



HAL
open science

Four contributions at the frontier of risk and uncertainty in finance and insurance

Silvia Faroni

► **To cite this version:**

Silvia Faroni. Four contributions at the frontier of risk and uncertainty in finance and insurance. Economics and Finance. Université Lumière - Lyon II, 2023. English. NNT : 2023LYO20030 . tel-04240608

HAL Id: tel-04240608

<https://theses.hal.science/tel-04240608>

Submitted on 13 Oct 2023

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.



N° d'ordre NNT : 2023LYO20030

THÈSE de DOCTORAT DE L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 486

Sciences économiques et de Gestion

Discipline : Sciences de gestion

Soutenue publiquement le 19 juin 2023, par :

Silvia FARONI

Four Contributions at the Frontier of Risk and Uncertainty in Finance and Insurance.

Devant le jury composé de :

Laurent VILANOVA, Professeur des Universités, Université Lumière Lyon 2, Président

Johanna ETNER, Professeure des Universités, Université Paris Nanterre, Rapporteur

Delphine LAUTIER, Professeure des Universités, Université Paris-Dauphine, Rapporteur

Krzysztof OSTASZEWSKI, Professeur, Université d'État d'Illinois, Examineur

Olivier LE COURTOIS, Professeur, EMLyon business school, Directeur de thèse

Éric ANDRÉ, Associate Professeur, EMLyon business school, Co-Directeur de thèse

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale - pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.



THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de

l'Université Lumière - Lyon II

**École Doctorale: ED 486
de Sciences Économiques et de Gestion**

Discipline: Sciences de Gestion

Four Contributions at the Frontier of Risk and Uncertainty in Finance and Insurance

Soutenance publique le 19 Juin 2023 à 14h par:

Silvia FARONI

Composition du Jury

Eric ANDRÉ	Professeur Associé	Emlyon Business School	<i>Co-directeur</i>
Johanna ETNER	Professeure des universités	Université de Paris Nanterre	<i>Rapportrice</i>
Delphine LAUTIER	Professeure des universités	Université de Paris Dauphine	<i>Rapportrice</i>
Olivier LE COURTOIS	Professeur	Emlyon Business School	<i>Directeur</i>
Krzysztof M. OSTASZEWSKI	Professeur distingué	Illinois State University	<i>Suffragant</i>
Laurent VILANOVA	Professeur des universités	Université Lumière Lyon 2	<i>Suffragant</i>

Acknowledgements

Three years ago, I had to decide whether I wanted to pursue a Ph.D. degree in Lyon. This decision was not easy, especially in a world turned upside down by the onset of the Covid pandemic. Back then, it seemed a rational choice to accept, and now, after three years, I can say that it was also the right choice that enabled me to overcome some of my limitations.

For this reason, I want to express my sincerest gratitude to my thesis advisor, professor Olivier Le Courtois, and co-advisor, professor Eric André, who guided me all along this project. Discussions with them were always helpful in my research work, and their advice and encouragement supported me a lot throughout my Ph.D. study.

I want to thank my committee members. I want to thank professor Johanna Etner and professor Delphine Lautier for kindly accepting to be my thesis reviewers. I thank Professor Laurent Vilanova for kindly accepting to participate in the committee.

Moreover, I would like to sincerely thank Professor Krzysztof Ostaszewski for his kind invitation and warm reception at Illinois State University. He made me feel welcome, and my experience in the U.S. would not have been the same without him and his generosity.

My thanks also go to all the professors and staff at emlyon business school with whom I worked these years for their stimulating talks, constructive comments, and encouraging words.

A special thanks go to my parents, who always support me even though they do not always agree with my choices, my brother, who is always there for me; and my boyfriend, for his endless encouragement and patience even in the most challenging times.

I want to thank all my friends that showed me that everything is surmountable. To all of you, you should know that your support and encouragement are worth more than I can express on paper.

Abstract

This dissertation focuses on the study of decision theory, a discipline that studies how choices are made and should be made and provides some valuable tools for decision-making. This thesis can be divided into two parts.

In the first part, we study decision theory in the risk management field and identify some instruments that can be helpful for risk managers. We concentrate on the equivalence between two popular risk measures, VaR and TCE. The Swiss Solvency Test is more risk-averse than Solvency II when losses are modeled using a generalized Pareto distribution. Further, we introduce a new indicator that extends TCE to consider high-order risks, and we compare the quantiles of this indicator to the quantile of VaR and TCE. However, the extended TCE does not respect the properties to be identified as a coherent risk measure. For this reason, we introduce a new variation of the extended TCE that is a coherent risk measure. Finally, we perform an empirical study over seventeen equity indices and the 3-month T-bill rate to identify which annual high-order moment is more stable over time. We discover that third-order moments are less stable than fourth-order moments. However, partial third-order moments are more stable than partial fourth-order moments. Moreover, raw moments are more stable than standardized moments over time. This new information can help a risk manager better identify which moments he or she should rely on when making a new strategic decision.

In the second part of this thesis, we study the home bias in a dynamic consumption and portfolio choice problem when the investor is ambiguous about stock returns but learns about their distribution. The investor's preference is modeled with the generalized recursive smooth ambiguity model, and it can separate among risk aversion, ambiguity aversion, and intertemporal substitution. We implement a numerical approach that extends the existing ones to tackle the specificities of our framework. Our results show that ambiguity-averse investors participate less in financial markets and have a higher home bias than an ambiguity-neutral decision maker. However, learning as well as a dynamic model reduces the home bias. Finally, we provide comparative statics for risk aversion, ambiguity aversion, and the learning process.

Keywords

Decision theory; Risk Management; Risk Measure; High-order risk; Ambiguity; Portfolio Choice; Home-bias Puzzle.

Résumé

Cette thèse porte sur l'étude de la théorie de la décision, une discipline qui étudie comment les choix sont faits ou devraient être faits, et qui fournit des outils précieux pour la prise de décision. Cette thèse peut être divisée en deux parties.

Dans la première partie, nous étudions la théorie de la décision dans le domaine de la gestion des risques et identifions certains instruments qui peuvent être utiles aux gestionnaires de risques. Nous nous concentrons sur l'équivalence entre deux mesures de risque populaires, la VaR et la TCE. Le Swiss Solvency Test est plus averse au risque que Solvency II lorsque les pertes sont modélisées à l'aide d'une distribution de Pareto généralisée. En outre, nous introduisons un nouvel indicateur qui étend la TCE pour prendre en compte les risques d'ordre élevé, et nous comparons les quantiles de cet indicateur au quantile de la VaR et de la TCE. Cependant, le TCE étendu ne respecte pas les propriétés nécessaires pour être identifiée comme une mesure de risque cohérente. C'est pourquoi nous introduisons une nouvelle variante de la TCE étendue qui est une mesure de risque cohérente. Enfin, nous réalisons une étude empirique sur dix-sept indices boursiers et sur le taux des bons du Trésor des États-Unis à trois mois afin d'identifier quel moment annuel d'ordre élevé est le plus stable dans le temps. Nous découvrons que les moments d'ordre trois sont moins stables que les moments d'ordre quatre. Cependant, les moments partiels de troisième ordre sont plus stables que les moments partiels de quatrième ordre. En outre, les moments bruts sont plus stables que les moments standardisés dans le temps. Ces nouvelles informations peuvent aider un gestionnaire des risques à mieux identifier les moments sur lesquels il doit s'appuyer pour prendre une nouvelle décision stratégique.

Dans la deuxième partie de cette thèse, nous étudions le biais domestique dans un problème dynamique de consommation et de choix de portefeuille lorsque l'investisseur est ambigu au sujet des rentabilités boursières mais qu'il connaît leur distribution. La préférence de l'investisseur est modélisée par le modèle d'ambiguïté différentiable récursif généralisé qui peut distinguer l'aversion au risque, l'aversion à l'ambiguïté et la substitution intertemporelle. Nous mettons en œuvre une approche numérique qui étend les approches existantes afin de prendre en compte les spécificités de notre cadre. Nos résultats montrent que les investisseurs ayant une aversion pour l'ambiguïté participent moins aux marchés financiers et ont un biais domestique plus important qu'un décideur neutre à l'égard de l'ambiguïté. Cependant, l'apprentissage ainsi qu'un modèle dynamique réduisent le biais domestique. Enfin, nous fournissons des statiques comparatives pour l'aversion au risque, l'aversion à l'ambiguïté et le processus d'apprentissage.

