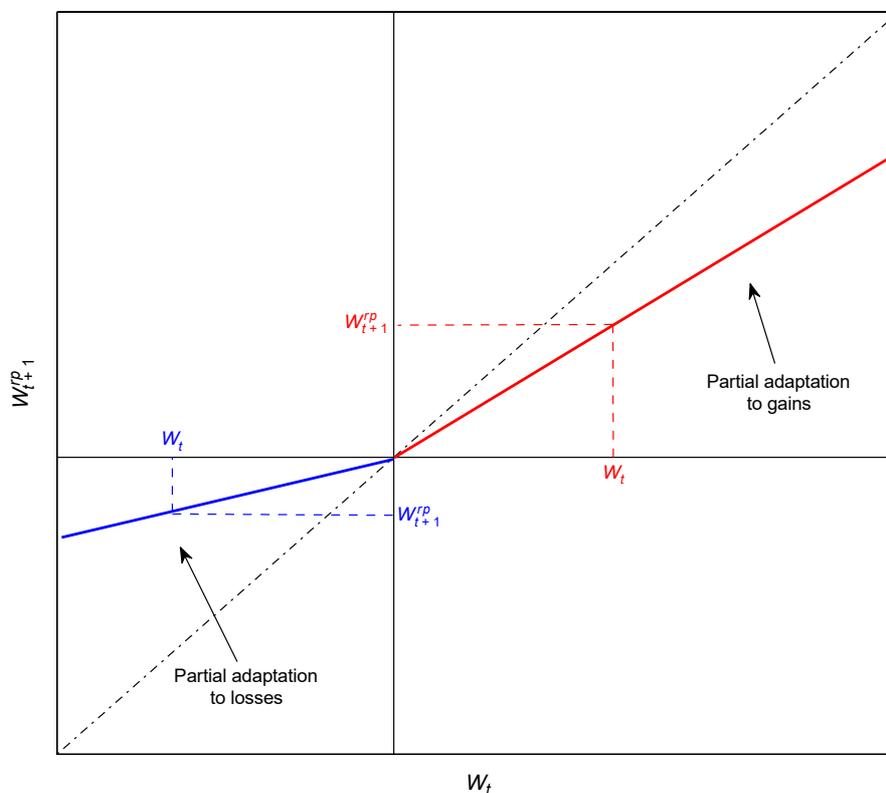


Figure 2.10. Représentation de l'adaptation du point de référence dans les pertes et dans les gains selon le modèle de Shi et al (2015).

b. Le modèle de Barberis et al (2001)

Le modèle de Barberis et al (2001) propose également une actualisation récursive du point de référence. A partir de ce modèle, les auteurs établissent des prédictions claires à propos de l'impact d'un gain ou d'une perte sur l'attitude face au risque subséquente. En particulier, les

auteurs parviennent à fournir capturer le MPwL et le HME. Le BEE n'est toutefois pas capturé par ce modèle.

Les auteurs se placent dans une configuration où l'agent opère sur le marché financier. On s'intéresse alors à l'utilité que l'individu dérive de l'évolution du prix d'un actif. Toutefois, le modèle peut être retranscrit dans un cadre plus général de variation de richesse. Nous décidons ici de traiter ce modèle dans cette perspective.

Dans ce modèle, Barberis et al (2001) présument que l'évaluation d'une richesse au temps $t + 1$ est comparé à deux points de référence. Supposons qu'un individu possède au temps t une richesse de $x_t = 100\text{€}$. Imaginons qu'au temps $t + 1$, sa richesse ait augmenté de 20€. Il possède au temps $t+1$ $x_{t+1} = 120\text{€}$. L'utilité qu'il dérive de cette nouvelle richesse dépend, en plus de la fonction d'utilité, d'une richesse de référence à laquelle il la compare. Le plus simple est de considérer que cette nouvelle richesse est comparée à la richesse au temps t , autrement dit 100€. Ainsi, la variation de richesse évaluée dans la fonction d'utilité est :

$$v(x_{t+1}) = v(s_{t+1}x_t - x_t) = v(x_t(s_{t+1} - 1))$$

Avec s_{t+1} le pourcentage d'augmentation (ou de diminution) entre x_t et x_{t+1} , d'où $s_{t+1} = (x_{t+1} - x_t)/x_t$. Ainsi, ici x_t , la richesse actuelle, fait office de premier référentiel. Pour capturer l'effet des richesses antérieures sur l'évaluation de x_{t+1} , les auteurs ajoutent un second référentiel, noté Z_t . Ce référentiel, que les auteurs appellent « benchmark », peut correspondre par exemple à une richesse saillante dans le passé dont l'individu se souvient ou encore à la richesse moyenne obtenu sur n périodes. Quoiqu'il en soit, lorsque l'agent juge sa richesse actuelle, il la compare à ce benchmark. Si $x_t > Z_t$ on considère que l'individu a obtenu un gain (appelé « prior gain ») et inversement si $x_t < Z_t$ l'individu a enduré une perte (appelé « prior perte »). Dès lors, x_t et Z_t sont deux points de références qui conditionnent comment la richesse

au temps $t + 1$ est perçu par l'agent. L'utilité de x_{t+1} s'exprime alors $v(x_{t+1}, x_t, Z_t)$. Par commodité pour le prochain développement, au lieu de Z_t , on utilisera plutôt $z_t = Z_t/x_t$. On peut alors dégager trois cas. La Figure 2.11 tiré de l'article de Barberis et al (2001) résume ces trois cas.

Si $z_t = 1$

Dans ce cas, l'individu n'a subi ni perte ni gain antérieurement. Pour simplifier la notation, nous supposerons une évaluation linéaire des variations de richesses tel que $v(x_{t+1}) = x_{t+1}$.

Dans ce cas, l'évaluation de la richesse au temps $t + 1$ s'écrit :

$$v(x_{t+1}) = \begin{cases} x_{t+1} & \text{si } x_{t+1} \geq 0 \\ -\lambda x_{t+1} & \text{si } x_{t+1} < 0 \end{cases}$$

Avec λ le paramètre d'aversion aux pertes. Nous sommes ici dans la situation la plus simple où l'individu juge la richesse ultérieure en fonction de sa richesse actuelle selon une fonction d'utilité de type prospect theory en deux parties.

Si $z_t < 1$

Dans cette situation l'individu a obtenu un gain antérieurement. Pour comprendre comment sera évalué la future richesse commençons par un exemple avant d'écrire la fonction d'utilité. Supposons que la richesse actuelle de l'individu soit de $x_t = 100\text{€}$ et que son référentiel antérieur saillant (Z_t) soit de 90€ . La différence entre x_t et Z_t est donc de 10€ . Ces 10€ constituent en quelque sorte le « coussin de sécurité » de l'individu. Autrement dit, une perte de 10€ au temps $t + 1$ sera ressentie comme moins douloureuse que si ce coussin n'existait pas (comme par exemple dans le cas 1. où l'individu n'a pas de gain antérieure). Supposons que l'individu ait subit une perte au temps $t+1$ et que sa richesse soit maintenant égale à 80€ . Sans coussin de sécurité et en supposant $\lambda = 2$, cette nouvelle richesse serait évaluée ainsi :

$-\lambda(x_{t+1} - x_t) = -\lambda(80 - 100) = -40$. En prenant en compte le coussin de sécurité, la richesse au temps $t + 1$ est maintenant évalué ainsi : $(Z_t - x_t) + (-\lambda(x_{t+1} - Z_t)) = (90 - 100) + (-\lambda(80 - 90)) = -30$. On remarque dans cette situation que le sentiment de gain de 10€ entre le benchmark Z_t et la richesse au temps t réduit la sensation de perte tout simplement parce que tant que la richesse n'est pas inférieure au benchmark, l'individu ne se sent pas en perte (et donc l'aversion aux pertes n'a pas d'effet). De façon générale, la fonction d'utilité s'écrit alors :

$$v(x_{t+1}) = \begin{cases} s_{t+1}x_t - x_t & \text{si } s_{t+1} \geq z_t \\ x_t(z_t - 1) + (-\lambda(s_{t+1}x_t - z_t)) & \text{si } s_{t+1} < z_t \end{cases}$$

Si $z_t > 1$

Dans cette situation l'individu a subi une perte antérieure. Dès lors, une perte supplémentaire au $t+1$ sera encore plus douloureuse que si l'individu n'avait pas subi de perte antérieurement ou s'il avait obtenu un gain antérieurement. Plus précisément, si la richesse au temps $t + 1$ est inférieure à la richesse au temps t et au benchmark Z_t , on peut considérer que la perte est pénalisée deux fois par le terme d'aversion aux pertes. En fait, les auteurs adoptent une formalisation qui pénalise moins fortement une telle situation mais l'intuition reste identique : plus la perte antérieure est forte, plus une future perte est ressentie comme douloureuse.

Pour finir on évoquera que les auteurs fournissent une formalisation du benchmark Z_t qui correspond basiquement à la moyenne des richesses obtenues au cours de n périodes sans toutefois justifier particulièrement ce choix.

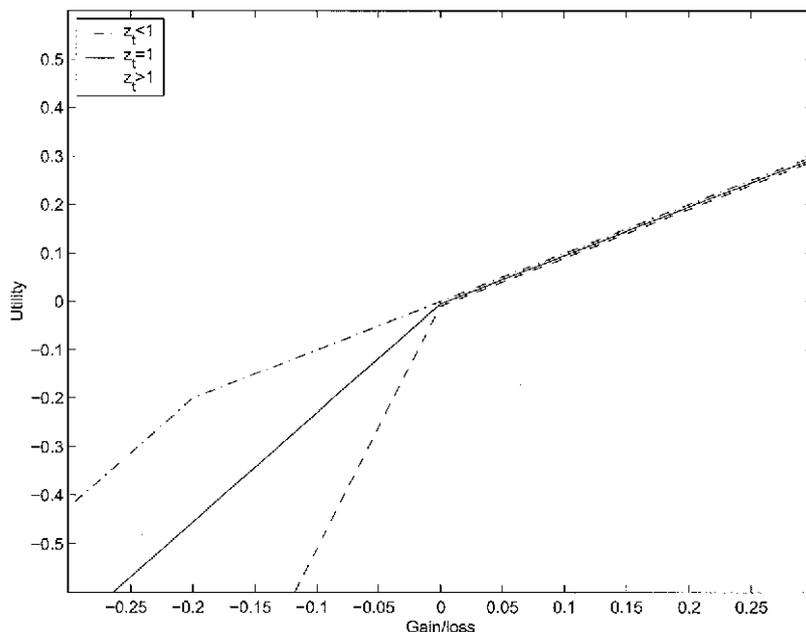


Figure 2.11. La fonction point-pointillé représente la fonction d'utilité prospect theory de la richesse au temps $t+1$ lorsque l'individu a connu un gain antérieurement. La fonction pointillée représente la fonction d'utilité prospect theory de la richesse au temps $t+1$ lorsque l'individu a connu une perte antérieurement et la fonction en train plein représente l'utilité de la richesse au temps $t+1$ dans le cas où l'individu n'a subi ni perte ni gain antérieurement.

4.4.2. Les modèles non récursifs

Baucells et al (2011) propose un modèle non-récursif basé sur leurs résultats expérimentaux. Comme nous l'avons évoqué, les auteurs révèlent qu'au sein d'une séquence d'évolution d'un prix d'un actif, le prix initial et final influencent davantage le prix de référence subséquent que les prix intermédiaires. D'une certaine manière Baucells et al (2011) révèle ici deux effets bien connue en psychologie l'effet de primauté (Anderson, 1965) et l'effet de récence (Murdock Jr, 1962). Pour accommoder ces résultats, les auteurs s'inspirent du modèle RDU, en formalisent le prix de référence au temps $t+1$, r_{i+1} , comme suit :

$$r_{i+1} = p + \sum_{i=1}^n \left[\varphi\left(\frac{i}{n}\right) - \varphi\left(\frac{i-1}{n}\right) \right] y_i$$

Avec y_i le prix au temps i , et $\varphi(x) = e^{\frac{-(-\ln x)^\gamma}{\delta}}$. On notera $\Phi_i = \sum_{i=1}^n \left[\varphi\left(\frac{i}{n}\right) - \varphi\left(\frac{i-1}{n}\right) \right]$ qui, comme dans la RDU peut s'interpréter comme le poids relatif d'une conséquence (ici le poids relatif d'un ancien prix). Ainsi, la dernière expression s'écrit explicitement :

$$r_{i+1} = p + \sum_{i=1}^n \left[e^{\frac{-(-\ln(\frac{i}{n}))^\gamma}{\delta}} - e^{\frac{-(-\ln(\frac{i-1}{n}))^\gamma}{\delta}} \right] y_i$$

La fonction $\varphi(x)$ est en fait la formalisation de la fonction de pondération des probabilités proposée par Prelec. Nous avons vu que cette fonction pouvait suivre une forme en S inversé en fonction des valeurs des paramètres libres δ et γ . Avec γ contrôlant la courbure de la fonction et δ l'élévation. Dès lors, nous avons montré que l'intégration de cette forme de fonction (concave puis convexe) dans la formalisation des poids décisionnels au sein de la RDU impliquait une surpondération des conséquences extrêmes (les pires et les meilleures conséquences) et une sous-pondération des conséquences intermédiaires.

Dans le modèle de Baucells et al (2011), il ne s'agit pas de conséquences extrêmes mais de la localisation temporelle des différents prix. Quand $\frac{i}{n}$ est petit, on parle d'un prix qui s'est produit au début d'une séquence d'évènements. Lorsque $\frac{i}{n}$ est grand on parle de prix qui se sont produits récemment. Ainsi, en appliquant la formalisation des poids décisionnels de la RDU à la situation de Baucells et al (2011), on obtient bien une surpondération des évènements initiaux et finaux et une sous pondération des évènements intermédiaires comme le montre la Figure 2.12.

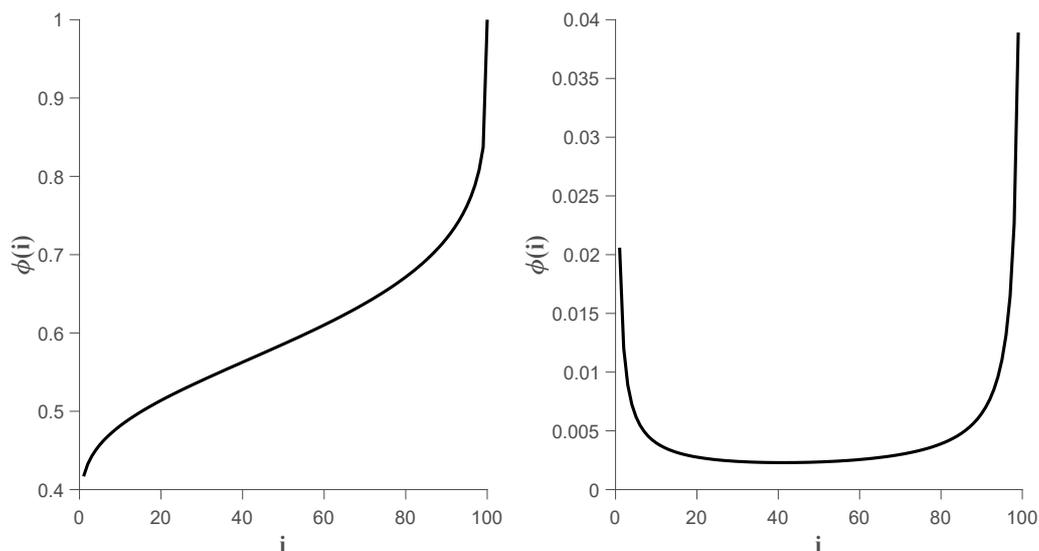


Figure 2.12 A gauche représentation de la fonction similaire à celle supposé dans la RDU. Lorsqu'elle a cette forme elle implique la fonction de droite qui représente la pondération des prix obtenu lors d'une séquence de 100 éléments. Les paramètres choisis pour construire ces fonctions sont tirés de Baucells et al (2011).

CONCLUSION PARTIE 2

Dans cette partie nous avons présenté la célèbre théorie de Kahneman et Tversky : La cumulative Prospect Theory. Nous avons montré en particulier que la principale force de ce modèle consistait à intégrer un référentiel qui permet d'appréhender le changement d'attitude face au risque des individus dans les pertes et dans les gains. Toutefois, ce référentiel constitue également sa principale faiblesse dans la mesure où sa spécification n'est pas encore établie. Ce manquement, accorde alors au modèle un degré de liberté trop important qui lui permet certes de capturer un grand nombre de phénomènes mais de manière ad hoc.

Dans des situations dynamiques, le même problème se pose avec en plus un questionnement vis-à-vis de son actualisation au cours du temps (ou des événements). Finalement, la principale difficulté réside dans la mesure de ce référentiel. Nous avons vu que l'inférence (indirecte) du référentiel à partir de choix se révèle particulièrement difficile. Finalement, de plus en plus de chercheurs ont adopté une méthodologie consistant à demander directement aux participants de

révéler eux-mêmes leur point de référence, ce qui jette malheureusement de fort doute vis-à-vis de ce qui est réellement mesuré.

La Prospect Theory constitue malgré tout à ce jour un bel exemple d'interdisciplinarité entre la psychologie et l'économie qui s'intègre dans le champ plus général de la théorie de la décision.

3. PROBABILITY DISCOUNTING ET TAILLE D'ENJEU DES LOTERIES

INTRODUCTION

Dans cette section nous abordons la question de l'impact de la taille d'enjeu des loteries sur l'attitude face au risque et sa modélisation. Cette problématique fut entre-autre étudiée par Markowitz (1952) qui montra que les agents se révèlent de plus en plus averse au risque à mesure que l'enjeu augmente. Sur la base de ces résultats (informels) il proposa alors sa célèbre fonction d'utilité à trois points d'inflexion : Convexe puis concave dans les pertes et convexe puis concave dans les gains.

Nous abordons cette question sous l'angle du modèle de « Probability Discounting » particulièrement utilisé par l'école comportementaliste en psychologie (Dai, Grace, & Kemp, 2009; Estle, Green, Myerson, & Holt, 2006; Kirby, 1997; McKerchar, Green, & Myerson, 2010; Myerson, Green, & Morris, 2011; Rachlin, 2006; Shead & Hodgins, 2009).

Après avoir décrit ce modèle, nous rendrons compte des principaux résultats expérimentaux révélés par ces études comportementales. Comme nous le verrons, l'ensemble des estimations statistiques du modèle opérées dans ces études montrent qu'un des paramètres libres du modèle n'est pas stable à travers le niveau d'enjeu. Plus spécifiquement, les chercheurs de ce domaine réalisent des régressions non-linéaires (par la méthode des moindres carrés non-linéaires) par niveau d'enjeu et révèlent que l'estimation d'un des paramètres libres augmente à mesure que l'enjeu augmente. A l'heure actuelle, les behavioristes ne parviennent pas à expliquer l'instabilité de ce paramètre.

En adoptant une perspective économique nous exposerons pourquoi ce paramètre n'est pas stable. Pour cela, nous montrons dans un premier temps que le modèle théorique développé par les comportementalistes n'est théoriquement pas en mesure de capturer les résultats expérimentaux. Nous verrons que le modèle prédit que l'aversion au risque relative reste stable à mesure que l'enjeu augmente alors que le pattern comportemental révélé par les expériences

montre que l'aversion au risque relative augmente avec le niveau d'enjeu. De plus, de façon collatérale, notre investigation nous a permis de révéler que les loteries utilisées par ces chercheurs ne permettent pas une estimation fiable des paramètres. Plus spécifiquement, ces loteries sont des loteries binaires et n'impliquent qu'un seul résultat différent de 0. Nous révélerons alors un problème d'indétermination lorsque l'on cherche à estimer les paramètres libres du modèle à partir de ces types loteries.

Nous proposons alors deux solutions. Au niveau théorique nous conseillons d'adopter une fonction d'utilité possédant une certaine propriété capable de capturer le pattern comportemental. Au niveau expérimental, nous conseillons d'utiliser des loteries à trois issues possibles différentes de 0. Une troisième voie est explorée dans l'annexe de ce chapitre qui consiste à recourir à l'estimation Bayésienne des paramètres du modèles lorsque des loteries binaires avec une seule conséquence différente de 0 est utilisé dans les expériences.

Enfin, dans une dernière section de ce chapitre nous verrons que le pattern comportemental révélé par les psychologues est également étudié par les économistes. Nous nous apercevons que ce pattern est en fait plus compliqué qu'il n'y paraît au premier regard. En particulier, il semble que l'aversion au risque relatif augmente à mesure que l'enjeu augmente uniquement pour des niveaux de probabilités faibles. Pour des niveaux de probabilités plus élevés, l'aversion au risque relatif reste constant à mesure que l'enjeu augmente.

Nous verrons alors qu'aucun modèle développé à ce jour en économie n'est en mesure de capturer ce pattern car, potentiellement, ce dernier semble remettre en question ce que l'on appelle le « principe de séparabilité » que nous définirons dans ce chapitre.

I. LE MODELE DE PROBABILITY DISCOUNTING

Dans cette section, nous présentons comment les behavioristes appréhendent la prise de décision en environnement risqué. Nous débuterons en présentant le contexte dans lequel s'inscrivent ces études, ce qui nous conduira d'abord à présenter le phénomène de « delay discounting » avant de présenter le phénomène de « probability discounting ».

Nous verrons alors que les deux phénomènes furent formalisés pas deux modèles très similaires. Cette similitude achemina les behavioristes à se demander si un processus sous-jacent commun aux deux phénomènes pouvait être l'œuvre. C'est à partir de cette question que fut révélé un pattern comportemental propre au phénomène de « probability discounting » et qui n'était pas présent dans les études portant sur « delay discounting ». Cette différence comportementale semble avoir tranché la question à propos de l'existence du « processus sous-jacent commun ». C'est dans ce contexte que les chercheurs du domaine se sont rendu compte que l'estimation d'un des paramètres libres de leur modèle de « Probability Discounting » semblait instable.

1.1. Le phénomène du Temporal (ou Delay) Discounting

Le phénomène du discounting fait référence à la tendance générale des individus à dévaluer subjectivement les conséquences différées dans le temps (delay discounting) ou incertaines (probability discounting). Autrement dit, ces phénomènes font référence à la préférence importante des individus pour le présent ou pour les issues certaines (Green & Myerson, 2004).

1.1.1. Définition et étude

Le delay discounting correspond spécifiquement à la dévaluation au moment présent d'une conséquence différé dans le temps (Ainslie, 1974; Mazur, 1987). La préférence pour le présent se révèle très simplement et correspond très bien à l'intuition que l'on pourrait se faire naturellement sur le sujet. En effet, il est clair que si nous avons le choix entre obtenir une

récompense maintenant ou bien la même récompense dans trois ans, nous préférons obtenir la récompense immédiatement. De façon plus intéressante, de nombreuses études montrent que cette préférence pour le présent est telle que les individus préfèrent obtenir une récompense immédiatement plutôt qu'une récompense plus importante en terme d'amplitude mais différée dans le temps. Par exemple, la plupart des agents préfèrent obtenir 10€ immédiatement que 20€ dans un an.

Ce phénomène est particulièrement visible chez les individus dépendant du tabac (Bickel, Odum, & Madden, 1999), de l'alcool (Petry, 2002) ou encore de l'héroïne (Madden, Bickel, & Jacobs, 1999). Ces études montrent que les individus dévaluent subjectivement plus rapidement l'acquisition futures de substances que l'argent lui-même.

Le phénomène de delay discounting est souvent investi expérimentalement de la manière suivante. On présente à un participant une situation de choix entre un grand renforçateur (grande récompense en termes d'amplitude) disponible après un certain délai ou bien un petit renforçateur disponible immédiatement. L'objectif est alors de trouver le point d'indifférence à travers une série de questions entre le grand renforçateur différé et le petit renforçateur. Pour cela on peut soit manipuler le délai d'obtention du petit renforçateur, soit manipuler la magnitude du petit renforçateur. Par exemple, pour connaître le point d'indifférence de 20€ dans 1 mois, on pourra confronter cette alternative à une série de plus petits gains accessibles immédiatement. Dans ce cas, on demandera pour chaque petit gain ce que l'individu préfère entre le petit renforçateur ou la conséquence plus importante mais différée dans le temps.

Pour construire une fonction de dévaluation, on répétera ce processus avec différents délais pour le grand renforçateur dont on conservera la même amplitude (voir Figure 3.1). Comme on le constate sur la Figure 3.1 le montant immédiat équivalant au grand renforçateur différé diminue à mesure que le délai d'obtention du grand renforçateur augmente. Autrement dit, sa valeur subjective actuelle diminue à mesure que son délai d'obtention augmente.

Les études portant sur le « delay discounting » se sont surtout focalisées sur le phénomène de « renversement des préférences » lorsqu'un délai supplémentaire est ajouté aux deux alternatives. Supposons qu'un agent préfère initialement obtenir un petit renforçateur immédiat plutôt qu'un plus grand renforçateur dans 1 mois. Il a souvent été montré que l'individu pouvait finir par préférer le grand renforçateur si on ajoute un délai suffisamment long identiquement aux deux alternatives (Ainslie, 1974; Mazur, 1987). Autrement dit, en reprenant notre dernier exemple, l'individu qui avait une préférence pour obtenir les 10€ immédiatement au lieu de 20€ dans deux mois, finit par préférer obtenir 20€ dans 4 mois plutôt que d'obtenir 10€ dans 2 mois. Ce type de renversement bien connu des économistes également (Laibson, 1997; Strotz, 1955). Ce renversement est toutefois une violation de l'axiome de stationnarité qui est en fait l'équivalent de l'axiome d'indépendance (Green & Myerson, 2004).

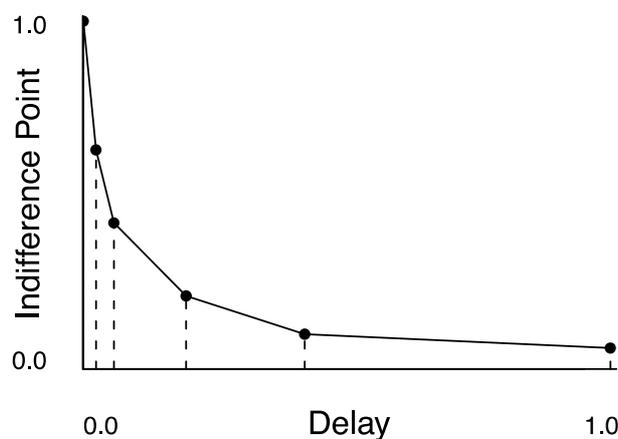


Figure 3.1 Point d'indifférences normalisé obtenus en fonction du délai d'obtention du renforçateur. (Tiré de Tsukayama et al, 2010)

1.1.2. Formalisation du Temporal Discounting

L'axiome de stationnarité suppose que les préférences ne se renversent pas si on ajoute un délai constant à deux alternatives possibles. Pour respecter cet axiome, il faut supposer que la dévaluation des renforçateurs diminue selon un taux constant par unité de temps. Une dévaluation selon une fonction exponentielle comporte cette propriété :

$$V_e(x) = x \exp(-\beta D) \quad (3.1)$$

Avec $V_e(x)$ dénotant la valeur subjective de l'amplitude du renforçateur x , D représente le délai d'obtention du renforçateur et β est un paramètre libre contrôlant le taux de dévaluation.

Les psychologues adoptent une formalisation capable de capturer ce renversement de préférence grâce à une fonction hyperbolique. La principale caractéristique de cette fonction est que le taux de dévaluation décélère progressivement avec le délai :

$$V_h(x) = x \frac{1}{1 + kD} \quad (3.2)$$

Avec k un paramètre libre représentant le taux de dévaluation du renforçateur. Une variante importante de ce modèle, qui semble mieux capturer les résultats expérimentaux consiste à rajouter un paramètre libre dénoté par la lettre s , à la fonction de dévaluation relative tel que $V_{hs}(x) = x \frac{1}{(1+kD)^s}$ (Myerson & Green, 1995). Dans ce cas, le paramètre s est interprété comme un facteur de mise à l'échelle non-linéaire du temps ET de la quantité (amplitude) de renforçateur (nous insistons sur le « et » car cela aura son importance dans le modèle du probability discounting). Il s'agit d'une incorporation explicite de la loi puissance suggérée par Stevens (1957). Ce dernier a clairement montré que pour une grande variété de modalités sensorielles, la relation entre l'amplitude objective d'un stimulus et sa perception subjective pouvait être décrite par une fonction puissance. L'exposant représente alors le degré pour lequel la mise à l'échelle de l'amplitude objective du stimulus dévie de la linéarité (McKerchar et al, 2010).

Bien entendu, d'autres modèles existent comme celui proposé par Laibson (1997), mais notre propos ici est de présenter le raisonnement qui a poussé les psychologues à adopter une formalisation similaire dans le cadre du probability discounting. Par conséquent, nous ne nous attarderons pas sur ces modèles.

1.2. Le phénomène du Probability Discounting

Similairement au temporal discounting, le probability discounting réfère à une diminution de la valeur subjective d'un renforçateur à mesure que sa probabilité d'occurrence diminue. On notera qu'implicitement on se place dans une situation où l'individu peut gagner le renforçateur avec une certaine probabilité ou ne rien gagner avec la probabilité complémentaire. Autrement dit, il s'agit de loteries binaires avec seulement un seul résultat différent de 0, tel que : $L \equiv (x_1 \neq 0, p_1, x_2 = 0, 1 - p_1)$. Nous verrons plus loin que cela a une grande importance pour comprendre les résultats d'estimation des paramètres du modèle de probability discounting. Dans le cadre du probability discounting on étudie donc comment évolue la valeur subjective d'un renforçateur probabiliste à mesure que sa probabilité diminue.

1.2.1. Définition et études

Les études qui portent sur le probability discounting au sein du behaviorsisme ont recours à une méthodologie classique pour déterminer la valeur subjective d'un renforçateur probabiliste. Il s'agit simplement d'une variante de la méthode de l'équivalent certain. Nous ne développerons pas ici la méthode même si certaines critiques pourraient lui être adressées (voir par exemple Estle, Green & Myerson, 2006).

Ces expériences montrent évidemment que la valeur subjective d'un renforçateur diminue à mesure que sa probabilité diminue. Le résultat intéressant révélé par ses études est similaire à celui obtenu avec le temporal discounting dans le sens où on observe également un renversement de préférence. Plus spécifiquement, un individu peut préférer initialement un petit

renforçateur certain plutôt qu'un renforçateur probabiliste plus grand. Toutefois, si on divise ces probabilités par un facteur commun, tel que $p \cdot r$, avec $0 < r < 1$, on observe que l'individu peut finir par préférer le plus grand renforçateur. Il s'agit classiquement du paradoxe d'Allais (Rachlin, Castrogiovanni, & Cross, 1987; Rachlin, Raineri, & Cross, 1991).

Dès lors, les chercheurs ont proposé un modèle similaire à celui du temporal discounting. C'est-à-dire une dévaluation subjective de la valeur d'un renforçateur selon une fonction hyperbolique.

1.2.2. Formalisation du Probability Discounting

Nous avons vu dans la chapitre 1 que la théorie de l'utilité espérée n'était pas en mesure de capturer le paradoxe d'Allais. Ces renversements de préférences furent finalement expliqués grâce à l'introduction d'une nouvelle fonction : la fonction de pondération des probabilités au sein de la RDUT. Le modèle de Probability Discounting, formulé par Rachlin (1991) intègre cette fonction de pondération des probabilités mais ne tient que pour des loteries binaires avec un seul résultat différents de 0. Nous dénotons ce type de loterie particulier ainsi : $L_D \equiv (x_1, p_1)$. La valeur de x_1 , noté $V_{ph}(x_1)$, s'écrit :

$$V_{ph}(L_D) \equiv V_{ph}(\tilde{x}_1) = x_1 \frac{1}{1 + h\Theta} \quad (3.3)$$

Où $\Theta = (1 - p)/p$ est la cote contre la réception de x_1 . Le paramètre h représente le taux de dévaluation du renforçateur. On pourrait également interpréter ce paramètre comme le degré de pessimisme (si $h > 1$, la fonction de pondération des probabilités est convexe et donc sous pondère toutes les probabilités différentes de 0 et de 1) ou d'optimisme (si $h < 1$, la fonction

est concave) de l'individu. Pour revenir à une forme d'expression proche des modèles que nous avons évoqués jusque-là, nous avons choisi de réécrire le modèle comme suit³¹ :

$$V_{ph}(L_D) \equiv V_{ph}(x_1) = x_1 \frac{p_1}{p_1 + h(1 - p_1)} \quad (3.4)$$

Le second terme du modèle peut donc s'assimiler à une fonction de pondération des probabilités, noté $\varphi(p)$. Ce modèle fut raffiné plus tard par Green et Myerson (2004), qui, comme pour le temporal discounting, ajoutent un paramètre, noté s . En l'intégrant dans l'Equation (3.4) on obtient :

$$V_{ph}(L_D) \equiv V_{ph}(x_1) = x_1 \left(\frac{p_1}{p_1 + h(1 - p_1)} \right)^s \quad (3.5)$$

Avec $\varphi(p) = \left(\frac{p_1}{p_1 + h(1 - p_1)} \right)^s$. Le paramètre " s " s'interprète comme dans le temporal discounting, c'est-à-dire une mise à l'échelle de l'amplitude du renforçateur ET des probabilités dans la logique de la théorie de Stevens (1957). Si h et s sont égaux à 1, on retrouve un modèle d'espérance mathématique tel que $V_{ph}(L_D) = E(L_D)$. Empiriquement, les estimations des paramètres libres donnent $h \approx 5$ et $s \approx 0.3$ (voir par exemple Mckerchar et al, 2010). Ce qui fournit une fonction de pondération des probabilités en forme de S inversé comme celles que nous avons présenté dans le chapitre 1.

1.3. Temporal Discounting et Probability Discounting : un processus sous-jacent commun ?

³¹ En effet, $\frac{1}{1+h\left(\frac{1-p_1}{p_1}\right)} = \frac{1}{\frac{p_1+h(1-p_1)}{p_1}} = \frac{1}{\frac{p_1+h(1-p_1)}{p_1}} = \frac{p_1}{p_1+h(1-p_1)}$.

La similarité des expressions mathématiques utilisées pour formaliser le temporal discounting et le probability discounting ont fait naître une interrogation chez ces chercheurs : existe-t-il un processus sous-jacent commun qui explique la dévaluation subjective par rapport au temps et par rapport aux probabilités (Green & Myerson, 1996; Rachlin, Logue, Gibbon, & Frankel, 1986; Rachlin & Siegel, 1994) ? Sans rentrer dans le détail de ces recherches, l'idée du processus commun, suppose implicitement que des variables affectant le temporal discounting et probability discounting devraient avoir le même effet sur les deux phénomènes³² (Green & Myerson, 2004). Or, de nombreuses études ont montré que l'augmentation de l'amplitude du renforçateur ne produisait pas le même effet dans le temporal discounting que dans le probability discounting (Christensen, Parker, Silberberg, & Hursh, 1998; Estle et al., 2006; Green, Myerson, & Ostaszewski, 1999).

Plus spécifiquement, il a été découvert que les grands renforçateurs probabilistes étaient dévalués plus fortement que les petits renforçateurs probabilistes (voir Figure 3.2). On appelle cela « l'effet de magnitude » (Johnson & Bickel, 2002; Kirby, 1997; Kirby & Maraković, 1996; Myerson & Green, 1995; Raineri & Rachlin, 1993; Thaler & Shefrin, 1981). Alors que concernant le temporal discounting, l'augmentation de l'amplitude des renforçateurs n'a pas d'effet sur le taux de dévaluation relatif (Green & Myerson, 2004; Richards, Mitchell, De Wit, & Seiden, 1997).

Pour finir, l'estimation des paramètres libres du modèle de probability discounting réalisé dans de nombreuses études montre que la valeur du paramètre « s » augmente à mesure que l'amplitude du renforçateur augmente. A l'inverse, les mêmes études montrent que l'estimation

³² Certains auteurs ont défendu l'idée que le risque sous tendait la dévaluation des renforçateurs en environnement probabiliste comme en environnement temporel. L'idée étant que plus on doit attendre pour obtenir un renforçateur, plus le risque de ne pas l'obtenir augmente (Myerson & Green, 1995). D'autres auteurs ont défendu l'idée que le temps était responsable de la dévaluation des renforçateurs dans les environnements temporels comme dans les environnements probabilistes. L'idée étant que plus la probabilité d'obtenir un renforçateur est faible, plus le temps pour l'obtenir sera grande (Rachlin, Raineri, & Cross, 1991).

des paramètres du modèle de temporal discounting restent stables à mesure que l'amplitude du renforçateur augmente (Estle et al., 2006; McKerchar et al., 2010; Myerson et al., 2011).

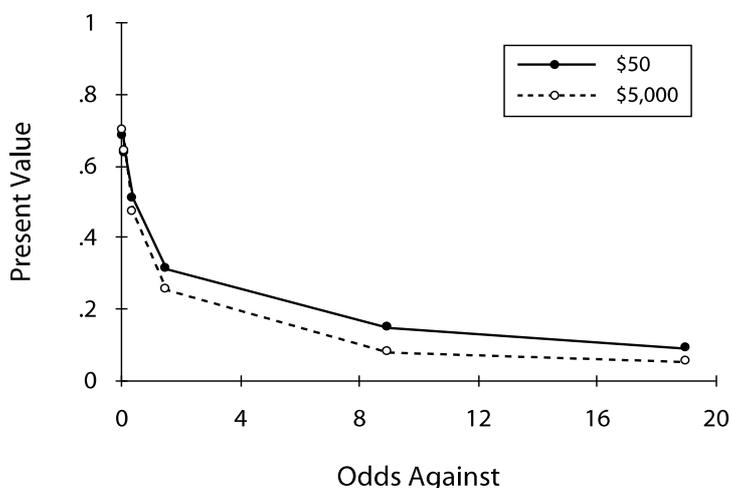


Figure 3.2 Équivalents certains normalisés en fonction de la cote contre l'obtention du renforçateur $x_1 = \$50$ (magnitude faible, trait plein) et $x_1 = \$5000$ (magnitude haute, trait pointillé). (Tiré de Dai et al, 2009)

La Figure 3.2 représente la dévaluation subjective de deux renforçateurs probabilistes n'ayant pas la même amplitude lorsque leur probabilité respective d'obtention diminue. Afin d'être en mesure de mesurer l'impact de la dévaluation subjective pour différents niveaux d'amplitude de renforcement, les auteurs ont simplement normalisé la valeur subjective (c'est-à-dire les montants financiers certains pour lesquels l'individu était indifférent entre la loterie et le montant certain, autrement dit, les équivalents certains) par x_1 , tel que $\frac{V_{ph}(x_1)}{x_1} = \left(\frac{1}{1+h\theta}\right)^s$. On remarque dans la Figure 3.4 que la dévaluation subjective « relative » est plus élevée pour le plus grand renforçateur

Ainsi, sur la base de ces résultats, les chercheurs du domaine concluent qu'il est très probable que des processus différents soient à l'œuvre dans la dévaluation des renforçateurs probabilistes et celle des renforçateurs temporels. En revanche, concernant l'estimation des paramètres du modèle de probability discounting, les chercheurs ne parviennent pas à expliquer pourquoi ce paramètre « s » augmente en fonction de l'enjeu (McKerchar et al., 2010; Myerson et al., 2011).

II. EXPLICATION DU POINT DE VUE DE L'ECONOMIE DE L'EFFET DE MAGNITUDE

L'effet de magnitude tel qu'il est reporté dans les études portant sur le Probability Discounting est en fait une manipulation du niveau d'enjeu des loteries. Cette manipulation a été réalisée de nombreuses fois dans la littérature économique et a donc fait l'objet d'analyses plus ou moins approfondies (Bouchouicha & Vieider, 2017; Fehr-Duda, Bruhin, Epper, & Schubert, 2010; Holt & Laury, 2002; Markowitz, 1952; Scholten & Read, 2014; Weber & Chapman, 2005). Cet effet de l'enjeu sur l'attitude face au risque a reçu différents noms dans la littérature comme le « peanuts effect » (Weber & Chapman, 2005) ou encore le « stake size effect » (Fehr-Duda et al., 2010). Dans cette section, nous reportons le point de vue de la littérature économique sur le sujet. Cette approche nous permettra de remettre en perspective les résultats trouvés par les chercheurs behavioristes et d'expliquer pourquoi le paramètre « s » de leur modèle est instable.