Mots clés

Théorie de la décision ; gestion du risque ; mesure du risque ; risque d'ordre élevé ; ambiguïté ; choix de portefeuille ; Biais domestique.

Contents

Acknowledgements	iii
Abstract	v
Résumé	vii
Introduction Générale (Version française)	1
General Introduction (English version)	17
1 Equivalent Risk Indicators: VaR, TCE, and Beyond	31
1.1 The Pareto distribution	32
1.1.1 Pareto Type I Distribution	33
1.1.2 Generalized Pareto distribution	35
1.2 Risk Measures	37
1.2.1 Value-at-Risk	38
1.2.2 Tail Conditional Expectation	41
1.2.3 High-order Tail Conditional Expectation	44
1.2.4 Examples	48
1.3 Risk Measures and Regulation.	49
1.4 Relation between Risk Measures	51
1.4.1 Relation between VaR and TCE	52
1.4.2 Relation between VaR and high-order TCE	54
1.4.3 Relation between high-order TCEs	60
1.5 Conclusion	68

CONTENTS

Appendix	71
2 Coherent Extended TCE	77
2.1 Coherent Measures of Risk.	77
2.2 <i>(Not) Coherent</i> VaR	80
2.3 <i>Coherent</i> TCE.	81
2.4 <i>(Not) Coherent</i> Extended TCE	81
2.5 Weakly Coherent Extended TCE	83
2.6 Conclusion	87
Appendix	89
3 Stability of High Order Moments: a Risk Management Approach for Assessing the Risk of Insurance Companies	93
3.1 The importance of high order moments	94
3.2 Key concept	96
3.3 Data	97
3.4 Empirical Study.	100
3.4.1 Computation	100
3.4.2 Results	101
3.5 Conclusion	105
4 Home Bias and Learning in a Dynamic Portfolio Choice under Smooth Ambiguity	107
4.1 Literature Review	108
4.1.1 Home Bias	108
4.1.2 Decision Theory	112
4.1.3 Home Bias and Decision theory	117
4.2 The model	118
4.2.1 Generalized Recursive smooth ambiguity preferences	118
4.2.2 Financial Assets	122
4.2.3 Learning	122

CONTENTS

4.3	The investor's problem	127
4.4	Numerical Implementation.	129
4.4.1	Solving without learning	129
4.4.2	Approximating the value function	130
4.4.3	Solving with learning	130
4.5	Results	132
4.5.1	No Rebalancing portfolio	133
4.5.2	Solving with rebalancing	139
4.6	Conclusion	147
	Appendix	149
	General Conclusion	151
	Bibliography	153

Introduction Générale

La vie quotidienne est remplie de décisions à prendre : que manger, comment aller au travail, comment s'habiller, quoi acheter, que faire le soir, quelles émissions regarder à la télévision, que faire pendant le week-end, et ainsi de suite. Même si ces décisions font partie de la vie quotidienne et ne prennent pas trop de place dans notre esprit, tout le monde est confronté à des questions difficiles qui peuvent nécessiter un long processus de décision avant d'être résolues : changements de carrière, demandes en mariage ou décisions d'investissement importantes. Lorsque nous prenons une décision, nous essayons généralement d'améliorer notre qualité de vie. Ensuite, une décision peut être prise dans le cadre de notre rôle de gestionnaire, par exemple en tant qu'entrepreneur ou manager. Dans ce cas, les décisions peuvent conduire l'entreprise à la faillite ou au succès. Une seule mauvaise décision peut changer l'orientation de l'entreprise. En outre, certaines personnes occupant une position particulière prennent des décisions qui peuvent influencer un pays ou le monde entier. Elles peuvent prendre des décisions seules, comme les dictateurs, ou en groupe, comme les gouvernements, les banques centrales, les conseils, etc. Les gens ont tendance à être indécis si la décision peut changer leur vie, ce qui peut également conduire à un processus de décision plus long et à des remises en question avant le choix final.

Un instrument couramment utilisé pour prendre des décisions est la "liste des avantages et des inconvénients", dans laquelle l'agent identifie tous les résultats positifs et négatifs de cette décision. Cependant, cette liste n'est pas toujours utile, car il n'est pas toujours facile de comparer les avantages et les inconvénients de différentes décisions. Lorsque les choix sont importants, il est nécessaire d'identifier les instruments qui nous permettent de prendre la meilleure décision compte tenu des informations disponibles et des convictions du décideur. Ce type de décision est appelé choix rationnel ; toutefois, le décideur ne découvre si cette décision est la meilleure, c'est-à-dire le bon choix, que lorsque les conséquences de cette décision se produisent.

L'étude du processus de prise de décision et de la manière de prendre des décisions rationnelles relève d'une discipline appelée théorie du choix ou théorie de la décision.

Théorie de la Décision

La théorie de la décision analyse le processus de prise de décision : comment les choix sont faits ou devraient être faits sur la base de l'attribution de probabilités et de valeurs, appelées utilité, à chaque résultat. Il s'agit d'un domaine interdisciplinaire auquel participent des psychologues, des économistes, des philosophes, des statisticiens et des politologues. La théorie de la décision peut être divisée en deux branches principales : la théorie descriptive de la décision et la théorie normative de la décision.

La théorie descriptive de la décision tente d'expliquer comment les gens prennent des décisions dans des situations réelles. Elle cherche à comprendre les processus cognitifs et les biais qui affectent la prise de décision humaine et comment les gens s'écartent du modèle idéal de prise de décision rationnelle. Cette discipline repose sur des études empiriques.

La théorie de la décision normative guide la prise de décisions optimales. L'objectif est de fournir un ensemble de lignes directrices ou de principes qui peuvent aider l'agent à faire des choix optimaux en fonction de ses préférences, de ses croyances et de ses valeurs. Il est possible de modéliser le comportement des individus comme s'ils respectaient certains axiomes et suivaient un modèle spécifique.

Il est essentiel de noter qu'un choix rationnel est une décision qui respecte les convictions du décideur et qui est considérée comme optimale, compte tenu de toutes les informations disponibles. En revanche, une bonne décision est une décision qui donne les meilleurs résultats, et elle n'est connue qu'après la prise de décision. La théorie de la décision se concentre sur les choix rationnels. En outre, les théories normatives et descriptives de la décision partagent certains éléments communs, car elles s'accordent sur le fait que les décisions sont prises en fonction des croyances et des désirs du décideur.

Formalisation du problème des agents

Selon Peterson (2017), il existe trois niveaux d'abstraction lorsqu'il s'agit de faire un choix :

- le problème de décision que l'agent doit résoudre exactement comme il est proposé au décideur ;
- une formalisation du problème de décision où l'agent identifie toutes les informations nécessaires pour prendre une décision ;
- Une visualisation de la formalisation est généralement effectuée dans une matrice de décision ou un arbre de décision s'il s'agit de choix séquentiels.

Introduction Générale

L'étape délicate consiste à identifier toutes les informations utiles pour prendre une décision éclairée. Ces informations peuvent être divisées en états, résultats et actes.

Un *état* peut être défini comme l'un des scénarios possibles qui pourraient se produire dans le futur, et l'état qui se produit dans le monde réel est révélé une fois que le choix est fait. L'agent doit soigneusement choisir uniquement les états qui sont causalement indépendants et pertinents pour son choix. Par exemple, si le problème de décision est "Dois-je prendre un parapluie aujourd'hui?", les états pourraient être (i) Aujourd'hui, il pleuvra, et (ii) Aujourd'hui, il ne pleuvra pas. Les états (iii) J'ai pris la bonne décision et (iv) J'ai pris la mauvaise décision ne sont pas indépendants du choix, ce qui signifie qu'ils ne peuvent pas être considérés comme des états du monde.

Le décideur prend sa décision en fonction des *résultats* qu'il recevra dans chaque état du monde. Le décideur doit être en mesure de classer la probabilité des différents résultats. Ce classement est entièrement subjectif et doit refléter son attitude à l'égard des résultats. Afin de mesurer la valeur d'un résultat, il doit attribuer une valeur à chaque résultat à l'aide d'une *fonction d'utilité*. Le décideur peut utiliser sa fonction d'utilité pour ordonner ses préférences sur un ensemble de choix. Toutefois, l'ordre des préférences n'est pas comparable d'un agent à l'autre. Kreps (1988) indique qu'une fonction d'utilité doit avoir les propriétés suivantes : (i) *complétude*, ce qui signifie que l'agent doit toujours être en mesure de classer les différents résultats, (ii) *transitivité* si un agent préfère A à B et B à C, il préférera A à C, et (iii) *séparabilité* qui est un concept technique. Nous nous référons à Debreu (1954) pour définir les propriétés des fonctions d'utilité continues. L'identification d'une fonction d'utilité spécifique (et éventuellement de nouvelles propriétés à respecter) dépend des caractéristiques de l'agent et du modèle adopté, comme nous le verrons dans les sections suivantes.

Les *actes* sont toutes les actions que le décideur peut faire, et il doit décider laquelle il va adopter. Une fois l'acte choisi et l'état révélé, il sait quel résultat a été obtenu. Les actes sont considérés comme mutuellement exclusifs, ce qui signifie qu'un décideur rationnel ne peut choisir qu'un seul acte.

Prenons l'exemple suivant. Laura paie un loyer à Paris et aimerait acheter une maison l'année prochaine. Cependant, elle ne sait pas si son entreprise la transférera dans les dix prochaines années. Supposons que, dans ce monde simpliste, il ne peut y avoir que trois états :

- Laura n'est pas transférée,
- Laura est transférée dans une nouvelle ville proche de Paris,
- Laura est transférée dans une nouvelle ville éloignée de Paris.

Laura a deux possibilités ou actes : (a) continuer à louer une maison ou (b) acheter une maison. Si elle achète la maison et reste à Paris, elle sera ravie. Cependant, si elle achète la maison et qu'elle est transférée loin, elle devra vendre la maison qu'elle vient d'acheter ou la louer à quelqu'un d'autre. En revanche, si elle continue à louer un logement, il lui sera plus facile d'être transférée ailleurs, mais si elle reste à Paris, elle continuera à payer un loyer sans être propriétaire d'un logement. Le tableau 1 identifie la matrice de décision qui met en évidence les états et les actes pour ce problème. En outre, le résultat est le bonheur de Laura, identifié par un nombre de 0 à 10, l'utilité attribuée à chaque résultat.

États \ Actes	Aucun transfert	Transfert vers une ville proche	Transfert vers une ville loin
Louer à Paris	1	5	8
Acheter à Paris	10	5	0

Table 1: *Matrice de décision pour le problème de Laura*

Nous avons vu comment un problème peut être réécrit pour identifier les états, les actes, les résultats et les utilités associées. En utilisant la théorie de la décision, nous pouvons donner un ordre à tous les actes que l'agent peut choisir. Pour ce faire, l'agent doit identifier la probabilité à attribuer à chaque état. Si la probabilité est identifiée comme objective ou si elle est donnée comme connue du décideur, la décision est définie comme étant sous risque. En revanche, si la probabilité n'est pas objectivement connue, la décision est dans l'incertitude.

Décision sous Risque

Les décisions sous risque définissent des situations dont les probabilités objectives sont connues ou données au décideur. Il s'agit généralement de situations telles que la roulette dans les casinos, les paris à pile ou face, mais aussi de situations où la probabilité est donnée, par exemple, décider de pratiquer une opération en connaissant le pourcentage de personnes qui ont survécu ou de pratiquer des sports extrêmes en connaissant le pourcentage de personnes qui ont subi des blessures après avoir pratiqué le sport.

Von Neumann and Morgenstern (1944) identifient une théorie dans laquelle le décideur rationnel devrait décider de l'acte qui maximise l'utilité attendue. Il y a n états du monde, et chaque acte donne un résultat, x_i , pour l'état i dont la probabilité est connue, p_{x_i} , la fonction d'utilité espérée est calculée comme suit :

$$\mathbb{E}[u(X)] = \sum_{i=1}^n u(x_i) \cdot p_{x_i},$$

Introduction Générale

où $u(\cdot)$ est défini comme une fonction d'utilité de Von Neumann-Morgenstern qui doit respecter certains axiomes spécifiques, mais qui n'identifient pas la fonction d'utilité à utiliser. Cette théorie est appelée utilité espérée objective et le décideur devant choisir l'acte X qui maximise son utilité espérée.

La partie la plus difficile des décisions sous risque est d'identifier la fonction d'utilité qui décrit le mieux les préférences du décideur car la probabilité est donnée. Un agent peut identifier sa fonction d'utilité en déterminant ses préférences à l'aide de loteries simples et en dérivant point par point sa fonction d'utilité. Nous renvoyons le lecteur à Gilboa (2010) pour suivre une procédure étape par étape afin de dériver sa fonction d'utilité. Une liste de fonctions d'utilité ayant différentes propriétés liées à l'assurance et à l'argent se trouve dans Gerber and Pafumi (1998).

Il est généralement possible d'identifier que $u(X)$ est une fonction croissante de x car plus d'argent est toujours perçu comme meilleur que moins d'argent. Cependant, la concavité de la fonction dépend de la préférence du décideur, qui peut avoir une aversion pour le risque ou rechercher le risque. Afin de mieux comprendre ce concept, nous donnons un exemple. Le décideur doit choisir entre ces deux loteries :

Loterie A donne 1 000 euros avec une probabilité de 0,5 et 0 sinon ;

Loterie B donne 500 d'euros à coup sûr.

La valeur attendue de la loterie A est de $1,000 \cdot 0,5 + 0 \cdot 0,5 = 500 \text{ €}$. On dit qu'un décideur a une aversion au risque s'il préfère la loterie B à la loterie A, car il préfère la valeur attendue de la loterie à la loterie elle-même. D'autre part, un décideur est défini comme aimant le risque ou recherchant le risque s'il préfère à coup sûr la loterie aux 500 euros. Un agent averse au risque a une fonction d'utilité concave, tandis qu'un agent amateur de risque a une fonction d'utilité convexe. En économie et en finance, on suppose que les agents ont une aversion pour le risque, même si leur fonction d'utilité est généralement concave lorsqu'il s'agit de faibles gains et convexe lorsqu'il s'agit de gains élevés (Peterson, 2017).

Revenons au problème de décision de Laura, où nous supposons qu'elle a demandé au service des ressources humaines de son entreprise quelle était la probabilité qu'elle soit mutée. Le service lui répond qu'il y a 50% de chances qu'elle soit mutée dans l'arrière-pays parisien, 30% de chances qu'elle soit mutée loin de Paris et 20% de chances qu'elle reste dans le même bureau. Afin de décider quelle action elle doit choisir, elle calcule son utilité espérée objective :

$$\begin{aligned}\mathbb{E}[u(\text{Louer})] &= 1 \cdot 0.2 + 5 \cdot 0.5 + 8 \cdot 0.3 = 5.1, \\ \mathbb{E}[u(\text{Acheter})] &= 10 \cdot 0.2 + 5 \cdot 0.5 + 0 \cdot 0.3 = 4.5.\end{aligned}$$

L'utilité de la location d'une maison à Paris est plus élevée que celle de l'achat d'une maison, ce qui conduit Laura à continuer à louer la maison pour le moment. Ce choix pourrait changer si le service des ressources humaines lui donnait des probabilités différentes ou si son bonheur, c'est-à-dire la valeur d'utilité, changeait.