2.1. Niveau d'enjeu et attitude face au risque

Dans cette section nous définissons la notion de taille (ou de niveau) d'enjeu d'une loterie. Dans un second temps nous présentons le pattern comportemental qui est classiquement observé dans la littérature concernant l'impact du niveau d'enjeu sur l'attitude face au risque.

Le niveau d'enjeu et mesure de la prime de risque

Il est difficile de définir exactement la notion d'enjeu. Classiquement, les études expérimentales augmentent le niveau d'enjeu d'une loterie en multipliant par le même facteur toutes les issues d'une loterie donnée (Binswanger, 1982; Bouchouicha & Vieider, 2017; Fehr-Duda et al., 2010; Holt & Laury, 2002). Par exemple, on peut augmenter le niveau d'enjeu d'une loterie noté $L^S \equiv (x_i, p_i)_{i=1, \dots, n}$ comme ceci : $L^H \equiv (\delta x_i, p_i)_{i=1, \dots, n}$, avec $\delta > 1$. Ainsi, on comprend que la notion de petit, moyen ou grand enjeu est une notion relative.

On notera, même si cela est trivial, qu'en multipliant ainsi les issues possibles d'une loterie on augmente alors sa variance mais également son espérance. Dès lors se pose un problème :

comment comparer l'attitude face au risque lorsque le niveau d'enjeu diffère ? Supposons que l'on demande à un agent son équivalent certain pour une loterie $L^S \equiv (10, 0.5 ; 1, 0.5)$ et que nous trouvions $Ce = 3 \sim L^S$. Maintenant supposons que pour le même agent nous recherchions son équivalent certain pour une loterie à plus fort enjeu tel que $L^H \equiv (1000, 0.5 ; 100, 0.5)$ et que nous trouvions $Ce = 300 \sim L^S$. Pour mesurer son degré d'aversion au risque, nous avons vu qu'il était pratique de calculer la prime de risque absolue ($\pi_a(L) = E(L) - Ce(L)$). Respectivement on obtient donc $\pi_a(L^S) = 2$ et $\pi_a(L^H) = 200$. Peut-on pour autant conclure que l'individu est plus averse au risque dans la seconde situation dans la mesure où l'espérance n'est pas la même dans les deux situations ?

On va donc préférer ici des mesures relatives. En l'occurrence, on s'intéressera aux primes de risque relatives :

$$\pi_r(L) = \frac{\pi_a(L)}{E(L)} \quad (3.6)$$

La prime de risque absolue est ici normalisée par l'espérance. Cette prime de risque représente alors la proportion de richesse finale espérée à laquelle l'individu renonce pour éviter le risque. Dans notre exemple, on observe que la prime de risque relative est de 40% dans les deux cas. Autrement dit, si la prime de risque relative étant constante, on pourrait conclure que son aversion au risque relative est constante à travers le niveau d'enjeu (nous y reviendrons plus loin dans ce chapitre).

2.1.1. L'étude de Markowitz (1952)

Markowitz (1952b) propose de façon informelle à ses collègues et amis une expérimentation consistant à observer comment évolue l'attitude face au risque lorsque le niveau d'enjeu d'une loterie augmente (voir Tableau 3.1).

Tableau 3.1. Choix de loteries proposé par Markowitz (1952)

Questions	Sans risque (A)	Loterie (B)	Choix
.1.	\$0.1	$L(1, 0.1 ; 0, 0.9)$	A
.2.	\$1	$L(10, 0.1 ; 0, 0.9)$	A
.3.	\$10	$L(100, 0.1 ; 0, 0.9)$	B
.4.	\$100	$L(1000, 0.1 ; 0, 0.9)$	B
.5.	\$1000	$L(10000, 0.1 ; 0, 0.9)$	B
.6.	\$10000	$L(100000, 0.1 ; 0, 0.9)$	B

Les résultats révèlent que la plupart des individus interrogés préfèrent la loterie pour les questions 1 à 2, c'est-à-dire lorsque l'enjeu n'est pas trop important. En revanche, à partir de la question 3, lorsque les enjeux deviennent plus conséquents, les agents préfèrent l'option A qui assure avec certitude une certaine quantité d'argent égale à l'espérance de la loterie. Dès lors, lorsque les enjeux sont peu importants, les individus préfèrent prendre des risques, du moins jusqu'à un certain seuil. Au-delà de ce « seuil » les agents deviennent riscophobes. Dans cette partie de la thèse, nous nous focalisons sur la section des gains, mais signalons tout de même que Markowitz reproduit l'expérience cette fois dans le domaine des pertes. Il montre que les agents sont averses au risque pour des petites pertes, mais chercheurs de risque lorsque les pertes sont plus substantielles.

Dès lors, sur la base de ces résultats, Markowitz propose une fonction d'utilité à trois points d'inflexions (voir Figure 3.3)

Dans le domaine des gains, la fonction d'utilité est convexe pour des petits gains et devient concave pour de plus grands gains. Dans les pertes, la fonction est d'abord concave pour les petites pertes puis convexe pour les grandes pertes. Le point w_0 signale la richesse initiale de l'individu qui sert ici à séparer le domaine des pertes du domaine des gains. On voit pour la première fois en économie l'apparition du point de référence attaché à la richesse initiale. Toutefois, comme nous l'avons évoqué dans le chapitre 2, Markowitz signalera lui-même que

le point de référence (qu'il appelle « customary wealth ») peut différer de la richesse initiale, mais qu'il n'a pas de formule pour lui donner une spécification précise (Markowitz, 1952).

Finalement l'un des plus importants points d'ombre dans cette fonction, au-delà du point de référence, concerne ce que l'on appelle précisément « petit » et « grand » gain ou « petite » et « grande » perte. Les spécifications des deux points d'inflexion qui servent à séparer ce qui est « petit » du « grand » n'est pas formalisé non plus par Markowitz. Nous présenterons dans la partie 4, un modèle proposé par Kontek et Lewandowski (2017) puis raffiné ensuite par Baucells, Kontek et Lewandowski (2018, working paper) pour approcher une formalisation plus précise de la notion de grand et petit dans une fonction d'utilité de Markowitz.

Depuis les travaux de Markowitz, plusieurs études ont confirmé ces résultats : à probabilité donnée, la prime de risque relative augmente à mesure que l'enjeu augmente (Binswanger, 1982; Bouchouicha & Vieider, 2017; Fehr-Duda et al., 2010; Weber & Chapman, 2005).

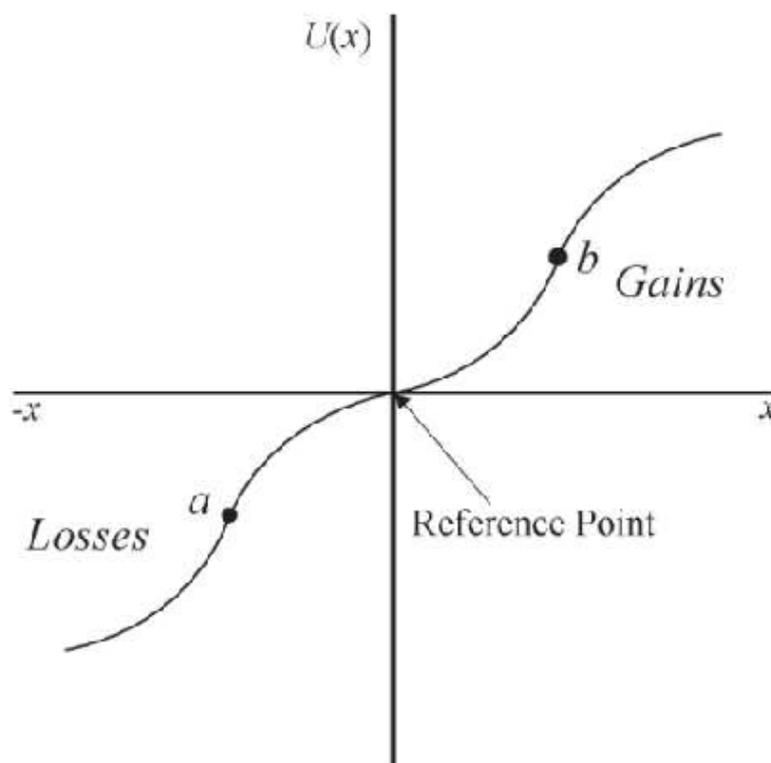


Figure 3.3 Fonction de Markowitz. (Tiré de Kammoun, 2007)

2.1.2. L'étude de Holt et Laury (2002)

Dans ce cadre on peut citer la célèbre étude de Holt et Laury (2002). Les participants devaient systématiquement choisir entre deux loteries binaires non-dégénérées A ou B. La loterie A était moins risquée que la loterie B car la variance était moins importante dans A, de sorte que $A(x_2, x_3)$ et $B(x_1, x_4)$ avec $x_1 > x_2 > x_3 > x_4$. Aucune loterie ne dominait stochastiquement l'autre au premier ordre. La probabilité p associée à la meilleure issue de A était toujours identique à celle associée à la meilleure issue de B. Pour chaque niveau de probabilité p de gagner x_1 ou x_2 , les participants devaient indiquer quelle loterie ils préféraient, avec $p = \{0.1, 0.2, \dots, 0.9\}$. Les issues étaient agencées de sorte que l'espérance de A étaient plus élevée que celle de B quand p était compris entre $p = 0.1$ et $p = 0.4$ compris. De plus, les auteurs ont élaboré plusieurs niveaux d'enjeux : en multipliant les issues du plus petit niveau d'enjeu par différents facteurs δ tel que $\delta = \{20 ; 50 ; 90\}$.

Les résultats ont révélé que les agents étaient averses au risque pour le plus petit enjeu puisque la majorité des participants préféraient A lorsque p était compris entre 0.1 et 0.5. Toutefois, lorsque le niveau d'enjeu augmentait, la proportion d'individu qui préférait A à B était de plus en plus importante. Par exemple, pour la condition $\delta = 90$, plus de 50% des participants préfèrent encore A à B lorsque $p = 0.8$. Ainsi, à mesure que l'enjeu augmentait, l'aversion au risque s'accroissait également.

2.2. Fonction d'utilité et aversion relative face au risque

Comme nous l'avons évoqué dans la section 2.1 de ce chapitre, une manière d'accroître la taille de l'enjeu d'une loterie consiste à multiplier l'ensemble des issues financières par un facteur $\delta > 1$. Si l'on en croit les résultats que nous avons reportés, les individus sont de plus en plus averses au risque à mesure que le niveau d'enjeu augmente. Dans cette section, nous allons détailler les relations entre fonction d'utilité, degré d'aversion au risque et prime de risque

absolue et relative. Cette partie est particulièrement importante car nous nous appuyons sur cette analyse pour expliquer pourquoi le modèle de Probability Discounting ne peut pas capturer les résultats IRRP que nous avons mis en avant dans la section antérieure. Elle nous permettra également de comprendre pourquoi le paramètre « s » augmente à mesure que l'enjeu augmente.

2.2.1. Cadre d'analyse

Nous avons vu dans la section 4.1 du chapitre 1 qu'il était possible de trouver une approximation locale de la prime de risque absolue grâce à l'expansion d'une suite de Taylor. La prime de risque absolue se comprend dans le cadre d'un risque additif, où l'individu possède une richesse certaine qui sera perturbée par les issues d'une variable aléatoire, c'est-à-dire les gains et pertes d'une loterie (une composante aléatoire s'ajoute à la richesse certaine). On formalise une situation de risque additif ainsi : $\tilde{W} = w_0 + \tilde{x}$, avec \tilde{x} une variable aléatoire comprenant n issues x_i . Par conséquent, \tilde{W} est également une variable aléatoire mais portant sur des richesses finales prospectives. Nous avons en particulier noté que la prime de risque dépendait de la variance de la loterie mais également de la courbure de la fonction d'utilité normalisée par sa dérivée première. Cette courbure étant capturée par ce que Arrow et Pratt ont appelé le coefficient d'aversion au risque absolue : $\Gamma_a(E(W)) = -u''(E(W))/u'(E(W))$, avec $E(W) = \sum_{i=1}^n p_i(w_0 + x_i)$. Plus la fonction est concave, plus ce coefficient est élevé (et indique donc que plus l'aversion au risque est importante).

L'hypothèse communément admise est que l'aversion au risque absolue diminue avec l'espérance de richesse (DARA³³). Autrement dit :

³³ Decreasing Absolute Risk Aversion

$$\frac{\partial \Gamma_a(E(W))}{\partial E(W)} < 0 \quad (3.7)$$

Cette hypothèse se comprend aisément. L'espérance de la richesse peut augmenter soit par le biais de w_0 soit par le biais de l'espérance de la loterie. A mesure que w_0 augmente et que la même loterie est présentée à l'individu, alors les richesses finales prospectives sont évaluées sur une section de moins en moins concave de la fonction d'utilité. Autrement dit, l'utilité marginale diminue à mesure que w_0 augmente. Si maintenant nous préservons la même richesse initiale mais que l'espérance de la loterie augmente, cela veut dire que nous n'avons pas présenté la même loterie à l'individu. Nous avons augmenté le montant des issues financière. Si l'augmentation des issues n'a pas modifié la variance de la loterie (en additionnant toutes les issues par le même nombre), seul l'espérance augmente. Là encore, l'aversion au risque absolue diminue selon le même raisonnement.

La prime de risque relative (noté PRR) demande une clarification particulière ici. Nous avons défini la PRR comme la proportion de richesse finale que l'individu est prêt à payer pour être indifférent entre la loterie et un montant certain. On peut également la définir comme le nombre de points de rendements (en % de la richesse total) auxquels un agent est prêt à renoncer pour se débarrasser du risque. On introduit par le biais de cette définition la notion de risque multiplicatif. Il s'agit d'une situation où la totalité de la richesse est soumise au risque, telle que : $\tilde{W} = w_0(1 + \tilde{Y})$. Y est ici une variable aléatoire des différents taux de rendements et w_0 une richesse certaine.

Dans cette thèse nous nous focalisons sur la méthode qui consiste à révéler les équivalents certains de loteries que l'on présente à un individu. On pourrait dans ce cadre penser qu'il n'y a pas de notion de risque multiplicatif dans ces situations. Pourtant, il existe bien une relation entre risque additif et multiplicatif. A titre d'illustration, la loterie $L(x_1, p; x_2, 1 - p)$ peut être perçue comme un risque additif puisque la richesse w_0 sera modifiée par x_1 ou x_2 tel que $W_1 =$

$w_0 + x_1$ ou $W_2 = w_0 + x_2$. On peut tout à fait réécrire la relation qui lie la richesse certaine à l'aléa mais cette fois de manière multiplicative. On procède ainsi : puisque l'espérance de la richesse finale est une richesse certaine, on peut alors écrire $W_c = E(w_0 + \tilde{x}) = E(\tilde{W})$. Dès lors, $\tilde{W} = E(W)(1 + \tilde{Y})$. Les richesses finales s'écrivent $W_1 = E(\tilde{W})(1 + y_1)$ avec $y_1 = (W_1 - E(\tilde{W})) / E(\tilde{W})$ et $W_2 = E(\tilde{W})(1 + y_2)$ avec $y_2 = (W_2 - E(\tilde{W})) / E(\tilde{W})$.

Dans le cas d'un risque multiplicatif, Arrow et Pratt donnent l'approximation autour du point $E(\tilde{W})$ de la prime de risque relative définit selon la seconde définition :

$$\pi_r \approx -\frac{1}{2} \sigma^2(\tilde{Y}) \frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} E(\tilde{W}) \quad (3.8)$$

Nous remarquons que la prime de risque relative dépend de la variance des rendements et du coefficient d'aversion au risque qui est cette fois relatif $\Gamma_r(E(\tilde{W})) = -\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} E(\tilde{W})$. On notera la relation qui existe entre aversion absolue et relative :

$$E(\tilde{W})\Gamma_a(E(\tilde{W})) = \Gamma_r(E(\tilde{W}))$$

Cet indice d'aversion relative mesure l'élasticité de l'utilité marginale u' au point $E(\tilde{W})$, c'est à dire le taux auquel l'utilité marginale décroît ou s'accroît lorsque la richesse certaine augmente de 1% (Cayatte, 2009).

On peut retrouver facilement l'approximation de la prime de risque absolue à partir de celle de la prime de risque relative facilement. Il suffit de multiplier l'approximation de la prime de risque relative par $E(\tilde{W})$:

$$\pi_a \approx -\frac{1}{2}\sigma^2(\tilde{Y}) \frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} (E(\tilde{W}))^2$$

Dès lors, de ces relations on déduit facilement l'Equation (3.8) puisqu'en divisant la prime de risque absolue par $E(\tilde{W})$ on obtient bien l'approximation de la prime de risque relative formalisé par l'Equation (3.8).

2.2.2. Propriétés des fonctions d'utilités et aversion relative face au risque

L'Equation (3.8) indique que la prime de risque relative dépend de la variance des rendements et du coefficient d'aversion au risque relative. La question qui se pose maintenant est celle-ci : comment devrait évoluer l'aversion au risque relative lorsque la richesse certaine (ou l'espérance de richesse finale) augmente d'une proportion donnée ? La difficulté réside dans le fait que lorsque la richesse certaine augmente, la variance des richesses finales augmentent également à cause du risque multiplicatif³⁴. Pour un individu riscophobe nous avons vu que l'hypothèse DARA semblait raisonnable. Lorsque l'on dérive l'aversion au risque relative par rapport à l'espérance de richesse finale, on obtient :

$$\frac{d\Gamma_r(E(\tilde{W}))}{dE(\tilde{W})} = \Gamma_a(E(\tilde{W})) + E(\tilde{W}) \frac{d\Gamma_a(E(\tilde{W}))}{dE(\tilde{W})}$$

Dès lors, pour un individu riscophobe le signe de la dérivée du coefficient d'aversion au risque relatif est difficile à déterminer car le premier terme est positif (puisque la fonction d'utilité est croissante concave), $E(\tilde{W})$ est positif également mais en vertu de l'Equation (3.7) la dérivée

³⁴ Nous avons évoqué qu'habituellement la taille d'enjeu était augmentée en multipliant les issues d'une loterie par un certain facteur. Il s'agit simplement d'une autre manière de d'augmenter le risque de manière multiplicative. L'effet est le même : la variance ainsi que l'espérance (donc la richesse certaine) augmentent.

du coefficient d'aversion au risque absolue est négative (l'aversion au risque absolue diminue avec la richesse pour des fonctions DARA). Par conséquent, selon cette équation, une augmentation de la richesse certaine diminue bien l'aversion au risque absolue, mais à cause du risque multiplicatif l'impact de l'augmentation de la richesse augmente la variance des richesses finales (Cayatte, 2009). En général, on pose comme hypothèse que l'aversion au risque relative ne diminue pas avec l'augmentation de la richesse. Autrement dit, seules les fonctions CRRA³⁵ et IRRA³⁶ sont considérées comme plausibles.

L'augmentation, la diminution ou la constance de l'aversion au risque relative (et de la prime de risque relative) dépendent de la fonction d'utilité supposée. Nous allons nous demander quelles fonctions d'utilités sont en mesure d'expliquer le pattern : l'aversion au risque relative augmente lorsque la taille de l'enjeu s'accroît. Autrement dit, nous recherchons les fonctions qui possèdent la propriété IRRA.

La fonction d'utilité la plus utilisée en économie est sans nul doute la fonction puissance :

$$u(x) = \frac{1}{1-\alpha} x^{1-\alpha} \quad (3.9)$$

Le coefficient d'aversion au risque relatif avec une fonction d'utilité puissance est constant ($\Gamma_r(E(\tilde{W})) = \alpha$) à mesure que la richesse augmente d'une certaine proportion (CRRA) alors que l'aversion au risque absolue décroît quand la richesse certaine augmente (DARA). Par conséquent, la prime de risque relative reste constante lorsque la richesse certaine (ou l'espérance de richesse) augmente d'une proportion constante.

³⁵ Constant Relative Risk Aversion

³⁶ Increasing Relative Risk Aversion

Les fonctions d'utilités dites IRRA³⁷ ont la propriété suivante : à mesure que la richesse certaine augmente d'une certaine proportion, l'utilité augmente d'une proportion de plus en plus faible.

Deux fonctions IRRA sont souvent utilisés dans la littérature. La fonction de Saha (1993) :

$$u(x) = \frac{1}{\alpha} (1 - \exp(-\alpha x^{1-r})) \quad (3.10)$$

Et la fonction de Rachlin (1992), avec le paramètre libre β qui contrôle la courbure de la fonction d'utilité:

$$u(x) = \frac{1}{\beta} \log(1 + \beta x) \quad (3.11)$$

Le coefficient d'aversion au risque relatifs $\Gamma_r(x)$ selon une fonction d'utilité Saha est :

$$\Gamma_r(x) = \frac{\partial^2 u(x) x}{\partial u(x)} = r + \alpha(1-r)x^{1-r}$$

Maintenant, si on dérive le coefficient d'aversion au risque relatif selon une fonction d'utilité de type Saha, celle-ci est positive par rapport à l'augmentation marginale de la richesse :

$$\forall r > 1; \frac{\partial \Gamma_r(x)}{\partial x} = \frac{\alpha(r-1)}{x^r} > 0$$

Autrement dit, l'aversion au risque relative augmente avec la richesse si $r > 1$ (si $r > 1$ la fonction d'utilité est concave).

Le coefficient d'aversion au risque relatif selon une fonction d'utilité de type Rachlin est :

³⁷ Increasing Relative Risk Aversion

$$\Gamma_r(x) = \frac{\partial^2 u(x)x}{\partial u(x)} = \frac{\beta x}{1 + \beta x}$$

Si on dérive le coefficient d'aversion au risque relatif selon une fonction d'utilité de type Rachlin, celle-ci est positive par rapport à l'augmentation de la richesse :

$$\forall \beta > 0; \Gamma_r(x) = \frac{\partial \Gamma_r(x)}{\partial x} = \frac{\beta}{(1 + \beta x)^2}$$

Autrement dit, l'aversion au risque relative augmente avec la richesse si $\beta > 0$.

Dans les deux cas, les dérivées des fonctions d'aversion au risque relative par rapport à la richesse sont positives. Par conséquent, la prime de risque relative augmente avec l'augmentation de la richesse.

Dès lors, en vertu des résultats révélés par Markowitz (1952), Holt & Laury (2002) une fonction d'utilité IRRA semble plus pertinente. En revanche, intégrer ces fonctions dans un modèle d'espérance d'utilité ne suffit pas à accommoder les comportements décisionnels. Comme nous l'avons exposé dans le chapitre 1, les individus sont plutôt chercheurs de risque lorsque la probabilité de gagner est faible et averse au risque lorsque la probabilité de gagner est plus élevée (voir le paradoxe d'Allais). Or l'EUT ne peut pas capturer ce pattern comportemental. Nous avons vu que le modèle RDU, grâce à l'incorporation d'une fonction de pondération des probabilités avait un meilleur pouvoir descriptif. Dès lors, un modèle RDU armé d'une fonction de pondération des probabilités non-linéaire et d'une fonction d'utilité IRRA devrait donc se révéler pertinente pour capturer les prises de décision des agents en environnement risqué.

2.2.3. Approximation de la prime de risque relative basé sur l'analyse de Hilton (1988)

Dans cette section, nous montrons que les propriétés des fonctions d'utilité vu précédemment restent opérantes mêmes lorsque le modèle de décision comporte une fonction de pondération des probabilités cumulées. Autrement dit, nous verrons qu'un modèle RDU armé d'une fonction d'utilité CRRA et d'une fonction pondération des probabilités induit toujours une prime de risque relative constante. De même un modèle RDUT armé d'une fonction IRRA et d'une fonction de pondération des probabilités induit toujours une prime de risque relative croissante.

Nous partons de l'approximation de la prime de risque absolue selon un modèle RDU donné par Hilton (1988) :

$$\pi_a(\tilde{W}) \approx \frac{1}{2} \sigma_{\Phi}^2(\tilde{W}) \left[-\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} \right] + (E(\tilde{W}) - E_{\Phi}(\tilde{W}))$$

On sait par ailleurs que la prime de risque relative est donnée par $\pi_r(\tilde{W}) = \pi_a(\tilde{W})/E(\tilde{W})$

$$\pi_r(\tilde{W}) \approx \frac{1}{E(\tilde{W})} \left[\frac{1}{2} \sigma_{\Phi}^2(\tilde{W}) \left[-\frac{u''(E(\tilde{W}))}{u'(E(\tilde{W}))} \right] + (E(\tilde{W}) - E_{\Phi}(\tilde{W})) \right]$$

Etudions maintenant comment évolue l'approximation de la prime de risque relative lorsque les issues d'une loterie sont multipliées par un facteur $\delta > 1$ en supposant une fonction d'utilité puissance (Equation (3.9) dans le cadre de la RDUT. Pour cela, nous prenons une loterie $L^S \equiv (x_i, p_i)_{i=1, \dots, n}$ et une loterie à plus fort enjeu $L^H \equiv (\delta x_i, p_i)_{i=1, \dots, n}$. Nous allons montrer que $\pi_r(L^S) = \pi_r(L^H)$ lorsqu'on suppose une fonction d'utilité puissance.

Fonction d'utilité puissance

Tout d'abord on sait que

$$E(L^H) = \delta E(L^S)$$

puisque,

$$E(L^S) = \sum_{i=1}^n p_i x_i$$

$$E(L^H) = \sum_{i=1}^n p_i \delta x_i = \delta \sum_{i=1}^n p_i x_i = \delta E(L^S)$$

De cette égalité on déduit que :

$$(E(L^H) - E_{\Phi}(L^H)) = \delta (E(L^S) - E_{\Phi}(L^S))$$

On procède de la même manière concernant le terme de la variance

$$\frac{1}{2} \sigma_{\Phi}^2(L^S) = \frac{1}{2} \sum_{i=1}^n \Phi_i (x_i - E_{\Phi}(L^S))^2$$

$$\begin{aligned} \frac{1}{2} \sigma_{\Phi}^2(L^H) &= \frac{1}{2} \sum_{i=1}^n \Phi_i (\delta x_i - \delta E_{\Phi}(L^S))^2 = \frac{1}{2} \sum_{i=1}^n \Phi_i (\delta (x_i - E_{\Phi}(L^S)))^2 \\ &= \delta^2 \frac{1}{2} \sum_{i=1}^n \Phi_i (x_i - E_{\Phi}(L^S))^2 = \delta^2 \frac{1}{2} \sigma_{\Phi}^2(L^S) \end{aligned}$$

Concernant le coefficient d'Arrow-Pratt représenté par $\Gamma_{\alpha}(E(L^S)) = -\frac{u''(E(L^S))}{u'(E(L^S))}$. Avec une fonction d'utilité puissance on obtient :

$$\Gamma_{\alpha}(E(L^S)) = \frac{\alpha}{E(L^S)}$$

On peut alors montrer que le coefficient d'aversion au risque pour la loterie L^H est :

$$\Gamma_a(E(L^H)) = \frac{1}{\delta} \frac{\alpha}{E(L^S)} = \frac{1}{\delta} \Gamma_a(E(L^S))$$

Ainsi on peut écrire la prime de risque relative de la loterie L^H comme suit :

$$\pi_r(L^H) = \frac{1}{\delta E(L^S)} \left[\delta^2 \frac{1}{2} \sigma_{\Phi}^2(L^S) \left[-\frac{1}{\delta} \frac{u''(E(L^S))}{u'(E(L^S))} \right] + \delta (E(L^S) - E_{\Phi}(L^S)) \right]$$

Ce qui après simplification nous donne une expression égale à la prime de risque relative de la loterie L^S :

$$\pi_r(L^H) = \frac{1}{E(L^S)} \left[\frac{1}{2} \sigma_{\Phi}^2(L^S) \left[-\frac{u''(E(L^S))}{u'(E(L^S))} \right] + (E(L^S) - E_{\Phi}(L^S)) \right] = \frac{\pi_a(L^S)}{E(L^S)} = \pi_r(L^S)$$

Fonction d'utilité Rachlin (IRRA)

Dans cette section nous montrons qu'en supposant une fonction d'utilité de Rachlin dans un modèle de décision RDUT, on obtient $\pi_r(L^S) < \pi_r(L^H)$. Le seul élément qui change par rapport à notre première analyse avec la fonction d'utilité puissance concerne le terme représentant le coefficient d'Arrow-Pratt. Avec une fonction d'utilité Rachlin on obtient pour la loterie L^S le coefficient d'aversion au risque absolue suivant :

$$\Gamma_a(L^S) = \frac{\beta}{\beta E(L^S) + 1}$$

Pour la loterie L^H , on sait que $E(L^H) = \delta E(L^S)$. Ainsi, le coefficient d'aversion au risque absolue pour la loterie L^H est :

$$\Gamma_a(L^H) = \frac{\beta}{\beta\delta E(L^S) + 1}$$

L'approximation de la prime de risque relative obtenue pour la loterie L^S selon une fonction d'utilité de Rachlin est :

$$\pi_r(L^S) = \frac{1}{E(L^S)} \left[\frac{1}{2} \sigma_{\Phi}^2(L^S) \frac{\beta}{\beta E(L^S) + 1} + (E(L^S) - E_{\Phi}(L^S)) \right]$$

Par corollaire, la prime de risque relative de la loterie L^H est :

$$\begin{aligned} \pi_r(L^H) &= \frac{1}{\delta E(L^S)} \left[\delta^2 \frac{1}{2} \sigma_{\Phi}^2(L^S) \frac{\beta}{\beta\delta E(L^S) + 1} + (E(L^S) - E_{\Phi}(L^S)) \right] \\ &= \frac{1}{E(L^S)} \left[\frac{1}{2} \sigma_{\Phi}^2(L^S) \frac{\delta\beta}{\beta\delta E(L^S) + 1} + (E(L^S) - E_{\Phi}(L^S)) \right] > \pi_r(L^S) \end{aligned}$$

Nous devons maintenant prouver que $\pi_r(L^H) > \pi_r(L^S)$. La seule différence entre la prime de risque relative de L^H et celle de L^S réside dans le coefficient d'aversion au risque absolue. Il nous suffit donc de montrer $\frac{\delta\beta}{\beta\delta E(L^S) + 1} > \frac{\beta}{\beta E(L^S) + 1}$ si $\delta > 1$.

Ce qui est effectivement le cas puisque :

$$\frac{\delta\beta}{\alpha\delta E(L^S) + 1} > \frac{\beta}{\alpha E(L^S) + 1}$$

$$\beta E(L^S) + 1 > \frac{\beta\delta E(L^S) + 1}{\delta}$$

$$\beta\delta E(L^S) + \delta > \beta\delta E(L^S) + 1$$

$$\delta > 1$$

La Figure 3.4 représente des primes de risque relatives simulées à partir d'un modèle RDU pour trois niveaux d'enjeux différents $L^{small} \equiv (100, p_1; 10, 1 - p_1)$, $L^{medium} \equiv (1000, p_1; 100, 1 - p_1)$, $L^{high} \equiv (10000, p_1; 1000, 1 - p_1)$ croisés avec neuf probabilités de gagner la meilleure issue : $p_1 = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. A gauche, nous représentons les primes de risque relatives en supposant une fonction d'utilité CRRA (fonction puissance avec $a = 0.4$). A droite nous représentons les primes de risques relatives en supposant une fonction d'utilité IRRA (fonction de Rachlin avec $\beta = 0.031$). Pour les deux modèles nous supposons une fonction de pondération de probabilité à deux paramètres de Green et Myerson (Equation 1.10, où $h = 4$ et $s = 0.4$).

Pour conclure, la théorie économique est donc bien en mesure d'expliquer que la prime de risque relative augmente à mesure que l'enjeu augmente. Il suffit pour cela d'armer n'importe quel modèle de décision d'une fonction d'utilité IRRA.

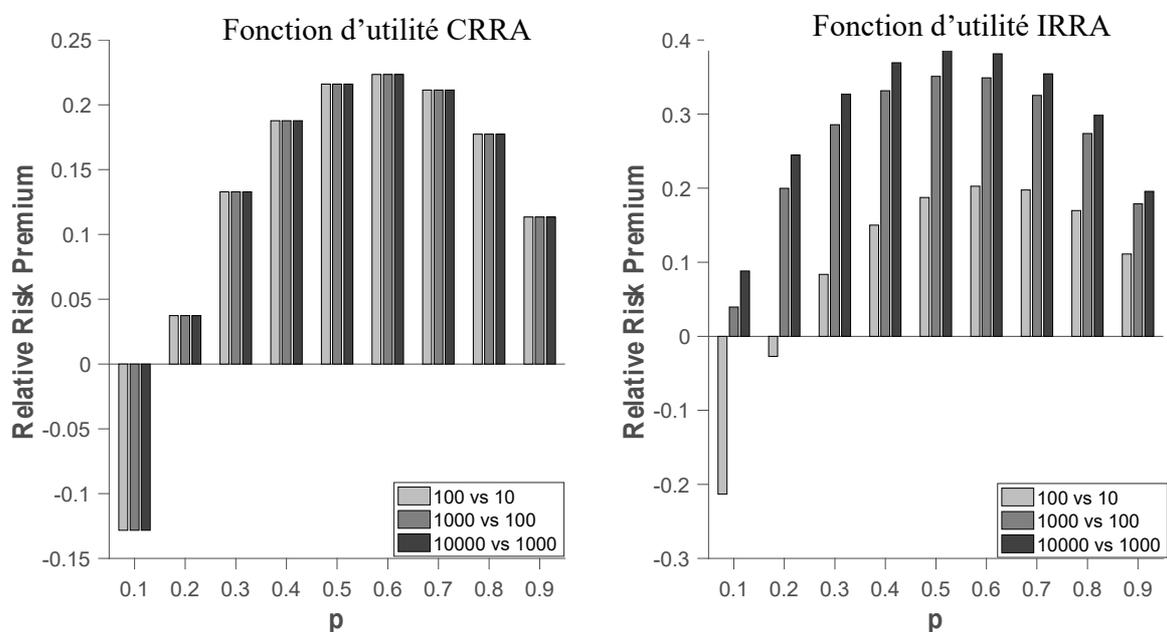


Figure 3.4 Représentation des primes de risques selon un modèle RDU armé d'une fonction d'utilité puissance (à gauche) ou d'une fonction d'utilité Rachlin (à droite).

2.2.4. Elasticité des fonctions d'utilité IRRA et CRRA

Ici nous complétons rapidement notre propos sur les implications des propriétés des fonctions d'utilité sur l'attitude face au risque. La différence entre ces deux types de fonction d'utilités s'analyse très bien en termes d'élasticité (Scholten & Read, 2014). L'élasticité correspond à la modification en pourcentage d'une variable dépendante causé par la modification en pourcentage d'une variable indépendante. Il existe plusieurs mesures théoriques de l'élasticité comme l'élasticité point ou encore l'élasticité arc. Une façon de connaître la fonction d'élasticité d'une fonction $u(x)$ est la suivante :

$$\varepsilon(u(x)) = \frac{\partial \log(u(x))}{\partial \log(x)} \quad (3.12)$$

Par exemple, une élasticité égale à 0.5 signifie que lorsque la richesse augmente de 1%, l'utilité augmente de 0.5%.

La différence fondamentale entre une fonction d'utilité CRRA et IRRA réside dans leurs élasticités respectives. L'élasticité d'une fonction puissance (Equation (3.9)) est constante. Autrement dit à mesure que la richesse augmente de 1% l'utilité augmente de C%, avec C une constante qui ne dépend donc pas de du niveau de richesse. En revanche, l'élasticité d'une fonction IRRA est décroissante avec l'augmentation en proportion de la richesse. Autrement dit, à mesure que la richesse augmente de 1% la fonction d'utilité augmente d'une proportion de plus en plus petite.

III. RETOUR SUR LE MODELE DE PROBABILITY DISCOUNTING

Le développement de la section II de ce chapitre nous a permis d'illustrer l'influence des propriétés des fonctions d'utilités sur l'attitude face au risque lorsque la richesse augmente. Sur la base de ces informations, nous ré-analysons le modèle de Probability Discounting (Equation

(3.5) et nous montrons que ce modèle n'est pas en mesure de capturer le pattern comportemental observé lorsque la magnitude d'un renforçateur probabiliste augmente.

3.1. Retour sur les résultats du Probability Discounting

3.1.1. La prime de risque relative et l'équivalent certain normalisé

Nous avons vu dans la section I, que les chercheurs reportaient sur les graphiques des valeurs d'équivalents certains normalisés : $V(x_1)/x_1$ selon leurs écritures. Nous pourrions plutôt l'écrire ainsi : Ce/x_1 . Précisons que cette formulation n'est valable que si la loterie ne comprend que deux issues possibles avec un seul résultat différent de 0. Les auteurs reportent que les équivalents certains normalisés par rapport à la probabilité de perdre diminuent moins rapidement à mesure que l'amplitude du renforçateur est élevé.

En termes économique, l'augmentation de l'amplitude du renforçateur équivaut à une augmentation de l'enjeu (En effet chaque issue de la loterie L_D est multipliée par un facteur δ . Seulement, étant donné que x_2 est égal à 0, ce dernier reste égal à 0 pour n'importe quelle valeur de x_1). Ainsi, les résultats du Probability Discounting indiquent que les équivalents certains normalisés diminuent moins rapidement au travers des probabilités de perdre à mesure que l'enjeu augmente d'une certaine proportion.

Traduit en termes de prime de risque relative, cela indique que cette dernière augmente à mesure que l'enjeu augmente (ou dit autrement, à mesure que la proportion d'augmentation de x_1 augmente). On note ce pattern par l'acronyme IRRP pour Increasing Relative Risk Premium. En effet, trivialement en reprenant l'équation de la prime de risque relative, cela veut dire que l'équivalent certain augmente en proportion moins vite que celle de l'espérance. Autrement dit, l'élasticité de l'équivalent certain est décroissante. Les agents semblent alors de plus en plus averse au risque à mesure que l'enjeu augmente.

3.1.2. Le modèle de Probability Discounting à l'épreuve des résultats

Dans cette section nous confrontons le modèle de Probability Discounting proposé par Myerson, Green et Myerson (2011), que nous dénotons par l'acronyme MGM, présenté par l'Equation (3.5) aux résultats obtenus dans la littérature.

a. Le modèle de probability Discounting : Equivalent certain de la RDU ou variante de la Dual theory ?

Le modèle de Probability Discounting est interprétable de deux façons différentes. Rappelons d'abord le modèle MGM: la valeur subjective d'un renforçateur probabiliste x_1 est donnée par :

$$V_{ph}(L_D) \equiv V_{ph}(x_1) = x_1 \left(\frac{p_1}{p_1 + h(1 - p_1)} \right)^s$$

On peut estimer qu'il s'agit d'un modèle s'intégrant dans le cadre de la Dual Theory puisque x_1 n'est pas affecté d'une fonction d'utilité non-linéaire. Par contre les probabilités étant transformées par une fonction de pondération, l'attitude face au risque s'exprime alors par le biais de cette déformation des probabilités comme dans la Dual Theory.