Décision dans l'incertitude

L'incertitude, également appelée ambiguïté, est un terme technique qui désigne les cas où le décideur connaît les actes et les résultats. Cependant, il ne peut pas attribuer de probabilités aux états. Par exemple, il est impossible d'estimer, de déduire en utilisant des données existantes ou de demander à quelqu'un la probabilité d'une guerre dans un pays donné au cours des dix prochaines années, la probabilité qu'un nouveau parcours professionnel soit couronné de succès ou d'échec, etc.

Dans un premier temps, les économistes ont tenté de traiter ce type de problèmes en essayant de ne pas utiliser des distributions de probabilité. Le choix du décideur devrait être fait uniquement en considérant les meilleurs ou les pires résultats de chaque acte sans impliquer la distribution de probabilité. Nous suivons Peterson (2017) pour identifier certains des différents principes utilisés.

Le principe de dominance stipule qu'un décideur ne doit jamais choisir un acte dont les résultats sont les pires, quel que soit l'état réel du monde. En d'autres termes, si un acte spécifique donne toujours un résultat pire qu'un autre, le premier aurait dû être exclu du problème de décision.

Le principe Maxmin selon lequel le décideur doit choisir l'acte qui donne les résultats les moins mauvais indépendamment de l'état dans lequel il se trouve. Dans le cas du problème de décision de Laura, Laura devrait décider de louer à Paris car le pire résultat donne une utilité égale à 1, alors que le pire résultat dans le cas de l'achat d'une maison a une utilité nulle.

La règle de Maximax selon laquelle le décideur doit choisir l'acte qui a les meilleurs résultats possibles par rapport aux meilleurs résultats possibles pour chaque acte. Dans le cas du problème décisionnel de Laura, celle-ci devrait décider d'acheter une maison car le meilleur résultat donne une utilité de 10, ce qui est plus élevé que le meilleur résultat de la location d'une maison.

La règle de l'optimiste et du pessimiste qui considère les meilleurs et les pires résultats possibles de chaque alternative et choisit une alternative

Introduction Générale

en fonction de son degré d'optimisme ou de pessimisme, et il devrait s'agir d'une moyenne pondérée de ces deux utilités où le poids dépend du degré d'optimisme ou de pessimisme.

Toutes ces méthodes ne se concentrent que sur les meilleurs ou les pires résultats, mais elles ne prennent pas en compte les résultats intermédiaires. Pour prendre une décision éclairée, tous les résultats et toutes les utilités doivent être pris en compte afin de faire un choix rationnel qui optimise la fonction d'utilité de l'agent.

Selon Savage (1972), une personne devrait être en mesure d'identifier une probabilité subjective qui garantit que ses croyances sont cohérentes au niveau interne. Ces probabilités subjectives doivent respecter certains axiomes pour pouvoir être utilisées dans les processus de prise de décision. Gilboa (2010) fournit une procédure simple, étape par étape, pour dériver une probabilité subjective en commençant par les loteries accessibles. Cette procédure est similaire à celle qui permet de créer sa fonction d'utilité de manière ponctuelle. Cependant, les décideurs ne peuvent pas toujours attribuer une probabilité spécifique ; par exemple, à l'état "demain il pleuvra", dois-je attribuer 39% ou 41% de chance ?

En contradiction avec cette théorie et parce que les gens ne peuvent pas toujours avoir des probabilités subjectives cohérentes, de nombreux paradoxes différents ont été proposés. L'un des paradoxes les plus populaires est le **Paradoxe d'Ellsberg (1961)**, présenté ci-dessous, dans lequel les gens préfèrent les paris avec une probabilité connue aux paris avec une probabilité inconnue. Ce comportement est qualifié d'aversion à l'incertitude, violant ainsi l'un des axiomes de Savage.

Supposons que deux urnes contiennent 100 boules. L'agent sait que l'urne A contient 50 boules noires et 50 boules blanches, alors qu'il ne connaît pas la proportion de boules noires et blanches dans l'urne B. Le décideur a alors le choix entre les paris suivants :

Loterie 1 Recevoir 100 euros si une boule noire est tirée de l'urne A.

Loterie 2 Recevez 100 euros si une boule noire est tirée de l'urne B.

À ce stade, le décideur peut raisonner de deux manières. Il peut préférer le pari 1 car il connaît la proportion de boules noires dans l'urne A, ou il peut préférer le pari 2 car il pense que la probabilité pourrait être plus élevée. Comme le raisonnement ne change pas, nous supposons que le décideur préfère l'urne connue, c'est-à-dire la loterie 1. Il a été prouvé empiriquement que les gens ont tendance à préférer la loterie 1.

Ensuite, une deuxième paire de paris leur est présentée.

Loterie 3 Recevez 100 euros si une boule blanche est tirée de l'urne A.

Loterie 4 Recevez 100 euros si une boule blanche est tirée de l'urne B.

Lorsqu'elle est confrontée à une nouvelle paire de loteries, une personne qui préfère le pari 1 au pari 2 préférera le pari 3 au pari 4.

Ceci est incompatible avec la théorie de l'utilité espérée car un décideur qui préfère le pari 1 au pari 2 pense que la probabilité d'avoir des boules noires dans l'urne A est plus élevée que la probabilité d'avoir des boules blanches dans l'urne B, c'est-à-dire que $P(\text{noir dans l'urne A}) = 0,5 > P(\text{noir dans l'urne B})$. Cependant, comme la même personne préfère le pari 3 au pari 4, elle pense que la probabilité d'avoir des boules blanches dans l'urne A est plus élevée que la probabilité d'avoir des boules blanches dans l'urne B, soit $P(\text{blanc dans l'urne A}) = 0,5 > P(\text{blanc dans l'urne B})$. Ces deux probabilités sont incohérentes l'une par rapport à l'autre, car la somme de $P(\text{noir dans l'urne B}) + P(\text{blanc dans l'urne B})$ ne serait pas égale à 1. Comme le décideur ne connaît pas la probabilité de chaque boule dans l'urne B, l'agent a tendance à préférer l'urne connue. C'est ce qu'on appelle l'aversion à l'ambiguïté, qui démontre qu'il est impossible d'utiliser le modèle de Savage pour trouver la décision optimale dans ce scénario.

Au cours des dernières décennies, les modèles de décision qui prennent en compte l'ambiguïté et l'aversion pour l'ambiguïté afin de résoudre le paradoxe d'Ellsberg ont suscité un grand intérêt théorique, comme Schmeidler (1989), Gilboa and Schmeidler (1989), Klibanoff et al. (2005), parmi d'autres. Une discussion plus détaillée est fournie ci-dessous.

Biais domestique et théorie de la décision

La théorie de la décision a été largement appliquée aux problèmes d'allocation optimale de portefeuille afin d'essayer d'expliquer certaines énigmes liées à l'allocation d'actifs. La théorie financière standard prévoit que les investisseurs devraient détenir un portefeuille diversifié d'actions dans le monde entier, en supposant que les capitaux sont totalement mobiles à travers les frontières. Les actions étrangères offrant de grandes possibilités de diversification, l'abaissement des barrières au commerce international d'actifs financiers aurait dû conduire les investisseurs du monde entier à rééquilibrer leur portefeuille des actifs nationaux vers les actifs étrangers (Coerdacier and Rey, 2013). Le fait que les gens détiennent des portefeuilles sous-diversifiés peut être lié au *biais domestique*. La préférence pour les actifs domestiques est un comportement d'investissement dans lequel les investisseurs ont tendance à surpondérer le marché de leur pays d'origine par rapport à sa part dans le portefeuille global du marché.

Coerdacier and Rey (2013) montrent que la préférence pour les actifs domestiques a diminué au cours de la période 1988-2008 dans les pays développés

Introduction Générale

à travers les régions du monde, mais qu'elle reste élevée dans la plupart des pays. En outre, les marchés émergents ont des portefeuilles d'actions moins diversifiés que les pays développés et n'affichent pas de tendance claire à la baisse du biais domestique. Le tableau 2, extrait de Coeurdacier and Rey (2013), montre le pourcentage d'actifs nationaux qu'un investisseur devrait avoir dans son portefeuille et le pourcentage qu'il a, ce qui démontre que le biais domestique existe toujours sur le marché.

Pays source	Marché national en % de la capitalisation boursière mondiale	Part de portefeuille en dans le marché national en %
Australie	1.8	76.1
Zone Euro	13.5	56.7
Suède	0.7	43.6
Suisse	2.3	50.9
États-Unis	32.6	77.2

Table 2: *Biais domestique pour les actions en 2008 pour une sélection de pays.*
Source : Coeurdacier and Rey (2013)

Le biais domestique a été confirmé par différentes études empiriques telles que Fidora et al. (2007), Lippi (2016), Lütje and Menkhoff (2007), Mishra (2008), Lin and Viswanathan (2016), parmi d'autres.

De nombreux chercheurs ont tenté de déterminer les raisons pour lesquelles les investisseurs ont une préférence pour le marché domestique, et ces raisons peuvent être classées en trois catégories principales :

- Des raisons institutionnelles telles que les taxes, les coûts de transaction et les barrières ;
- Les asymétries d'information entre les investisseurs ;
- Les raisons comportementales telles que la familiarité, le patriotisme, l'optimisme, l'aversion à l'ambiguïté, etc.

Le lecteur est invité à consulter Gaar et al. (2020) pour une analyse plus exhaustive de la littérature.

Différentes études ont examiné la répartition optimale des actifs du portefeuille afin d'étudier les éléments susceptibles de générer un biais domestique.

De nombreux modèles de théorie de la décision sont utilisés pour étudier les biais cognitifs d'un investisseur lors de la répartition de son portefeuille. Les modèles de théorie de la décision peuvent s'inscrire dans un cadre statique, comme Von Neumann and Morgenstern (1944) qui utilisent des distributions de probabilité objectives, Schmeidler (1989) dont la mesure de probabilité n'est pas nécessairement additive, Gilboa and Schmeidler (1989) qui maximisent

l'utilité attendue dans le pire des cas sur un ensemble de probabilités, Klibanoff et al. (2005) qui identifient un modèle de préférence pour l'ambiguïté lisse. Ces modèles ont également été développés dans un cadre dynamique, comme Epstein and Schneider (2003) qui ont étendu le modèle de Gilboa and Schmeidler (1989), Klibanoff et al. (2005) et Hayashi and Miao (2011) qui ont étendu le modèle de Klibanoff et al. (2005). Il ne s'agit là que de quelques exemples des différents modèles de théorie de la décision développés pour étudier l'allocation optimale des actifs.

Les chercheurs ont appliqué certains de ces modèles à l'étude du biais domestique en analysant l'allocation optimale des actifs lorsque deux actions risquées sont présentes sur le marché. Nous n'évoquerons ici que quelques-unes des études récentes sur ce sujet. Guidolin and Liu (2016) étudient le biais domestique à l'aide d'un modèle qui prend en compte les croyances incertaines sur un modèle d'évaluation des actifs, l'incertitude des paramètres et l'aversion à l'ambiguïté. Ces auteurs ont utilisé un modèle d'ambiguïté statique et lisse proposé par Klibanoff et al. (2005) pour caractériser les attitudes d'ambiguïté et inclure un cadre bayésien. En utilisant leur approche, ils constatent que l'aversion à l'ambiguïté peut conduire à un biais domestique solide dans la détention d'actions, indépendamment de la croyance d'un investisseur dans le CAPM domestique. Peijnenburg (2018) étudie l'évolution du biais domestique dans un cadre dynamique à l'aide d'un modèle d'utilité attendue max-min. Elle constate également que l'ambiguïté contribue à expliquer la sous-diversification et le biais domestique. Cependant, le modèle adopté ne sépare pas l'ambiguïté et les attitudes ambiguës. Un modèle d'ambiguïté lisse dans un cadre dynamique est également étudié dans Yu et al. (2022), où ils se concentrent uniquement sur l'analyse de sensibilité des paramètres de risque et d'ambiguïté. Ils ont constaté que lorsqu'un individu devient plus averse à l'ambiguïté, il consomme moins, achète plus d'assurance-vie et investit moins dans des actifs risqués.

Gestion des risques et la théorie de la décision

Le risque et la prise de décision sont deux facteurs interdépendants de la gestion organisationnelle, et ils sont tous deux liés à la présence d'incertitude dans l'organisation (Lu et al., 2012). Un gestionnaire de risque doit utiliser les instruments fournis par la théorie de la décision pour faire des choix qui peuvent rendre l'organisation rentable et éviter la faillite.

La gestion des risques est définie comme une série d'activités visant à diriger et à contrôler une organisation en tenant compte des risques. Il est nécessaire d'identifier et d'évaluer les risques potentiels, puis de les gérer en prenant les décisions appropriées. Le processus de gestion des risques peut être divisé en quatre étapes différentes :

Introduction Générale

1. Identification des risques à l'aide d'un brainstorming, d'une structure, d'entretiens, de questionnaires et d'enquêtes, d'une analyse des données relatives aux pertes, d'une analyse des hypothèses, etc ;
2. Évaluer la vulnérabilité des biens essentiels face à des menaces spécifiques ;
3. Analyse et mesure du risque ;
4. La gestion des risques, qui peut consister à (i) éviter le risque, (ii) conserver le risque par le biais d'un mécanisme tel que l'allocation de capital-risque, (iii) atténuer le risque en réduisant l'exposition, la fréquence et la gravité, ou (iv) transférer le risque à une tierce partie.

Il est difficile de prendre des décisions appropriées dans des scénarios réels en raison des risques et des incertitudes. Un gestionnaire de risques doit être capable d'analyser les problèmes en utilisant des critères de décision pour identifier les résultats de chaque scénario, l'utilité de chaque résultat et attribuer la probabilité de chaque résultat. Borgonovo et al. (2018) étudient le lien entre l'analyse du risque et la théorie de la décision dans le but de permettre une intégration étroite de la partie gestion du risque de l'analyse de la décision au résultat de l'évaluation du risque. La plupart des chercheurs définissent l'analyse de risque comme la quantification du risque, l'identification des options de gestion du risque et la communication des résultats au gestionnaire du risque, c'est-à-dire les étapes 1 à 3 du processus de gestion du risque. Même si l'analyse de risque et l'analyse de décision présentent certaines similitudes et sont souvent complémentaires, l'analyse de risque est souvent réalisée en l'absence d'un décideur connu pour évaluer les probabilités et choisir les options de gestion du risque sur la base de ses préférences et de son attitude face au risque (Paté-Cornell and Dillon, 2006). Le résultat de l'analyse des risques est ensuite communiqué au gestionnaire des risques, qui décide de la manière de gérer le risque global (étape 4 du processus de gestion des risques). Le gestionnaire de risque doit choisir l'option qui maximise sa fonction d'utilité en fonction de ses préférences et de ses caractéristiques.

Borgonovo et al. (2018) relie la gestion du risque opérationnel aux fondements théoriques de la théorie de la décision. Ils montrent un lien direct entre l'analyse des risques et la théorie de la décision lorsque l'on utilise des probabilités objectives ou subjectives. Une reformulation différente est nécessaire en cas d'incertitude. De manière générale, ils ont prouvé que les gestionnaires de risques peuvent accéder à la boîte à outils de la théorie de la décision pour définir leurs stratégies.

Régulateurs et mesures des risques

La gestion des risques a été jugée nécessaire, de même que l'introduction d'une réglementation sur le marché dans les années 1990 pour protéger un pays ou un système économique des risques trop élevés pris par les institutions financières pour obtenir des profits importants. Les régulateurs sont des comités qui définissent les différentes règles qu'une institution financière doit respecter pour entrer et rester sur le marché. Pour éviter la faillite, les régulateurs définissent généralement une mesure du risque, c'est-à-dire une exigence de capital que l'entreprise doit conserver pour faire face à des pertes importantes.

En 1990, J.P. Morgan a mis au point la valeur à risque (VaR), un indicateur de risque global qui mesure la pire perte attendue à un horizon donné dans des conditions de marché normales et à un niveau de confiance donné. Supposons, par exemple, que la VaR de 95% à un an soit égale à 2 millions d'euros. Cela signifie que, dans des conditions de marché normales, il y a 5% de chances que les pertes soient supérieures à 2 millions d'euros pendant l'année à venir. Une analyse plus détaillée de la VaR peut être trouvée dans Klugman et al. (2012), Jorion (2007), Linsmeier and Pearson (2000), parmi d'autres. Toutefois, le principal problème de la VaR est qu'elle ne permet pas de quantifier la part de risque dans les queues de la distribution.