Ou bien, on peut également concevoir ce modèle comme l'équivalent certain d'un modèle RDU. L'évaluation d'une loterie binaire sous RDU s'exprime ainsi :

$$RDU(L_D) = \varphi(p_1)u(x_1) + (1 - \varphi(p_1))u(x_2)$$

Comme $x_2 = 0$, l'expression se réduit à $u(x_1)\varphi(p_1)$. Si on suppose une fonction d'utilité puissance (Equation (3.9)), l'équivalent certain s'écrit alors :

$$Ce(L_D) = u^{-1}[RDU(L_D)] = \left[1 - \alpha \frac{x_1^{1-\alpha}}{1-\alpha} \left(\frac{p_1}{p_1 + h(1-p_1)} \right)^s \right]^{1/(1-\alpha)}$$

En simplifiant cette dernière expression on a alors :

$$Ce(L_D) = x_1 \left(\frac{p_1}{p_1 + h(1-p_1)} \right)^\zeta \quad (3.13)$$

Avec $\zeta \equiv s/(1-\alpha)$. On retrouve alors ici le modèle de Probability Discounting décrit dans l'Equation (3.5). Avec toutefois une petite subtilité. En effet, on retrouve dans cette dernière expression une formalisation explicite de l'interprétation donnée par Green et Myerson (2004) à propos du paramètre "s". Pour rappel, les auteurs déclarent que "s" est un paramètre contrôlant la mise à l'échelle psychophysique des probabilités ET de l'amplitude de renforcement. Toutefois, au lieu de représenter le modèle de Probability Discounting avec le ratio $s/(1-\alpha)$, les auteurs le réduisent à un seul paramètre « s ». Les auteurs n'expliquent pas pourquoi ils opèrent de cette manière, et finalement la traduction que nous faisons de ce modèle n'est jamais explicité par les auteurs. Nous verrons plus loin dans le chapitre qu'une raison possible de cette « réduction » du modèle à deux paramètres au lieu de trois paramètres consiste à éviter un problème d'indétermination, même si cela n'est pas précisée par les auteurs.

En ce qui nous concerne, pour référer au modèle à deux paramètres tel qu'il est écrit par Myerson, Green et Morris (2011), nous notons ce ratio par la lettre grecque ζ pour éviter toute confusion avec le modèle à trois paramètres.

Inconsistance entre le modèle et le pattern IRRP

Le modèle MGM semble donc supposer une fonction de pondération des probabilités à deux paramètres et une fonction d'utilité puissance. Nous avons vu dans la section II de ce chapitre,

qu'une fonction d'utilité puissance implique que la prime de risque relative reste constante lorsque le niveau d'enjeu (ou l'amplitude de x_1) augmente d'une certaine proportion. Dit autrement, l'élasticité de l'équivalent par rapport à l'augmentation en proportion de x_1 est constante si on suppose une fonction d'utilité puissance. Donc, en vertu de l'Equation (3.6) de la prime de risque relative on remarque que la prime de risque reste constante. Or, les résultats révélés par les expériences correspondent à un pattern IRRP.

Selon nous, cette inconsistance explique l'augmentation du paramètre ζ du modèle MGM à deux paramètres. Rappelons que plus le paramètre alpha de la fonction puissance est grand, plus la courbure concave de la fonction d'utilité est importante, ce qui veut dire que plus le degré d'aversion au risque est élevé. Dès lors, une façon d'accommoder le pattern IRRP avec une fonction d'utilité puissance consiste à augmenter la valeur de ce paramètre à mesure que l'enjeu augmente. Or, dans le modèle à trois paramètres, plus le paramètre alpha augmente plus le ratio est grand. Par conséquent, dans le modèle à deux paramètres le paramètre ζ augmente. Cette hypothèse qui consiste à supposer que seule la fonction d'utilité réagit et non la fonction de pondération des probabilités lorsque les enjeux sont modifiés respecte le principe de séparabilité (Bouchouicha & Vieider, 2017; Fehr-Duda et al., 2010) : « principle according to which changes in preferences over outcomes ought to be reflected purely in utility, while changes in preferences over probabilities ought to be reflected in probability weighting » (Bouchouicha & Vieider, 2017, p.2). Pour vérifier notre hypothèse, nous procédons à des simulations-estimations que nous développons dans la partie suivante.

3.2. Simulation-Estimation du modèle MGM à trois paramètres

Pour procéder à l'estimation du modèle MGM à trois paramètres, nous allons simuler des équivalents certains à partir d'un modèle que nous spécifions plus loin. Toutefois, le premier élément que nous devons signaler ici est qu'en l'état, le modèle d'équivalent certain provenant du modèle MGM à trois paramètres souffre d'un problème d'indétermination à cause du ratio

$s/(1 - \alpha)$. En effet, un même équivalent certain peut être capturé par une infinité de valeurs possibles de s et α .

Pour éviter ce problème d'indétermination, la procédure classique consiste à construire des loteries à deux issues financières possibles différentes de 0, tel que $L \equiv (x_1, p_1; x_2, 1 - p_1)$. Dans ce cas, l'équation de l'équivalent certain dérivé d'un modèle RDU pour une loterie à deux issues différentes de 0 ne semble plus souffrir d'un problème d'indétermination :

$$CE(L, \theta) = \left[x_1^{1-\alpha} \left(\frac{p}{p + h(1-p)} \right)^s + x_2^{1-\alpha} \left(1 - \left(\frac{p}{p + h(1-p)} \right)^s \right) \right]^{\frac{1}{1-\alpha}} \quad (3.14)$$

Avec $\theta = (\alpha, s, h)$.

3.2.1. Loteries

Dans notre procédure, nous cherchons à estimer les paramètres du modèle MGM à trois paramètres pour trois niveaux d'enjeux différents. Nous suivons alors la procédure de Fehr-Duda et al (2010) qui consiste à établir trois loteries différentes par niveau d'enjeu afin de rendre la base de données plus riche et favoriser la convergence de l'estimation des paramètres du modèle.

Spécifiquement, pour le petit enjeu nous choisissons $L_1^s \equiv (x_1 = 100; x_2 = 0)$, $L_2^s \equiv (x_1 = 110; x_2 = 20)$ and $L_3^s \equiv (x_1 = 120; x_2 = 30)$. Pour l'enjeu moyen, nous avons $L_1^m \equiv (x_1 = 1,000; x_2 = 0)$, $L_2^m \equiv (x_1 = 1,100; x_2 = 200)$ and, $L_3^m \equiv (x_1 = 1,200; x_2 = 300)$. Finalement pour le grand enjeu nous choisissons $L_1^h \equiv (x_1 = 10,000; x_2 = 0)$, $L_2^h \equiv (x_1 = 11,000; x_2 = 2,000)$ and $L_3^h \equiv (x_1 = 12,000; x_2 = 3,000)$. Par ailleurs, chaque loterie est croisée avec huit probabilités p_1 de gagner x_1 . En définitive, nous obtenons donc 72 jeux, indexé

par la lettre $g = 1, \dots, 72$, pour lesquels nous simulons les équivalents certains (à raison de 200 agents artificiels par jeu).

3.2.2. Simulation des Equivalents Certains

Pour chaque jeu g nous simulons son équivalent certain selon une fonction d'utilité Rachlin (Equation (3.11), noté R et une fonction de pondération des probabilités donné par le modèle MGM (Equation 1.10). Nous notons l'équivalent certain d'un jeu selon les spécifications indiqués ci-dessus ainsi³⁸ : \widehat{CE}_g^R :

$$\widehat{CE}_g^R = u^{-1}[w(p_g)u(x_{1g}) + (1 - w(p_g))u(x_{2g})] \quad (3.15)$$

Nous intégrons une structure stochastique à ce modèle théorique représenté par ε_{ig} , avec i représentant un individu. Pour chaque jeu nous procédons à 200 tirages, donc $i = 1, \dots, n$, avec $n = 200$. Ce terme reflète n'importe quelle source d'erreur que l'on retrouve dans le monde réel et que nous avons évoqué dans le chapitre 1 section V. Dès lors l'équivalent certain « réel » s'écrit :

$$CE_{ig} = \widehat{CE}_g^R + \varepsilon_{ig} \quad (3.16)$$

Dans la lignée de Fehr-Duda et al (2010) et Bouchouicha et al (2017), nous supposons que ces termes d'erreurs, pour chaque i et g sont indépendamment tirés d'une loi normale de moyenne 0 et d'écart type $\sigma_g = \eta|x_{1g} - x_{2g}|$ lequel est proportionnel à la taille de l'enjeu. Donc, nous

³⁸ Explicitement on obtient le modèle d'équivalent certain suivant: $\frac{1}{\beta} \log(1 + \beta x_1) \left(\frac{p_1}{p_1 + h(1-p_1)} \right)^s + \frac{1}{\beta} \log(1 + \beta x_2) \left[1 - \left(\frac{p_1}{p_1 + h(1-p_1)} \right)^s \right]$.

supposons que les erreurs sont homoscédastiques sans perte de généralité vis à vis de notre propos.

Concernant les valeurs des paramètres des fonctions d'utilité et de pondération des probabilités, nous établissons des valeurs tirées de la littérature tel que : Pour la fonction de Rachlin $\alpha = 0.031$, $s = 0.4$ et $h = 4$ et $\eta = 0.01$.

3.2.3. Estimation

Rappelons que notre objectif est d'ajuster les données simulées par un modèle IRRP (c'est-à-dire des CE_{ig} simulés par l'Equation (3.16) aux équivalents certains prévus par le modèle MGM à trois paramètres. Ainsi, nous estimons l'Equation (3.16) dans laquelle nous remplaçons \widehat{CE}_g^R par la prédiction MGM, désignée par \widehat{CE}_g^M et décrite dans l'Equation (3.14). Ce modèle est estimé par Estimation du Maximum de vraisemblance à l'aide du logiciel R. En effet, nous maximisons la log-vraisemblance de l'échantillon simulé, ce qui revient à résoudre le programme suivant :

$$\max_{\theta} \sum_i \sum_g \log \left[\frac{1}{\sigma_g} \phi \left(\frac{CE_{ig} - \widehat{CE}_g^M}{\sigma_g} \right) \right] \quad (3.17)$$

Avec $\theta = (\alpha, s, h, \eta)$ et ϕ la densité de la distribution normal standard. Puisque nous nous intéressons à la stabilité des paramètres à travers les niveaux d'enjeux, nous utilisons deux dummy variable dans notre modèle, appelé « MEDIUM » ET « HIGH », où HIGH prend la valeur 1 quand $x_1 \geq 10000$ et 0 sinon. La variable MEDIUM prend la valeur 1 quand $120 < x_1 < 10000$ et 0 sinon. Le niveau d'enjeu « SMALL » est donc notre groupe de référence. A travers cette spécification, nous supposons que les paramètres du modèle MGM dépendent linéairement des « dummies » variables HIGH et MEDIUM tel que :

$$\theta = \theta_0 + \theta_{MEDIUM}MEDIUM + \theta_{HIGH}HIGH \quad (3.18)$$

Avec θ_0 représente les paramètres du modèle MGM pour le plus petit enjeu. Si l'attitude face au risque augmente avec le niveau d'enjeu, au moins un des paramètres devrait être significativement différent de 0, suggérant ainsi que le modèle est mal spécifié ou bien ne peut pas expliquer le pattern IRRP. Sous le principe de séparabilité, nous prédisons que seul le paramètre alpha de la fonction d'utilité puissance devrait être sensible (devrait augmenter) au niveau d'enjeu.

3.3. Résultats

Les estimations des paramètres sont regroupées dans le Tableau 3.2. Le panneau supérieur "simulation" affiche les valeurs des paramètres que nous utilisons pour générer des équivalents certains à partir du modèle de Rachlin. Le panneau inférieur "estimation" est divisé en deux parties. La partie supérieure présente les résultats du « parameters recovery » du modèle de Rachlin pour évaluer la validité de notre procédure d'estimation. En effet, nous trouvons que les vraies valeurs des paramètres sont estimées assez précisément ce qui signifie que notre protocole de simulation et d'estimation est fiable.

Considérons maintenant les estimations du modèle MGM. Tout d'abord, nous remarquons que l'estimation de α diminue avec la taille de l'enjeu puisque l'inégalité $0 < \hat{\alpha}_{MEDIUM} < \hat{\alpha}_{HIGH}$ est statistiquement significative (les intervalles de confiance ne se chevauchent pas et ne contiennent pas 0). Deuxièmement, nous observons que les paramètres de la fonction de pondération des probabilités varient légèrement. En effet, l'estimation de h ne dépend pas de la taille de l'enjeu puisque \hat{h}_{MEDIUM} et \hat{h}_{HIGH} ne sont pas significativement différentes de 0, mais nous notons que $\hat{\gamma}_{HIGH}$ est négatif et significatif mais tout à fait négligeable par rapport à $\hat{\gamma}_0$.

Ainsi, la variation par rapport à la taille de l'enjeu nécessaire pour accommoder le modèle IRRP est principalement attribuable à la variation de la fonction d'utilité par le paramètre α .

Tableau 3.2. Valeurs des paramètres ayant été utilisés pour la simulation des équivalents certains et résultats des estimations

SIMULATION WITH THE RACHLIN MODEL				
Parameters	β	s	h	η
Input simulation	0.031	0.400	4.000	0.100

ESTIMATION				
Parameters	β	s	h	η
	$\hat{\theta}_0$	0.401	3.860	0.099
Estimation Rachlin model	$\hat{\theta}_{Medium}$	[0.391 0.412]	[3.680 4.041]	[0.098 0.101]
	$\hat{\theta}_{High}$	0.005	0.050	0.001
		[-0.010 0.020]	[-0.211 0.310]	[-0.001 0.003]
		-0.004	0.178	0.001
		[-0.019 0.012]	[-0.093 0.450]	[-0.001 0.003]

Parameters	α	s	h	η
	$\hat{\theta}_0$	0.434	5.187	0.099
Estimation MGM model	$\hat{\theta}_{Medium}$	[0.423 0.444]	[4.956 5.418]	[0.098 0.101]
		-0.012	0.241	0.001
		[0.324 0.360]	[-0.107 0.588]	[0.000 0.003]
		0.531	-0.177	0.001
		[-0.043 -0.014]	[-0.526 0.173]	[-0.001 0.003]

Si l'on revient à la configuration d'une loterie binaire avec un seul résultat différent de zéro, nous avons vu à la section 2 que le modèle de décision décrit par l'Equation (3.13) implique un ratio contenant le paramètre α de la fonction d'utilité au dénominateur et le paramètre s de la fonction de pondération des probabilités au numérateur. MGM adopte une version réduite de ce modèle dans laquelle ce rapport est incorporé dans un seul paramètre ζ . Selon nos résultats, puisque le paramètre α augmente à mesure que la taille de l'enjeu augmente et que le paramètre

s diminue légèrement à mesure que la taille de la mise augmente (en tout cas « s » diminue moins qu'augmente α), alors lorsque la taille de l'enjeu augmente, le ratio ζ augmente également. Ce résultat montre une fois de plus que la fonction d'utilité ne peut s'accommoder avec le pattern IRRP qu'au détriment d'un paramètre instable. Ceci explique pourquoi dans le modèle MGM à deux paramètres, on observe généralement que l'exposant s augmente avec la taille de l'enjeu.

Nos résultats sont robustes aux changements de valeurs réelles des paramètres. Cependant, nous remarquons qu'en augmentant le nombre d'agents simulés (et donc en rendant les estimations plus précises), nous observons que l'estimation de s et h varie significativement selon la taille de l'enjeu, ce qui signifie que la fonction de pondération de probabilité estimée semble dépendre du montant tandis que le pattern d'estimation du paramètre α demeure inchangé. Plus précisément, s diminue de façon significative à mesure que la taille de la mise augmente, alors que nous n'identifions pas de modèle clair pour le paramètre h . Bien que la véritable fonction de pondération de la probabilité ne soit pas dépendante de l'enjeu, cela soulève un problème d'identifiabilité lors de l'estimation de la fonction d'utilité avec la fonction de pondération de la probabilité. Dans ce cas, le modèle IRRP est pris en compte à la fois par les fonctions de pondération de l'utilité et de la probabilité grâce à des paramètres instables³⁹.

3.3.1. Conclusion

Dans la section III de ce chapitre nous avons expliqué pourquoi le paramètre « s » du modèle MGM à deux paramètres augmente lorsque l'enjeu augmente. En particulier, nous montrons que la fonction d'utilité puissance ne pouvant que prédire un pattern RRP constant, alors la valeur du paramètre alpha devrait augmenter à mesure que l'enjeu augmente si le pattern

³⁹ Notons toutefois que nous avons effectué des analyses supplémentaires avec des loteries à trois issues possibles différentes de 0. Dans ce cas nous n'observons plus de problème d'identifiabilité. Plus spécifiquement, les paramètres de la fonction de pondération de probabilité ne varient pas avec la taille d'enjeu. Seule le paramètre de la fonction d'utilité puissance augmente lorsque l'enjeu augmente.

observé est un pattern IRRP. Dans la configuration d'une loterie binaire avec un seul résultat différent de 0, le paramètre alpha de la fonction d'utilité puissance se trouve au dénominateur : du ratio $\xi = s/1-\alpha$.

Dès lors si α augmente et que s reste stable, la valeur du ratio augmente.

Ainsi, pour capturer le pattern IRRP, un modèle de décision armé d'une fonction d'utilité IRRA semble plus pertinente pour accommoder les résultats. Pourtant, nous allons voir dans la section IV, que les prédictions d'un modèle intégrant une fonction IRRA ne permettent pas non plus d'accommoder le pattern. Nous verrons dans un premier temps que le pattern révélé n'est pas exclusivement IRRP. Il semble que la prime de risque augmente lorsque l'enjeu augmente uniquement pour des valeurs de probabilités de gagner faible. Par contre, lorsque la probabilité de gagner augmente, on observe plutôt un pattern CRRP.

IV. Limites des modèles économiques face au pattern IRRP-CRRP

Dans cette section nous révélons dans un premier temps le pattern IRRP-CRRP et mettons en évidence qu'il est robuste puisque plusieurs études l'ont révélé. Dans un second temps, nous montrons que ce pattern ne peut pas être accommodé par les modèles existants puisqu'il semble que le principe de séparabilité soit ici violé. Nous proposerons pour finir une possible modélisation de ce phénomène où la séparabilité des fonctions n'est pas respectée.

4.1. Le pattern IRRP-CRRP

Nous avons décompté quatre articles scientifiques qui révèlent l'existence du pattern IRRP-CRRP. L'article de Fehr-Duda et al (2010), l'article de Bouchouicha et al (2017), l'article de Weber et Chapman (2005) et l'article de Estle, Green et Myerson (2006).

L'article de Estle et al (2006) ne reporte pas clairement le pattern IRRP-CRRP, mais nous avons pu récupérer les données et l'identifié nous-même en calculant les primes de risque relatives observées. La Figure 3. montre les médianes de prime de risque relative dans l'expérience dans l'expérience de Estle et al (2006) (haut de la figure) et de Fehr-Duda et al (2010) (bas de la figure).

Nous analysons dans un premier temps les résultats de Estle et al (2006). Ces derniers, ont mesuré les équivalents certains pour 3 niveaux d'enjeux différents, petit, moyen et grand (c'est-à-dire trois amplitudes de x_1 différentes) croisés avec sept niveaux de probabilités. Autrement dit 8*3 équivalents certains furent révélés pour chaque participants. Les trois valeurs de x_1 étaient $x_1^{small} = \$100$; $x_1^{medium} = \$20000$; $x_1^{high} = \$60000$. La seconde issue possible était toujours égale à 0. Les huit probabilités de gagner étaient {0.05; 0.1; 0.2; 0.33; 0.5; 0.75; 0.9; 0.95}.

Nous réalisons une série de comparaison statistiques non-paramétrique avec le test des rangs signés de Wilcoxon. Avec ces tests, nous comparons si les primes de risques relatives sont significativement différentes entre les niveaux d'enjeux, pour chaque niveau de probabilité. Etant donné l'usage de comparaison multiple, nous appliquons une correction Benjamini et Hochberg sur les p-values. Comme le montre la Figure 3., on observe que les primes de risque relatives ne sont jamais différentes entre l'enjeu moyen et grand et ce pour n'importe quel niveau de probabilité. Cependant, les primes de risque relatives pour le petit enjeu diffèrent significativement des enjeux moyen et grand pour les probabilités 0.05, 0.1, 0.2 et 0.33. Pour des valeurs plus élevées de probabilités, nous n'observons aucune différence significative.

Ces observations révèlent un effet d'interaction particulier entre valeurs des probabilités et niveau d'enjeu. Plus spécifiquement, la sensibilité au niveau d'enjeu est plus forte lorsque les probabilités de gagner sont faibles que lorsqu'elles sont importantes. Ces résultats corroborent parfaitement avec l'étude Fehr-Duda et al (2010) (voir Figure 3.).

En ce qui nous concerne, nous ne pouvons pas estimer un modèle de décision RDU armé d'une fonction de pondération de probabilité et d'une fonction d'utilité IRRA pour vérifier si les paramètres seraient stables lorsque le niveau d'enjeu augmente⁴⁰. En effet, étant donné que la seconde issue est toujours égale à 0, nous sommes exposés à un problème d'indétermination. Toutefois, Bouchouicha et al (2017) ont réalisé cette analyse sur leurs données expérimentales. Ils révèlent que les paramètres de la fonction de pondération des probabilités varient significativement avec la taille d'enjeu même lorsque le modèle de décision est armé d'une fonction d'utilité IRRA.

L'ensemble de ces résultats semblent donc remettre en question le principe de séparabilité que nous avons évoqué antérieurement. Cependant, nous noterons que ces résultats d'estimation sont obtenus avec des loteries binaires. Or nous avons vu dans la section III un possible problème d'identifiabilité avec ces loteries. Nous recommandons des loteries à trois issues possibles différentes de 0 pour assurer une estimation fiable.

⁴⁰ Dans la partie annexe de cette thèse nous montrons que cela est possible avec le modèle de Baucells et al (2018) mais sous certaines conditions.

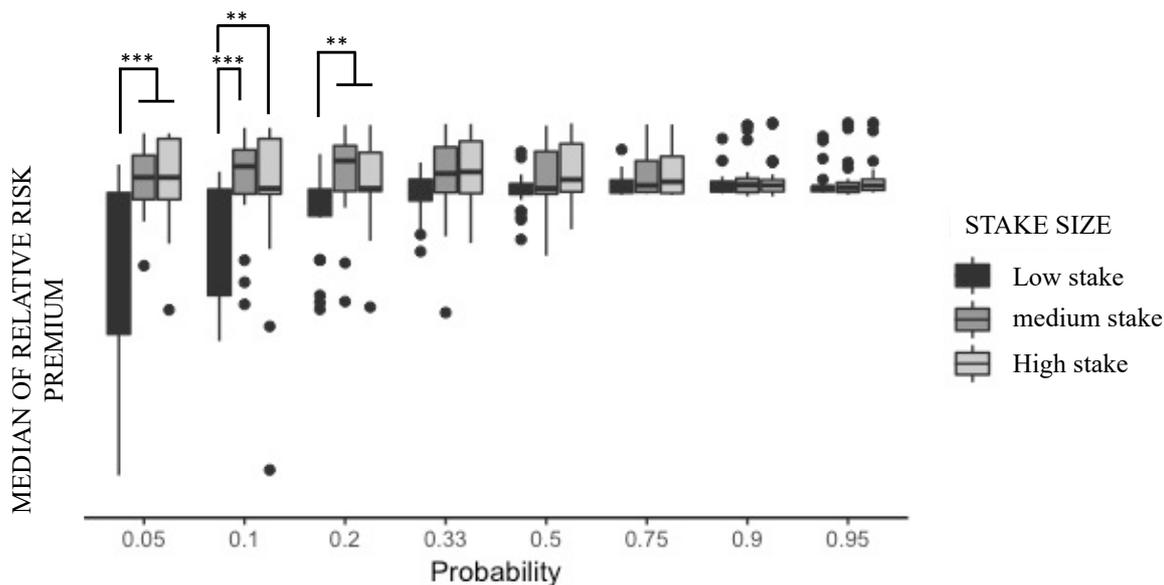


Figure 3.5 Médiane des primes de risques obtenues à partir des données de Estle et al (2006) pour chaque niveau d'enjeu et chaque niveau de probabilités. *** $p < 0.001$, ** $p < 0.01$

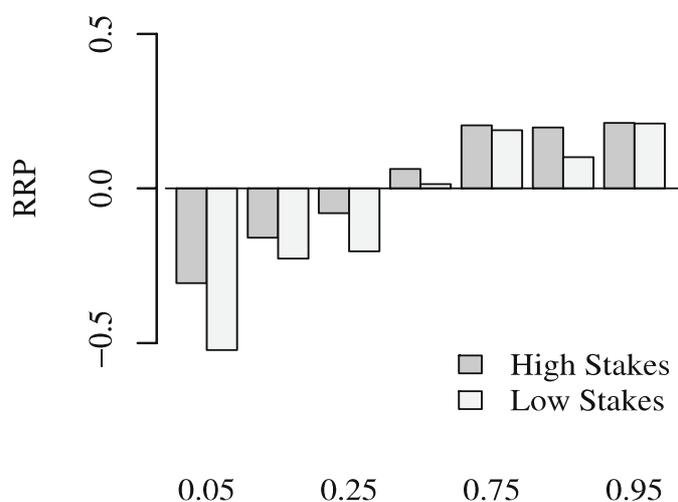


Figure 3.6. Médiane des primes de risque relatives obtenues dans l'étude de Fehr-Duda et al (2010)

4.2. Tentative de capture du pattern IRRP-CRRP

4.2.1. Les modèles économiques

A ce jour aucun modèle n'est en mesure de capturer le pattern particulier IRRP-CRRP. La littérature est donc ouverte à ce sujet et différentes pistes peuvent être explorées. Notons que

les modèles de Cumulative Prospect Theory, de Rank-dependent Utility ou encore de Dual theory formalisent bien une dépendance entre probabilités et issues financières. Toutefois, pour la RDUT et la DT, il s'agit d'une dépendance entre les rangs des issues et leurs probabilités, et non d'une dépendance des probabilités à l'amplitude des issues financières (Wakker, Erev, & Weber, 1994). En effet, le poids subjectif associé à une issue de 10€ dépend de sa place (son rang) par rapport aux autres issues de la loterie : le poids associé à 10% de chance de gagner 10€ n'est pas le même si 10€ est la meilleure issue possible ou une issue intermédiaire. Par contre, les poids associés à une issue de 10€ ou de 1000€ sont les mêmes si ces deux issues ont le même rang dans leur loterie respective et que leurs probabilités ainsi que celles des autres issues sont identiques. La CPT ajoute une dépendance supplémentaire entre les probabilités et les issues financières d'une loterie en prenant en compte leurs signes (pertes ou gains). En effet, le poids associé à une perte de 10€ n'est pas le même que le poids associé à un gain de 10€, et ce même si leurs rangs sont identiques (ce qui a conduit Kahneman et Tversky à supposer l'existence de deux fonctions de pondération des probabilités différentes selon le signe considéré, gain ou perte : φ^+ , φ^-).

Comme indiqué par Sanjit Dahmi (2016) l'effet d'interaction entre le niveau de probabilité et la taille d'enjeu sur l'aversion relative face au risque ne peut pas être capturé par les théories qui séparent les utilités des probabilités. Cela comprend les théories comme l'EUT, la RDUT, la CPT, la DT et même la regret theory ou la disappointment theory (Dahmi, 2016, p. 113). Cela suggère que des théories qui relâchent le principe de séparabilité entre probabilité et issues financières constituent un champ de recherche prometteur pour l'avenir. Dans cette perspective, nous avons relevé un modèle existant mais incomplet, celui de Myerson, Green et Morris (2011).

4.2.2. La tentative de Myerson, Green et Morris (2011)

Le modèle initial MGM retranscrit dans l'Equation (3.5) n'est pas en mesure de capturer le pattern IRRP puisqu'il intègre une fonction d'utilité puissance. Le nouveau modèle proposé par MGM (2011) est une tentative ad-hoc d'accommoder leurs résultats. Précisons que cette tentative est destinée simplement à capturer le pattern IRRP. Ils ne font en effet jamais référence au pattern IRRP-CRRP. Seulement, au lieu de remplacer la fonction puissance (CRRA) par une fonction IRRA (comme celle de Rachlin ou celle de Saha) les auteurs proposent plutôt un modèle de décision qui ne respecte pas le principe de séparabilité. Ils proposent alors d'intégrer dans la fonction de pondération des probabilités un argument de richesse. Dès lors, il s'agit d'insérer x_1 dans la fonction de pondération des probabilités tel que (rappelons que les auteurs s'intéressent à des loteries binaires avec un seul résultat différent de 0. Par conséquent, leur modèle s'en trouve limité à ce cas figure particulier) :

$$\varphi(p, x_1) = \left(\frac{p}{p + h(1 - p)} \right)^{\lambda x_1^s} \quad (3.19)$$

La valeur subjective de L_d est alors :

$$V(L_d) = x_1^\alpha \left(\frac{p}{p + h(1 - p)} \right)^{\lambda x_1^s}$$

Et l'équivalent certain s'écrit donc :

$$Ce(L_d) = u^{-1}[V(L_d)] = x_1 \left(\frac{p}{p + h(1 - p)} \right)^{\frac{\lambda x_1^s}{\alpha}} \quad (3.20)$$

Ce modèle de décision comprend l'incorporation d'un nouveau paramètre par rapport au modèle original, λ . Les auteurs ne donnent aucune signification particulière à ces paramètres. Littéralement, ce modèle est effectivement en mesure de capturer le pattern IRRP. En effet, à mesure que x_1 augmente (autrement dit à mesure que l'enjeu augmente, puisque x_2 reste égal à 0), la fonction $\varphi(p, x_1)$ devient de plus en plus convexe. Autrement dit, l'individu devient de plus en plus averse au risque car il devient de plus en plus pessimiste. Cette solution consistant à intégrer les arguments de richesse au sein de la fonction de pondération des probabilités rejette clairement le principe de séparabilité. Notons que cette solution avait également été envisagée par Scholten et Read (2014). Une autre solution aurait consisté à intégrer l'argument de la probabilité au sein de la fonction d'utilité comme évoqué (mais pas développé) par Weber et Chapman (2005).

Toutefois, un certain nombre de limites peuvent être abordées concernant ce modèle. Premièrement, on peut facilement redouter un problème d'indétermination des paramètres du modèle à cause du ratio $\frac{\lambda x_1^\gamma}{\alpha}$ en exposant de l'Equation (3.20). Deuxièmement, la façon dont le modèle doit être implémenté lorsque le nombre d'issues financières d'une loterie est supérieur à deux n'est pas précisé par les auteurs. Si on applique par exemple la fonctionnelle de la RDUT armée de cette nouvelle fonction de pondération des probabilités, la façon d'intégrer les issues financières n'est pas triviale et mériterait d'être discuté. Plaçons-nous dans la configuration établi dans le cadre de la RDUT : Supposons que les issues financières d'une loterie L soient ordonnées : $x_1 < \dots < x_n$, lesquelles sont attachées aux probabilités p_1, \dots, p_n respectivement. Une façon d'implémenter le modèle proposé par MGM au sein de la RDUT pourrait être :

$$SD_{RDUT}(L) = \varphi(p_n, x_n)u(x_n) + \sum_{k=1}^{n-1} \left[\varphi \left(\sum_{j=0}^k p_{n-j}, x_k \right) - \varphi \left(\sum_{j=0}^{k-1} p_{n-j}, x_{k-1} \right) \right] u(x_{n-k})$$

Si on adapte la formalisation proposée par MGM selon la fonctionnelle RDUT on peut néanmoins considérer qu'il s'agit d'une généralisation de l'EUT si $h = 1, \lambda = 1$ et $\gamma = 0$, et ce peu importe la manière dont on intègre l'argument de richesse dans la fonction de pondération des probabilités. En effet dans ce cas $\varphi(p) = p$ puisque l'exposant de la fonction $\varphi(p)$ sera égal à 1 (et qu'on a préalablement supposé $h = 1$). On peut également établir qu'il s'agit d'une généralisation de la RDUT si $s = 0$, dans ce cas l'exposant de l'Equation (3.19) est égal à λ .

Troisièmement, cette formalisation plus complexe rend le modèle difficilement traçable.

Dans ce cadre, étant donné qu'il s'agit d'une modification ad-hoc qui a pour but de mieux capturer les données, il est difficile de donner une interprétation claire à ces paramètres. De plus, comme toute hypothèse ad-hoc, par nature, il en existe potentiellement une infinité. En effet, ici les auteurs choisissent d'intégrer l'argument x_1 dans la fonction de pondération des probabilités, mais on pourrait tout aussi bien intégrer la variance de la loterie ou encore son écart type comme cela est suggéré par Schultz (2015). Finalement, même si le modèle permet de mieux capturer les données, aucune axiomatique n'a encore été développée qui permettrait pourtant de dégager logiquement une hypothèse plutôt qu'une autre.

4.2.3. Critique du modèle à fonctions non-séparables

Bien que l'idée de rendre non séparable l'évaluation des probabilités de leurs issues financières associées semble une voie prometteuse, nous pouvons nous interroger sur la nature de cette interaction, le sens à lui accorder et à la formalisation proposée par MGM (2011). De façon purement littérale, le modèle de MGM indique que la perception d'une probabilité est fonction de l'amplitude de l'issue financière à laquelle elle est attachée. Concrètement, on pourrait déclarer que plus l'issue est importante plus la probabilité subjective qui lui est attaché sera inférieur à la vraie probabilité. En effet comme nous l'avons indiqué, plus x_1 est grand plus la fonction $\varphi(p)$ est convexe.

Toutefois, une traduction aussi littérale a peu de chance de trouver une réalité empirique. A ce jour aucune étude n'a montré que les probabilités subjectives étaient fonction de l'amplitude des gains associés (Krizan & Windschitl, 2007). D'ailleurs Kahneman et Tversky n'évoquent pas la notion de « probabilités subjectives » en parlant de la fonction de pondération des probabilités. Ils parlent plutôt des « poids décisionnels ». Plus exactement, « les probabilités transformées ne sont pas des probabilités subjectives car elles ne sont pas interprétées comme des mesures de croyance en la réalisation d'un événement. Elles ne mesurent pas la probabilité d'occurrence perçue de l'évènement auquel elles sont associées. Elles mesurent plutôt l'impact qu'a cet évènement sur l'attractivité de l'alternative offerte » (Pffiffelman & Roger, 2013, p.91-92).

Néanmoins, une telle définition n'est pas véritablement rigoureuse non plus puisque nous ne savons pas de quoi dépend cet « impact ». Dans le cadre de la CPT ou de la RDUT, l'impact d'un événement sur l'attractivité d'une alternative dépend explicitement de sa probabilité d'occurrence et du rang de l'évènement lui-même par rapport aux autres événements de la loterie. Il s'agit d'une spécification possible du terme « impact », mais d'autres propriétés de l'environnement peuvent potentiellement être prises en considération. La prise en compte du rang est une manière de maintenir séparée l'évaluation des issues financières de leurs probabilités. Mais on peut supposer également que « l'impact » d'un événement dépend de son amplitude (ou sa désirabilité) absolue, ou bien de son amplitude relative. Par amplitude relative on pourrait entendre que le poids d'une issue financière sur l'attractivité d'une loterie dépend de son amplitude par rapport aux autres amplitudes financières engagées dans la loterie. Selon cette dernière spécification, on introduit à nouveau la notion de point de référence mais cette fois dans la fonction de pondération des probabilités et qui pourrait constituer une piste à poursuivre à l'avenir.

CONCLUSION PARTIE 3

Dans ce chapitre nous avons étudié en profondeur le modèle de Probability Discounting proposé par l'école comportementaliste. Nous avons entre-autres, mis en évidence l'incapacité de leur modèle à accommoder un pattern comportemental IRRP. L'analyse de ce phénomène nous a toutefois amené à montrer que le pattern se révèle finalement plus complexe. En effet, les primes de risque relatives augmentent à mesure que l'enjeu augmente uniquement lorsque la probabilité de gagner est faible. L'impact de l'enjeu sur la prime de risque relative n'a ensuite plus d'influence puisqu'on observe un pattern CRRP pour des niveaux de probabilités modérées à élevées.

Ce pattern particulier constitue à ce jour un véritable challenge pour les économistes qui ne possèdent pas les modèles pouvant le capturer. Ce pattern semble remettre en cause le principe de séparabilité entre fonction d'utilité et fonction de pondération des probabilités décumulées. Pour le moment, seules des modélisations *ad hoc* sont fournis, mais il est clair que cela ne peut constituer une fin en soi.

ANNEXE

A. Analyse complémentaire pour contourner le problème d'indétermination – L'estimation Bayésienne

Nous avons vu que le modèle de Probability Discounting fournit par Myerson, Green et Morris (2011) souffrait d'un problème d'indétermination en particulier lorsque l'on prend en compte des loteries à deux issues possibles dont une seule est différente de 0. Dès lors, l'estimation des paramètres par la technique du maximum de vraisemblance ne peut pas être fiable puisque plusieurs paramétrisations différentes du modèle peuvent fournir les mêmes résultats⁴¹. Une façon de contourner le problème lors de l'estimation consiste à contraindre en amont les paramètres eux-mêmes. C'est-à-dire que l'on définit à l'avance l'intervalle dans lequel nous recherchons les valeurs des paramètres. Cela permet souvent de favoriser la convergence du modèle puisqu'on empêche l'algorithme de chercher des valeurs considérées comme improbables ou impossibles dans la littérature concernée.

Pour cela deux techniques sont possibles : avoir recours à la méthode du maximum de vraisemblance pénalisée ou bien l'inférence Bayésienne. Nous avons choisi d'essayer d'estimer les paramètres libres du modèle MGM à partir des données fournies par Estle et al (2006) par la méthode de l'inférence Bayésienne. Dans un premier temps nous expliquons brièvement en quoi consiste l'estimation Bayésienne. Puis nous expliquons comment nous l'avons appliqué à notre situation. Enfin nous reportons les résultats et concluons.

Le théorème de Bayes

⁴¹ Le problème d'indétermination dans le modèle MGM est particulièrement visible puisqu'on remarque que le calcul de l'équivalent certain à partir de ce modèle (voir Equation (3.13)) implique un ratio entre deux paramètres libres. Dès lors une infinité de valeurs des paramètres du ratio peuvent être trouvées pour une valeur donnée de C_e .

Le théorème de Bayes permet d'utiliser les connaissances ou les croyances que l'on a sur un phénomène ou un évènement puis de calculer la probabilité conditionnelle d'un évènement relié au premier évènement. Intuitivement, la connaissance que l'on peut avoir sur un phénomène sera actualisée par la prise en compte d'informations qui permettent de modifier la connaissance initiale. La connaissance que l'on porte a priori sur un évènement est appelé « prior ». Un prior est en fait une distribution de probabilité portant sur un évènement et la connaissance actualisée est une distribution de probabilités postérieure sur l'évènement suite à la prise en compte d'informations « nouvelles ». Le théorème de Bayes est défini ainsi :

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

La probabilité conditionnelle que l'évènement A (B) survienne sachant que l'évènement B (A) est présent est donné par $p(A|B)$ ($p(B|A)$). $p(A)$ et $p(B)$ sont les probabilités marginales (ou à priori) des évènements A et B.

A partir de ce théorème nous pouvons maintenant l'appliquer dans le domaine de l'inférence statistique. L'inférence consiste à déduire des informations sur une population à partir données. L'estimation par maximum de vraisemblance est bien une méthode d'inférence qui permet d'estimer des paramètres d'un modèle qui maximise la vraisemblance de ces paramètres sachant un ensemble de données. L'inférence Bayésienne réalise la même chose mais cette fois en appliquant la règle de Bayes.

L'inférence Bayésienne

Le théorème de Bayes appliqué à l'inférence Bayésienne s'écrit de la façon suivante :

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{P(Y)}$$

Y est une matrice représentant les données y_i , avec $i = 1, \dots, n$, dont on dispose par exemple à la suite d'une expérience. θ représente un ensemble des paramètres d'un modèle que l'on suppose comme étant potentiellement génératif des données observées. Par exemple si l'on cherche à estimer des paramètres à partir de données que l'on suppose pouvant être générées selon une distribution gaussienne, alors on cherchera à estimer la moyenne (μ) et l'écart type (σ) contenu dans la fonction de densité de probabilité de la loi normale⁴².