Une autre mesure de risque célèbre est l'espérance conditionnelle de queue (Tail Conditional Expectation - TCE), également appelée Expected Shortfall, Conditional Tail Expectation ou Tail-VaR, définie comme la moyenne de toutes les pertes supérieures à un seuil spécifique, généralement la VaR au quantile p . Par construction, la TCE sera toujours au moins égal à la VaR calculée pour le même quantile.

Même si tous les régulateurs ont le même objectif, ils ne sont pas d'accord sur les mesures de risque qu'ils devraient utiliser, et tous n'utilisent pas les deux mesures de risque mises en évidence précédemment. La réglementation canadienne des assurances et les États-Unis préfèrent les modèles qui exigent une plus grande implication des régulateurs dans la supervision. En revanche, Solvency II adopte une VaR de 99,5% calculée sur un horizon d'un an, et le Swiss Solvency Test utilise une TCE de 99% calculé sur le même horizon. On trouvera une comparaison des différents régimes réglementaires d'assurance dans Comité Européen des Assurances and Mercer Oliver Wyman Limited (2005), ou une comparaison entre un régime d'assurance (Solvency II) et un régime bancaire (Basel III) dans Gatzert and Wesker (2011). Il n'existe pas de mesure générale du risque qui soit adoptée dans le monde entier ; les régulateurs décident des mesures du risque à adopter en fonction de leur niveau d'aversion pour le risque et en suivant un certain processus politique.

Les deux mesures de risque adoptées par Solvency II et le Swiss Solvency Test sont basées sur la distribution des pertes. Elles utilisent des quanti-

tés statistiques pour décrire la distribution conditionnelle ou inconditionnelle des pertes du portefeuille sur un horizon prédéterminé T à un intervalle de confiance spécifique p .

Dans cette thèse, nous supposons que les pertes suivent une distribution de Pareto de type I ou une distribution de Pareto généralisée qui est une distribution qui peut considérer un grand nombre de valeurs extrêmes et dont la fonction de densité de probabilité présente des queues épaisses. Vilfredo Pareto a introduit la distribution de Pareto pour la première fois en 1895 pour étudier les revenus. Il a identifié le “principe de Pareto”, qui stipule qu’un plus petit pourcentage de personnes dans la société possède une grande part de la richesse de la société. La distribution de Pareto de type I est caractérisée par α , qui identifie le risque sur le marché, et x_0 , qui est la valeur minimale possible de la variable aléatoire étudiée. D’autre part, la distribution de Pareto généralisée a trois paramètres : ξ , qui identifie le degré de risque sur le marché et peut être interprété comme l’inverse de α , le paramètre de localisation μ et le paramètre d’échelle σ . Nous modélisons les sinistres à l’aide de cette distribution spécifique, car ils sont cohérents avec ce qui est observé pour les sinistres dont les queues de distribution sont importantes distribuées, et cela nous permet d’obtenir des résultats lisibles.

Mesures cohérentes de risque

Artzner et al. (1999) identifient une série de propriétés qu’une mesure de risque devrait avoir pour être définie comme cohérente. Ils précisent qu’une mesure de risque cohérente doit satisfaire aux propriétés de monotonie, de sous-additivité, d’homogénéité positive et d’invariance de translation.

La propriété de *monotonie* signifie que si les pertes d’une variable aléatoire sont toujours plus élevées que les pertes d’une autre variable aléatoire dans chaque état du monde, il est logique que la mesure du risque de la première soit supérieure à la mesure du risque de la seconde.

La *sous-additivité* reflète le fait que la combinaison des risques conduit à une diversification et, par conséquent, à une réduction du risque global total. Elle stipule que la mesure du risque liée à un portefeuille agrégé doit être égale ou inférieure à la somme des mesures du risque calculées pour chaque élément.

La propriété d’*homogénéité positive* spécifie que la mesure du risque d’un multiple constant de la perte inattendue doit être le multiple constant de la mesure du risque, tandis que *invariance de translation* dit que la mesure du risque de la combinaison d’une perte inattendue et d’une perte fixe doit être la mesure du risque de la perte inattendue plus cette perte fixe.

Il a été prouvé par Artzner et al. (1999) que la Value-at-Risk n’est pas une mesure de risque cohérente, alors que la TCE est une mesure de risque cohérente. Cependant, il existe différentes critiques sur les mérites des quatre

propriétés, voir par exemple Rau-Bredow (2019), Daniélsson et al. (2005), Cont et al. (2010), parmi d'autres.

L'importance des moments d'ordre élevé

Un instrument utile que le gestionnaire de risque peut utiliser pour évaluer le risque de marché est le calcul des moments d'ordre élevé de la distribution des rendements qui peut capturer l'asymétrie et le caractère épais de la queue des rendements. En outre, plus les moments sont connus, meilleure est l'estimation de la distribution qui peut aider le décideur à identifier la stratégie optimale pour éviter la faillite. Par exemple, selon Dai et al. (2021), la dynamique des moments d'ordre supérieur est considérée comme une mesure du niveau de risque partagé entre les classes d'actifs. En outre, De Clerk and Savel'ev (2022) indique que nous pouvons utiliser l'analyse des moments d'ordre supérieur comme outil pour indiquer les changements dans le risque d'un actif financier.

De nombreuses études soutiennent qu'en plus de la moyenne et de la volatilité, les moments d'ordre supérieur peuvent être inclus comme facteurs de risque dans l'évaluation des actifs, comme Sihem and Slaheddine (2014), Jurczenko and Maillet (2006), Ahmed and Al Mafrachi (2021), Chen et al. (2021), parmi d'autres.

Il existe également une vaste littérature qui révèle que les moments de distribution d'ordre élevé sont utilisés pour prédire les rendements du marché, comme Jondeau et al. (2019), Doan et al. (2010), Amaya et al. (2015), Mei et al. (2017), parmi d'autres.

Cette série d'articles montre que les moments d'ordre élevé sont un outil utile pour le gestionnaire de risque afin d'identifier la stratégie la plus efficace.

Structure de la thèse

Cette thèse se concentre sur l'étude du processus de prise de décision et fournit quelques outils qui peuvent aider à la prise de décision dans le domaine de la gestion du risque ainsi que l'étude du biais cognitif des investisseurs en utilisant les modèles adoptés dans les décisions sous incertitude. Cette thèse peut être divisée en deux parties : la première se concentre sur la gestion du risque et la décision sous risque, et la seconde concerne les décisions sous incertitude et les investisseurs.

La première partie est composée des chapitres 1, 2, et 3. Elle traite de la théorie de la décision du point de vue de la gestion des risques.

Le chapitre 1 étudie l'équivalence entre différentes mesures de risque. De nombreuses recherches se concentrent sur les mérites de la VaR et de la TCE, les deux indicateurs de risque les plus classiques utilisés par les institutions

financières. Cependant, nous avons décidé d'axer notre recherche sur l'équivalence entre ces indicateurs afin d'identifier une fonction d'utilité implicite qui nous aide à comprendre comment les régulateurs choisissent tel ou tel indicateur de risque pour les calculs de solvabilité. Nous savons que les régulateurs peuvent prendre une décision à l'issue d'un processus politique. Cependant, rendre explicite le fonctionnement implicite des réglementations actuelles devrait être une contribution précieuse pour garantir que les réglementations fonctionnent de manière efficace et efficiente. En outre, la TCE n'est peut-être pas l'indicateur le plus précis pour tenir compte de la nature des queues de distribution de probabilité. C'est pourquoi nous introduisons un nouvel indicateur de risque qui étend la TCE, appelée TCE d'ordre supérieur ou TCE étendue, qui peut prendre en compte des risques d'ordre supérieur. Notre étude part du principe que les demandes d'indemnisation suivent un cadre de Pareto simple, puis un cadre de Pareto généralisé. Nous examinons également les résultats d'équivalence entre les quantiles des TCEs d'ordre supérieur. Ce chapitre est basé sur l'article publié par Faroni et al. (2022).

Ensuite, nous identifions les quatre propriétés qu'une mesure de risque devrait avoir pour être définie comme une mesure de risque cohérente dans le chapitre 2. Ce concept a eu une grande influence sur l'évolution des mesures de risque et sur la définition de nouveaux instruments de gestion du risque. Nous expliquons que la VaR n'est pas une mesure de risque cohérente, alors que la TCE l'est. Nous mettons également en évidence certaines critiques de ce concept. En outre, nous expliquons que la TCE étendue présentée dans le chapitre précédent n'est pas une mesure de risque cohérente. Nous présentons également une variante de la TCE étendue qui peut respecter les quatre propriétés identifiées par Artzner et al. (1999).

Dans le chapitre 3, nous réalisons une étude empirique sur la stabilité des moments annuels d'ordre élevé des indices d'actions entre les marchés. Nous étendons notre étude aux moments annuels partiels d'ordre élevé, c'est-à-dire aux moments calculés en utilisant uniquement les valeurs supérieures ou inférieures à un seuil spécifié. Nous utilisons dix-sept indices boursiers pour prendre en compte les régions les plus représentatives du monde, ainsi que le taux des bons du Trésor à trois mois, et nous collectons les prix de clôture quotidiens de 2000 à 2022. Cette étude vise à aider les gestionnaires de risques à identifier le moment le plus stable dans le temps, ce qui les conduira à une évaluation plus fiable des risques futurs du marché.

La deuxième partie de cette thèse comprend le chapitre 4 dans lequel nous utilisons des modèles de décision en situation d'incertitude. Nous étudions l'allocation optimale dynamique des actifs et la consommation optimale que l'agent devrait adopter lorsqu'il y a deux types d'actions sur le marché : les actions locales et les actions étrangères. Nous supposons que les deux actions

sont identiques sur le marché. Cependant, l'agent les perçoit différemment. Il pense mieux connaître l'action locale et est plus incertain quant au rendement réel de l'actif étranger. L'objectif est d'étudier le biais domestique et son évolution dans le temps lorsque nous supposons que le décideur est à la fois averse au risque et à l'ambiguïté. Nous utilisons le modèle d'ambiguïté différentiable récursif généralisé axiomatisé par Hayashi and Miao (2011) qui peut distinguer l'aversion au risque, l'aversion à l'ambiguïté et la substitution intertemporelle. Cette séparabilité nous permet de mieux comprendre quels sont les paramètres qui déterminent le biais domestique de l'agent. Nous réalisons une implémentation numérique qui inclut l'ambiguïté sur le marché et le fait que le décideur apprend au fil du temps chaque année à réduire son incertitude. Notre implémentation est explicitement calculée pour deux actions risquées et ambiguës sur le marché. Cependant, elle peut être généralisée à n actifs risqués et ambigus.

General Introduction

Everyday life is packed with decisions: what to eat, how to go to work, what to wear, what to buy, what to do in the evening, which shows to watch on TV, what to do during the weekend, and so on. Even though these decisions are part of daily life and do not take up too much space in our minds, everyone has to face some difficult questions that can take a long decision-making process before answering: career changes, wedding proposals, or significant investment decisions. When we make a decision, we usually try to improve our quality of life. Then, a decision can be made as part of our managerial role, such as entrepreneurs and managers. In this case, decisions can lead the enterprise to bankruptcy or success. Even only one wrong decision could change the direction of the firm. Moreover, some people in a particular position make decisions that can influence a country or the entire world. They can make decisions by themselves, such as dictators, or as a group, such as governments, central banks, councils, etc. People tend to be indecisive if the decision can change their lives, which could also lead to a longer decision-making process and second thoughts before the final choice.

A popular instrument used to make decisions is the “pro and cons list” where the agent identifies all the positive and adverse outcomes of that decision. However, this list is only sometimes helpful as it is not always easy to compare the pros and cons of different decisions. As choices are powerful, it is necessary to identify instruments that allow us to make the best decision given the available information and the decision maker’s beliefs. This type of decision is called a rational choice; however, the decision maker discovers whether this decision is the best one, i.e., the right choice, only when the consequences of that decision happen.

The study of the decision-making process and how to make rational decisions is done by a discipline called the theory of choice or decision theory.

Decision Theory

Decision theory analyzes the decision-making process: how choices are made and should be made based on assigning probabilities and values, called

utility, to each outcome. It is an interdisciplinary field involving psychologists, economists, philosophers, statisticians, and political scientists. Decision theory can be divided into two main branches: descriptive decision theory and normative decision theory.

Descriptive decision theory tries to explain how people make decisions in real-world situations. It seeks to understand the cognitive processes and biases that affect human decision-making and how people deviate from the ideal rational decision-making model. This discipline is based on empirical studies.

Normative decision theory guides on making optimal decisions. The purpose is to provide a set of guidelines or principles that can help the agent to make optimal choices based on their preferences, beliefs, and values. It is possible to model individuals' behavior as if they respect some axioms and follow a specific model.

It is essential to notice that a rational choice is a decision that respects the decision maker's belief and is considered optimal, giving all the information available. However, a right decision is a decision that gives the best outcomes, and it is only known after the decision is made. Decision theory concentrates on rational choices. Furthermore, normative and descriptive decision theories share some common elements, as they all agree that decisions are made according to the decision maker's beliefs and desires.

Formalization of the agent's problem

Following Peterson (2017), there are three levels of abstractions when making a choice:

- the decision problem that the agent needs to solve exactly how it is proposed to the decision maker;
- a formalization of the decision problem where the agent identifies all the information necessary to make a decision;
- A visualization of the formalization is usually done in a decision matrix or a decision tree if it involves sequential choices.

The tricky step is identifying all the useful information to make an informed decision. This information can be divided into states, outcomes, and acts.

A *state* can be defined as one of the possible scenarios that could happen in the future, and the state that happens in the real world is revealed once the choice is made. The agent should carefully choose only the states that are causally independent and relevant to her choice. For example, if the decision problem is "Should I take the umbrella today?", the states could be (i) Today it will be rainy, and (ii) Today it will not be rainy. The states (iii) I took the

right decision and (iv) I took the wrong decision are not independent of the choice, meaning they cannot be considered states of the world.

The decision maker decides based on the *outcomes* she will receive in each state of the world. A decision maker should be able to rank the likelihood of different outcomes. This ranking is entirely subjective and should reflect her attitudes toward the outcomes. In order to measure the value of an outcome, she should assign a value at each outcome using a *utility function*. The decision maker can use her utility function to order her preferences over a choice set. However, the preference ordering is not comparable across agents. Kreps (1988) states that a utility function should have the following properties: (i) *completeness*, which means that the agent should always be able to rank the different outcomes, (ii) *transitivity* if an agent prefers A to B and B to C, she will prefer A to C, and (iii) *separability* which is a technical concept. We refer to Debreu (1954) for defining properties of continuous utility functions. Identifying a specific utility function (and eventually new properties to respect) depends on the agent's characteristics, and the model adopted, as discussed in the following sections.

Acts are all the actions the decision-maker can do, and she has to decide which one she will adopt. Once chosen the act and revealed the state, she knows which outcome she will receive. Acts are considered alternative, meaning a rational decision-maker must choose only one act.

Let us consider the following example. Laura is paying rent in Paris and would like to buy a house next year. However, she does not know if her company will transfer her in the next ten years. Suppose that in this simplistic world, there can only be three states:

- Laura is not transferred,
- Laura is transferred to a new city close to Paris,
- Laura is transferred to a new city far from Paris.

Laura can have two options, acts, (a) continue to rent a house or (b) buy a house. If she buys the house and she stays in Paris, she will be delighted. However, if she buys the house and gets transferred far away, she will have to sell the house she just bought or rent it to someone else. On the other hand, if she continues to rent a house, it will be easier to transfer somewhere else, but if she stays in Paris, she will continue to pay rent without owning a house. Table 3 identifies the decision matrix that highlights states and acts for this problem. Moreover, the outcome is Laura's happiness, identified by a number from 0 to 10, the utility assigned at each outcome.