Le terme $P(\theta)$ représente la distribution de probabilité à priori portant sur les paramètres d'un modèle. Autrement dit, ce terme représente nos connaissances à priori sur les valeurs des paramètres du modèle. Ainsi, le terme $P(\theta|Y)$ représente la distribution des probabilités postérieure portant sur les paramètres du modèle après avoir pris en compte les données observées. C'est-à-dire après avoir pris en compte la densité de probabilité d'observer les données sachant les paramètres du modèle $P(Y|\theta)$. Les priors sont donc actualisés par la vraisemblance (ou plausibilité) que les paramètres prennent certaines valeurs sachant les données observées⁴³.

Ainsi pour procéder à l'inférence Bayésienne, nous avons besoin de quatre éléments : 1- des données, 2- un modèle supposé génératif des données, 3- de priors sur les paramètres du modèle génératif, 4- un méthode computationnelle pour calculer la distribution de probabilité postérieure portant sur les valeurs des paramètres sachant les données. Cette méthode computationnelle a plus exactement pour objectif de maximiser la probabilité d'observer les paramètres du modèle sachant les données. Ou dit autrement, maximiser la vraisemblance d'observer les données sachant les paramètres. Cette méthode qui vise à résoudre un problème

⁴² $f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

⁴³ On sait que $\mathcal{L}(\theta|Y) = P(Y|\theta)$. C'est-à-dire que la probabilité d'observer Y (les données) sachant les paramètres du modèle est égale à la vraisemblance (ou la plausibilité) d'observer les paramètres sachant les données

d'optimisation peut permettre parfois de trouver des solutions analytiques. Toutefois, très souvent, étant donné la complexité des modèles, les conditions ne sont pas toujours réunies pour réaliser les calculs menant à une solution analytique. Dans ce cas on aura recours des méthodes de descentes de gradients.

Dans le cas de l'estimation Bayésienne, la distribution à posteriori est obtenue via une méthode de Monte Carlo par chaîne de Markov (méthode MCMC). Il s'agit d'un algorithme stochastique (simulation de Monte Carlo) qui permet de simuler une distribution à l'aide d'une chaîne de Markov.

Application au modèle MGM

On comprend qu'avec l'inférence Bayésienne, il est possible de contraindre l'espace des valeurs que peuvent prendre les paramètres du modèle MGM. Pour cela, il suffit d'établir des priors sur les paramètres qui ne peuvent prendre que certaines valeurs selon une fonction de densité que l'on choisit à priori également. Si on suppose une distribution de probabilité uniforme sur les valeurs que peuvent prendre un paramètre dans un intervalle borné, l'estimation revient à réaliser la méthode par maximum de vraisemblance puisque la distribution de probabilité portant sur le paramètre, $P(\theta)$, est une constante dans ce cas. Comme nous l'avions donc évoqué, il est donc tout à fait possible de contraindre les paramètres à ne prendre que certaines valeurs possibles grâce à la méthode du maximum de vraisemblance. Les avantages et inconvénients de cette méthode par rapport à l'inférence Bayésienne est en dehors du champ de notre étude. On notera toutefois que travailler sur des distributions de probabilités postérieures est en soi un avantage par rapport à l'estimation ponctuelle donnée par la méthode du maximum de vraisemblance puisque cela nous permet de quantifier clairement notre degré d'incertitude vis-à-vis de la valeur d'un paramètre dans un espace contraint.

Nous pensons qu'en contraignant suffisamment l'espace des paramètres libre du modèle MGM nous pourrions potentiellement lever le problème d'indétermination. Pour cela, nous

renseignons dans un premier nos priors vis-à-vis des paramètres du modèle MGM, puis nous révélerons leurs distributions de probabilité postérieure que nous obtenons par la méthode MCMC.

LES PRIORS

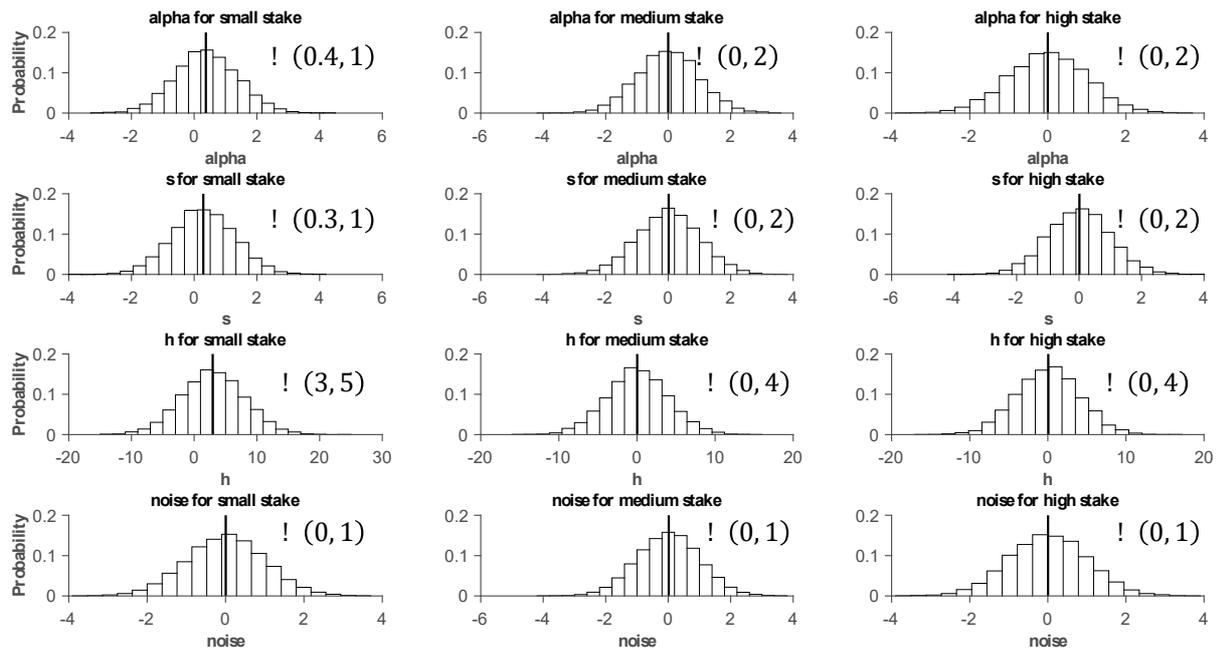


Figure A.1. Priors des paramètres libres du modèle de MGM suivant une loi normale.

Ici nous définissons les priors que nous avons choisi pour chacun des paramètres du modèle avant l'estimation Bayésienne. La figure A.1 montre les distributions pour chaque paramètre ainsi que les paramètres de moyenne et d'écart type de la loi normale qui ont servi à former ces distributions. Comme pour la méthode économétrique que nous avons utilisé dans la partie 3.2 de ce chapitre, le niveau d'enjeu « small » est notre groupe de référence concernant l'estimation. Dès lors, nous déterminons les priors des paramètres du modèle sur 0 selon une loi normale concernant l'enjeu « medium » et l'enjeu « high ».

Les distributions postérieures⁴⁴ ont été approximées par un total de 30 000 échantillons de MCMC obtenus à partir de trois chaînes, après un rodage (« burn-in » ou « warm-up ») de 3000 échantillons (pour chaque chaîne). Les convergences des chaînes de MCMC sont inspectées visuellement

RESULTATS

Les résultats reportés sur la figure A.2 de notre procédure montrent qu'il est possible d'effectuer une estimation des paramètres libre du modèle MGM même avec des loteries binaires avec une seul résultat différent de 0. Toutefois, une analyse fine des résultats jette un doute sur la fiabilité des estimations. On notera dans un premier temps que les intervalles de crédibilités à 90% sont particulièrement élevés pour l'ensemble des paramètres estimés ce qui révèle un grand degré d'incertitude au regard des distribution postérieurs.

La figure A.2 montre les fonctions d'utilités puissances et les fonctions de pondération des probabilités construites à partir des paramètres estimés pour chaque niveau d'enjeu. On constate que les estimations impliquent des fonctions d'utilité particulièrement convexes tout comme les fonctions des pondérations des probabilités. Ce résultat peut à nouveau être le signe d'un manque d'identifiabilité des paramètres. En effet, la convexité des fonctions d'utilités qui impliquent la recherche de risque sont compensés par à un haut de degré de convexité des fonctions de pondération des probabilités qui traduit en théorie le pessimisme (et donc l'aversion face au risque).

Evidemment, il est toujours possible de contraindre encore davantage l'espace des paramètres libres au niveau de nos priors. Toutefois, nous pensons qu'une telle manipulation consiste plutôt à faire « coller » à tout prix la réalité à ce à quoi l'expérimentateur souhaite obtenir de son modèle.

⁴⁴ Nous avons réalisé cette estimation grâce au logiciel libre R avec le package « rstan »

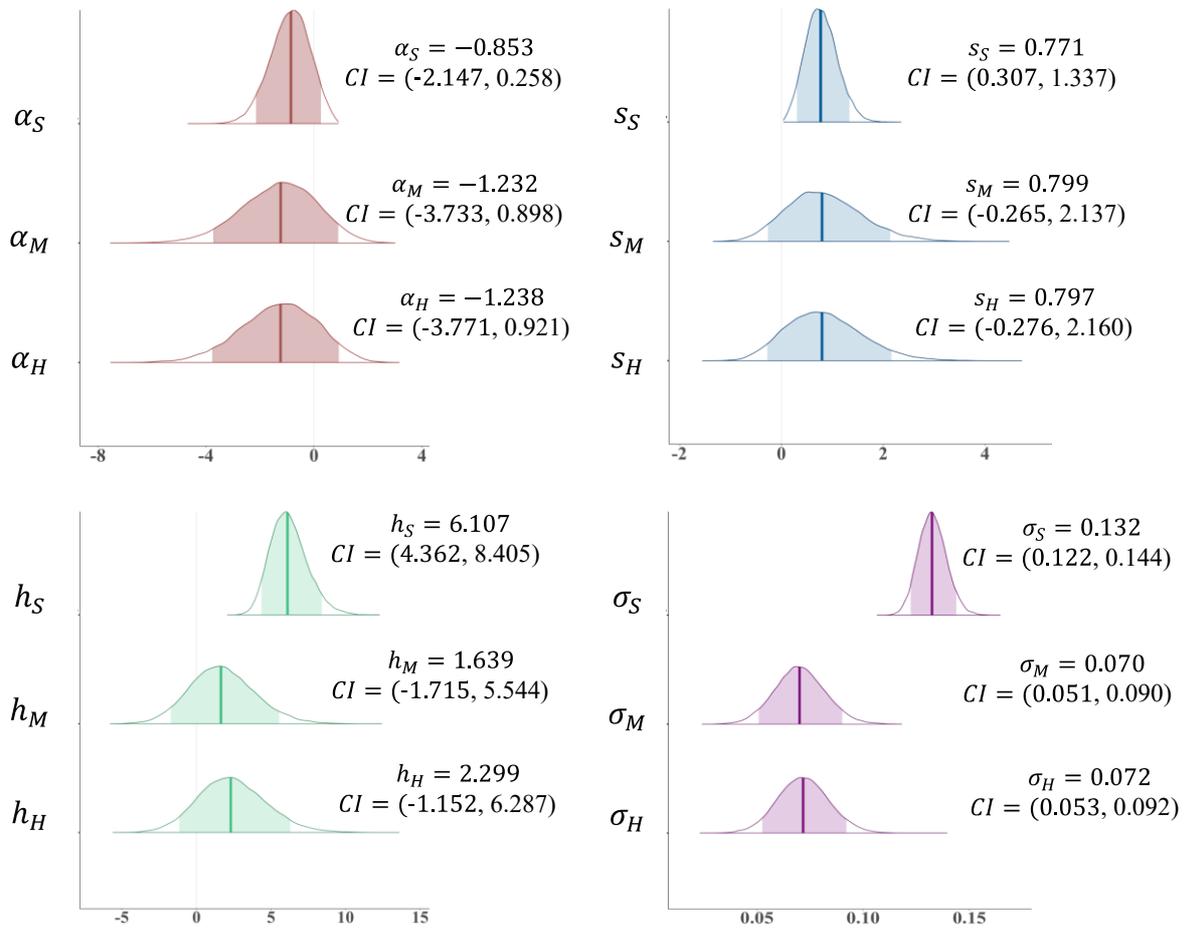


Figure A.2. Résultats des estimations des coefficients du modèle MGM

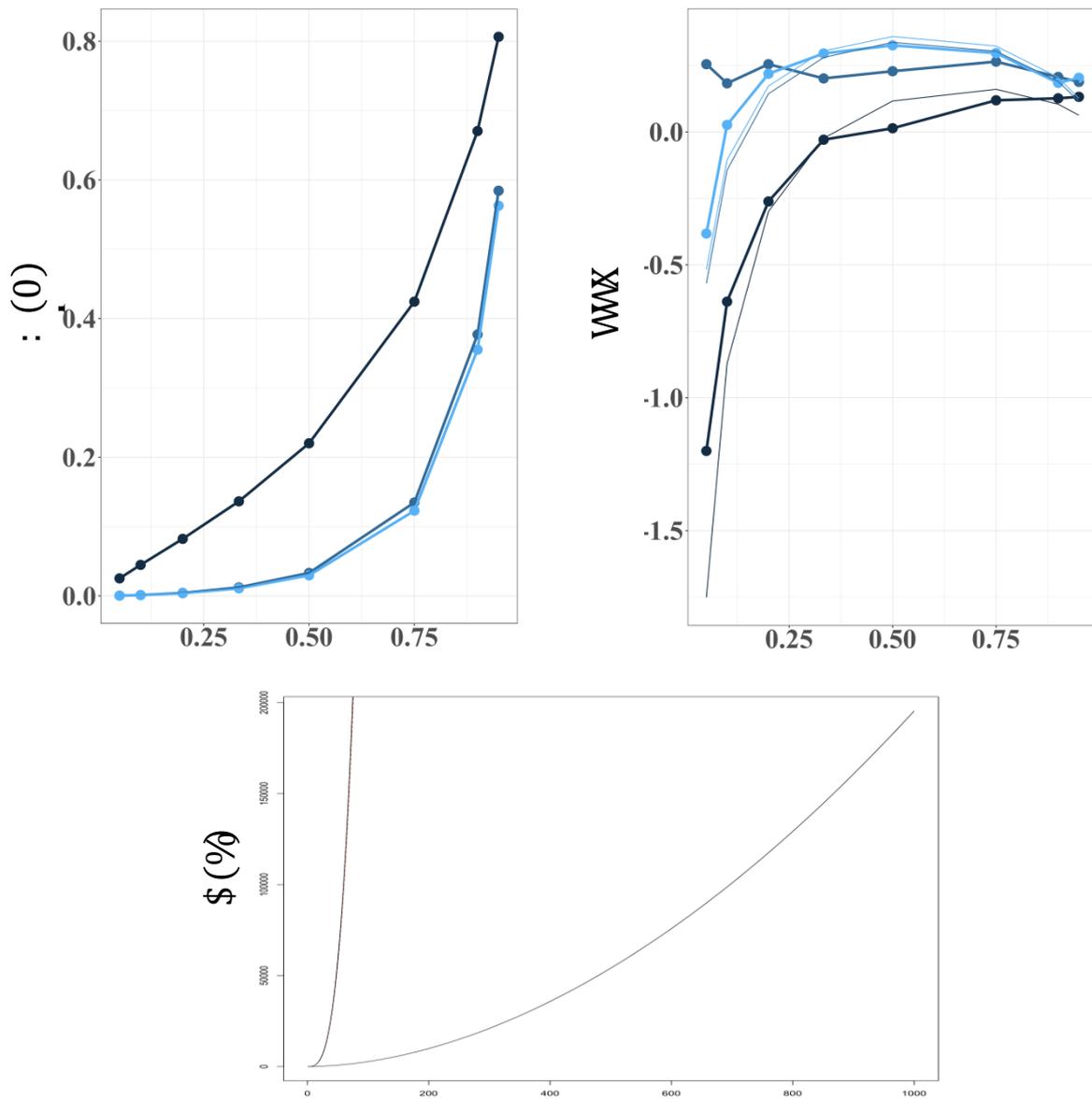


Figure A.3. (Haut gauche) Représentation de la fonction de pondération des probabilités selon le modèle MGM pour les trois niveaux d'enjeux selon les paramètres estimés répertoriés dans la figure A.2. (Haut droite) Fit du modèle MGM aux primes de risques relatives observés pour les niveaux d'enjeux. (Bas) Représentation des fonctions d'utilité pour les trois niveaux d'enjeux, selon les paramètres estimés du modèle MGM répertoriés dans la figure A.2.

4 EFFET DE L'ETENDUE SUR L'ATTITUDE FACE AU RISQUE

INTRODUCTION

Dans ce dernier chapitre, nous recherchons si la prise de risque à un moment t peut être influencée par des éléments de situations antérieures qui sont pourtant déconnectés fonctionnellement de la situation de choix en cours. En outre, nous étudierons l'impact de l'exposition antérieure à différents niveaux d'enjeux de loteries sur la prise de risque subséquente.

Nous avons évoqué lors du chapitre 2, que la théorie de l'espérance d'utilité est en mesure d'expliquer des variations (et non des renversements) d'attitudes face au risque sur plusieurs périodes. Dans ce cas, l'influence du passé sur la prise de risque subséquente passe simplement par le vecteur de la « richesse actuelle » qui se modifie à chaque période sous l'effet des décisions de l'individu et de la réalisation de l'aléa. Par exemple sous hypothèse DARA, plus l'individu accumule de la richesse, plus son aversion absolue au risque devrait diminuer.

De même, concernant la Prospect Theory, nous avons vu que les variations d'attitudes face au risque sur plusieurs périodes pouvaient s'expliquer par l'adaptation partielle et asymétrique du point de référence. Cette adaptation, précisons-le, s'effectue sur des éléments objectifs de la situation risquée, comme par exemple la richesse actuelle, l'espérance, ou autre. Dès lors, comme pour l'EUT, l'influence du passé sur la future décision est envisagée à travers l'intermédiaire d'un vecteur reliant objectivement plusieurs périodes de décisions.

Dans cette étude, nous avons réalisé une expérience impliquant deux phases de décisions. Ces phases n'étaient aucunement reliées entre elles par un vecteur objectif. En effet, l'individu n'accumulait pas de gains ou de pertes et n'avait jamais connaissance des conséquences de ses choix. Dès lors, étant donné la déconnexion entre les deux phases, nous ne devrions pas observer d'effet de la première phase sur la seconde.

Pourtant, de nombreux travaux en psychologie montrent que des éléments de contexte antérieure influencent notre perception des valeurs de stimuli sur une plus ou moins longue période de temps. Un effet bien connu des psychologues est celui de l'étendue : le « range effect » (Parducci, 1965; Pollack, 1952). Comme nous le verrons, l'étendue fait référence à la distance entre la valeur minimum et maximum d'une propriété d'un stimulus (par exemple un son, une température, une couleur, une taille). Or, cette étendue influence directement la perception des valeurs intermédiaires du stimulus, un peu comme un processus de « mise à l'échelle ».

Un modèle appliqué à la prise de décision dans le risque intégrant le « range effect » a été proposé par Baucells, Kontek et Lewandowski (2018, working paper). En particulier, ils définissent une fonction d'utilité normalisée par rapport à l'étendue des issues financières de loteries afin d'opérationnaliser les notions de « petite » et de « grande » issue financière. Selon leur modèle, une issue financière est perçue comme « petite » ou « grande » en fonction de sa localisation par rapport au maximum et minimum des résultats possibles d'une loterie. Autrement dit, le maximum et le minimum constituent l'étendue (ou le « range ») sur lequel se fonde le jugement ou la perception d'un stimulus. Sur la base de cette étendue, les autres issues sont normalisées selon une fonction en forme de S, c'est-à-dire une section convexe pour ce qui est considéré comme de petites issues et une section concave pour ce qui est considéré comme de grandes issues. Il s'agit en définitive d'une version normalisée de la fonction d'utilité de Markowitz (1952).

Sur la base de ce modèle et sur ce que nous savons du « range effect » en psychologie, nous pensons que l'étendue d'une loterie présentée lors d'une première phase influencera la perception des issues d'une autre loterie lors d'une seconde phase. Cette perception altérée des issues sous l'effet du « range » induit lors la première phase, est retranscrite dans la fonction d'utilité formalisée par Baucells et collaborateurs (2018). Par conséquent, étant donné la

relation entre forme de la fonction d'utilité et attitude face au risque, la prise de décision en seconde phase devrait être influencée par l'étendue (ou l'enjeu) de la loterie présentée en première phase, même si ces deux phases ne sont pas connectées par un élément objectif tel que la richesse actuelle ou autre chose.

A travers cette étude, nous montrons que les stimuli sont donc jugés de façon relative et que l'étendue joue en quelque sorte le rôle d'une mise à l'échelle des valeurs d'un stimulus. Cette idée de relativité de la perception, bien que déjà évoquée par la Prospect Theory, a le mérite d'être davantage spécifiée dans le modèle de Baucells et al (2018) et permet donc d'induire des hypothèses plus précises que la Prospect Theory.

Nous verrons toutefois qu'un élément de ce modèle est laissé volontairement flou par les auteurs. Il s'agit de la notion de « contexte décisionnel ». Ce contexte est primordial puisque les valeurs minimum et maximum qui définissent l'étendue des issues financières sont tirées de ce contexte décisionnel. Ainsi, si cette notion n'est pas suffisamment bien spécifiée, on se retrouve dans une situation proche de la Prospect Theory : c'est à dire qu'un degré de liberté trop important est laissé au modèle, le rendant ainsi ad hoc dans ses explications des phénomènes.

En particulier, nous montrerons dans notre situation expérimentale que le contexte décisionnel doit intégrer les loteries présentées antérieurement et non uniquement les loteries en présence au moment du choix pour pouvoir expliquer l'influence du passé sur la prise de risque subséquent

I. LE RANGE EFFECT

Les premières études portant sur le « range effect » remontent aux années 50 avec les recherches de Pollack (1952) ou encore de Parducci (1965, 1968). Ces études s'intègrent dans le champ de la psychophysique et portent sur le jugement d'attributs de stimuli.

Dans les années 2000, ces études ont notamment été reprises par des chercheurs dans le champ du marketing qui y voyaient un intérêt certain pour étudier la mise en valeur de différents produits commerciaux, ou le jugement de prix de biens. Dans cette section nous présenterons les bases théoriques du « range effect », ainsi que quelques études qui mettent en évidence cet effet.

1.1. Le modèle psychophysique du « range effect »

Afin de poser le cadre, nous débutons par une définition du « range effect » données par Yeung et Soman (2005) : « [...] when people evaluate a stimulus, they identify what they believe to be the most and least extreme values and thus determine a range that serves as a context for evaluating the stimulus. *«The particular range used might be drawn from past experience or from the range of stimuli in the current context. Judgment of any stimuli along this range is function of their relative location within the range»* (Ostrom & Upshaw, 1968; Parducci, 1968, 1968). De cette définition, Yeung et Soman (2005) en déduisent ceci: *«Therefore, the perceived difference between two stimulus values is smaller when they are evaluated in the context of a wide than a narrow range»* (Yeung & Soman, 2005, p.363-364).

Illustrons cette définition avec un exemple purement spéculatif. Supposons un agent *A* à qui on présente un stimulus sonore dont on fait varier l'intensité. On lui présente une étendue « large » allant de 0 à 100 dB. Pour chaque intensité, on lui demande de juger l'intensité du son sur une échelle allant de « extrêmement faible » à « extrêmement fort ». Supposons, pour simplifier, une relation linéaire entre le son et la perception du son. Selon la théorie du « range », les

intensités intermédiaires vont être jugées et normalisées selon les extrêmes (ici 0dB est le minimum et 100dB le maximum). Parducci (1965) formalise la perception d'un stimulus en fonction de l'étendue de la façon suivante :

$$\varphi(s_i) = \frac{s_i - s_{min}}{s_{max} - s_{min}} \quad (4.1)$$

Avec s_{max} et s_{min} les valeurs extrêmes que peut prendre l'intensité d'un stimulus et s_i une valeur située dans l'intervalle $[s_{max}, s_{min}]$ dont l'évaluation $\varphi(s_i)$ est donc affectée par les valeurs extrêmes. Dans ce modèle, la distance de chaque valeur du stimulus par rapport au minimum est évaluée ($s_i - s_{min}$), puis normalisée par rapport à la distance (l'étendue) totale entre la valeur minimum et maximum du stimulus, $\left(\frac{1}{s_{max}-s_{min}}\right)$. Nous représentons graphiquement le modèle (4.1) dans la Figure 4.1. La fonction en noir représente la perception « normalisée » de l'agent A pour différentes amplitudes de son et est obtenue pour $s_{max} = 100dB$ et $s_{min} = 0dB$ avec $s_{min} \leq s_i \leq s_{max}$. Comme indiqué sur la Figure 4.1, supposons que l'expérimentateur s'intéresse à la distance de perception entre deux sons $s_1=40$ dB et $s_2=60$ dB. Il observe que cette variation du stimulus produit la variation d'estimation (ou de perception) $\Delta\varphi_{Large_range}$ chez l'agent A.

Prenons maintenant un second agent, l'Agent B, à qui on présente une étendue plus restreinte de volume du son. La fonction en rouge de la Figure 4.1 est calculée pour $s_{max} = 65dB$ et $s_{min} = 35dB$ avec $s_{min} \leq s_i \leq s_{max}$. La théorie du « range » induit que la perception des volumes est maintenant évaluée en fonction d'une étendue plus restreinte que l'Agent A. Notamment, le volume le plus faible (fort) qui était 0 (100) pour l'agent A est maintenant de 35 (60) pour l'agent B.

Dès lors, selon le modèle de Parducci, une intensité objective considérée comme « faible » par l'agent B pourrait être jugée comme plus forte par l'agent A car la localisation de cette intensité

au sein de l'étendue est plus proche du minimum pour le second agent que pour le premier agent. A l'inverse une intensité jugée comme forte par l'agent B (comme pour s_2 par exemple) pourrait être jugée comme plus faible par le premier agent car la localisation au sein de l'étendue de cette intensité de son est plus proche du maximum pour l'agent B que pour l'agent A. Les sons intermédiaires étant donc associés à des normalisations différentes, la différence perçue entre les deux intensités s_1 et s_2 est donc plus grande lorsque l'étendue est faible que lorsqu'elle est plus substantielle, d'où $\Delta\varphi_{Large\ range}(s) < \Delta\varphi_{Small\ range}(s)$.

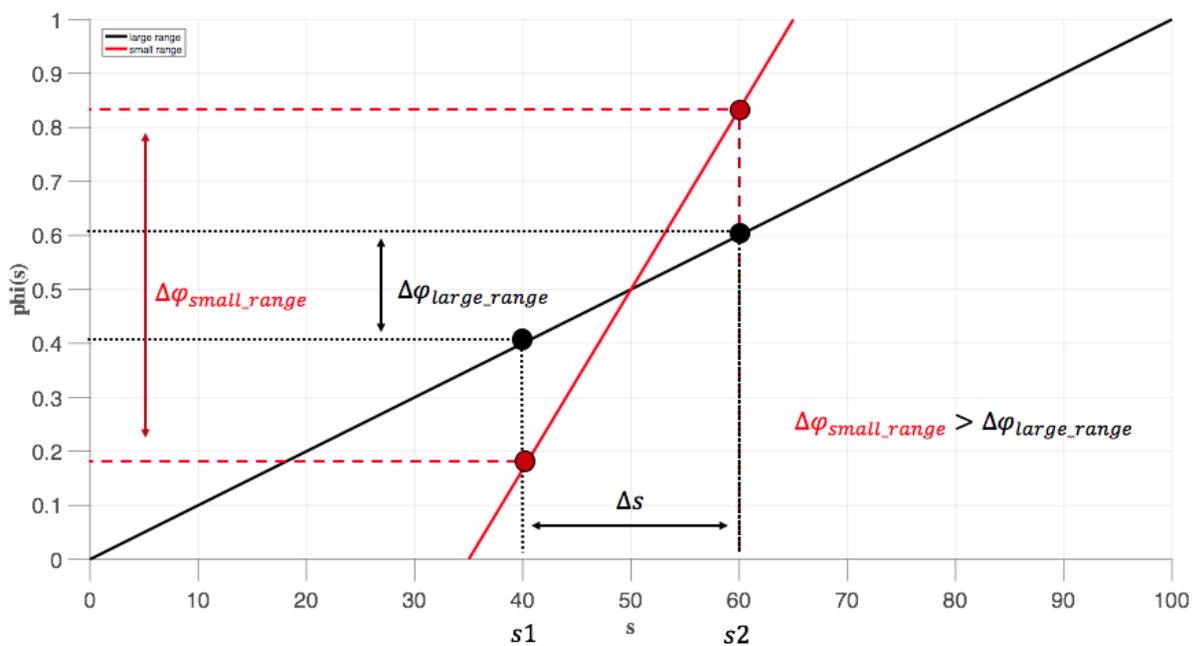


Figure 4.1. Représentation du modèle de Parducci (1965) selon l'Equation (4.1) pour les agents A et B exposés à différentes étendues de volume d'un son.

On comprend avec cet exemple que dans des plans expérimentaux « within-subject », il est très important d'être conscient de l'impact des précédentes modalités expérimentales auxquelles l'individu a été exposé, car les précédentes expositions peuvent influencer, sous l'effet de l'étendue, la perception des stimuli subséquents et donc avoir un impact sur les réponses du sujet dans les modalités expérimentales qui suivent.

Cette dernière considération est précisément ce que nous allons investir dans notre étude expérimentale. En particulier, nous verrons que le modèle de Baucells, Kontek et Lewandowski (2018) ne spécifie pas clairement le contexte dans lequel se forme le minimum et le maximum dans une situation de choix de loterie. Par exemple, le minimum et le maximum proviennent-ils des loteries qui sont présentes à un moment t au moment du choix ? ou bien sont-ils constitués sur la base de toutes les loteries rencontrées au cours d'une session expérimentale ? D'ailleurs, la définition donnée au début de cette partie évoque subtilement cette interrogation : “[...] *The particular range used might be drawn from past experience or from the range of stimuli in the current context*” ?

1.2. Les études en marketing

L'effet du « range » en marketing est bien connu. Il a fait l'objet d'un nombre substantiel d'études. Nous relatons ici rapidement quelques-unes d'entre elles. L'investigation menée par Janiszewski et Lichtenstein (1999) est une excellente illustration du range effect. Les auteurs se demandent si l'attractivité du prix d'un produit dépend de sa comparaison aux prix extrêmes (minimum et maximum) d'autres produits. Les auteurs montrent que le prix d'un produit est vu comme plus (moins) cher quand des produits moins (plus) chers sont ajoutés au contexte. Cet effet est également visible lorsqu'il s'agit de produits similaires entre eux comme des billets d'avion. En l'occurrence, présenter des billets plus (moins) chers à un acheteur lui fera apparaître le prix d'un certain billet en particulier comme moins (plus) cher (Niedrich, Sharma, & Wedell, 2001).

Bennet, Brennan et Kearns (2003) établissent une situation expérimentale où les individus doivent indiquer leur propension à payer (« willingness to pay ») pour un certain produit. Pour cela, les auteurs présentent une liste de prix pour le produit et l'individu doit indiquer le prix qu'il serait prêt à payer. Différents facteurs sont manipulés concernant la présentation de la liste. Les listes présentent les prix dans un ordre soit croissant, soit décroissant soit aléatoire.

L'étendue des prix est également variée. Certaines listes présentent une gamme de prix très étendue d'autres peu étendue. Les résultats montrent que les individus reportent un prix plus élevé lorsque les prix sont présentés dans un ordre décroissant et lorsque l'étendue des prix est élevée.

Yeung et Soman (2005) examinent une situation où deux attributs d'un produit varient. Les produits sont des ordinateurs et les auteurs font varier, entre autres, l'étendue des prix des ordinateurs ainsi que l'étendue des notes de qualité des ordinateurs. L'une des variables dépendantes consiste pour les individus à indiquer leur perception vis-à-vis de différences de prix entre deux options en notant, sur une échelle de 1 à 9, dans quelle mesure l'option la plus chère est plus chère que l'option la moins chère, avec 1 : légèrement plus cher et 9 : beaucoup plus cher. La même question est posée mais cette fois vis-à-vis de la différence de qualité entre le produit ayant la meilleure note de qualité et le produit ayant la moins bonne note. Il est important de préciser que les options à comparer concernent les ordinateurs dont les prix ou les qualités se situent à un niveau intermédiaire par rapport à l'étendue globale de prix/qualité. Les résultats montrent qu'à mesure que l'étendue pour un attribut est élevée (prix ou qualité) la différence de prix (ou qualité) perçue entre deux produits diminue, même si les prix (ou qualités) comparés sont objectivement strictement identiques (Yeung & Soman, 2005).

Cette introduction nous a permis de présenter et définir l'effet de l'étendue sur le jugement et la perception des stimuli. Nous allons maintenant aborder le modèle de Baucells et collaborateurs (2018) qui intègre cet effet au sein de la formalisation de la fonction d'utilité.

II. MODELE DE BAUCELLS, KONTEK ET LEWANDOWSKI (2018)

Dans cette section nous présentons formellement le modèle de BKL, qui précisons-le, n'est pas encore publié à ce jour. Toutefois, sa première version fut publiée en 2017 par Kontek et Lewandowski dans la revue *Management Science* et fut affinée ensuite par Baucells, Kontek et Lewandowski (2018). Par la suite, nous dégagerons une prédiction qu'établit ce modèle sur la base d'une spécification précise du « contexte de décision » évoquée en introduction de ce chapitre. Nous testerons ensuite expérimentalement cette prédiction afin de valider ou non notre spécification du « contexte décisionnel ». Concrètement, pour répondre à notre problématique, nous nous demanderons si l'exposition antérieure à une loterie d'un certain niveau d'enjeu influence l'attitude face au risque subséquente pour une loterie à enjeu différent de la première.

Formalisation du modèle BKL

2.1. Avant-propos

Le modèle de BKL peut fondamentalement être vu comme une extension de la Prospect Theory dans sa première version plutôt que de la RDUT ou de la CPT. En effet, les auteurs intègrent la notion de point de référence dans leur modèle, ainsi qu'une fonction d'utilité en deux parties : concave dans les gains, convexe dans les pertes et un « coude » au niveau du point de référence représentant l'aversion aux pertes. La différence fondamentale du modèle BKL avec la Prospect Theory concerne la prise en compte de l'étendue des issues financières pour normaliser la fonction d'utilité. En particulier, alors que la PT suppose que les individus évaluent la différence absolue entre une issue financière et un point de référence, le modèle BKL suppose que les individus normalisent cette différence par rapport à l'étendue des issues possibles contenues dans un contexte de décision donné.

Ainsi, selon les auteurs, dans une situation de choix en environnement risqué, les individus perçoivent les issues financières selon trois points d'ancrage : le point de référence et l'étendue des issues financières déterminée par la plus petite issue et la plus grande issue d'un ensemble de loterie dans un « contexte de décision donné ».

Les auteurs adoptent une certaine spécification du point de référence que nous avons évoqué dans la Partie 2 section III. Il s'agit de la règle MaxMin qui garantit un certain niveau de sécurité aux individus qui choisissent cette stratégie. Les auteurs appuient leurs choix de spécification du référentiel, entre autres, sur l'étude de Baillon et al. (2015) que nous avons également relaté dans le chapitre 2 et qui semble effectivement montrer que ce point de référence est adopté par une majorité des participants.

Descriptivement, le modèle procède en deux étapes. Premièrement l'individu classe les issues des loteries en termes de gains et des pertes (déterminés comme tel par rapport au point de référence) et la partie de la fonction d'utilité de type Prospect Theory correspondant à son domaine de définition est appliquée (fonction concave dans les gains et fonction convexe dans les pertes). Ces valeurs sont ensuite normalisées par rapport à l'étendue des gains si ces valeurs sont perçues comme des gains, ou par rapport à l'étendue des pertes si ces valeurs sont perçues comme des pertes. Ensuite, une distorsion selon une certaine fonction, noté $D(\cdot)$ est appliquée à ces utilités normalisées. Deuxièmement, on calcule l'espérance d'utilité séparément pour les pertes et pour les gains ce qui fournit ce que l'on appelle une probabilité équivalente spécifique à chaque domaine. Pour finir, ces probabilités équivalentes sont remises à l'échelle pour finalement obtenir la valeur subjective de la loterie.

Le point critique ici concerne la distorsion appliquée aux utilités normalisées par la fonction $D(\cdot)$. Les auteurs choisissent d'appliquer l'inverse de la fonction de pondération des probabilités de Prelec (1998) dans sa version à un paramètre libre. En fonction de la valeur du paramètre libre de la fonction de Prelec, la transformation non-linéaire de l'utilité normalisée peut produire différentes distorsions. La distorsion qui semble la plus vraisemblable aux auteurs est celle qui produit une fonction d'utilité normalisée en forme de S de part et d'autre du point de référence. C'est-à-dire une section convexe puis une section concave dans les gains. La même fonction est appliquée dans les pertes (voir Figure 4.2). On retrouve dès lors une fonction avec trois

points d'inflexions similaire à celle proposée par Markowitz (1952a). La différence fondamentale avec la fonction de Markowitz est que cette fois, les localisations des points d'inflexions sont déterminées par trois éléments : 1) l'étendue des gains et des pertes, 2) le point de référence, 3) les valeurs des paramètres libres de la fonction d'utilité et de la fonction de distorsion. Ainsi, contrairement à Markowitz qui ne parvenait pas à trouver une spécification claire des points d'inflexions, BKL en proposent une formalisation qui fait intervenir la notion d'étendue des issues.

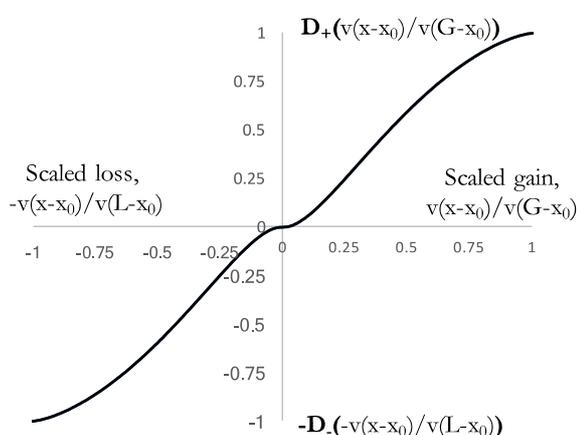


Figure 4.2. Représentation de la fonction d'utilité normalisée-transformée de Baucells et al (2018) dans les pertes et les gains

2.2. Le modèle

Une loterie $L(x_i, p_i)_{i=1, \dots, n}$ avec $x_i \in \mathbb{R}$ dénotant une issue de la loterie, est une distribution de probabilités sur un ensemble fini de taille n . Nous devons maintenant spécifier la notion de « contexte décisionnel » à partir duquel sont déterminés le minimum, le maximum et le point de référence.

Un contexte de décision $\mathbb{C} = \{L^1, L^2, \dots, L^m\}$ dénote simplement un ensemble de loteries (dégénérées ou non) pris en considération par un agent à un moment donné. Ici, Baucells et al (2018) ne définissent pas ce qui est intégré dans ce contexte de décision : « A theory of how \mathbb{C}

comes about is outside the scope of this paper » (Baucells et al, 2018, p5). Selon nous, et en vertu des études menées en psychophysique, un contexte de décision n'intègre pas forcément que les loteries en présence au moment d'un choix. Nous pensons que le contexte de décision est défini sur l'ensemble des loteries qui peuvent être pris en considération au moment d'un choix, ce qui inclut potentiellement des loteries antérieures auxquelles l'individu a pu être confronté dans le passé mais qui ne sont pas présentes au moment d'un choix particulier. Par exemple, l'individu peut avoir le choix entre une loterie L^2 et L^5 mais sa décision peut être influencée par d'autres loteries auxquelles il a été exposé par le passé.

De ce contexte de décision est tiré un « cadre » dénoté \mathcal{U} qui est un triplet composé de $\{T, R, G\}$ dénotant respectivement le minimum, le point de référence et le maximum d'un contexte de décision \mathcal{C} . Supposons $\mathcal{C} = \{L^1, \dots, L^j, \dots, L^m\}$, avec $L^j = (x_1^j, p_1^j; x_2^j, p_2^j; \dots; x_{n(j)}^j, p_{n(j)}^j)$, $j = 1, \dots, m$ et $p_i^j > 0 \forall i, j$. Le triplet est défini de la façon suivante :

$$G = \max_{j=1, \dots, m} \max_{i=1, \dots, n(j)} x_i^j$$

$$T = \min_{j=1, \dots, m} \min_{i=1, \dots, n(j)} x_i^j$$

$$R = \max \left\{ \max_{j=1, \dots, m} \min_{i=1, \dots, n(j)} x_i^j, \min\{0, G\} \right\} \in [L, G]$$

Ainsi, le maximum G de l'étendue est obtenu en prenant l'issue maximum des issues maximums des loteries présentes dans le contexte de décision. Le minimum T de l'étendue est obtenu en prenant le minimum des minimums des loteries présentes dans le contexte de décision. Le point de référence est spécifié comme étant l'issue maximum des issues minimums des loteries du contexte de décision.

Une fois le triplet déterminé, nous pouvons passer à la manière dont sont évaluées les loteries et leurs issues respectives. Dans un premier temps, une fonction d'utilité de type Prospect

Theory (en deux parties) est définie sur les variations de richesses par rapport au point de référence. L'utilité de chaque issue est ensuite normalisée par rapport à l'étendue spécifiquement à la région des pertes ou des gains. Ainsi :

$$u_{norm}(x) = \begin{cases} \frac{u_+(x-R)}{u_+(G-R)} & \forall x > R \\ R & \\ \frac{u_-(x-R)}{u_-(T-R)} & \forall x < R \end{cases} \quad (4.2)$$

La fonction $u(\cdot)$ est une fonction monotone croissante. On applique ensuite une fonction de distorsion $D(\cdot)$ à la fonction d'utilité normalisée afin d'obtenir ce que l'on appelle la fonction d'utilité normalisée-transformée. La fonction $D(\cdot)$ est définie comme l'inverse de la fonction de Prelec (1998), telle que :

$$D(x) = \exp\left(-\left(\log\left(\frac{1}{x}\right)\right)^{1/\gamma}\right) \quad (4.3)$$

Avec γ capturant la courbure de la fonction $D(\cdot)$. La fonction $D(\cdot)[0,1] \rightarrow [0,1]$ est strictement croissante et surjective. On a également $D(0) = 0$ et $D(1) = 1$.

L'évaluation de la loterie se réalise ensuite par domaine, c'est-à-dire pour les gains, dénoté par p^+ , et pour les pertes, dénotée par p^- .

$$\begin{aligned} p^+ &= \sum_{\{i:x_i>R\}} p_i D_+ \left(\frac{u_+(x_i-R)}{u_+(G-R)} \right) \\ p^- &= \sum_{\{i:x_i<R\}} p_i D_- \left(\frac{u_-(x_i-R)}{u_-(T-R)} \right) \end{aligned} \quad (4.4)$$

La spécification D_+ ou D_- signale que le paramètre de la fonction peut différer selon le domaine considéré. Autrement dit, la distorsion opérée dans les pertes peut différer de la distorsion opérée dans les gains. A partir de ces probabilités équivalentes on peut calculer la valeur subjective de la loterie, dénotée $U(L)$, grâce à une remise à l'échelle des probabilités équivalentes p^+ et p^- de la façon suivante :

$$U(L) = u_+(G - R)D_+^{-1}(p^+) + u_-(T - R)D_-^{-1}(p^-) \quad (4.5)$$

On peut ensuite calculer l'équivalent certain de la loterie classiquement en prenant soin toutefois d'appliquer la fonction inverse de la partie de la fonction d'utilité concernée (si la valeur subjective de la loterie est positive, on applique l'inverse de la fonction d'utilité dans les gains à $U(L)$, si elle est négative, on applique l'inverse de la fonction d'utilité dans les pertes à $U(L)$).

$$Ce(L) = \begin{cases} u_+^{-1}(U(L)) & \forall U(L) \geq 0 \\ u_-^{-1}(U(L)) & \forall U(L) < 0 \end{cases} \quad (4.6)$$

Attitude face au risque dans le modèle BKL

2.3. Le fourfold pattern

A première vue, le modèle BKL ne semble pas être en mesure d'accommoder le fourfold pattern évoqué dans le chapitre II. Rappelons que le fourfold pattern établit que les individus sont averses au risque pour les moyennes/grandes probabilités de gagner et chercheurs de risque pour les petites probabilités de gagner. Dans les pertes, les individus sont chercheurs de risque pour les moyennes/grandes probabilités de perdre et averses au risque pour les petites probabilités de perdre. Nous avons vu que la RDUT et la CPT parviennent à capturer ce pattern

en supposant une fonction de pondération des probabilités non-linéaire, en l'occurrence en forme de S inversé et une fonction d'utilité convexe dans les pertes et concave dans les gains. Or, étant donné que le modèle BKL ne fait pas intervenir de distorsion des probabilités à travers une fonction $\varphi(p)$, nous pourrions penser que ce modèle n'est pas en mesure de capturer ce pattern. Pourtant il n'en est rien. Pour le voir, considérons une loterie binaire $L(x_1, p_1; x_2, (1 - p_1))$ où l'on suppose sans perte de généralité que $x_1 > x_2 \geq 0$. Selon le modèle BKL, le minimum T est x_2 , le maximum est x_1 et le point de référence est x_2 . Selon les Equations (4.2) et (4.3) les utilités normalisées et transformées de x_1 et x_2 sont respectivement 1 et 0. Dès lors, selon l'Equation (4.4) qui détermine p^+ , on obtiendra $p^+ = p_1$. En vertu de l'Equation (4.5) on calculera ensuite la valeur subjective de la loterie en appliquant l'inverse de la fonction $D(\cdot)$ à p_1 , qui est normalisée par $u(G - R)$.

On peut alors observer que l'équivalent certain obtenu par l'intermédiaire du modèle BKL est similaire au modèle CPT. Or puisque le modèle CPT est en mesure de capturer le fourfold pattern et que les calculs d'équivalents certains à partir des modèles BKL et CPT sont similaires, alors le modèle BKL peut accommoder les mêmes résultats. Les équivalents certains obtenus avec le modèle BKL et avec le modèle CPT pour une loterie binaire sont respectivement :

$$Ce_{BKL}(L) = x_2 + u_+^{-1}(u_+(x_1 - x_2)D_+^{-1}(p_1))$$

$$Ce_{CPT}(L) = x_2 + u_+^{-1}(u_+(x_1 - x_2)\varphi_+(p_1))$$

Pour le modèle CPT, nous avons supposé $R = x_2$. La différence entre les deux modèles réside dans la fonction pondérant la probabilité p_1 . Or ces fonctions sont similaires puisque $D^{-1}(p_1)$ est en fait la fonction de pondération des probabilités $\varphi_+(p_1)$ donnée par Prelec (1998). Par conséquent, $Ce_{BKL}(L) = Ce_{CPT}(L)$. Dès lors, le modèle CPT et BKL sont équivalents au moins lorsque nous étudions des loteries binaires. Ils sont toutefois distinguables lorsque les loteries

impliquent plus de deux issues. On notera aussi que le modèle BKL est donc capable d'accommoder le fourfold pattern sans supposer l'existence d'une fonction de pondération des probabilités contrairement à la CPT.

2.4. Range-effect et attitude face au risque

Dans le but de tester notre hypothèse vis-à-vis de la spécification du contexte décisionnel du modèle BKL, nous avons d'abord besoin de dériver certaines prédictions du modèle concernant l'évolution de l'attitude face au risque en fonction de l'étendue. Le raisonnement est le suivant : Si l'attitude face au risque dépend de l'étendue, et que nous parvenons justement à manipuler l'étendue lors de phase de préexposition, alors la prise de décision subséquente devrait logiquement être influencée par la phase de préexposition. Si tel est le cas, nous aurons montré que l'étendue est dérivée d'un contexte décisionnel intégrant des loteries antérieures et non uniquement des loteries en présence au moment du choix.

La prédiction intéressante de ce modèle vis-à-vis de notre problématique concerne donc l'impact de l'étendue des issues financières sur l'attitude face au risque en fonction des loteries que l'on considère présentes dans le contexte \mathbb{C} de décision. Pour comprendre la problématique que nous soulevons, représentons tout d'abord graphiquement la fonction d'utilité normalisée-transformée (voir Figure 4.3). Nous restreignons ici notre analyse à la section des gains puisque nous focalisons notre présente étude dans ce domaine. Considérons l'utilité de chaque élément x contenu dans un intervalle tel que $x \in [1, 50]$. Dans ce cadre, le triplet est le suivant : $T = 1, G = 50, R = 1$. Nous supposons une fonction d'utilité normalisée puissance telle que $u(x) = \frac{(x-R)^\alpha}{(G-R)^\alpha} = \left(\frac{x-R}{G-R}\right)^\alpha$, avec $\alpha = 0.5$ et une fonction $D(\cdot)$ telle que spécifiée par l'Equation (4.3) qui transforme la fonction d'utilité normalisée, avec $\gamma = 0.6$ (une valeur empirique standard).

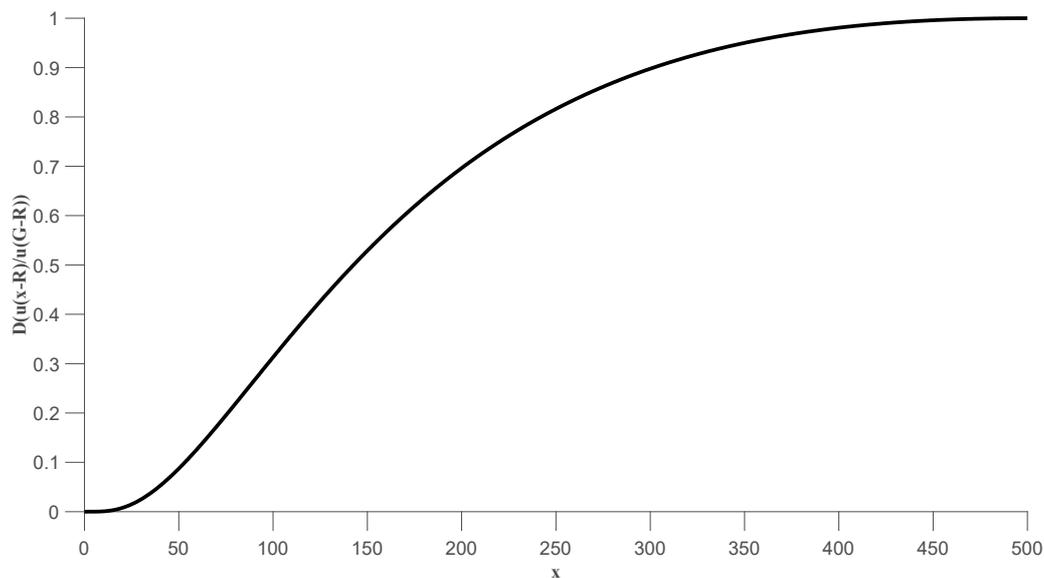


Figure 4.3 Fonction d'utilité normalisée de Baucells et al (2018) dans l'intervalle $x \in [0, 500]$

Comme le montre la Figure 4.3 nous remarquons une section convexe puis une section concave similairement à la fonction de Markowitz. Les issues considérées comme petites relativement à l'étendue, se trouvent dans la section convexe alors que les issues perçues comme plus substantielles sont évaluées dans la section concave de la fonction d'utilité. Par conséquent, on comprend intuitivement que, en fonction de la taille de l'étendue, les mêmes issues d'une loterie peuvent être évaluées sur une section convexe et/ou concave. Or, étant donné que la forme de la fonction d'utilité détermine l'attitude face au risque, on comprend que l'étendue des issues influence directement la prise de décision en environnement risqué.

Bien entendu, d'autres formes peuvent émerger en fonction des paramètres choisis. La fonction peut alors être purement concave, linéaire ou convexe, mais empiriquement, une fonction en forme de S semble plus appropriée selon les auteurs. Entre autres, cela permet de capturer le fameux twofold pattern de Markowitz à savoir : « les agents sont chercheurs de risque pour de petits gains (section convexe) et averses au risque pour des gains plus importants (section

concave) ». Avec la nouvelle formalisation de Baucells et al (2018), la perception de ce qui est ressenti comme « petit » ou « grand » dépend directement de l'étendue.

2.4.1. Analyse graphique de l'impact de G sur l'aversion au risque

Nous allons maintenant chercher à évaluer l'impact d'une augmentation de G sur l'attitude face au risque. Autrement dit, comment sont évaluées les issues financières d'une loterie donnée lorsque le maximum est plus important ? Nous procédons d'abord à une analyse graphique afin de fournir l'intuition de la relation puis nous évaluerons ensuite mathématiquement son impact sur le degré d'aversion au risque.

Supposons une fonction d'utilité u_1 définie dans l'intervalle $x \in [10, 500]$ et une autre, u_2 , définie sur $x \in [10, 100]$. Nous conservons ici les mêmes valeurs de paramètres libres que précédemment. Le triplet de la première fonction d'utilité est $\mathcal{U}_1 = \{10, 10, 500\}$ et celui de la seconde fonction d'utilité est : $\mathcal{U}_2 = \{10, 10, 100\}$. Les fonctions d'utilité sont représentées dans la Figure 4.4 (en bas à droite courbes noires).

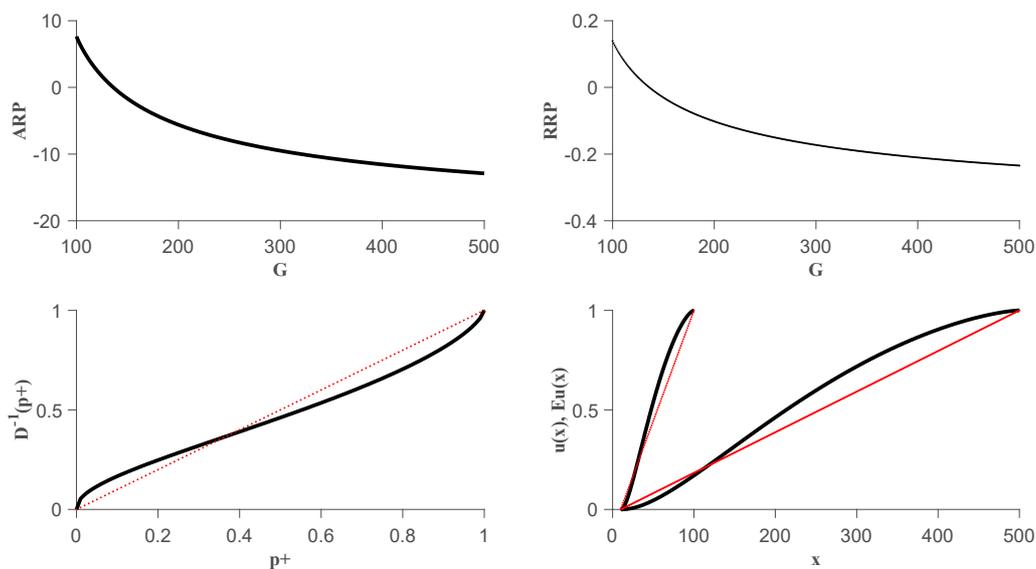


Figure 4.4(En haut) : Représentation de la fonction de la prime de risque absolue (gauche) et relative (droite) en fonction de G pour la loterie $L(x_1 = 100, p_1 = 0.5 ; x_2 = 10, p_2 = 1 - p_1)$. (En bas) Représentation de la fonction de Prelec appliquée aux probabilités équivalentes (gauche). Représentation des fonctions (en noir) u_1 et u_2 définies sur l'intervalle $[10, 100]$ et $[10, 500]$ respectivement et de $p+$ en rouge (droite).

Il est alors intéressant d'observer sur quelle section de u_1 sont évaluées les issues financières du second vecteur. On remarque que les issues contenues dans l'intervalle du second vecteur sont évaluées sur la section convexe de u_1 alors que ces mêmes issues sont évaluées sur une section convexe et concave par u_2 . Cette analyse nous indique donc qu'en fonction de l'étendue, pour un intervalle d'issues données, ces dernières ne sont pas évaluées de la même façon.

Dans la mesure où ce qui est considéré comme le maximum dans un contexte de décision donné modifie drastiquement l'évaluation d'issues financières engagées dans un ensemble de loterie, on peut par corolaire prédire un fort impact du maximum sur l'attitude face au risque⁴⁵. Pour l'illustrer, nous avons simulé les primes de risques absolues (ARP) et relatives (RRP) d'une loterie $L(x_1 = 100, p_1 = 0.5 ; x_2 = 10, p_2 = 1 - p_1)$ pour différentes valeurs de G tout en conservant constants les autres paramètres : $\alpha = 0.5$, $\gamma = 0.6$, $R = 10$ et $T = 10$. Nous spécifions un vecteur $G \in [x_1, 5x_1]$. Pour chaque valeur de G , nous calculons les primes de risques absolues et relatives. Nous reportons les résultats de la simulation dans la Figure 4.4. On observe que plus G augmente, plus les primes de risques absolues et relatives diminuent. Autrement dit, l'agent devient de moins en moins averse au risque, voire change d'attitude face au risque en passant d'averse à chercheur de risque à mesure que G augmente.

2.4.2. Analyse mathématique de l'impact de G sur l'aversion au risque absolue

Afin d'analyser l'impact d'une augmentation de l'étendue uniquement par l'accroissement de G sur l'aversion au risque, nous allons simplement analyser comment évolue le coefficient d'aversion au risque absolue d'Arrow-Pratt, noté Γ_A , lorsque G augmente. Pour conduire notre

⁴⁵ Les relations entre concavité/convexité de la fonction d'utilité et attitude face au risque ont été précisé dans le chapitre 1 : concavité implique l'aversion au risque, la convexité implique la recherche de risque. Le degré de courbure de la fonction détermine quant à lui le degré d'aversion ou de recherche de risque.

analyse, nous supposons une fonction d'utilité puissance dans le domaine des gains. La fonction d'utilité normalisée et transformée est alors :

$$u(x) = \exp \left[- \left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} \right]$$

Le coefficient d'aversion au risque absolue d'Arrow-Pratt, $\Gamma_A = -\frac{u''(x)}{u'(x)}$, nécessite de calculer la dérivée première et seconde de la fonction d'utilité. La dérivée première est donnée par :

$$\begin{aligned} \forall x \in]R, G[, G > R \text{ et } \forall \alpha, \gamma > 0, \frac{\partial u(x)}{\partial x} \\ = - \frac{\exp \left[- \left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} \right] \left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma}}{\gamma (x-R) \log \left(\frac{x-R}{G-R} \right)} > 0 \end{aligned}$$

La dérivée seconde est donnée par :

$$\begin{aligned} \frac{\partial^2 u(x)}{\partial x^2} \\ = \frac{\exp \left[- \left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} \right] \left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} \left(\left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} + \gamma \log \left(\frac{x-R}{G-R} \right) + \gamma - 1 \right)}{\gamma^2 (x-R)^2 \log^2 \left(\frac{x-R}{G-R} \right)} \end{aligned}$$

Le coefficient d'aversion au risque absolue d'Arrow-Pratt est alors :

$$\Gamma_A = -\frac{u''(x)}{u'(x)} = \frac{\left(-\alpha \log \left(\frac{x-R}{G-R} \right) \right)^{1/\gamma} + \gamma \log \left(\frac{x-R}{G-R} \right) + \gamma - 1}{\gamma (x-R) \log \left(\frac{x-R}{G-R} \right)}$$

Nous pouvons maintenant identifier comment évolue le coefficient d'aversion au risque d'Arrow-Pratt lorsque G augmente. Pour cela nous dérivons la fonction Γ_A par rapport à G . On obtient :

$$\forall \gamma \in]0, 1[; \forall \alpha \in]0, +\infty[; \forall x \in]R, G[\text{ avec } G > R, \frac{\partial \Gamma_A}{\partial G} = \frac{(\gamma-1) \left[-\alpha \log\left(\frac{x-R}{G-R}\right) \right]^{1/\gamma} + \gamma}{\gamma^2 (x-R)(G-R) \log^2\left(\frac{x-R}{G-R}\right)} < 0$$

L'analyse de variation du coefficient d'aversion au risque d'Arrow-Pratt nous indique que ce coefficient diminue à mesure que G augmente. Toutefois, cette relation n'est vraie que sous certaines restrictions sur les paramètres de la fonction Γ_A . Ces restrictions ne nous semblent pas dérangement ici puisque les intervalles que nous avons définis pour chaque paramètre correspondent aux valeurs que peuvent prendre empiriquement ces paramètres.

III. Problématique et procédure expérimentale

Nous avons exposé dans la section précédente l'influence de l'étendue sur l'attitude face au risque en manipulant uniquement le paramètre G et en conservant constants les autres éléments du triplet \mathcal{U} . Nous avons vu qu'à mesure que G augmente, l'aversion au risque absolue diminue. Cette prédiction nous permet de tester expérimentalement notre spécification du contexte décisionnel d'où est tiré le triplet. Dans cette section nous présentons comment nous avons procédé expérimentalement.

L'objectif de l'étude consistait à spécifier davantage ce que les auteurs du modèle BKL ont appelé le « contexte de décision » puisqu'ils ont volontairement laissé cette notion floue. En vertu des études menées notamment en psychologie, nous pensons que l'exposition antérieure à certaines valeurs de stimuli influencent la perception d'autres valeurs du stimulus sous l'effet de l'étendue. Appliquée à des situations de choix en univers risquée, nous pensons que

l'exposition antérieure à certaines loteries caractérisées par leurs étendues (ou de façon équivalente par leurs niveaux d'enjeux) influencera la prise de risque pour des loteries subséquentes ayant des étendues différentes. Dès lors, si tel est le cas, nous devons supposer que le contexte de décision évoqué par BKL peut comprendre des loteries incluant dans des situations de choix antérieurs qui n'ont pourtant plus cours dans une situation de choix actuel. Dans la section 2.2 de ce chapitre, nous avons montré que pour un intervalle de richesse donné, la fonction d'utilité liée à cet intervalle devient de plus en plus convexe⁴⁶ (ou de moins en moins concave) à mesure que G augmente. Par conséquent, nous établissons qu'un individu exposé à des loteries à fort enjeu (étendue large) lors d'une phase préliminaire devrait subséquemment se montrer moins averse au risque (ou davantage chercheur de risque) face à une loterie à plus faible enjeu que des individus n'ayant pas été exposés préalablement à de forts enjeux (étendue faible).

De plus, nous montrerons également que le modèle BKL prédit que l'exposition antérieure à des loteries à plus faible enjeu n'a pas d'impact sur la prise de risque subséquente face à des loteries à plus fort enjeu. Autrement dit, un individu qui n'a pas subi de préexposition préalable devrait se comporter similairement à un individu ayant été exposé à une loterie à faible enjeu. Nous testerons également cette prédiction expérimentalement.

3.1. Méthodologie

3.1.1. Groupes expérimentaux

Afin d'évaluer l'impact d'une exposition préalable à des loteries à plus ou moins fort enjeu sur l'attitude subséquente face au risque, nous constituons quatre groupes indépendants. Les conditions expérimentales sont résumées dans le Tableau 4.1. Notre méthode générale pour

⁴⁶ Ceci est vrai uniquement si on suppose une fonction d'utilité en forme de S. À l'inverse, si la fonction d'utilité est en forme de S inversé (concave puis convexe), dans ce cas, pour un intervalle de richesse donné et à mesure que G augmente, alors la section de la fonction d'utilité liée à cette section de richesse devient de plus en plus concave (ou de moins en moins convexe).

mesurer les attitudes des participants face au risque pour différentes loteries consiste en une procédure d'élicitation d'équivalents certains.

L'expérience est alors découpée en deux phases. Durant la phase 1, les quatre groupes sont exposés à une loterie binaire à enjeu plus ou moins élevé, tel que : {Groupe 1 : « petit enjeu » ; Groupe 2 : « grand enjeux » ; Groupe 3 : « moyen enjeux » ; Groupe 4 : « moyen enjeux »}. Durant cette première phase d'exposition, nous mesurons pour chaque individu les équivalents certains de chaque loterie pour différentes valeurs de probabilités de gagner la meilleure issue. Lors d'une seconde phase les participants sont ensuite exposés à une deuxième loterie telle que : {Groupe 1 : « moyen enjeu » ; Groupe 2 : « moyen enjeu » ; Groupe 3 : « grand enjeu » ; Groupe 4 : « petit enjeu »}. Pour simplifier l'écriture nous dénoterons dorénavant la variable « taille d'enjeu » sous la lettre grecque Ψ , tel que $\{S, M, H\} \in \Psi$, représentant respectivement un enjeu faible, moyen et élevé. On notera donc que $S < M < H$. Par exemple, la loterie présente en phase 2 pour les groupes 1 et 2 est caractérisée par l'enjeu ψ_M , c'est-à-dire un enjeu moyen. En phase 2, comme pour la phase 1, nous recherchons les équivalents certains d'une loterie binaire pour différentes probabilités de gagner.

Pour évaluer le degré d'aversion au risque des individus, nous renseignerons la prime de risque absolue pour chaque niveau d'enjeu et chaque probabilité. La prime de risque absolue est notée π .

Pour la suite de cette section nous introduisons le formalisme suivant. A probabilité donnée, la prime de risque π d'un individu du groupe k , avec $k = 1, 2, 3, 4$ dans la phase j , avec $j = 1, 2$, confronté à une loterie L à l'enjeu ψ , est désigné par $\pi_k^j(L^\psi)$. Afin de spécifier si nous parlons d'une loterie en phase 2 ou en phase 1, nous établissons le formalisme ainsi : S'il s'agit de la prime de risque d'une loterie de la phase 2, alors nous précisons quelle loterie fut exposée en phase 1 comme suit : $\pi_k^j(L^\psi | L^{\neq\psi})$. Par exemple, la prime de risque $\pi_k^2(L^M | L^H)$ dénote la prime de risque pour la loterie L^M en phase 2 sachant que l'individu a été exposé à la loterie L^H durant

la phase 1. Une loterie de la phase 1 n'ayant pas de loterie antérieure est simplement dénotée $\pi_k^j(L^\psi | \cdot)$.

Toutes les loteries auxquelles sont exposés les participants sont des loteries binaires procurant le gain x_1 ou x_2 avec $x_1 > x_2$ et où $x_2 = 0$. La probabilité p_1 de gagner x_1 est variée de sorte que $p_1 = \{0.05, 0.1, 0.4, 0.7, 0.9, 0.95\}$, et ce pour chaque niveau d'enjeu.

Nous définissons maintenant nos différents enjeux. « Petit enjeu » : $L^S \equiv (x_1 = 50\text{€}, p_1 ; x_2 = 0\text{€}, 1 - p_1)$. Pour établir les deux niveaux d'enjeux supérieurs nous multiplions par un facteur 10 puis 100 les conséquences monétaires du petit enjeu. Ainsi, « Moyen enjeu » $L^M \equiv (x_1 = 500\text{€}, p_1 ; x_2 = 0\text{€}, 1 - p_1)$ et « Grand enjeu » $L^H \equiv (x_1 = 5000\text{€}, p_1 ; x_2 = 0\text{€}, 1 - p_1)$. Etant donné que chaque participant est exposé à deux niveaux d'enjeux différents et que chaque niveau d'enjeu est croisé avec les six niveaux de probabilités, on obtient donc douze équivalents certains par individu (Six équivalents certains par phase).

Tableau 4.1 Conditions expérimentales

p \ K	Groupe 1	Groupe 2	Groupe 3	Groupe 4
0.05				
0.1				
0.4	Phase 1 → Phase 2			
0.7	(L^S) (L^M)	(L^H) (L^M)	(L^M) (L^H)	(L^M) (L^S)
0.9				
0.95				
$\pi(\text{Phase 1})$	$\pi_1^1(L^S \cdot)$	$\pi_2^1(L^H \cdot)$	$\pi_1^1(L^M \cdot)$	$\pi_2^1(L^M \cdot)$
$\pi(\text{Phase 2})$	$\pi_1^2(L^M L^S)$	$\pi_2^2(L^M L^H)$	$\pi_1^2(L^H L^M)$	$\pi_2^2(L^S L^M)$

p dénote la probabilité de la meilleure issue. Les loteries situées à gauche des flèches sont les loteries qui sont présentées en première phase. Les loteries à droite de la flèche sont les loteries qui sont présentées en seconde phase.

3.1.2. Procédure

L'expérience est effectuée par internet. Le lien internet de l'expérience a été posté sur le site du CRNS : <https://www.risc.cnrs.fr/> qui transmet le lien à des participants volontaires inscrits sur leur base de données. Les participants ouvraient le lien internet et lisaient les consignes (voir appendix B). Les volontaires étaient directement informés que l'expérience impliquait des gains financiers hypothétiques. Il leur était demandé de réaliser leurs choix en faisant comme si ces choix impliquaient de l'argent réel. Durant cette phase d'instruction, nous leur avons expliqué qu'ils seraient exposés à différents choix entre une alternative risquée et une alternative non risquée. Ils étaient également informés qu'ils ne connaîtraient jamais les conséquences de leurs choix puisque nous nous intéressons seulement à leurs préférences entre des alternatives. Une fois les instructions lues par les participants, ces derniers cliquaient sur l'onglet « suivant » pour commencer l'expérience. Celle-ci consistait à révéler les équivalents certains des différentes loteries présentées dans la section 3.1.1 de ce chapitre et qui étaient au nombre de 12 loteries (2 niveaux d'enjeux croisés avec 6 probabilités). Pour chaque loterie, une liste de montants certains espacés selon un intervalle constant était affichée (design 1).

L'étendue des montants certains allaient du paiement minimum de la loterie au paiement maximum de la loterie. Pour chaque montant certain, le sujet devait indiquer s'il préférerait la loterie ou le montant certain. Pour le premier choix (la première ligne du Design 1), l'individu avait le choix entre la loterie et un montant certain égal au minimum de la loterie. Selon le principe de la dominance stochastique de premier ordre, l'agent devrait préférer la loterie.

Lorsque l'agent finissait par préférer le montant certain plutôt que la loterie, une deuxième question était posée afin de préciser pour quel montant certain minimal il préférerait le montant certain plutôt que la loterie. Pour cela, une deuxième liste de montants certains également espacés était affichée (Design 2). L'étendue de ces montants certains de la seconde liste était déterminée de la façon suivante : le plus grand montant certain était égal au montant certain de la première liste pour lequel l'individu avait préféré pour la première fois un montant certain

plutôt que la loterie. Le montant certain le plus petit était égal au montant précédant le changement de préférence. L'équivalent certain était calculé à partir de la deuxième liste en utilisant la moyenne arithmétique du plus petit montant certain préféré à la loterie et du montant certain précédant ce plus petit montant certain. Notons qu'afin d'assurer la cohérence des réponses et d'éviter les changements de préférences multiples de la part des participants, un seul changement de préférence entre la loterie et le montant certain était autorisé.

Design 1: First list of certain outcomes

ROW (n)	Option A (Lottery)	CHOICE		Option B (sure outcome)
1	$L^{small}(50€, 0.4 ; 0€, 0.6)$	A	B	0
2		A	B	5
3		A	B	10
4		A	B	15
5		A	B	20
6		A	B	25
7		A	B	30
8		A	B	35
9		A	B	40
10		A	B	45
11		A	B	50

In this example, the subject switches at the fifth row. His certainty equivalent is between 20€ and 25€

Design 2: Second list of certain outcomes

ROW	Option A (Lottery)	CHOICE		Option B (sure outcome)
1	$L^{small}(50€, 0.4 ; 0€, 0.6)$	A	B	20
2		A	B	20.5
3		A	B	21
4		A	B	21.5
5		A	B	22
6		A	B	22.5
7		A	B	23
8		A	B	23.5
9		A	B	24
10		A	B	24.5
11		A	B	25

In this example, the certainty equivalent is equal to $(21.5+22)/2=21.75$

3.2. Hypothèses selon le modèle BKL

3.2.1. Le Contexte de décision

Tableau 4.2. Conditions expérimentales et détermination des triplets (=cadre : \mathcal{U}_k^j)

Phase (j)	PHASE 1 $\mathcal{U}_k^1 = \{T_k^1, R_k^1, G_k^1\}$		PHASE 2 $\mathcal{U}_k^2 = \{T_k^2, R_k^2, G_k^2\}$	
GROUPE 1 ET 2				
Group (k)	Group 1	Group 2	Group 1	Group 2
Lottery (L^ψ)	$L^S(50\text{€}; 0\text{€})$	$L^H(5000\text{€}; 0\text{€})$	$L^M(500\text{€}; 0\text{€})$	$L^M(500\text{€}; 0\text{€})$
Frame (\mathcal{U}_k^j)	$\mathcal{U}_1^1 = \{0, 0, 50\}$	$\mathcal{U}_2^1 = \{0, 0, 5000\}$	$\mathcal{U}_1^2 = \{0, 0, 500\}$	$\mathcal{U}_2^2 = \{0, 0, 5000\}$
Risk-Premium (π_k^j)	$\pi_1^1(L^S \cdot)$	$\pi_2^1(L^H \cdot)$	$\pi_1^2(L^M L^S)$	$\pi_2^2(L^M L^H)$
GROUPE 3 ET 4				
Group (k)	Group 3	Group 4	Group 3	Group 4
Lottery (L^ψ)	$L^M(500\text{€}; 0\text{€})$	$L^M(500\text{€}; 0\text{€})$	$L^H(5000\text{€}; 0\text{€})$	$L^S(50\text{€}; 0\text{€})$
Frame (\mathcal{U}_k^j)	$\mathcal{U}_3^1 = \{0, 0, 500\}$	$\mathcal{U}_4^1 = \{0, 0, 500\}$	$\mathcal{U}_3^2 = \{0, 0, 5000\}$	$\mathcal{U}_4^2 = \{0, 0, 500\}$
Risk premium (π_k^j)	$\pi_3^1(L^M \cdot)$	$\pi_4^1(L^M \cdot)$	$\pi_3^2(L^H L^M)$	$\pi_4^2(L^S L^M)$

Dans cette section, nous présentons notre spécification du contexte de décision \mathbb{C} duquel nous tirons le cadre $\mathcal{U} = \{T, R, G\}$. De cette spécification nous en déduisons ensuite des hypothèses concernant les attitudes face au risque des différents groupes expérimentaux.

Dans les sections précédentes, nous avons expliqué que l'étendue d'un stimulus était évaluée selon les valeurs considérées comme la plus élevée et la moins élevée par un individu. De plus, dans la définition donnée par Yeung et Soman (2005), il est également précisé que ces valeurs extrêmes peuvent s'établir à partir des expériences passées de l'individu. Par conséquent, nous établissons qu'en phase 2, les extrêmes ainsi que le point de référence (c'est à dire le cadre)

sont déterminés sur l'ensemble des issues financières des deux phases, alors qu'en phase 1 le cadre est seulement déterminé par les issues présentes durant cette phase préliminaire. Par cette spécification, nous établissons que le « contexte décisionnel » intègre donc l'ensemble des loteries auxquelles l'individu a été exposé à un moment donné de l'expérience.

Formellement, nous notons \mathcal{U}_k^j le cadre de la phase j de l'expérience, avec $j = 1, 2$ pour le groupe k , avec $k = 1, 2, 3, 4$. Chaque cadre est un triplet constitué d'un minimum, d'un point de référence et d'un maximum, tel que $\mathcal{U}_k^j = \{T_k^j, R_k^j, G_k^j\}$. Selon ces notations, dans le Tableau 4.2 nous répertorions les triplets pour chaque groupe et chaque phase. Pour faciliter la lecture nous renseignons également les loteries auxquelles sont exposées chaque groupe dans chaque phase ainsi que les notations des primes de risque absolues selon chaque groupe et chaque phase.

3.2.2. Hypothèses

Dans notre analyse évoquée dans la partie 2.2 nous avons établi que, sous certaines restrictions, à mesure que G augmente (tout en conservant T et R constant), la section de la fonction d'utilité sur laquelle un certain intervalle d'issues financière est évalué sera de plus en plus convexe (ou de moins en moins concave). Dès lors, à loterie donnée, à mesure que G augmente, l'aversion au risque devrait diminuer (ou la recherche de risque devrait s'accroître).

En accord avec cette assertion, et selon notre spécification du contexte de décision intégrant les loteries de la phase 1 et 2, nous prédisons que lors de la phase 2, les participants du groupe 2 ($\mathcal{U}_2^2 = \{0, 0, 5000\}$) seront moins averses au risque que les sujets du groupe 1 ($\mathcal{U}_1^2 = \{0, 0, 500\}$) quel que soit le niveau de probabilité car le paramètre G du groupe 2 est plus grand que celui du groupe 1.

Focalisons-nous maintenant sur les groupes 3 et 4. En phase 1, le maximum G est alors de $G=500$ car les participants de ces groupes sont exposés à la loterie L^M . Il est intéressant de noter

que le cadre \mathcal{U} du groupe 1 en phase 2 est identique à celui des groupes 3 et 4. Il est donc implicitement supposé ici que G s'adapte au nouveau maximum entre la phase 1 et la phase 2 lorsque l'enjeu augmente. Dès lors, étant donné que le cadre et la loterie sont similaires entre les groupes 1 phase 2, et 3 4 phase 1, nous prédisons que nous ne devrions observer aucune différence dans les primes de risque de ces groupes. Autrement dit, le fait d'être exposé d'abord à un plus petit enjeu ne devrait pas influencer l'attitude subséquente face au risque de ces agents pour une loterie à enjeu plus important. Par corollaire, la prime de risque du groupe 2 phase 2 devrait être inférieure à celle du groupes 1 phase 2 et des groupes 3 et 4 phase 1.

Proposition 1 : A loterie donnée $L(x_1, p_1; x_2, 1 - p_1)$ et T, R donnés, la prime de risque absolue d'un agent A est plus petite qu'un agent B si le maximum G est plus important pour l'agent A que pour l'agent B . D'où :

$$\text{H1.} \quad \pi_2^2(L^M | L^H) < \pi_1^2(L^M | L^S) = \pi_{\{3,4\}}^2(L^M | L^S) \quad \begin{array}{l} \mathcal{U}_2^2 = \{0, 0, 5000\} \\ \mathcal{U}_1^2 = \{0, 0, 500\} \\ \mathcal{U}_{\{3,4\}}^1 = \{0, 0, 500\} \end{array}$$

En appliquant ces mêmes raisonnements nous dérivons les prédictions supplémentaires suivantes :

Le groupe 2 en phase 1 est exposé à L^H . Le groupe 3 est aussi exposé à L^H mais après avoir été exposé à L^M . Si conformément à notre supposition, G s'adapte parfaitement entre la phase 1 et 2 lorsque l'enjeu augmente, alors le maximum du groupe 3 devrait être égal à 5000 en phase 2. Dès lors, le cadre étant identique pour ces deux groupes, on prédit que leur prime de risque pour L^H ne devrait pas différer, et ce, pour n'importe quel niveau de probabilités.

$$\text{H2.} \quad \pi_2^1(L^H |) = \pi_3^2(L^H | L^M) \quad \begin{array}{l} \mathcal{U}_2^1 = \{0, 0, 5000\} \\ \mathcal{U}_3^2 = \{0, 0, 5000\} \end{array}$$

Centrons-nous maintenant sur la loterie L^S . Le groupe 4 en phase 2 est exposé à la loterie L^S après avoir été exposé à la loterie L^M . Ainsi, lors de la phase 2, selon notre spécification du contexte décisionnel, le maximum G est donc 500. Le groupe 1 en phase 1 est également exposé à la loterie L^S mais sans préexposition ; le maximum G est donc égal à 50. Au regard de notre spécification du contexte décisionnel et de notre analyse concernant l'impact de G sur l'aversion au risque, nous prédisons que les primes de risque du groupe 1 en phase 1 seront plus importantes que celles du groupe 4 en phase 2.

$$\text{H3.} \quad \pi_4^2(L^S | L^M) < \pi_1^1(L^S |) \quad \begin{array}{l} U_3^2 = \{0, 0, 500\} \\ U_1^1 = \{0, 0, 50\} \end{array}$$

IV. Résultats

4.1. Base de données et analyses statistiques

4.1.1. Base de données

Sur un total de 2388 participants ayant répondu à notre questionnaire, seules les données de 215 participants ayant terminé l'expérience ont été conservées. Ces participants étaient répartis entre les quatre groupes de la manière suivante : Groupe 1 (n=54), Groupe 2 (n=46), Groupe 3 (n=59), Groupe 4 (n=56). Ainsi, au total nous avons obtenu 2580 réponses à raison de 12 réponses (équivalents certains) par participant.

Dans un second temps, nous avons également rejeté les participants pour qui les équivalents certains étaient toujours le montant maximum ou le montant minimum des listes. Ces participants n'ont pas d'intérêt pour notre analyse puisque leurs comportements ne sont même pas influencés par la probabilité de gagner. Nous rejetons également les participants qui faisaient preuves d'une inconsistance trop importante (il s'agit d'une procédure classique notamment employée par Holt et Laury, 2002). Par inconsistance, nous entendons les individus qui ne respectaient pas, de façon exagéré, l'axiome de monotonicit . Selon cet axiome,  

mesure que la probabilité de gagner augmente, l'équivalent certain ne diminue pas. Dans notre expérience, étant donné que l'ordre d'occurrence des probabilités était aléatoire, il est clair que nous favorisons des comportements inconsistants. Dès lors, nous rejetons les sujets pour lesquels nous relevions au moins deux fois un problème d'inconsistance au cours d'une des deux phases. Nous obtenions alors 148 participants restant répartis entre les quatre groupes : (groupe 1, $n=39$; groupe 2, $n = 32$; groupe 3, $n=38$; groupe 4, $n=36$).

Par la suite, nous avons réalisé un filtrage des données selon la règle « 1.5*interquartile » en ne conservant que les réponses des participants présentes dans l'intervalle $[Q1-1.5*\text{écart interquartile} ; Q3+ 1.5*\text{écart interquartile}]$.

Nous avons appliqué cette méthode séparément pour chaque niveau de probabilité de chaque phase et de chaque groupe. Notons que les réponses des groupes 3 et 4 en phase 1 étaient réunies en un seul groupe, noté groupe $\{3,4\}$, puisque les conditions expérimentales étaient strictement identiques durant cette phase. En effet, ces deux groupes étaient exposés à la loterie L^M produisant 500 ou 0 en première phase, sans préexposition préalable. A la suite de cette procédure de nettoyage de données (suppression des outliers), nous avons supprimé 2.1% des données.

4.1.2. Analyse des données

Dans cette étude nous nous sommes focalisés sur les primes de risque absolues des individus. A loterie donnée, cette mesure permet de quantifier le degré d'aversion au risque des participants : plus la prime de risque est élevée, plus l'aversion au risque est grande.

Nous avons cherché à manipuler le degré d'aversion au risque des agents par le biais de préexpositions à certaines loteries, comme nous l'avons exposé dans la partie « hypothèses ».

Afin de tester l'impact d'une préexposition sur la loterie L^M , nous avons comparé les primes de risque du groupe 2 phase 2, du groupe 1 phase 2 et du groupe $\{3,4\}$ phase 1. De la même manière, l'impact d'une préexposition sur la loterie L^H a été évalué en comparant les primes de

risque du groupe 2 phase 1 à celles du groupe 3 phase 2. Enfin, pour la loterie L^S , nous avons comparé les primes de risque du groupe 2 phase 1 à celles du groupe 4 phase 2. L'effet de la préexposition a été évalué pour les différents niveaux de probabilités de gagner auxquels les agents ont été confrontés durant l'expérience.

Nous avons réalisé une ANOVA à deux facteurs dans le but de tester les effets principaux des variables GROUPE et PROBA (probabilité de gagner la meilleure issue), ainsi que l'effet d'interaction entre ces deux variables (GROUPE X PROBA) sur les primes de risque absolues. Des analyses post-hoc ont été réalisées afin de comparer les groupes d'intérêts deux à deux en appliquant une correction de Bonferonni sur les probabilités critiques. Les résultats sont reportés dans le Tableau 4.3 (après vérification de l'hypothèse de normalité des résidus)

4.2. Résultats de l'expérience

Tableau 4.3 Prime de risque absolue en fonction des variables GROUPE, PROBA et de leurs interactions

	Hypothèse 1 (L^M)		Hypothèse 2 (L^H)		Hypothèse 3 (L^S)	
	Gr2_Phase2 (n=32) Gr1_Phase 2 (n=39) Gr{3,4}_Phase 1 (n=74)		Gr2_Phase 1 (n=32) Gr3_Phase 2 (n=38)		Gr1_Phase 1 (n=38) Gr4_Phase 2 (n=36)	
	F	p	F	p	F	p
GROUPE	23.221	0.000	10.005	0.002	0.003	0.951
PROBA	73.352	0.000	31.032	0.000	56.812	0.000
GROUPE X PROBA	1.984	0.032	2.178	0.06	2.229	0.05

4.2.1. HYPOTHESE 1

$$H1. \quad \pi_2^2(L^M | L^H) < \pi_1^2(L^M | L^S) = \pi_{\{3,4\}}^1(L^M | L^S)$$

$$U_2^2 = \{0, 0, 5000\}$$

$$U_1^2 = \{0, 0, 500\}$$

$$U_{(3,4)}^1 = \{0, 0, 500\}$$

L'analyse statistique a révélé un effet principal de la variable GROUPE ($F(2) = 23.221$, $p < 0.0001$), de la variable PROBA ($F(5) = 73.352$, $p < 0.0001$) ainsi qu'un effet d'interaction entre ces deux variables ($F(10) = 1.984$, $p < 0.05$). On remarque que les primes de risque absolues des trois groupes sont proches ou inférieures à 0 pour les probabilités 0.05 et 0.1 signalant une attitude de recherche ou de neutralité face au risque. Pour les probabilités supérieures, les primes de risque augmentent et sont positives montrant que les agents étaient en moyenne averses au risque. Nous retrouvons ici un pattern très classique où les individus sont chercheurs de risque pour les petites probabilités de gagner et averses au risque pour les probabilités plus importantes (conformément au paradoxe d'Allais).

L'analyse post hoc⁴⁷ (voir Figure 4.5 représentant les moyennes et erreur standard pour chaque groupe et chaque niveau de probabilité de gagner) avec correction de Bonferonni sur les valeurs p ⁴⁸ montre que la moyenne des primes de risque du groupe 2 phase 2 (moyenne = 19.447, $sd = 77.166$) était significativement inférieure à celle du groupe 1 phase 2 (moyenne = 74.081, $sd = 104.969$) ($t(422) = -5.9806$, $p_{adj} < 0.0001$) et à celle du groupe {3,4} phase 1 (moyenne = 49.686, $sd = 100.945$) ($t(634) = -5.9806$, $p_{adj} < 0.0001$). Ces résultats confirment donc une partie de notre hypothèse 1. Nous avons également émis l'hypothèse que les moyennes des primes de risque des groupe 1 phase 2 et {3,4} phase 1 ne devraient pas différer significativement. Or, nous trouvons que la moyenne des primes de risque du groupe {3,4} phase 1 est inférieure à celle du groupe 1 phase 2 ($t(678) = -2.952$, $p_{adj} = 0.0048$). Ce dernier résultat semble donc montrer que l'exposition à une loterie à enjeux plus faibles influence la prise de risque subséquente face à une loterie à plus forts enjeux.

Enfin, puisque l'interaction entre les variables GROUPE et PROBA est significative nous effectuons pour chaque niveau de probabilité de gagner des comparaisons deux à deux entre les

⁴⁷ L'ensemble des comparaisons deux à deux a été réalisée avec le test t de student. Toutefois, nous avons complété l'analyse à l'aide de tests non-paramétriques en utilisant le test des rangs signés Wilcoxon Mann Whitney. Les résultats sont similaires.

⁴⁸ Nous effectuons ici trois comparaisons, les valeurs p sont alors multipliées par 3.

trois groupes. Autrement dit, pour chaque probabilité nous réalisons trois comparaisons⁴⁹. L'analyse montre que les moyennes de prime de risque du groupe 2 phase 2 sont inférieures à celles du groupe 1 phase 2 pour les probabilités 0.7 ($p_{\text{adj}} < 0.0001$), 0.9 ($p_{\text{adj}} < 0.05$) et 0.95 ($p_{\text{adj}} < 0.0001$) et ne sont jamais différentes de celles du groupe {3,4} phase 1. Enfin les moyennes de primes de risque du groupe {3,4} sont significativement inférieures à celles du groupe 1 phase 2 pour les probabilités 0.7 ($p_{\text{adj}} = 0.004$) et 0.95 ($p_{\text{adj}} = 0.028$)

Dès lors, les résultats semblent indiquer que l'effet de la préexposition n'est effectif que pour les grandes probabilités de gagner.

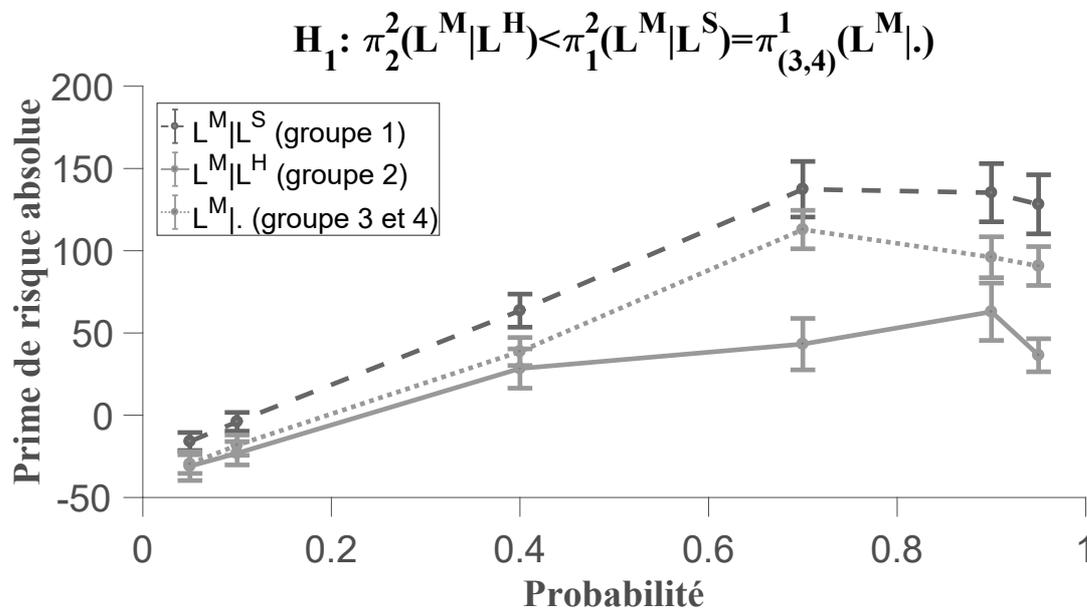


Figure 4.5 Moyenne des primes des risque absolues du groupe 2 phase 2 (traits pleins), groupe 1 phase 2 (traits pointillés) et groupe {3,4} phase 1 (pointillé) pour les différents niveaux de probabilités de gagner la meilleure issue. Les barres représentent les erreurs standards.

4.2.2. HYPOTHESE 2

$$H2. \quad \pi_2^1(L^H|.) = \pi_3^2(L^H|L^M) \quad \begin{aligned} \mathcal{U}_2^1 &= \{0, 0, 5000\} \\ \mathcal{U}_3^2 &= \{0, 0, 5000\} \end{aligned}$$

⁴⁹ Voir dans l'annexe C pour une représentation en boxplot des primes de risques absolues des trois groupes pour chaque niveau de probabilité).

L'analyse statistique a révélé un effet principal de la variable GROUPE ($F(1) = 10.005$, $p=0.002$), de la variable PROBA ($F(5) = 31.032$, $p<0.0001$) mais pas d'effet d'interaction entre ces deux variables ($F(10) = 2.178$, $p=0.06$). On observe à nouveau une attitude face au risque proche de la neutralité lorsque la probabilité de gagner est faible puisque les primes de risque sont proches de 0, et une plus forte aversion au risque lorsque la probabilité augmente puisque les primes de risques deviennent de plus en plus positives.

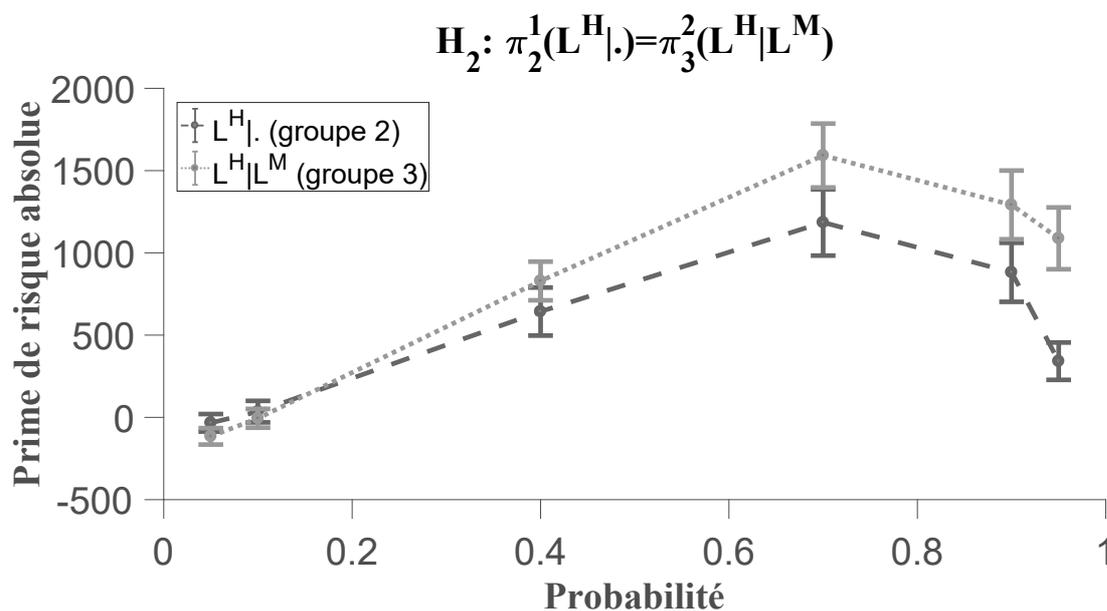


Figure 4.6 Moyenne des primes des risque absolues du groupe 2 phase 1 (traits pointillés) et du groupe 3 phase 2 (pointillés) pour les différents niveaux de probabilités de gagner la meilleure issue. Les barres représentent les erreurs standards.

L'analyse post-hoc montre que la moyenne des primes de risque du groupe 2 phase 1 (moyenne = 521.631, $sd = 891.825$) était significativement inférieure à celle du groupe 3 phase 1 (moyenne = 797.224, $sd = 1126.5$) ($t(415) = -2.7309$, $p = 0.003$) (voir Figure 4.6).

Ce résultat ne confirme donc pas notre hypothèse 2 selon laquelle nous suggérions que ces deux groupes ne devaient pas différer vis-à-vis de leur aversion face au risque. Similairement aux résultats obtenus pour l'hypothèse 1, il semble que l'exposition antérieure à une loterie à faible enjeu augmente l'aversion au risque subséquente face à une loterie à un plus fort enjeu.

Enfin, puisque l'interaction entre les variables GROUPE et PROBA n'est pas significative, nous n'effectuons pas d'analyse supplémentaire pour chaque niveau de probabilité de gagner.

4.2.3. HYPOTHESE 3

$$H_3: \pi_4^2(L^S | L^M) < \pi_1^1(L^S | \cdot)$$

$$U_3^2 = \{0, 0, 500\}$$

$$U_1^1 = \{0, 0, 50\}$$

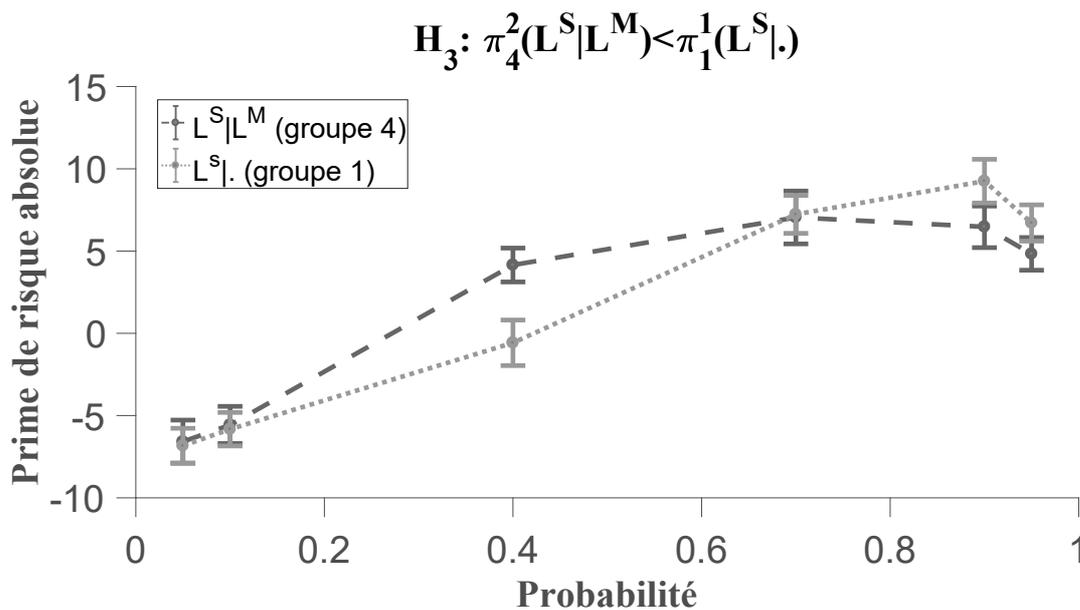


Figure 4.7 Moyenne des primes des risque absolues du groupe 4 phase 2 (traits pointillés) et du groupe 1 phase 1 (pointillés) pour les différents niveaux de probabilités de gagner la meilleure issue. Les barres représentent les erreurs standards.

L'analyse statistique (voir Figure 4.7) n'a pas révélé d'effet principal de la variable GROUPE ($F(1) = 0.003$, $p=0.951$). On observe cependant évidemment un effet principal de la variable PROBA ($F(5) = 56.812$, $p<0.0001$) mais pas d'effet d'interaction entre ces deux variables ($F(10) = 2.229$, $p=0.05$). A nouveau, les individus sont chercheurs ou neutres face au risque pour les petites probabilités de gagner et averses au risque pour les probabilités plus substantielles.

Ce résultat ne confirme donc pas notre hypothèse 3 selon laquelle nous suggérions que ces deux groupes devaient différer vis-à-vis de leur aversion face au risque. En particulier, nous avons

émis l'hypothèse que le groupe 4 en phase 2 devrait se montrer moins averse au risque que le groupe 1 en phase 1 sous l'effet de la supériorité de l'étendue du groupe 4 par rapport au groupe 1.

4.3. Analyse économétrique

4.3.1. Spécification du modèle BKL

Dans cette section nous précisons notre méthode d'estimation des paramètres libres du modèle BKL. Nous focalisons notre analyse sur les groupes 1 et 2 phase 2 et le groupe {3,4} phase 1 puisque ces groupes sont intégré dans notre hypothèse principale. Par cette analyse nous voulons observer si les paramètres libres du modèle BKL-Rachlin et BKL-puissance sont stables entre les groupes. Ce critère nous permet d'analyser si le modèle BKL est bien spécifié. Bien entendu, il est probable que cela ne sera pas le cas dans la mesure où certaines hypothèses découlant du modèle BKL n'ont pas été vérifiées dans l'analyse précédente.

Dans un premier temps, rappelons le modèle de décision de BKL. Selon ce modèle, les individus évaluent une loterie binaire où les issues sont dans le domaine des gains ainsi :

$$U_{BKL}(L) = u_+(G - R)D_+^{-1} \left[p_1 D_+ \left(\frac{u_+(x_1 - R)}{u_+(G - R)} \right) + (1 - p_1) D_+ \left(\frac{u_+(x_2 - R)}{u_+(G - R)} \right) \right]$$

Comme dans notre situation expérimentale $x_2 = 0$ et $R = 0$ le modèle se simplifie comme ceci :

$$U_{BKL}(L) = u_+(G)D_+^{-1} \left[p_1 D_+ \left(\frac{u_+(x_1)}{u_+(G)} \right) \right]$$

La fonction d'équivalent certain s'écrit alors :

$$Ce = u^{-1}[U_{BKL}(L)]$$

Dans le but d'opérationnaliser ce modèle, nous avons supposé une fonction d'utilité puissance $u_+(z) = z^\alpha$. La fonction $D(\cdot)$ est la fonction inverse de Prelec comme spécifié par l'Equation (4.3). Dès lors, le modèle BKL comprend deux paramètres libres $\theta = \{\alpha, \gamma\}$. Nous avons également spécifié une autre fonction d'utilité, celle de Rachlin, présentée dans la partie 3 telle que $u_+(z) = \frac{1}{\beta} \log(1 + \beta z)$. Nous comparerons ces deux modèles vis-à-vis de la stabilité des paramètres mais également vis-à-vis du BIC afin de comparer la capacité des modèles à capturer les résultats.

4.3.2. Dépendance à l'étendue

Dans le but d'étudier l'impact de l'effet de l'étendue selon notre spécification du contexte décisionnel pour les différents groupes, nous introduisons deux variables indicatrices GROUPE_1 et GROUPE_3_4 dans le modèle décisionnel tel que GROUPE_1=1 si le groupe en considération est le groupe 1 en phase 2, et 0 sinon. GROUPE_3_4=1 si les groupes en considération sont les groupes 3 ou 4, et 0 sinon. Le groupe 2 en phase 2 est donc notre groupe de référence. Chacun des paramètres libres du modèle de BKL est présumé dépendre linéairement des variables indicatrices, telle que :

$$\theta = \theta_0 + \theta_{GROUPE_1}GROUPE_1 + \theta_{GROUPE_{3_4}}GROUPE_3_4$$

Avec θ_0 représente les paramètres du modèle BKL du groupe 2 phase 2. Dans ce modèle de décision rappelons que le paramètre G change en fonction du groupe considéré. Pour le groupe 2, G est égal à 5000, pour le groupe 1, 3 et 4, G est égal à 500. Si cette variation de paramètre est le seul facteur influençant les variations d'attitudes face au risque, alors les paramètres libres

devraient rester stables entre les groupes. Autrement dit, on s'attend à ce que les coefficients estimés des variables indicatrices *GROUPE_1* et *GROUPE_3_4* ne soient pas significativement différents de 0.

4.3.3. Spécification de l'erreur

Dans cette expérience, les comportements de prise de risque d'un individu était mesuré par ses équivalents certains pour différentes loteries. Le modèle comportemental pour expliquer les équivalents certains est déterministe. Or nous avons évoqué dans le chapitre 1 qu'il existe de nombreuses sources d'erreurs (inattention, main tremblante...). Dès lors, nous introduisons une structure stochastique à notre modèle d'équivalent certain tel que : $Ce_{i,L} = \widehat{Ce}_L + \varepsilon_{ig}$. Dans la lignée de Fehr-Duda et al (2010) et Bouchouicha et al (2017), nous supposons que ces termes d'erreurs, pour chaque i et L sont indépendamment tirés d'une loi normale de moyenne 0 et d'écart type $\sigma_L = \eta|x_{1L} - x_{2L}|$ lequel est proportionnel à la taille de l'enjeu. Donc, nous supposons que les erreurs sont homoscédastiques sans perte de généralité vis-à-vis de notre propos.

4.3.4. Estimation

Les paramètres du modèle de décision sont estimés par Estimation du Maximum de vraisemblance à l'aide du logiciel R. L'estimation s'effectue par « pooling » et non par individu car peu de données sont disponibles par sujet, ce qui rendrait l'estimation peu fiable. Nous maximisons la log-vraisemblance de nos échantillons, ce qui revient à résoudre le programme suivant

$$\max_{\theta} \sum_i \sum_L \log \left[\frac{1}{\sigma_L} \phi \left(\frac{Ce_{i,L} - \widehat{Ce}_L}{\sigma_L} \right) \right]$$

Avec $\theta = (\alpha, \gamma, \eta)$ pour le modèle comprenant la fonction puissance ou bien $\theta = (\beta, \gamma, \eta)$ pour le modèle intégrant la fonction de Rachlin. ϕ la densité de la distribution normale standard.

4.3.5. Résultats de l'estimation

Tableau 4.4 Estimation des paramètres du modèle BKL armé d'une fonction d'utilité de Rachlin ou d'une fonction d'utilité puissance

ESTIMATION				
Paramètres	β	γ	η	
	$\hat{\theta}_0$	0.022	0.446	0.140
Estimation BKL-Rachlin model	$\hat{\theta}_{GROUPE_1}$	[0.014 0.030]	[0.317 0.574]	[0.126 0.154]
		-0.017	0.088	0.026
		[-0.025 -0.009]	[-0.056 0.232]	[0.005 0.046]
	$\hat{\theta}_{GROUPE_3_4}$	-0.020	0.129	0.028
		[-0.028 -0.012]	[-0.009 0.266]	[0.011 0.046]
Paramètres	α	γ	η	
	$\hat{\theta}_0$	0.310	0.265	0.140
Estimation BKL- puissance model	$\hat{\theta}_{GROUPE_1}$	[0.295 0.324]	[0.201 0.329]	[0.126 0.154]
		0.396	0.216	0.026
		[0.335 0.458]	[0.131 0.301]	[0.005 0.046]
	$\hat{\theta}_{GROUPE_3_4}$	0.500	0.271	0.029
		[0.449 0.551]	[0.194 0.349]	[0.011 0.047]
BIC BKL_Rachlin	17936.99			
Nombre de paramètres	9			
BIC BKL_Puissance	17935.18			
Nombre de paramètres	9			

Les estimations des paramètres sont regroupées dans le Tableau 4.4. Concernant l'estimation des paramètres libres du modèle BKL-Rachlin (haut du Tableau 4.4), on remarque que les coefficients des variables indicatrices associés au paramètre de la fonction d'utilité de Rachlin diminuent significativement puisque 0 n'est pas incluse dans l'intervalle de confiance. On notera

toutefois que ces estimations sont malgré tout très proches de 0 ce qui indique une certaine stabilité du paramètre de la fonction d'utilité Rachlin à travers les groupes. De plus, l'estimation des paramètres liés à la fonction $D(\cdot)$ ne diffère pas significativement de 0, ce qui indique que le paramètre γ est stable entre les groupes.

Passons maintenant à l'analyse des paramètres du modèle BKL-puissance. On constate que les paramètres libres du modèle BKL-puissance ne sont pas stables à travers les groupes expérimentaux. En effet, systématiquement, l'estimation des coefficients des variables indicatrices $\hat{\theta}_{\text{GROUPE}_1}$ et $\hat{\theta}_{\text{GROUPE}_{3,4}}$ n'intègre pas 0 dans l'intervalle de confiance. Autrement dit, les paramètres estimés de la fonction d'utilité puissance et de la fonction $D(\cdot)$ des groupes 1 et {3,4} diffèrent significativement de ceux estimés pour le groupe 2.

La partie gauche de la Figure 4.8 représente les fonctions d'utilités normalisées et transformées pour le groupe 1 phase 2 (courbe rouge), le groupe 2 phase 2 (courbe noire) et le groupe {3,4} phase 1 (courbe verte) selon les paramètres obtenus suite à l'estimation par MLE pour les modèles BKL-puissance (haut gauche de la Figure 4.8) et BKL-Rachlin (bas-gauche de la Figure 4.8). Si les paramètres étaient stables entre les groupes, alors les fonctions d'utilités devraient se confondre puisque l'échelle de l'axe des abscisses est normalisée tel que $u(x - R)/u(G - R)$.

Au regard de cette Figure 4.8, on pourrait conclure alors que le modèle BKL-Rachlin semble mieux adapté que le modèle BKL-puissance pour décrire les données puisque les paramètres estimés des fonctions sont bien plus stables entre les groupes. Toutefois, un regard sur le BIC des deux modèles nous indique que modèle BKL-puissance semble mieux décrire les données que le modèle BKL-Rachlin.

Dès lors, puisque les paramètres du modèle BKL-puissance et BKL-Rachlin ne sont pas stables à travers les groupes, nous pouvons conclure que : soit le modèle n'est pas bien spécifié pour

rendre compte de l'effet de l'étendue sur le comportement décisionnel, soit l'effet de l'étendue n'est pas le seul responsable de l'évolution de l'aversion au risque entre les groupes.

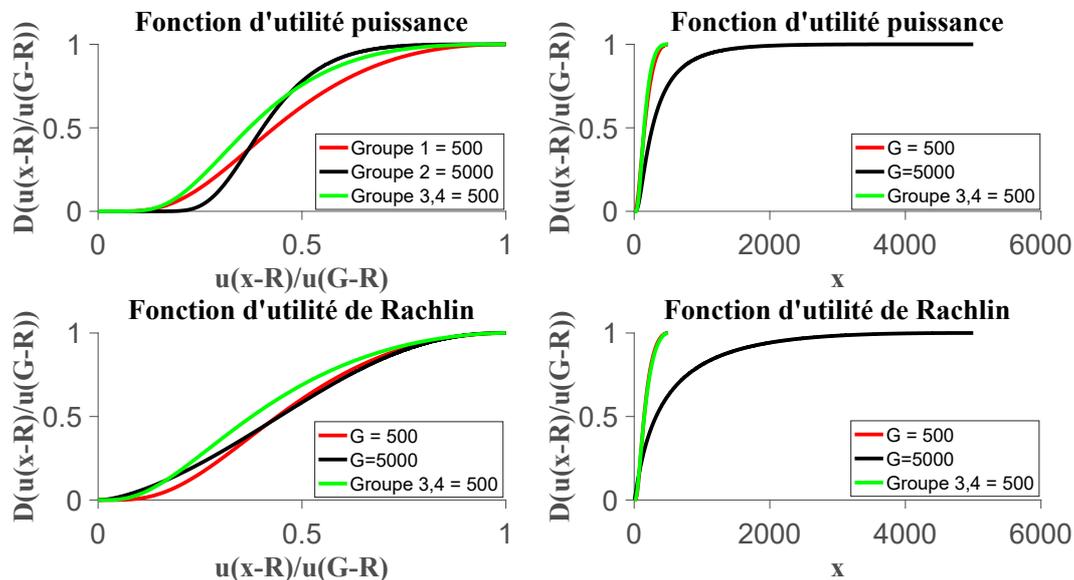


Figure 4.8 Représentation des fonctions d'utilité du groupe 1 phase 2 (courbe rouge), du groupe 2 phase 2 (courbe noir) et du groupe {3,4} phase 1 (courbe verte) selon les paramètres estimés par MLE du modèle BKL-puissance (haut) et du modèle BKL-Rachlin (bas). A gauche, les fonctions d'utilité transformée sont représentées en fonction de l'utilité normalisée par rapport à l'étendue de la richesse. A droite, les fonctions d'utilité normalisées-transformées sont représentées en fonction de la richesse.

Toutefois, on notera que l'on peut douter de la fiabilité de nos estimations. Nous avons procédé à une analyse des corrélations entre les paramètres des modèles. En effet, des paramètres corrélés indique que la solution trouvée par l'algorithme d'optimisation n'est pas unique. Ce problème d'identification des paramètres peut être soit le fruit d'un problème structurel du modèle (comme nous l'avons vu dans le chapitre 3 avec le modèle Myerson Green est Morris (2011) où nous devons estimer un ratio entre deux paramètres) soit le fruit d'un problème pratique dû à un manque suffisamment important de données (« identifiabilité pratique ») (Li et Vu, 2013). L'analyse des corrélations est représentée dans la Figure 4.9. Ces figures représentent les contours de la fonction objectif (ou fonction économique) pour l'estimation de deux paramètres. Les différents cercles indiquent les valeurs les plus probables que peuvent prendre ces paramètres à différents niveaux d'intervalles de confiance. Une forme en ellipse

révèle une corrélation entre les paramètres. La taille des « cercles » par rapport à l'échelle indique le degré d'incertitude quant à la valeur des paramètres estimés.

Pour chaque groupe nous représentons donc les intervalles de confiance en deux dimensions du couple de paramètres $(\hat{\alpha}, \hat{\gamma})$ pour le modèle BKL-puissance et le couple $(\hat{\beta}, \hat{\gamma})$ pour le modèle BKL-Rachlin.

Les résultats présentés dans la Figure 4.9 montrent que les paramètres $(\hat{\alpha}, \hat{\gamma})$ du modèle BKL-puissance pour le groupe 2 sont corrélés à hauteur 0.746, alors que les autres couples de paramètres ne sont pas corrélés. En revanche les couples de paramètres $(\hat{\beta}, \hat{\gamma})$ du modèle BKL-Rachlin sont fortement corrélés indiquant ainsi que la solution de l'algorithme d'optimisation n'est pas unique. Nous pourrions évidemment tester s'il existe des corrélations pour l'ensemble des combinaisons possibles de paramètre, mais cela apparaît inutile puisque nous cherchons simplement à déterminer si les paramètres du modèle sont identifiables. Les analyses conduites ici suffisent à répondre à cette question.

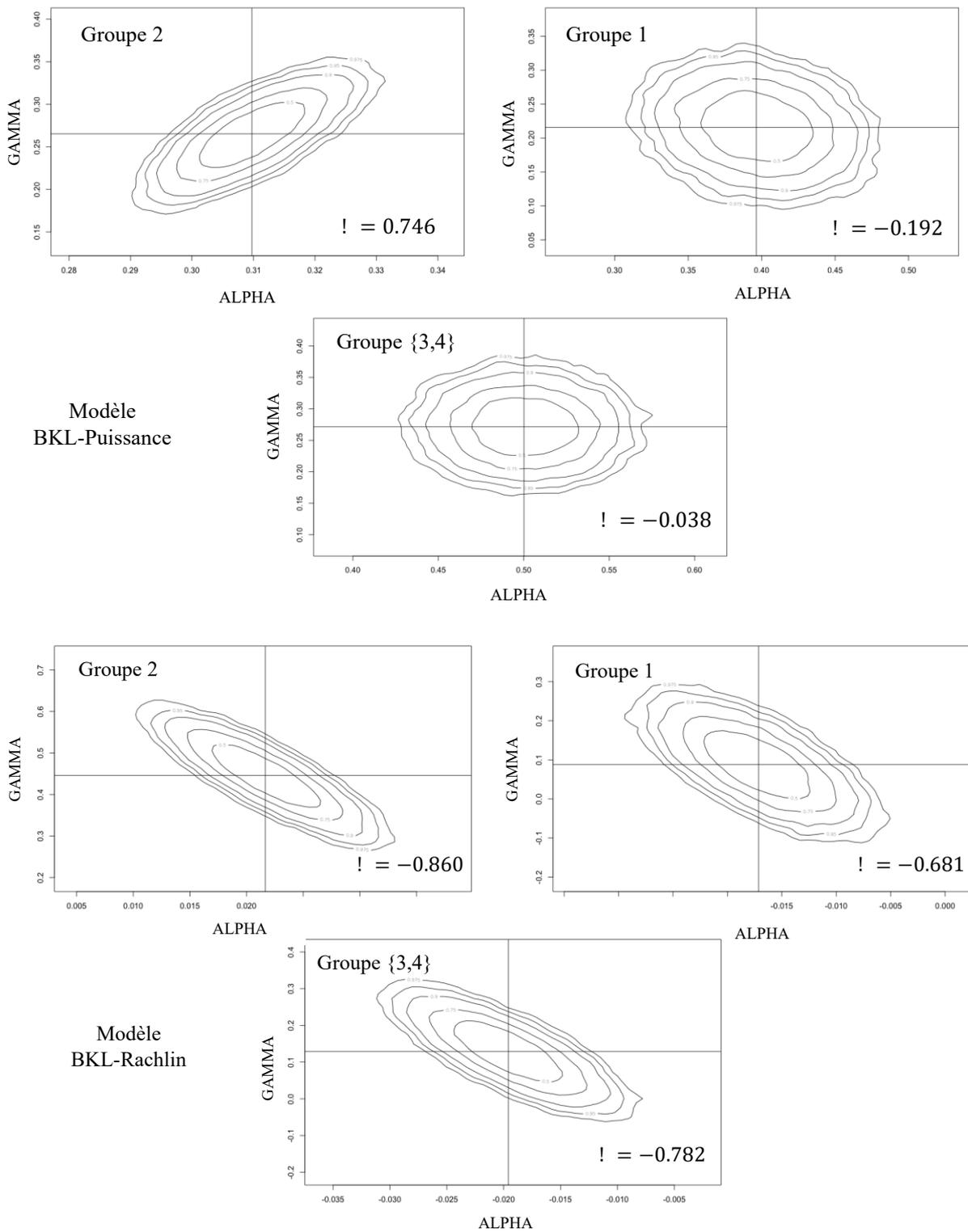


Figure 4.9.(Haut) Contours de la fonction objectif au voisinage de sa valeur optimale tracée pour les paramètres du modèle BKL-puissance: $\hat{\alpha}$ (abscisse) et $\hat{\gamma}$ (ordonnée) (haut gauche), $\hat{\alpha}_{\text{groupe1}}$ (abscisse) et $\hat{\gamma}_{\text{groupe1}}$ (ordonnée) (haut droite), $\hat{\alpha}_{\text{groupe}\{3,4\}}$ (abscisse) et $\hat{\gamma}_{\text{groupe}\{3,4\}}$ (ordonnée) (haut centre).

(Bas) Contours de la fonction objectif au voisinage de sa valeur optimale tracée pour les paramètres du modèle BKL-Rachlin : $\hat{\alpha}$ (abscisse) et $\hat{\gamma}$ (ordonnée) (bas gauche), $\hat{\alpha}_{\text{groupe1}}$ (abscisse) et $\hat{\gamma}_{\text{groupe1}}$ (ordonnée) (bas droite), $\hat{\alpha}_{\text{groupe}\{3,4\}}$ (abscisse) et $\hat{\gamma}_{\text{groupe}\{3,4\}}$ (ordonnée) (bas centre). En partant du centre vers l'extérieur, chaque cercle d'une figure représente un intervalle de confiance à 0.5, 0.75, 0.9, 0.95 et 0.975. Les coefficients de corrélation entre les deux paramètres (alpha, gamma) pour chaque groupe sont donnés par ρ .

4.3.6. Out of sample

Pour montrer explicitement l'inadéquation du modèle BKL avec les données observées, nous établissons une procédure d'out of sample. Cette procédure permet d'établir ce que prédirait le modèle pour une situation donnée, sachant l'estimation des paramètres obtenus avec les données expérimentales. Cette analyse est représentée par la Figure 4.10. Afin de ne pas alourdir l'analyse, nous focalisons notre analyse sur les groupes 1 phase 2 et 2 phase 2. En particulier, à partir de l'estimation des paramètres libres obtenue pour le groupe 1 phase 2, nous rendons compte des prédictions du modèle BKL pour le groupe 2 phase 2.

La Figure 4.10 affiche les moyennes de primes de risque absolues pour chaque niveau de probabilité de gagner la meilleure conséquence de la loterie L^M pour le groupe 2 phase 2 (ligne pointillée bleue avec marqueur rond) et le groupe 1 phase 2 (ligne pointillée rouge avec marqueur triangle). Les traits pleins rouges et bleus représentent les fonctions de primes de risques absolues simulées à partir des paramètres estimés de chaque groupe avec le modèle BKL-Rachlin (haut de la Figure 4.10) et BKL-puissance (bas de la Figure 4.10). Les primes de risques absolues estimées pour le groupe 1 phase 2 correspondent à la fonction en rouge, celles du groupe 2 phase 2 sont en bleue.

La fonction en vert est le résultat de notre procédure « out of sample ». Cette fonction représente la prime de risque absolue pour le groupe 2 phase 2 lorsque les paramètres libres du modèle BKL sont égaux à ceux estimés pour le groupe 1 phase 2. Si le modèle était bien spécifié, la fonction verte se confondrait avec la fonction bleue. Or le modèle prédit des primes de risques absolues bien inférieures à la réalité.

Par ce biais, nous voulons montrer que même si le modèle BKL produit des prédictions qui vont dans le sens de nos hypothèses, il surestime largement la prise de risque des agents. Au vu des résultats exposés dans la Figure 4.10, nous pouvons néanmoins établir que les prédictions fournies par le modèle BKL-Rachlin semblent plus proches des données observées que les

prédictions fournies par le modèle BKL-Puissance. Pour le vérifier, nous comparons l'erreur moyenne absolue (Mean Absolute Error, MAE) du modèle BKL-Rachlin et celle du modèle BKL-puissance. La MAE est une mesure de différence entre deux variables quantitatives qui permet de calculer un écart absolu moyen entre ces deux variables. Ici, nous nous intéressons à l'écart entre les moyennes de primes de risques absolues observé du groupe 2 phase 2 pour chaque niveau de probabilité et les primes de risques prédites par nos deux modèles. Les prédictions des deux modèles sont formées à partir des estimations des paramètres obtenues pour le groupe 1 phase 2. Graphiquement, on calcule la distance absolue moyenne entre la courbe en vert et les marqueurs bleus ronds. Formellement, on procède ainsi :

$$MAE_{model} = \frac{\sum_{i=1}^n |\hat{\pi}_i(L^M | L^H) - \bar{\pi}_i(L^M | L^H)|}{n}$$

Avec $i = 1, \dots, 6$ l'indexation sur les probabilités auxquelles les individus ont été exposés durant l'expérience. $\hat{\pi}_i$ représente la prime de risque prédite par le modèle BKL pour la probabilité p_i et $\bar{\pi}_i$ représente la prime de risque moyenne observé du groupe 2 phase 2 pour la probabilité p_i . On obtient un MAE de 127.29 pour le modèle BKL-puissance et un MAE de 61.74 pour le modèle BKL-Rachlin^{50 51}. Un test des rangs signés de Wilcoxon Mann-Whitney nous indique que cette différence est significative ($p < 0.05$) spécifiant bien que les prédictions du modèle BKL-Rachlin sont plus proches des moyennes de primes de risques observés.

⁵⁰ Notons que le calcul de l'erreur quadratique moyenne nous fournit la même relation, à savoir que, les prédictions du modèle BKL-Rachlin (paramétré avec les valeurs de paramètres obtenues avec les données du groupe 1 phase 2) sont plus proches des primes de risques moyennes observées avec le groupe 2 phase 2 que le modèle BKL-puissance.

⁵¹ Nous avons également calculé la différence absolue moyenne entre les prédictions des modèles et les données individuelles. Dans ce cas $MAE_{model} = \frac{\sum_{i=1}^n \sum_{j=1}^m |\hat{\pi}_i(L^M | L^H) - \pi_{i,j}(L^M | L^H)|}{m}$ avec $j = 1 \dots, m$ l'indexation sur les individus. On obtient un MAE de 131.87 pour le modèle BKL-puissance et un MAE de 78.56 pour le modèle BKL-Rachlin.

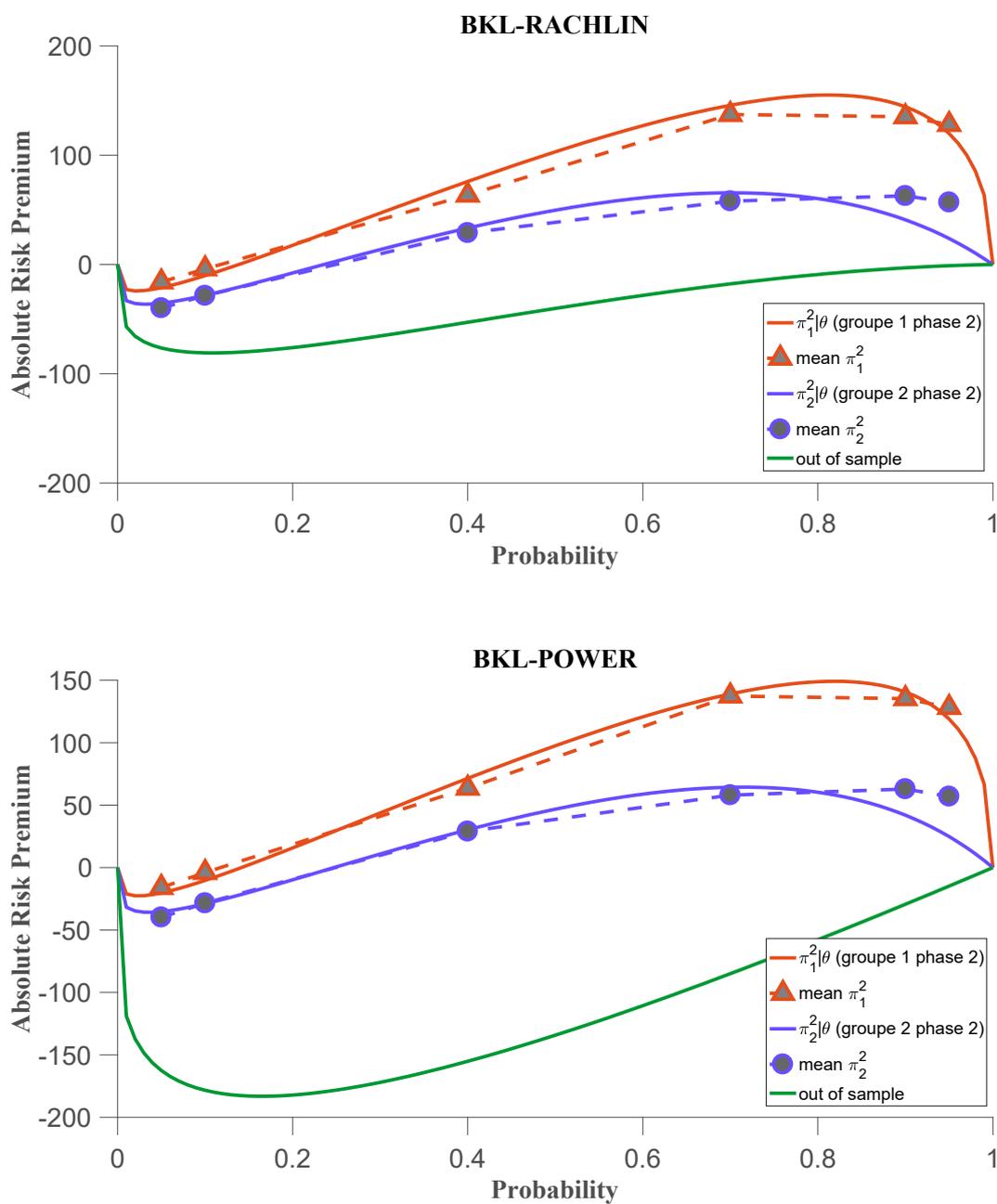


Figure 4.10 Les traits pleins représentent le fit du modèle BKL-puissance (haut) et du modèle BKL-Rachlin (bas) aux primes de risques moyennes pour la loterie L^M des groupes 1 (traits pointillés rouges) et 2 (traits pointillés bleu) en phase 2. Les courbes vertes représentent les primes de risques simulées pour la loterie L^M à partir des paramètres estimés du modèle BKL-puissance (haut) et BKL-Rachlin (bas) du groupe 1.

4.3.7. Nouvelles pistes de spécifications du modèle BKL

Dans le but d'améliorer la précision du modèle BKL, nous pourrions envisager soit de laisser le point de référence comme un paramètre libre à estimer, soit de fournir une spécification adaptative ou quasi-adaptative au point de référence comme nous l'avons évoqué dans le chapitre 2 (voir le modèle de Shi et al, 2015 ; Baucells et al, 2011 ; Barberis et al, 2001). Une autre solution consisterait à spécifier autrement le point de référence : comme par exemple le statu-quo ou l'espérance de la loterie plutôt que d'avoir recours à la règle MaxMin.

Nous avons dans un premier temps cherché à estimer le point de référence. Pour vérifier si les paramètres du modèle BKL sont identifiables par la méthode du maximum de vraisemblance lorsque le référentiel est un paramètre libre, nous avons réalisé une procédure de parameters recovery. Pour cela nous avons procédé comme dans notre précédente étude référencée dans la partie 3. Nous avons reproduit les loteries impliquées dans notre procédure expérimentale. Pour chacune de ces loteries nous avons simulé 200 équivalents certains à partir du modèle BKL associé à une structure stochastique comme celle décrite précédemment. Les paramètres du modèle BKL-puissance étaient fixés tel que : $\alpha = 0.4$ et $\gamma = 0.6$. De plus, étant donné que désormais les issues des loteries peuvent potentiellement être inférieures au point de référence, un nouveau paramètre doit être pris en compte : le paramètre d'aversion aux pertes : λ qui est spécifié comme dans la prospect theory. Ici $\lambda = -2.25$. Pour le modèle BKL-Rachlin nous avons : $\beta = 0.031$, $\gamma = 0.6$ et $\lambda = -2.25$. Les paramètres des triplets étaient quant à eux spécifiés comme indiqué par le modèle BKL.

Nous avons ensuite estimé les paramètres libres du modèle BKL ainsi que le paramètre du point de référence, R . L'estimation par maximum de vraisemblance n'a pas permis de retrouver les paramètres fixés pour la simulation des données, que ce soit pour le modèle BKL-puissance ou le modèle BKL-Rachlin. Nous en déduisons que dans notre situation expérimentale (c'est-à-dire étant donné les caractéristiques des loteries en présence) les paramètres du modèle ne sont pas identifiables lorsque R est un paramètre libre.

Une autre possibilité consiste à spécifier le référentiel différemment. Rappelons que dans le modèle BKL, les auteurs choisissent une spécification MaxMin du référentiel. Nous pourrions envisager un référentiel égal au statu quo, mais ce dernier rend le référentiel égal à 0, comme l'implique déjà la spécification MaxMin du référentiel dans notre situation expérimentale. Nous pouvons envisager de déterminer un référentiel égal à l'espérance de la loterie en jeu. Toutefois, une nouvelle fois, nous nous retrouvons face à un problème d'identification des paramètres. En effet, dans le cas où le référentiel est égal à l'espérance, l'issue la plus défavorable (x_2) est alors inférieure au référentiel. Or, dans les pertes, deux nouveaux paramètres doivent être pris en considération selon le modèle BKL. Le paramètre d'aversion aux pertes : λ et un paramètre γ_- jouant sur la courbure de la fonction $D(\cdot)$ dans les pertes. En effet, Baucells et al (2018) indiquent que le paramètre libre γ de la fonction $D(\cdot)$ peut différer entre le domaine des gains et le domaine des pertes. En procédant une nouvelle fois à une procédure de parameters recovery, il est apparu que les paramètres n'étaient pas retrouvés lors de l'estimation par maximum de vraisemblance.

Toutefois, précisons que l'estimation est fiable lorsque des loteries plus « complexes » sont instituées. Par « complexe », nous entendons des loteries impliquant aux moins deux issues possibles différentes de 0. En l'état, les loteries que nous avons choisies dans cette expérience ne permettent donc pas d'estimer les paramètres du modèle complet de BKL.

V. Discussion

Dans cette étude, nous nous sommes demandés si l'exposition antérieure à des loteries à plus ou moins fort enjeu influençait la prise de risque subséquente. Nous avons montré que, lorsque l'enjeu d'une loterie au temps $t - 1$ était supérieur à celui d'une loterie au temps t , la prise de risque au temps t était alors plus importante que lorsque l'enjeu de la loterie au temps $t - 1$ était inférieur à celui de la loterie au temps t . Ce résultat était correctement prédit par le modèle

BKL à la condition de spécifier un contexte décisionnel pouvant intégrer des loteries antérieures n'ayant plus cours au moment t .

En revanche, nous avons constaté à deux reprises que l'exposition antérieure à de faible enjeux semblait accroître l'aversion au risque subséquente face à des loteries à plus fort enjeux. Ce résultat, devra être confirmé par de futures études. En l'état, ces résultats vont à l'encontre de nos hypothèses.

Les modèles théoriques comme la théorie de l'espérance d'utilité (Von-Neumann et Morgenstern, 1945), la rank-dependent utility (Quiggin, 1982), la regret theory (Loomes et Sugden, 1982), la théorie de l'aversion désappointement (Gul, 1991), la dual theory (Yaari, 1987) ne permettent pas d'expliquer ces résultats. Pour cela, il faudrait que ces modèles intègrent une dépendance aux états passés d'une façon particulière. La seule façon dont le passé influence la future prise de décision selon ces théories passe par un vecteur objectif qui lie les périodes entre elles, comme par exemple la richesse actuelle qui se modifie sous l'impact des décisions et de la réalisation de l'aléa. Dans notre expérience cette explication n'est pas viable puisque les phases n'étaient pas connectées par un élément objectif.

La Prospect Theory quant à elle pourrait proposer une solution en spécifiant un point de référence adaptatif déconnecté des valeurs objectives de la situation. Dans cette étude nous n'avons pas suivi ce cadre théorique étant donné l'aporie que constitue la question du référentiel à ce jour.

Le modèle BKL offre un cadre théorique permettant d'expliquer ce résultat en associant la notion de taille d'enjeu à celle de l'étendue. L'étendue étant définie par la distance entre une issue financière minimum et une issues financière maximum. La question qui s'est alors posée est la suivante : lorsque l'individu opère un choix entre plusieurs loteries, de quel contexte décisionnel sont dérivées les valeurs maximum et minimum ? Ici deux solutions étaient envisageables. Soit nous considérons qu'elles étaient dérivées uniquement des loteries en

présence au moment du choix (pas de dépendance au passé dans ce cas), soit nous considérons qu'elles pouvaient être dérivées également des loteries auxquelles les individus ont été antérieurement exposés. En construisant d'une certaine façon nos conditions expérimentales, nous avons pu établir que la deuxième solution semblait correcte au vu de nos résultats et des prédictions du modèle.

Nous pensons que le modèle de BKL formalise, à travers l'effet de l'étendue, un effet de contraste. Un effet de contraste implique que la perception d'un stimulus dépend du contexte dans lequel il est intégré. Si l'on reprend nos résultats expérimentaux, nous pensons que les individus ayant été confrontés à un fort enjeu en première phase, perçoivent les issues de la loterie « à enjeu moyen » comme plus « petites » comparativement aux individus d'abord exposés à un faible enjeu (ou bien par rapport aux individus n'ayant pas subi de préexposition du tout). Dès lors, en spécifiant une fonction d'utilité convexe pour ce qui est considéré comme des petits gains et concave pour ce qui est considéré comme des grands gains, le modèle BKL est capable de capturer ces variations d'attitude face au risque en fonction du contexte. Plus spécifiquement, nous suggérons que les individus du groupe 2 préalablement exposé à un fort enjeu, perçoivent les issues de la loterie L^M comme plus petites que les individus du groupe 1 exposés préalablement à un enjeu plus faible ou que les individus du groupe 3 et 4 n'ayant pas subi de préexposition avant L^M . Dès lors, selon la spécification du modèle BKL, les issues de la loterie L^M devraient donc être évaluées sur une section moins concave par les individus du groupe 2.

A l'avenir, nous pensons que d'autres questions doivent être soulevées et investies concernant l'effet de l'étendue. Par exemple, nos résultats montrent que l'hypothèse 3 n'est pas vérifiée. Selon cette hypothèse, les individus ayant été exposé à une loterie à « moyen » enjeu devraient être moins averse au risque lorsqu'ils sont ensuite confronté à une loterie à faible enjeu. Pourtant nous n'avons observé aucune différence entre un groupe exposé antérieurement à la loterie à

enjeu moyen et un groupe directement exposé directement à la loterie à faible enjeu. Il semble donc que l'effet de l'étendue n'a pas eu d'impact sur la prise de risque. Cela pose la question d'un effet de seuil potentiel pour que l'effet de l'étendue se réalise : Il est possible que l'effet de l'étendue ne soit efficient que si la différence d'enjeu entre deux loteries réparties sur deux phases est suffisamment importante. On peut également se demander quelle est la forme de la relation entre l'aversion au risque et l'augmentation de l'étendue ? Est-ce que le degré d'aversion diminue linéairement avec l'augmentation de l'étendue ou non ? Ces questions pourront être investies dans de futures investigations expérimentales.

Les limites du modèles BKL :

Premièrement, l'effet d'interaction que nous observons entre la préexposition et le niveau de probabilité de gagner est inexplicable par le modèle BKL ni par aucune autre théorie. Nous avons observé que les primes de risque absolues du groupe 2 phase 2 étaient différentes du groupe 1 phase 2 et du groupe {3,4} phase 1 uniquement pour les probabilités supérieures ou égales à 0.7. Ainsi, la préexposition ne semble pas avoir d'impact sur l'aversion au risque pour des probabilités plus petites. Cette relation est sans doute à approfondir et des réplifications de l'expérience sont nécessaires pour confirmer ce résultat.

Deuxièmement, le modèle BKL, étant donné la spécification donnée par les auteurs à propos du cadre, ne permet pas de capturer le fait que l'exposition antérieure à un faible enjeu de loterie puisse influencer la prise de risque subséquente face à une loterie à plus fort enjeu. Si ce résultat est confirmé, il semble donc qu'un effet de contraste soit à nouveau à l'œuvre. Plus spécifiquement, il est envisageable que les issues de la loterie à plus fort enjeu soient perçues comme plus substantielles dans ce cas. Ainsi, les individus percevraient un enjeu plus important que s'ils n'avaient pas été confronté antérieurement à une loterie à faible enjeu. En l'état, le modèle BKL n'est pas en mesure de capturer un tel résultat. En effet, notre spécification du contexte décisionnel impliquait naturellement et implicitement une adaptation parfaite du

maximum G entre les deux phases lorsque l'enjeu de la première phase était inférieur à celui de la seconde phase. Dès lors, le cadre pour un individu ayant été confronté antérieurement à un faible enjeu était identique à celui d'un individu n'ayant subi aucune préexposition.

Pour finir, les résultats provenant de l'out of sample montrent que le modèle prédit un degré de recherche au risque beaucoup trop élevé. En effet, le modèle BKL prédit que l'individu sera chercheur de risque pour tous les niveaux de probabilités de gagner comme indiqué dans la Figure 4.10. On perçoit ici un potentiel problème inhérent à notre spécification du contexte décisionnel : imaginons des individus confrontés à un enjeu très élevé en première phase, par exemple $L1(1.000.000, p ; 0, 1 - p)$. Supposons ensuite que l'individu soit confronté à une loterie à plus faible enjeu tel que $L2(1.000, p ; 0, 1 - p)$. Face à cette seconde loterie il est évident que la plupart des individus devaient se montrer averse au risque pour des probabilités de gagner moyennes ou élevées même en si il y a eu préexposition à $L1$ lors d'une première phase. Pourtant, si l'on reprend les paramètres estimés du modèle BKL-Rachlin du groupe 1 phase 2, le modèle prédit que face à la loterie $L2$ les individus seront chercheurs de risque pour tous les niveaux de probabilités de gagner.

Conclusion

Même si le modèle BKL est en mesure de produire des hypothèses allant dans le sens de certains résultats obtenus dans cette expérience, il reste pour le moment très imprécis au regard de l'instabilité des paramètres libres entre les groupes expérimentaux que nous avons observé, que ce soit en supposant une fonction d'utilité puissance ou une fonction d'utilité de Rachlin. La procédure d'out of sample illustre également l'imprécision de ce modèle malgré une certaine supériorité du modèle BKL-Rachlin sur le modèle Modèle BKL-puissance.

Afin d'améliorer la qualité prédictive du modèle BKL, nous suggérons que les futures études intègrent des loteries plus complexes pour obtenir des estimations fiables des paramètres libres.

En outre, cela devrait permettre de tester différentes spécifications du point de référence. Nous pensons en effet qu'une piste possible pour améliorer la qualité du modèle BKL consiste à intégrer une spécification plus flexible du point de référence en adoptant des formalisations adaptatives comme celles de Baucells et al (2011) ou celle de Shi et al (2015). Toutefois, les spécifications de ces modèles doivent être adaptés à notre situation, ce qui n'est pas forcément réalisable de façon triviale. En l'état, notre méthodologie ne nous a pas permis d'investir plus en profondeur ces questions.

ANNEXE

A. Estimation avec un référentiel égal à l'espérance

Nous avons cherché à estimer les paramètres du modèle BKL en spécifiant un point de référence égal à l'espérance de la loterie. Cette spécification intégrant à la fois normalisation des valeurs par rapport à l'étendu et un référentiel basé sur l'espérance est appuyé par quelques études récentes en neuroéconomie (Khaw, Glimcher, & Louie, 2017; Tobler, Fiorillo, & Schultz, 2005).

Ainsi, nous remplaçons la spécification « MaxMin » du référentiel proposé par BKL par une spécification « Espérance » de la loterie, noté $E(L)$. Toutefois, avec cette formalisation, la meilleure issue est alors perçue dans le domaine des gains ($x_1 > R$) et la pire issue dans le domaine des pertes ($x_2 < R$). Or puisque selon le modèle BKL, l'évaluation des issues considérées comme des pertes et différentes de l'évaluation des issues considérées comme des gains. En particulier, BKL intègre un paramètre d'aversion aux pertes, et suppose également que le paramètre de courbure de la fonction $D(\cdot)$ peut différer entre le domaine des pertes et le domaine des gains. Nous définissons alors la fonction d'utilité puissance suivante :

$$u(x) = \begin{cases} u_+(x) = \frac{(x - E(L))^\alpha}{(G - E(L))^\alpha} & \forall x > E(L) \\ u_-(x) = \frac{-\lambda(E(L) - x)^\alpha}{-\lambda(E(L) - L)^\alpha} & \forall x < E(L) \end{cases} \quad (\text{A.1})$$

Avec λ , le paramètre d'aversion aux pertes. La fonction $D(\cdot)$ est défini comme l'inverse de la fonction de Prelec (1998) et peut différer entre le domaine des gains et le domaine des pertes, telle que :

$$D_+(u_+(x)) = \exp\left(-\left(\log\left(\frac{1}{u_+(x)}\right)\right)^{1/\gamma_+}\right) \quad (\text{A.2})$$

$$D_-(u_-(x)) = \exp\left(-\left(\log\left(\frac{1}{u_-(x)}\right)\right)^{1/\gamma_-}\right)$$

La suite de l'évaluation des loteries est spécifiée comme indiqué dans la section 2.1.2.

Le modèle BKL-EV-Puissance comprend donc maintenant quatre paramètres libres : $\theta = \{\alpha, \lambda, \gamma_+, \gamma_-\}$.

Nous avons indiqué dans la partie Résultat de ce chapitre, que le modèle souffrait d'un problème d'identification des paramètres puisque nos loteries n'étaient pas assez complexes. Toutefois, comme indiqué dans la partie 3 de cette thèse, il est possible de contraindre suffisamment l'espace des paramètres du modèle afin de faciliter la convergence du modèle. Toutefois, les contraintes que nous imposons étant particulièrement fortes, les résultats doivent être analysés avec beaucoup de prudence. Le tableau A.1 reporte les contraintes sur les paramètres que nous imposons pour réaliser l'estimation par MLE.

Cette manière de procéder a fonctionné pour un modèle BKL-puissance, mais n'a pas été fructueuse pour le modèle BKL-Rachlin. Nous reportons ici les résultats obtenus pour le modèle BKL-EV-Puissance. Nous reportons les résultats dans le tableau A.2. L'analyse des estimations par groupe montre que les coefficients sont stables. En effet, les coefficients estimés des variables indicatrices $\hat{\theta}_{\text{GROUPE}_1}$ et $\hat{\theta}_{\text{GROUPE}_{3,4}}$ intègrent 0 dans les intervalles de confiance. On notera toutefois que le BIC est supérieur aux modèles BKL-Rachlin et BKL-Puissance.

Tableau A.1. Contraintes sur les paramètres

CONTRAINTES					
Parameters	α	λ	γ_+	γ_-	η
Estimation BKL-EV- puissance model	$\hat{\theta}_0$	[0, 1]	[1, 5]	[0, 1]	[0, 1]
	$\hat{\theta}_{GROUPE_1}$	[-1, 1]		[-1, 1]	[-1, 1]
	$\hat{\theta}_{GROUPE_3_4}$	[-1, 1]		[-1, 1]	[-1, 1]

Tableau A.2. Estimation des paramètres du modèle BKL armé d'une fonction d'utilité puissance.

ESTIMATION						
Paramètres	α	λ	γ_+	γ_-	η	
Estimation BKL-EV- puissance model	$\hat{\theta}_0$	0.955 [0.622 1.288]	2.247 [0.614 3.879]	0.771 [0.436 1.105]	0.653 [0.309 0.996]	0.138 [0.124 0.152]
	$\hat{\theta}_{GROUPE_1}$	0.479 [-0.019 0.938]	-0.272 [-1.892 2.435]	0.321 [-0.184 0.825]	-0.089 [-0.532 0.354]	0.027 [0.007 0.048]
	$\hat{\theta}_{GROUPE_3_4}$	0.441 [-0.004 0.879]	-0.573 [-2.500 1.354]	0.314 [-0.141 0.769]	0.097 [-0.338 0.533]	0.031 [0.013 0.049]
BIC BKL-EV-puissance	23150.01					
Nombre de paramètres	15					

Out of sample

Nous procédons à nouveau à une procédure d'out of sample selon la même méthodologie décrite dans les résultats précédemment. La fonction en vert de la Figure A.1 est le résultat de cette procédure. Il s'agit de la prime de risque simulée à partir des coefficients estimés du

groupe 1 phase 2 sachant $G=5000$. Autrement dit, nous simulons les résultats du groupe 2 phase 2 à partir des coefficients estimés du groupe 1 phase 2. Comme nous pouvons le constater, le modèle prédit une aversion au risque trop élevée par rapport aux données observées. Cependant, le MAE par rapport à la moyenne des primes de risque observées (MAE=49.863) est inférieur à ceux obtenus avec les modèles BKL Rachlin et BKL-Puissance. Par contre le MAE entre le modèle BKL-EV-puissance et les données individuelles est équivalent à celui obtenu avec le modèle de Rachlin (MAE=90.98).

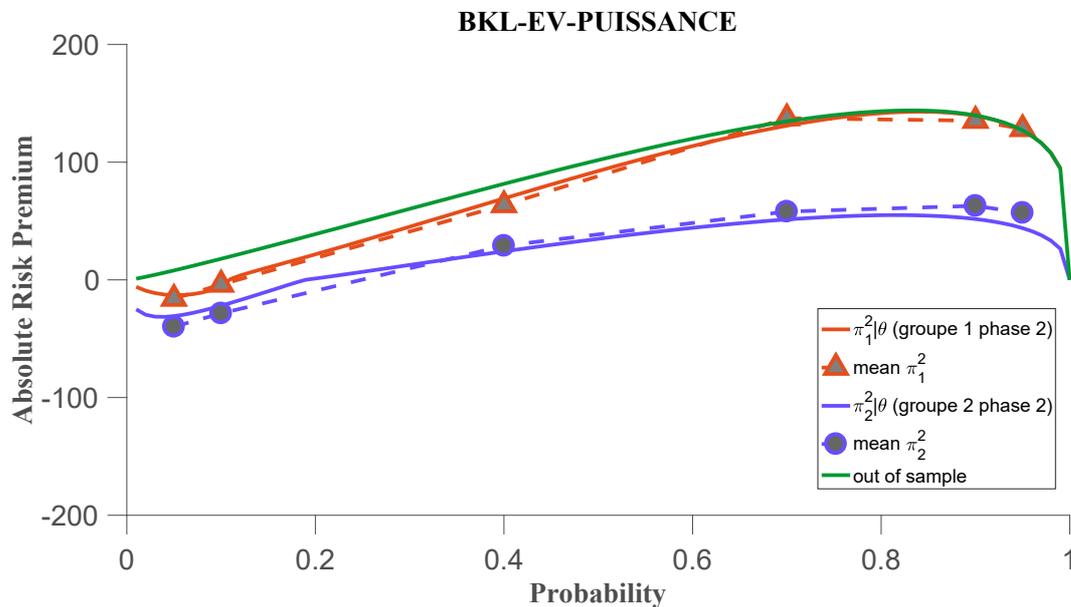


Figure A.1. Les traits pleins représentent le fit du modèle BKL-EV-puissance aux primes de risques moyennes pour la loterie L^M des groupes 1 (traits pointillés rouge) et 2 (traits pointillés bleue) en phase 2. La courbe verte représente les primes de risques simulées pour la loterie L^M du groupe 2 phase 2 à partir des paramètres estimés du modèle BKL-EV-puissance) du groupe 1 phase 2.

B. Consignes et screenshot de l'expérience

Bonjour,

Vous allez participer à une expérience d'économie expérimentale. Vous allez devoir opérer un certain nombre de choix entre deux alternatives impliquant des sommes d'argent. Ces montants financiers sont fictifs, mais nous vous demandons de réaliser vos choix comme s'il s'agissait d'argent réel. Autrement dit, comme si on vous proposait réellement ces choix dans la vraie vie.

Vous devez systématiquement choisir entre 2 alternatives :

1. Une loterie
- Ou
2. Un gain certain

Les choix seront toujours présentés de la manière suivante :

En haut de l'écran vous verrez affichée une loterie. Cette loterie vous donne l'opportunité de gagner une somme d'argent avec une certaine probabilité ou ne rien gagner avec la probabilité complémentaire. Par exemple :

En choisissant la loterie, vous avez 90% de chance d'obtenir $x\text{€}$ (90 chances sur 100) ou 10% de chance d'obtenir $y\text{€}$ (10 chances sur 100).
(les x et y seront remplacés par des montants financiers dans l'expérience)

Face à cette loterie, une liste de gains certains vous sera présentée. Ces sommes d'argent sont gagnables avec **certitude** : ils sont appelés « gains certains » dans l'expérience. Pour chacun de ses gains financiers certains vous devrez indiquer si vous préférez jouer à la loterie ou bien obtenir le gain certain.

Les alternatives vous seront présentées dans un tableau. La colonne la plus à droite présente tous les gains certains les uns à la suite de des autres.

Pour chacun de ses gains certains, vous devez indiquer si vous préférez la loterie ou le montant certain

1. Si vous préférez la loterie cliquez dans le rond correspondant à la ligne du gain certain dans la colonne la plus à gauche nommée « choisir la loterie »
2. Si vous préférez le gain certain, cliquez dans le rond correspondant à la ligne du gain certain dans la colonne du milieu nommée « choisir le gain certain ».

Une fois tous vos choix effectués, une autre question vous sera posée. Il s'agit en fait de préciser votre réponse. Vous devrez indiquer **pour quel gain certain minimal vous préférez obtenir le gain certain plutôt que la loterie.**

Après avoir répondu à cette question, une autre loterie apparaîtra et vous reproduire le processus un certain nombre de fois.

Faites toujours très attention à la loterie présentée en haut à gauche car elle change à chaque nouvelle question. Les probabilités peuvent changer mais aussi les gains financiers.

L'expérience dure en moyenne 10 minutes. Merci pour votre participation.

[Commencer la passation](#)

En choisissant la loterie, Vous avez 5% de chance d'obtenir 500€ ou 95% de chance d'obtenir 0€

Choisir la loterie	Choisir le gain certain	GAIN CERTAIN
<input type="radio"/>	<input type="radio"/>	0€
<input type="radio"/>	<input type="radio"/>	25€
<input type="radio"/>	<input type="radio"/>	50€
<input type="radio"/>	<input type="radio"/>	75€
<input type="radio"/>	<input type="radio"/>	100€
<input type="radio"/>	<input type="radio"/>	125€
<input type="radio"/>	<input type="radio"/>	150€
<input type="radio"/>	<input type="radio"/>	175€
<input type="radio"/>	<input type="radio"/>	200€
<input type="radio"/>	<input type="radio"/>	225€
<input type="radio"/>	<input type="radio"/>	250€
<input type="radio"/>	<input type="radio"/>	275€
<input type="radio"/>	<input type="radio"/>	300€
<input type="radio"/>	<input type="radio"/>	325€
<input type="radio"/>	<input type="radio"/>	350€
<input type="radio"/>	<input type="radio"/>	375€
<input type="radio"/>	<input type="radio"/>	400€
<input type="radio"/>	<input type="radio"/>	425€
<input type="radio"/>	<input type="radio"/>	450€
<input type="radio"/>	<input type="radio"/>	475€
<input type="radio"/>	<input type="radio"/>	500€

[Valider vos choix](#)

C. Médianes et boxplot

Tableau C.1. Médianes et écart-interquartiles

Prime de risque	Groupe 2 phase 2 (n=32)	Groupe 1 phase 2 (n=39)	Groupe (3,4) phase 1 (n=74)
Médiane	0	50	25
IQR	75	150	105

Tableau C.2. Médianes et écarts-interquartiles : H2

Prime de risque	Groupe 2 phase 1 (n=32)	Groupe 3 phase 2 (n=38)
Médiane	250	400
IQR	1000	1637

Tableau C.3. Médianes et écarts-interquartiles : H3

Prime de risque	Groupe 1 phase 1 (n=38)	Groupe 4 phase 2 (n=36)
Médiane	1.25	0
IQR	12.5	10

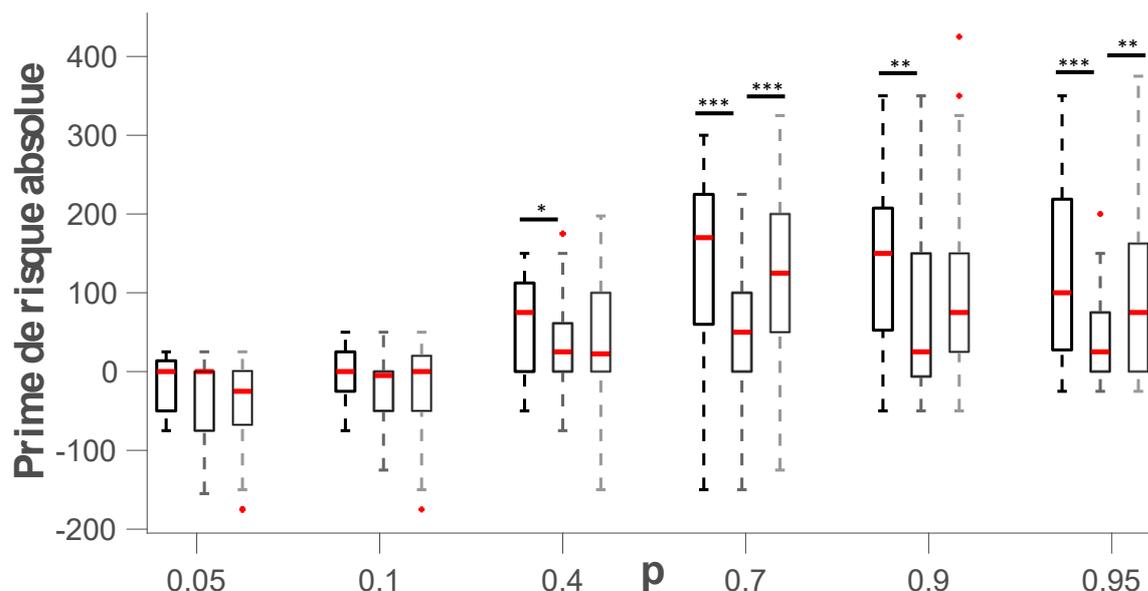


Figure C.1. Médianes (traits rouges) des distributions de primes de risque absolues des groupes 1 phase 2 (boîtes noires), groupe 2 phase 2 (boîtes grises) et groupe {3,4} phase 1 (boîtes blanches) pour chaque niveau de probabilité de gagner la meilleure issue ($x_1=500$) de la loterie L^M . Les boîtes représentent l'étendue interquartile des primes de risques. *** $p_{\text{adjusted}} < 0.001$, ** $p_{\text{adjusted}} < 0.01$, * $p_{\text{adjusted}} < 0.05$.

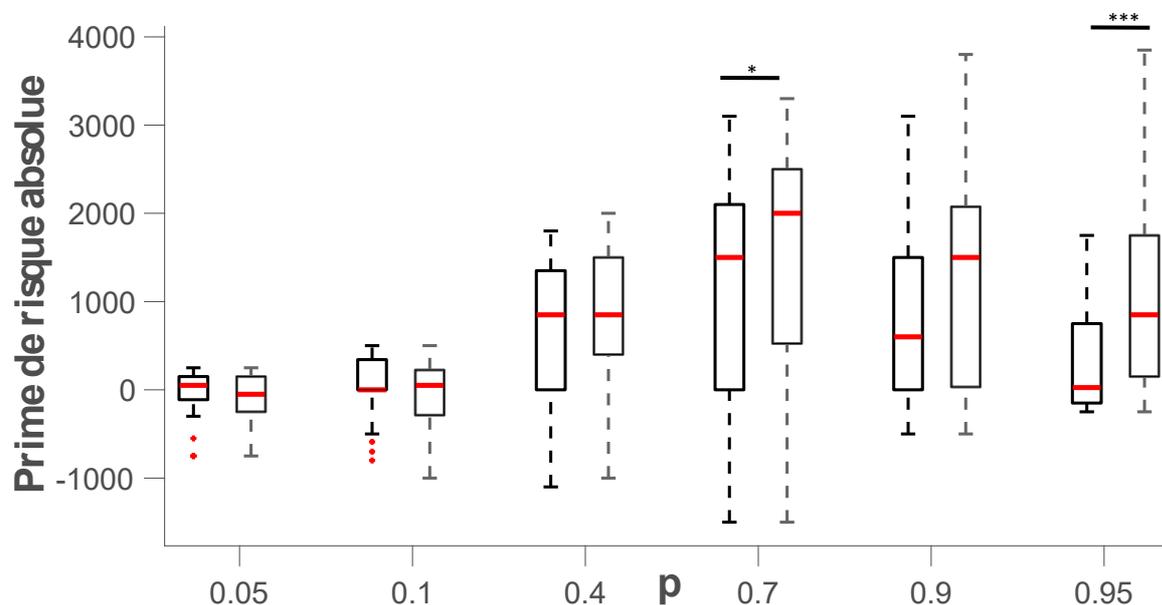


Figure C.2. Médianes (traits rouges) des distributions de primes de risque absolues des groupes 2 phase 1 (boîtes noires) et groupe 3 phase 2 (boîtes grises) pour chaque niveau de probabilité de gagner la meilleure issue ($x_1=5000$) de la loterie L^H . Les boîtes représentent l'étendue interquartile des primes de risques. *** $p_{\text{adjusted}} < 0.001$, ** $p_{\text{adjusted}} < 0.01$, * $p_{\text{adjusted}} < 0.05$.

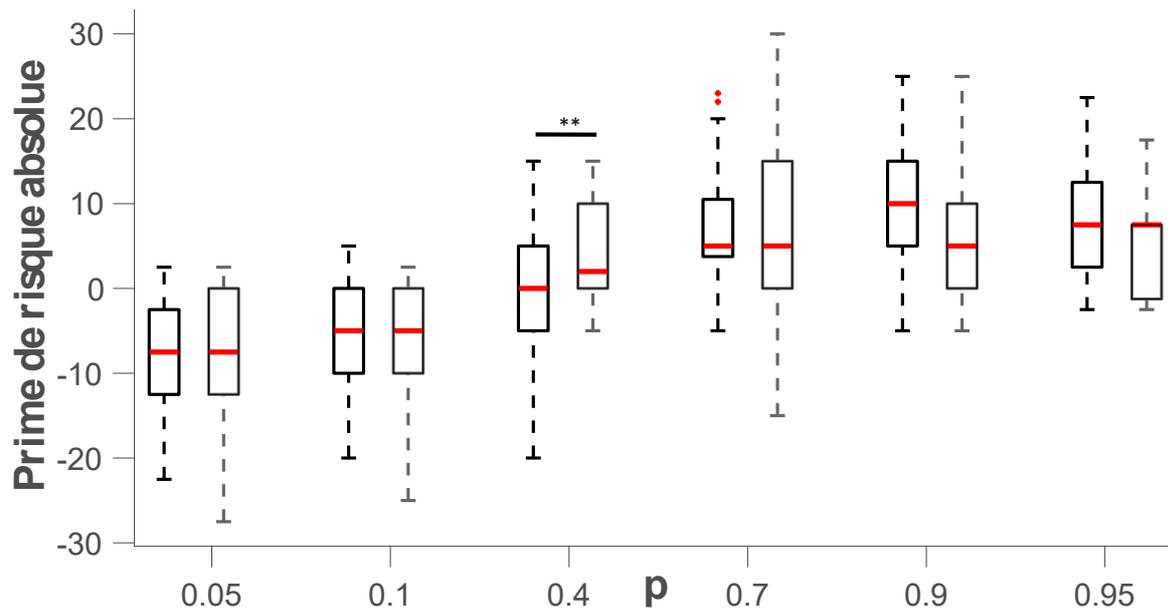


Figure C.3. Médianes (traits rouges) des distributions de primes de risque absolues des groupes 1 phase 1 (boîtes noires) et groupe 4 phase 2 (boîtes grises) pour chaque niveau de probabilité de gagner la meilleure issue ($x_1=50$) de la loterie L^S . Les boîtes représentent l'étendue interquartile des primes de risques. *** $p_{adjusted}<0.01$, ** $p_{adjusted}<0.05$, * $p_{adjusted}<0.1$

DISCUSSION GÉNÉRALE

Au cours de cette thèse nous avons dans un premier temps relaté l'histoire des théories économiques dans le risque. Nous avons vu que la théorie normative de l'utilité espérée a souvent été confrontée à son incapacité à capturer certains comportements exhibés par les agents (le paradoxe d'Allais, le reflection effect ...). Son manque de pouvoir descriptif n'a toutefois pas mené à son abandon pur et simple. Ce modèle fut dans un premier raffiné au prix de l'affaiblissement de l'axiome d'indépendance. Cette nouvelle théorie, appelée la « Rank Dependent Utility Theory » est l'une des alternatives les plus célèbres à l'EUT aujourd'hui. Elle permet notamment d'accommoder le fait que les préférences sont non-linéaires dans les probabilités.

D'autres théories comme la Regret Theory (Loomes et Sugden, 1982), la Disappointment Theory (Gul, 1991), la Saillance theory (Bordalo, 2012) ont également été proposées dans le but d'améliorer le pouvoir descriptif des modèles. Pour y parvenir, chacune d'elles a incorporé des concepts étudiés par le champ de la psychologie. Par exemple, la Regret Theory et la disappointment theory intègre dans leur formalisme un aspect « émotionnelle ». Selon ces théories, les individus anticipent avant de faire un choix, l'émotion de « joie » et de « regret » qu'ils ressentiront une fois le résultat obtenu de leur décision. Dans cette perspective, il est supposé que les individus chercheront à minimiser le regret ou le désappointement qu'ils pourraient ressentir après réalisation de l'aléa.

La théorie de la saillance, quant à elle, met en avant que les individus sont sensibles à des caractéristiques bien spécifiques de la situation risquée. Nous avons vu que la RDUT induit également cela. En effet, selon la RDUT, les issues extrêmes sont surpondérées si la fonction de pondération des probabilités est en forme de S inversé. La théorie de la saillance spécifie autrement cette surpondération mais souscrit à la même idée (voir Bardolo, 2012).

Nous voyons à travers ces exemples que les modèles économiques ont intégré de plus en plus de composants étudiés par la psychologie afin d'améliorer le pouvoir descriptifs des modèles. Que ce soit l'incorporation de l'émotion avec la Regret theory ou la disappointment theory, ou des heuristiques de jugement avec la RDUT ou la théorie de la saillance. Toutefois, ces ajouts ont un prix puisque ces théories reposent sur des axiomatiques affaiblies par rapport à celle de l'EUT et prédisent donc des comportements considérés comme incohérents.

Dans cette perspective, nous avons présenté en profondeur la Prospect Theory. Cette dernière est sans doute l'alternative la plus connue à l'EUT. Elle a véritablement été construite pour accommoder le mieux possible les comportements observés tout en s'affranchissant de considérations normatives. Ce n'est d'ailleurs pas un hasard si ce sont deux psychologues qui l'ont élaboré à l'origine. Les auteurs, Kahneman et Tversky, ont révélé que les préférences face au risque s'inversaient entre le domaine des gains et le domaine des pertes, et ce, même si les richesses finales prospectives étaient identiques. C'est dans cette perspective que les auteurs ont alors supposé l'existence d'un point de référence.

Cet ajout a constitué un pas fondamental car pour la première fois, il était pris en considération que les individus jugent les stimulus d'un environnement de façon relative et non de façon absolue. On notera toutefois que la PT présume que la distance qui sépare la valeur du stimulus de celle du référentiel est absolue. C'est précisément ce qui a été remis en question par Kontek et Lewandowski (2017) qui proposèrent de normaliser cette distance par celle de l'étendue. Dès lors la distance entre la valeur d'un stimulus et celle du référentiel est relative à la taille de l'étendue des valeurs d'un stimulus.

L'intégration du référentiel renvoie directement aux études menées dans le champ de psychophysique qui s'intéressent à la perception/jugement des stimuli. Dès lors, en couplant les études menées dans le cadre de la perception au formalisme en économie, il a été possible

de relié « sensation » et prise de décision d'une façon sans doute plus proche de la réalité que ne le faisait l'EUT.

Toutefois, nous avons vu à quel point identifier un référentiel se révèle difficile. Cette difficulté s'exprime à plusieurs niveaux.

- A un niveau théorique dans un premier temps. En effet, sur la base d'une spécification précise d'un référentiel, il faut alors en déduire ce que cela implique en termes de préférences face au risque dans des contextes statiques et dynamique. Il faut également dégager les différentes prédictions qu'induisent différentes spécifications du référentiel afin de déterminer celui qui semble être utilisé par les agents.
- A un niveau méthodologique dans second temps, il est particulièrement difficile d'inférer un point de référence sur la base de simple comportement décisionnel entre des loteries. Une des solutions qui consisterait à laisser le référentiel comme paramètre libre et à inférer sa valeur n'est pas véritablement réalisable en pratique. Notamment parce que le nombre de paramètres libres étant déjà élevé dans le modèle de la prospect theory, on se retrouve souvent face à des problèmes d'identifiabilité des paramètres.

Toutefois, un certain nombre de candidats au poste de « référentiel » ont été dégagé dans la littérature. Nous avons évoqué le statu quo, le MaxMin, le X at max P, le MinMax ou encore un point de référence stochastique comme dans le modèle de Koszegi et Rabin. Malheureusement, la plupart des études ne mettent pas en concurrence les prédictions de chacun de ces référentiels afin de trancher celui qui semble le mieux expliquer les comportements des agents. Une seule étude à notre connaissance a réalisé ce travail : l'étude de Baillon et al (2015) et semble avoir mis en évidence que le statu et la règle MaxMin étaient les référentiels les plus probablement utilisé par les agents.

A cette question du référentiel s'ajoute la question de son adaptation dans des contextes dynamiques. La encore des difficultés d'ordres théoriques et méthodologiques rendent son

étude difficile. Toutefois, certaines recherches révèlent que son actualisation semble asymétrique entre les pertes et les gains (Arkes et al, 2008, 2010) et est surtout influencé par les premiers et derniers événements d'une séquence de plusieurs périodes (Baucells et al, 2011).

En définitive, rechercher le point de référence revient à se demander comme les issues des loteries sont perçues par les agents. Une question très souvent abordé dans le champ de la psychophysique est d'étudier comment évolue la perception d'un stimulus lorsque l'amplitude de cette dernière augmente. A cette question nous connaissons la réponse prodiguée par Weber et Fechner qui formalisent une relation logarithmique entre la « sensation » et « l' amplitude ». La réponse de Stevens y voit plutôt une relation « puissance ». Cette question est primordiale dans notre recherche puisque la perception des conséquences détermine l'attitude face au risque. Nous avons notamment vu que l'augmentation de l'amplitude d'un renforçateur probabiliste induit une aversion relative face au risque de plus en plus forte. Le modèle de décision suggérée par les psychologues de l'école comportementaliste reprend la loi puissance de Stevens pour expliquer ces comportements. Pourtant, nous avons vu que ce modèle n'était pas en mesure d'accommoder les résultats.

Dès lors, nous nous sommes tournés vers des fonctions d'utilité dites « IRRA ». Ces fonctions d'utilités suggèrent que le pourcentage d'augmentation de l'utilité augmente à un taux décroissant à mesure que la richesse augmente d'une certaine proportion. On dit alors que l'élasticité de ces fonctions sont décroissantes. Nous avons alors vu que cette propriété permettait de mieux expliquer les comportements décisionnels des agents. En revanche nous en avons également éprouvé les limites de modèle puisque les individus ne présentent pas une aversion au risque IRRA pour tous les niveaux de probabilités. En effet, les primes de risques relatives augmentaient avec l'augmentation de l'enjeu des loteries uniquement pour de faibles

probabilités de gagner la meilleure issue. Pour des probabilités plus importantes, l'aversion au risque semblait constante à mesure que le niveau d'enjeu des loteries augmentait.

A ce jour, ce résultat constitue un véritable challenge. Aucun modèle existant ne semble pouvoir accommoder ce pattern comportemental. Certaines pistes sont suivies comme par exemple rendre la fonction de pondération des probabilités dépendante de l'amplitude des issues impliquées dans les loteries. D'autres pistes purement spéculatives consistent à rechercher du côté du référentiel.

Prendre en considération le niveau d'enjeu dans la perception des issues des loteries semble être une voie prometteuse à ce titre. La question qui se pose alors est : comment spécifier ce niveau d'enjeu ? Dans cette perspective, le modèle proposé par Baucells et al (2018) propose une solution envisageable. Les auteurs associent à la notion « d'enjeu » la notion « d'étendue » qui caractérise une grandeur de dispersion d'une distribution statistique. Selon eux, les issues financières engagées dans une loterie sont comparées à un référentiel comme le suppose la Prospect Theory. Toutefois, la distance séparant une issue et le référentiel est normalisé par rapport à l'étendue des issues engagées dans un contexte décisionnel particulier. En l'état, ce modèle n'est pas en mesure pour le moment d'accommoder le fameux pattern IRRA CRRA que nous évoquions. Mais elle reste néanmoins une voie prometteuse.

Ce modèle présente en effet différents attributs attrayants : en plus d'avoir été axiomatisé par les auteurs, elle repose sur de nombreuses études menées dans le champ de la psychophysique et de la neuroscience portant sur le « range effect ». L'idée sous-jacente au « range effect » est que les individus perçoivent les stimuli relativement à l'étendue des valeurs de ces stimuli (Parducci, 1965). Ce type de formalisation est également calqué sur la spécification moderne du codage neuronale : « *OFC value coding adapts to the local statistics of recently experienced rewards [38-40]. For example, in monkeys experiencing different ranges of reward magnitudes, narrow reward ranges produce steep firing rate functions of value while wide*

ranges elicit shallow ones. Similar adaptive value coding has been observed in other monkey brain regions including the anterior cingulate cortex and midbrain dopaminergic nuclei, as well as in reward-related circuits in the human brain. Such adaptive changes in neural response functions are widespread in sensory processing, and may be crucial for neurons with limited dynamic ranges facing changing contexts in a non-stationary world » (Louie, Glimcher & Webb, 2015, p. 94).

On notera que le modèle de Baucells et al (2018) n'est pas la seule tentative existante de formalisation faisant intervenir l'étendue. Récemment, une recherche menée par Bavard et al (2018) a poursuivi le même objectif : *« At the behavioral level, it has long been known that our sensitivity to sensory stimuli or monetary amounts is not the same across different ranges of intensity/magnitude. These findings have recently paralleled with the description of neuronal range adaptation: in short, the need to provide efficient coding of information in various ranges of situations entails that the firing rate of neuron adapts to the distributional properties of the variable being encoded. Converging pieces of evidence have recently confirmed neuronal range-adaptation in economic and perceptual decision-making, although its exact implementation remains debated » (Bavard et al, 2018, p2).*

CONCLUSION GENERALE

Cette thèse s'est inscrite au carrefour de la psychologie et l'économie réunies autour de l'étude de la prise de décision en environnement risqué. Nous avons pu exposer comment se sont imbriquées leurs influences mutuelles pour donner finalement naissance à l'économie expérimentale et à l'économie comportementale. Certains chercheurs différencient ces deux domaines en attribuant à l'économie expérimentale le soin de tester en laboratoire les théories économiques, et à l'économie comportementale le soin d'incorporer des notions

« psychologiques » aux modèles économiques. En réalité, la frontière est bien plus floue et l'hétérogénéité des approches méthodologiques et théoriques dans ce domaine de recherche en témoignent.

L'épistémologie de cette nouvelle science est en revanche encore en construction. Ces dernières décennies ont été marquées par un véritable jeu d'aller et retour entre la psychologie et l'économie : L'économie prodiguant une assise formelle à des modèles permettant d'induire des hypothèses précises et falsifiables et de l'autre côté la psychologie proposant une approche inductive basée sur les comportements des organismes quitte à s'affranchir d'un formalisme rigoureux. Certains y voient donc un terrain d'affrontement entre une science rigide et hypothético-déductive que constitue l'économie contre une science inductive plus proche de la réalité et moins à même de vouloir ériger à tout prix en loi la moindre régularité observée expérimentalement.

En définitive, le combat a souvent été réduit à la question « rationalité versus irrationalité ». Heureusement, aujourd'hui la conjugaison des deux champs a fini par exploiter des voies leur permettant d'exprimer chacun son domaine d'expertise. A ce titre, des éléments/processus cognitifs ont été intégrés dans les modèles économiques, qui en retour, en formalisant ces processus, ont permis de maintenir une certaine qualité normative et prédictive des modèles. On assiste aujourd'hui véritablement à une tension entre la volonté de mieux « coller » aux comportements observés et l'exigence de cohérence interne que l'on attend des modèles. Plus que cela, on exige une axiomatique suffisamment simple et réaliste comme base des modèles théoriques. Cette exigence est salutaire car elle permet, en théorie, de ne pas multiplier les modèles ad hoc, que l'on érige simplement pour capturer les données sans s'interroger sur les sous-bassements de ces modèles et de ce qu'ils impliquent normativement. L'exemple le plus évident à ce titre est le modèle 1.3 que nous avons exposé dans notre chapitre 1. Nous avons vu que les préférences n'étaient pas linéaires dans les probabilités. Il a donc été supposé

simplement que les individus distordaient les probabilités individuelles, ce qui permettait d'expliquer les comportements. Cependant, une analyse simple du modèle a montré son absurdité puisque le modèle ne respectait pas l'axiome de monotonicité. De l'autre côté, une théorie trop rigide ne prenant pas en compte la réalité des comportements risque d'être purement cohérente mathématiquement mais incapable de capturer le réel. Même si cela est exagéré, la théorie de l'espérance d'utilité est souvent associée à ce dernier cas de figure.

Toutefois, les données expérimentales ne doivent pas être vues comme l'alpha et l'oméga permettant d'appuyer ou rejeter une théorie ou un modèle. Si l'on se focalise uniquement sur les données, on risque de produire des modèles *ad hoc* potentiellement incohérents normativement et même de rejeter des théories sur la base de méthodologies imparfaites. Dans cette perspective, une vision Bayésienne appliquée à nos connaissances et à nos décisions vis-à-vis des modèles à adopter devrait être mise à l'œuvre. Une théorie n'est pas rejetée parce qu'une expérience a montré son invalidité, elle le devient si suffisamment de preuves sont réunies pour l'invalidiser. Il faudra d'autant plus de preuves si la théorie a, par le passé, montré son efficacité à capturer des résultats et sa robustesse normative. Dès lors, face à des théories efficaces mais qui ne parviennent pas à capturer certains phénomènes, on sera partagé entre « sauver » ces théories au prix de raffinements *ad hoc* sans modifier drastiquement les axiomatiques sur lesquels ces théories reposent. Soit on changera radicalement de théorie et d'axiomatique.

Ce genre de tension n'est pas propre à l'économie. La physique y est confrontée tout autant. Par exemple, prenons la vitesse de rotation des étoiles en bordure de galaxie, la théorie de la relativité générale prédit une vitesse de rotation bien plus élevée que la réalité. Or, cette théorie étant particulièrement efficace pour expliquer de nombreux phénomènes, on rechigne alors à en changer. Les chercheurs du domaine ont alors suggéré de façon purement *ad hoc* l'existence d'une matière invisible, la matière noire, qui en lui assignant certaines propriétés, permet de

capturer la vitesse de rotation des étoiles en bordure de galaxie. Reste maintenant à prouver l'existence de cette matière, sans quoi la théorie de la relativité générale devra être remplacée.

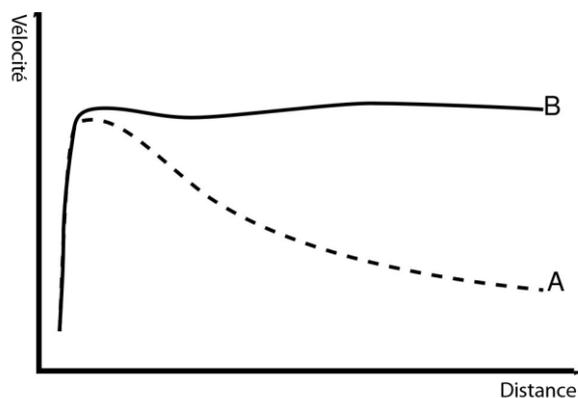


Figure 5.1 Vitesse de rotation des étoiles en fonction de leurs distances par rapport au centre d'une galaxie. En trait plein il s'agit de la vitesse prédite par la théorie de la relativité générale, en pointillé il s'agit de la vitesse observée.

L'histoire de la psychologie et de l'économie a finalement produit continuellement ce genre de tension. Cette thèse a montré à travers l'exemple de l'intégration du référentiel dans les modèles économiques ce parcours sinueux entre volonté de décrire les comportements et théories non *ad hoc*.

L'intégration du référentiel pose en fait la question fondamentale de la « perception » des stimuli la part des organismes. Il semble clair, voire évident, que l'évaluation d'un stimulus est relative plutôt qu'absolue. Toutefois, il reste à définir et spécifier ce que l'on entend par « relatif ». C'est-à-dire par rapport à quoi sont comparés les stimuli pour être évalués ?

BIBLIOGRAPHIE

- Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management science*, *46*(11), 1497–1512.
- Abdellaoui, M., l'Haridon, O., & Zank, H. (2010). Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty*, *41*(1), 39–65.
- Abdellaoui, M., & Munier, B. (1998). The risk-structure dependence effect: Experimenting with an eye to decision-aiding. *Annals of Operations Research*, *80*, 237–252.
- Abdellaoui, M., & Munier, B. (2001). Substitutions probabilistiques et décision individuelle devant le risque. *Revue d'économie politique*, *111*(1), 29–39.
- Ainslie, G. W. (1974). Impulse contrôle in pigeons. *Journal of the experimental analysis of behavior*, *21*(3), 485–489.
- Allais, M. (1953). L'extension des théories de l'équilibre économique général et du rendement social au cas du risque. *Econometrica, Journal of the Econometric Society*, 269–290.
- Andersen, S., Harrison, G. W., & Rutström, E. E. (2006). *Choice behavior, asset integration and natural reference points*. Working Paper 06.
- Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of personality and social psychology*, *2*(1), 1.
- Andreoni, J., & Sprenger, C. (2010). Certain and uncertain utility: The allais paradox and five decision theory phenomena. *Levine's Working Paper Archive*, 926159295.
- Arkes, H. R., Hirshleifer, D., Jiang, D., & Lim, S. (2008). Reference point adaptation: Tests in the domain of security trading. *Organizational Behavior and Human Decision Processes*, *105*(1), 67–81.

-
- Arkes, H. R., Hirshleifer, D., Jiang, D., & Lim, S. S. (2010). A cross-cultural study of reference point adaptation: Evidence from China, Korea, and the US. *Organizational Behavior and Human Decision Processes*, 112(2), 99–111.
- Arrow, K. J. (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnssoinin Säätiö.
- Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & cognition*, 32(8), 1369–1378.
- Baillon A., Bleichrodt, H., & Spinu, V. (in press). Searching for the Reference Point. *Management Science*.
- Barberis, N., & Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5), 2066–2100.
- Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The quarterly journal of economics*, 116(1), 1–53.
- Barkan, R., & Busemeyer, J. R. (1999). Changing plans: dynamic inconsistency and the effect of experience on the reference point. *Psychonomic Bulletin & Review*, 6(4), 547–554.
- Barkan, R., & Busemeyer, J. R. (2003). Modeling dynamic inconsistency with a changing reference point. *Journal of Behavioral Decision Making*, 16(4), 235–255.
- Baucells, M., Kontek, K., & Lewandowski, M. (2018). Range and Sign Dependent Utility for Risk and Time.
- Baucells, M., Weber, M., & Welfens, F. (2011). Reference-point formation and updating. *Management Science*, 57(3), 506–519.
- Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature communications*, 9(1), 4503.
- Bayes, T., Price, R., & Canton, J. (1763). An essay towards solving a problem in the doctrine of chances.

- Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations research*, 33(1), 1–27.
- Bennett, P., Brennan, M., & Kearns, Z. (2003). Psychological aspects of price: An empirical test of order and range effects. *Marketing Bulletin*, 14(1), 1–8.
- Bickel, W. K., Odum, A. L., & Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology*, 146(4), 447–454.
- Binmore, K., & Shaked, A. (2007). Experimental economics: science or what?
- Binswanger, H. P. (1982). Empirical estimation and use of risk preferences: Discussion. *American Journal of Agricultural Economics*, 64(2), 391–393.
- Birnbaum, M. H., & Schmidt, U. (2010). Testing transitivity in choice under risk. *Theory and Decision*, 69(4), 599–614.
- Bouchouicha, R., & Vieider, F. M. (2017). Accommodating stake effects under prospect theory. *Journal of Risk and Uncertainty*, 55(1), 1–28.
- Busemeyer, J. R., & Rieskamp, J. (2014). Psychological research and theories on preferential choice. *Handbook of Choice Modeling*. Cheltenham: Edward Elgar Publication, 49–72.
- Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, 50(3), 220–241.
- Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2), 407–441.
- Camerer, C. F. (1992). Recent tests of generalizations of expected utility theory. In *Utility theories: Measurements and applications* (p. 207–251). Springer.
- Camerer, C. F., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2), 167–196.

-
- Carbone, E., & Hey, J. D. (1994). Estimation of expected utility and non-expected utility preference functionals using complete ranking data. In *Models and experiments in risk and rationality* (p. 119–139). Springer.
- Cavagnaro, D. R., Gonzalez, R., Myung, J. I., & Pitt, M. A. (2013). Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management science*, *59*(2), 358–375.
- Cayatte, J.-L. (2009). *Microéconomie de l'incertitude*. De Boeck Supérieur.
- Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, *87*, 43–51.
- Chen, H., & Rao, A. R. (2002). Close encounters of two kinds: False alarms and dashed hopes. *Marketing Science*, *21*(2), 178–196.
- Chichilnisky, G. (2010). The foundations of statistics with black swans. *Mathematical Social Sciences*, *59*(2), 184–192.
- Christensen, J., Parker, S., Silberberg, A., & Hursh, S. (1998). Trade-offs in choice between risk and delay depend on monetary amounts. *Journal of the Experimental Analysis of Behavior*, *69*(2), 123–139.
- Cohen, M., Etner, J., & Jeleva, M. (2008). Dynamic decision making when risk perception depends on past experience. *Theory and Decision*, *64*(2-3), 173–192.
- Crosetto, P., & Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, *47*(1), 31–65.
- Dai, Z., Grace, R. C., & Kemp, S. (2009). Reward contrast in delay and probability discounting. *Learning & behavior*, *37*(3), 281–288.
- Dhmi, S. (2016). *The foundations of behavioral economic analysis*. Oxford University Press.
- Diecidue, E., & Wakker, P. P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty*, *23*(3), 281–298.

-
- Dillenberger, D., & Rozen, K. (2015). History-dependent risk attitude. *Journal of Economic Theory*, *157*, 445–477.
- Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses reflect marginal utility. *Current biology*, *24*(21), 2491–2500.
- Eil, D., & Lien, J. W. (2014). Staying ahead and getting even: Risk attitudes of experienced poker players. *Games and Economic Behavior*, *87*, 50–69.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics*, *643–669*.
- Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Memory & Cognition*, *34*(4), 914–928.
- Fafchamps, M., Kebede, B., & Zizzo, D. J. (2015). Keep up with the winners: Experimental evidence on risk taking, asset integration, and peer effects. *European Economic Review*, *79*, 59–79.
- Fechner, G. (1860). *Éléments de psychophysique*. Leipzig: Breitkopf and härtel.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817–868.
- Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty*, *40*(2), 147–180.
- Fehr-Duda, H., & Epper, T. (2012). Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences. *Annu. Rev. Econ.*, *4*(1), 567–593.
- Fiegenbaum, A., & Thomas, H. (1988). Attitudes toward risk and the risk–return paradox: prospect theory explanations. *Academy of Management journal*, *31*(1), 85–106.
- Franken, I. H., Georgieva, I., & Muris, P. (2006). The rich get richer and the poor get poorer: On risk aversion in behavioral decision-making.
- Gayant, J.-P. (1995). Généralisation de l'espérance d'utilité en univers risqué : représentation et estimation. *Revue économique*, 1047–1061.

-
- Gerhard, P., Hoffmann, A. O., & Post, T. (2015). How Past Performance Framing Impacts Investors' Belief Updating.
- Gneezy, U. (2005). Updating the reference level: Experimental evidence. In *Experimental business research* (p. 263–284). Springer.
- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, *112*(2), 631–645.
- Goette, L., Cerulli-Harms, A., & Sprenger, C. (s. d.). Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences. *American Economic Journal: Microeconomics*.
- Gómez-Mejía, L. R., Haynes, K. T., Núñez-Nickel, M., Jacobson, K. J., & Moyano-Fuentes, J. (2007). Socioemotional wealth and business risks in family-controlled firms: Evidence from Spanish olive oil mills. *Administrative science quarterly*, *52*(1), 106–137.
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive psychology*, *38*(1), 129–166.
- Green, L., & Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *American Zoologist*, *36*(4), 496–505.
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological bulletin*, *130*(5), 769.
- Green, L., Myerson, J., Holt, D. D., Slevin, J. R., & Estle, S. J. (2004). Discounting of delayed food rewards in pigeons and rats: is there a magnitude effect? *Journal of the experimental analysis of behavior*, *81*(1), 39–50.
- Green, L., Myerson, J., & O'Connell, P. (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 418.

- Grosshans, D., & Zeisberger, S. (2018). All's well that ends well? On the importance of how returns are achieved. *Journal of Banking & Finance*, *87*, 397–410.
- Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667–686.
- Haag, G. (1989). Spatial Interaction Models and their Micro-Foundation. In *Dynamic Decision Theory* (p. 165–190). Springer.
- Hammond, P. J. (1988). Consequentialist foundations for expected utility. *Theory and decision*, *25*(1), 25–78.
- Hanemann, W. M. (1991). Willingness to pay and willingness to accept: how much can they differ? *The American Economic Review*, *81*(3), 635–647.
- Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica: Journal of the Econometric Society*, 1251–1289.
- Hershey, J. C., & Schoemaker, P. J. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science*, *31*(10), 1213–1231.
- Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, 1291–1326.
- Hilton, R. W. (1988). Risk attitude under two alternative theories of choice under risk. *Journal of Economic Behavior & Organization*, *9*(2), 119–136.
- Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics*, *36*(1-3), 337–386.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644–1655.
- Janiszewski, C., & Lichtenstein, D. R. (1999). A range theory account of price perception. *Journal of Consumer Research*, *25*(4), 353–368.

-
- Johnson, J. G., & Busemeyer, J. R. (2001). Multiple-stage decision-making: the effect of planning horizon length on dynamic consistency. *Theory and Decision*, 51(2-4), 217–246.
- Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the experimental analysis of behavior*, 77(2), 129–146.
- Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational behavior and human decision processes*, 51(2), 296–312.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1), 193–206.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292.
- Kameda, T., & Davis, J. H. (1990). The function of the reference point in individual and group risk decision making. *Organizational Behavior and Human Decision Processes*, 46(1), 55–76.
- Kammoun, H. (2007). *Elicitation non-paramétrique de la fonction d'utilité et de l'aversion aux pertes sous l'hypothèse "prospect theory"* (PhD Thesis). Arts et Métiers ParisTech.
- Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational behavior and human performance*, 21(1), 61–72.
- Karmarkar, U. S. (1979). Subjectively weighted utility and the Allais paradox. *Organizational Behavior and Human Performance*, 24(1), 67–72.
- Karni, E., & Schmeidler, D. (1991). Utility theory with uncertainty. *Handbook of mathematical economics*, 4, 1763–1831.
- Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, 201715293.

-
- Khöszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.
- Khöszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, *97*(4), 1047–1073.
- Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, *126*(1), 54.
- Kirby, K. N. (2011). An empirical assessment of the form of utility functions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 461.
- Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic bulletin & review*, *3*(1), 100–104.
- Kontek, K. (2015). Continuous or discontinuous? Estimating indifference curves inside the Marschak-Machina triangle using certainty equivalents.
- Kontek, K., & Lewandowski, M. (2017). Range-dependent utility. *Management Science*, *64*(6), 2812–2832.
- Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological bulletin*, *133*(1), 95.
- Kunreuther, H. (1996). Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty*, *12*(2-3), 171–187.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478.
- Lainé, M. (2014). *Quelle rationalité pour les esprits animaux?: étude sur le comportement d'investissement des entrepreneurs en incertitude non probabilisable* (PhD Thesis). Université de Bordeaux.
- Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization*, *17*(3), 377–400.

-
- Lehner, J. M. (2000). Shifts of reference points for framing of strategic decisions and changing risk-return associations. *Management Science*, 46(1), 63–76.
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8(2), 75.
- Li, P., & Vu, Q. D. (2013). Identification of parameter correlations for parameter estimation in dynamic biological models. *BMC systems biology*, 7(1), 91.
- Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24(2), 103–130.
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368), 805–824.
- Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, 53(2), 271–282.
- Loomes, G., & Sugden, R. (1987). Some implications of a more general form of regret theory. *Journal of Economic Theory*, 41(2), 270–287.
- Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39(3-4), 641–648.
- Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: from biological to behavioral decision-making. *Current opinion in behavioral sciences*, 5, 91–99.
- Machina, M. J. (1982). « Expected Utility » Analysis without the Independence Axiom. *Econometrica: Journal of the Econometric Society*, 277–323.
- Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, 27(4), 1622–1668.

- Madden, G. J., Bickel, W. K., & Jacobs, E. A. (1999). Discounting of delayed rewards in opioid-dependent outpatients: exponential or hyperbolic discounting functions? *Experimental and clinical psychopharmacology*, 7(3), 284.
- Markowitz, H. (1952a). Portfolio selection. *The journal of finance*, 7(1), 77–91.
- Markowitz, H. (1952b). The utility of wealth. *Journal of political Economy*, 60(2), 151–158.
- Marschak, J. (1950). Rational behavior, uncertain prospects, and measurable utility. *Econometrica: Journal of the Econometric Society*, 111–141.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. *Commons, ML.; Mazur, JE.; Nevin, JA*, 55–73.
- McClelland, G. H., Schulze, W. D., & Coursey, D. L. (1993). Insurance for low-probability hazards: A bimodal response to unlikely events. In *Making Decisions About Liability And Insurance* (p. 95–116). Springer.
- McCord, M., & De Neufville, R. (1986). “Lottery equivalents”: Reduction of the certainty effect problem in utility assessment. *Management Science*, 32(1), 56–60.
- McKerchar, T. L., Green, L., & Myerson, J. (2010). On the scaling interpretation of exponents in hyperboloid models of delay and probability discounting. *Behavioural processes*, 84(1), 440–444.
- Morgenstern, O., & Von Neumann, J. (1953). *Theory of games and economic behavior*. Princeton university press.
- Munier, B. (1995). Entre rationalités instrumentale et cognitive: contributions de la dernière décennie à la modélisation du risque/From Instrumental to Cognitive Rationality: Contributions of the Last Decade to Risk Theory. *Revue d'économie politique*, 5–70.
- Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of experimental psychology*, 64(5), 482.

-
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the experimental analysis of behavior*, 64(3), 263–276.
- Myerson, J., Green, L., & Morris, J. (2011). Modeling the effect of reward amount on probability discounting. *Journal of the experimental analysis of behavior*, 95(2), 175–187.
- Niedrich, R. W., Sharma, S., & Wedell, D. H. (2001). Reference price and price perceptions: A comparison of alternative models. *Journal of Consumer Research*, 28(3), 339–354.
- Nilsson, H. akan, Rieskamp, J., & Wagenmakers, E.-J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, 55(1), 84–93.
- Nolte, S., & Schneider, J. C. (2018). How price path characteristics shape investment behavior. *Journal of Economic Behavior & Organization*, 154, 33–59.
- Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of finance*, 53(5), 1775–1798.
- Ostrom, T. M., & Upshaw, H. S. (1968). Psychological perspective and attitude change. In *Psychological foundations of attitudes* (p. 217–242). Elsevier.
- Papon, T. (2004). L'influence de la durée d'engagement et du vécu dans les décisions d'assurance: deux études expérimentales. *Cahiers de la MSE*, 40.
- Parducci, A. (1965). Category judgment: a range-frequency model. *Psychological review*, 72(6), 407.
- Parducci, A. (1968). The relativism of absolute judgments. *Scientific American*, 219(6), 84–93.
- Payne, J. W., Laughhunn, D. J., & Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26(10), 1039–1060.
- Petry, N. M. (2002). Discounting of delayed rewards in substance abusers: relationship to antisocial personality disorder. *Psychopharmacology*, 162(4), 425–432.
- Pollack, I. (1952). The information of elementary auditory displays. *The Journal of the Acoustical Society of America*, 24(6), 745–749.

-
- Post, T., Van den Assem, M. J., Baltussen, G., & Thaler, R. H. (2008). Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review*, 98(1), 38–71.
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32, 122–136.
- Pratt, J. W., & Zeckhauser, R. J. (1987). Proper risk aversion. *Econometrica: Journal of the Econometric Society*, 143–154.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 497–527.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4), 323–343.
- Rachlin, H. (1992). Diminishing marginal value as delay discounting. *Journal of the Experimental Analysis of Behavior*, 57(3), 407–415.
- Rachlin, H. (2006). Notes on discounting. *Journal of the experimental analysis of behavior*, 85(3), 425–435.
- Rachlin, H., Castrogiovanni, A., & Cross, D. (1987). Probability and delay in commitment. *Journal of the Experimental Analysis of Behavior*, 48(3), 347–353.
- Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice. *Psychological review*, 93(1), 33.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the experimental analysis of behavior*, 55(2), 233–244.
- Rachlin, H., & Siegel, E. (1994). Temporal patterning in probabilistic choice. *Organizational Behavior and Human Decision Processes*, 59(2), 161–176.
- Raineri, A., & Rachlin, H. (1993). The effect of temporal constraints on the value of money and other commodities. *Journal of Behavioral Decision Making*, 6(2), 77–94.
- Richards, J. B., Mitchell, S. H., De Wit, H., & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the experimental analysis of behavior*, 67(3), 353–366.

-
- Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic theory*, 2(3), 225–243.
- Saha, A. (1993). Expo-power utility: A ‘flexible’ form for absolute and relative risk aversion. *American Journal of Agricultural Economics*, 75(4), 905–913.
- Scholten, M., & Read, D. (2014). Prospect theory and the “forgotten” fourfold pattern of risk preferences. *Journal of Risk and Uncertainty*, 48(1), 67–83.
- Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. *Physiological reviews*, 95(3), 853–951.
- Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences*, 21(2), 153–167.
- Shead, N. W., & Hodgins, D. C. (2009). Probability discounting of gains and losses: Implications for risk attitudes and impulsivity. *Journal of the experimental analysis of behavior*, 92(1), 1–16.
- Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, 40(3), 777–790.
- Shi, Y., Cui, X., Yao, J., & Li, D. (2015). Dynamic trading with reference point adaptation and loss aversion. *Operations Research*, 63(4), 789–806.
- Slovic, P. (2000). Rational actors and rational fools: The influence of affect on judgment and decision-making. *Roger Williams UL Rev.*, 6, 163.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological review*, 64(3), 153.
- Stevenson, M. K. (1986). A discounting model for decisions with delayed positive or negative outcomes. *Journal of Experimental Psychology: General*, 115(2), 131.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165–180.

- Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, 36(6), 643–660.
- Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of political Economy*, 89(2), 392–406.
- Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715), 1642–1645.
- Tserenjigmid, G. (2015). *History-Dependent Risk Aversion and the Reinforcement Effect*. working paper.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124–1131.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *science*, 211(4481), 453–458.
- Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica: Journal of the Econometric Society*, 1255–1280.
- Volij, O. (1994). Dynamic consistency, consequentialism and reduction of compound lotteries. *Economics letters*, 46(2), 121–129.
- Wakker, P. (1994). Separating marginal utility and probabilistic risk aversion. *Theory and decision*, 36(1), 1–44.
- Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management science*, 42(8), 1131–1150.
- Wakker, P., Erev, I., & Weber, E. U. (1994). Comonotonic independence: The critical test between classical and rank-dependent utility theories. *Journal of Risk and Uncertainty*, 9(3), 195–230.
- Weber, B. J., & Chapman, G. B. (2005). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes*, 97(1), 31–46.

-
- Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin*, *115*(2), 228.
- Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, *33*(2), 167–184.
- Weber, M., & Zuchel, H. (2005). How do prior outcomes affect risk attitude? Comparing escalation of commitment and the house-money effect. *Decision Analysis*, *2*(1), 30–43.
- Willinger, M. (1990). Irréversibilité et cohérence dynamique des choix. *Revue d'économie politique*, 808–832.
- Wu, G., & Gonzalez, R. (1998). Common consequence conditions in decision making under risk. *Journal of Risk and Uncertainty*, *16*(1), 115–139.
- Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica: Journal of the Econometric Society*, 95–115.
- Yang, J.-B. (2000). Minimax reference point approach and its application for multiobjective optimisation. *European Journal of Operational Research*, *126*(3), 541–556.
- Yeung, C. W., & Soman, D. (2005). Attribute evaluability and the range effect. *Journal of Consumer Research*, *32*(3), 363–369.
- Yukalov, V. I., & Sornette, D. (2009). Physics of risk and uncertainty in quantum decision making. *The European Physical Journal B*, *71*(4), 533–548.
- Yukalov, V., & Sornette, D. (2017). Quantum probabilities as behavioral probabilities. *Entropy*, *19*(3), 112.
- Yukalov, Vyacheslav I., Yukalova, E. P., & Sornette, D. (2018). Information processing by networks of quantum decision makers. *Physica A: Statistical Mechanics and its Applications*, *492*, 747–766.