We have now discussed how a problem can be rewritten to identify states, acts, outcomes, and the associated utilities. Using decision theory, we need

States \ Acts	No Transfer	Transfer to a close city	Transfer to a far city
Rent in Paris	1	5	8
Buy in Paris	10	5	0

Table 3: *Decision Matrix of Laura’s decision problem*

to give an order to all the acts the agent can choose. In order to do so, the agent should identify the probability to assign at each state. If the probability is identified as objective or it is given as known to the decision maker, the decision is defined to be under risk. On the other hand, if the probability is not objectively known, the decision is under uncertainty.

Decision under Risk

Decisions under risk delineate situations with known objective probabilities or probabilities given to the decision maker. This is usually related to situation such as roulette in casinos, bets on flipping a coin, but also situations where the probability is given, for example, deciding whether to do an operation by knowing the percentage of people that survived or to do extreme sports by knowing the percentage of people that had injuries afterward.

Von Neumann and Morgenstern (1944) identify a theory where the rational decision maker should decide the act that maximizes the expected utility. There are n states of the world, and each act gives an outcome, x_i , for the state i whose probability is known, p_{x_i} , the expected utility function is computed as follows:

$$\mathbb{E}[u(X)] = \sum_{i=1}^n u(x_i) \cdot p_{x_i},$$

where $u(\cdot)$ is defined as a Von Neumann-Morgenstern utility function that has to respect some specific axioms, but it does not identify which utility function to use. This theory is called objective expected utility, where the decision maker should choose the act X that maximizes her expected utility.

As the probability is given, the most challenging part of decisions under risk is to identify the utility function that best describes the decision maker’s preferences. An agent can identify her utility function by identifying her preferences using simple lotteries and deriving point by point her utility function. We refer the reader to Gilboa (2010) to follow a step-by-step procedure to derive his or her utility function. A list of different utility functions with different properties related to insurance and money can be found in Gerber and Pafumi (1998).

It is generally possible to identify that $u(X)$ is an increasing function of

General Introduction

x as more money is always perceived better than less money. However, the concavity of the function depends on the decision maker's preference which can be averse to risk or seeking risk. In order to better understand this concept, we provide an example. The decision maker should choose between these two lotteries:

Lottery A gives 1,000€ with 0.5 probability and 0 otherwise;

Lottery B gives 500€ for sure.

The expected value of lottery A is $1,000 \cdot 0.5 + 0 \cdot 0.5 = 500\text{€}$. A decision-maker is said to be risk averse if she would prefer lottery B to lottery A as she prefers the expected value of the lottery to the lottery itself. On the other hand, a decision maker is defined risk lover or risk-seeking if she would prefer the lottery over the 500€ for sure. A risk-averse agent has a concave utility function, whereas a risk-lover agent has a convex utility function. In economics and finance, agents are assumed to be risk averse, even though their utility function is usually concave regarding low payoffs and convex when it comes to high ones (Peterson, 2017).

Let us go back to Laura's decision problem, where we assume that she asked the human resources department of her enterprise what is the probability that she will be transferred. They say that there is a 50% probability that she will be transferred to the hinterland of Paris, a 30% probability that she will be transferred far from Paris, and a 20% chance that she will stay in the same office. In order to decide which action she should choose, she computes her objective expected utility:

$$\begin{aligned}\mathbb{E}[u(\text{Rent})] &= 1 \cdot 0.2 + 5 \cdot 0.5 + 8 \cdot 0.3 = 5.1, \\ \mathbb{E}[u(\text{Buy})] &= 10 \cdot 0.2 + 5 \cdot 0.5 + 0 \cdot 0.3 = 4.5.\end{aligned}$$

The utility of renting a house in Paris is higher than that of buying a house, leading Laura to continue to rent the house for the moment. This choice could change if the human resources department gave her different probabilities or her happiness, i.e., the utility value, changes.

Decision under Uncertainty

Uncertainty, also called ambiguity, is a technical term that refers to cases in which the decision-maker knows acts and outcomes. However, she cannot assign probabilities to the states corresponding to the outcomes. For example, it is impossible to estimate, infer from existing data or ask someone the probability of war in a specific country in the next ten years, the probability that a new career path will end up in success or failure, etc.

At first, economists tried to deal with these kinds of problems by trying not to use any probability distribution. The decision maker's choice should be made only considering each act's best or worst outcomes without involving any probability distribution. We follow Peterson (2017) to identify some of the different principles used.

Dominance principle states that a decision maker should never choose an act whose outcomes are worst, no matter which state is the true state of the world. In other words, if the specific act always gives a worse outcome than another, the former should have been excluded from the decision problem.

Maxmin principle where the decision maker should choose the act that gives the least bad outcomes independently of the state in this happens. In the case of Laura's decision problem, Laura should decide to rent in Paris as the worst outcome gives a utility equal to 1, whereas the worst outcome in the case of buying a house has a null utility.

Maximax rule where the decision maker should choose the act that has the best possible outcomes compared to the best possible outcomes for each act. In the case of Laura's decision problem, Laura should decide to buy a house as the best outcome gives a utility of 10 which is higher than the best outcome of renting a house.

Optimist-Pessimist Rule that considers the best and the worst possible outcomes of each alternative and chooses an alternative according to her degree of optimism or pessimism, and it should be a weighted average of these two utilities where the weight depends on the degree of optimism or pessimism.

All this method only concentrates on the best or worst outcomes, but they do not consider intermediate outcomes. To make an informed decision, all outcomes and utilities should be considered to make a rational choice that optimizes the agent's utility function.

According to Savage (1972), a person should be able to identify a subjective probability that guarantees that her beliefs are internally coherent. These subjective probabilities have to follow some axioms to be used in decision-making processes. Gilboa (2010) provides an easy step-by-step procedure for deriving a subjective probability by starting from accessible lotteries. This procedure is similar to the one to create its utility function pointwise. However, decision-makers cannot always assign a specific probability; for example, to the state "tomorrow it will rain," should I assign 39% or 41% of chance?

In contrast to this theory and the fact that people cannot always have coherent subjective probabilities, many different paradoxes have been proposed.

One of the most popular paradoxes is the **Ellsberg Paradox** by Ellsberg (1961), presented below, where he argues that people prefer bets with known probability to bets with unknown probability. This behavior is classified as uncertainty aversion, violating one of Savage's axioms.

Suppose that two urns contain 100 balls. The agent knows that urn A has 50 black balls and 50 white balls, whereas she does not know the proportion of black and white balls inside Urn B. The decision maker is then offered a choice between the following gambles:

Lottery 1 Receive 100€ if a black ball is drawn from Urn A.

Lottery 2 Receive 100€ if a black ball is drawn from Urn B.

At this point, the decision-maker can reason in two ways. She could prefer gamble 1 as she knows the proportion of black balls in Urn A, or she could prefer gamble 2 as she thinks the probability could be higher. As the reasoning does not change, we assume that the decision-maker prefers the known urn, i.e., lottery 1. It has been empirically proven that people tend to prefer lottery 1.

Then, they are presented with a second pair of gambles.

Lottery 3 Receive 100€ if a white ball is drawn from Urn A.

Lottery 4 Receive 100€ if a white ball is drawn from Urn B.

When confronted with a new pair of lotteries, a person that prefers gamble 1 to gamble 2 would prefer gamble 3 to gamble 4.

This is inconsistent with the expected utility theory as a decision maker that prefers gamble 1 to gamble 2 would believe that the probability of black balls in urn A is higher than the probability of white balls in urn B, i.e., $P(\text{black in urn A}) = 0.5 > P(\text{black in urn B})$. However, as the same person prefers gamble 3 to gamble 4, she would believe that the probability of white balls in urn A is higher than the probability of white balls in urn B, i.e., $P(\text{white in urn A}) = 0.5 > P(\text{white in urn B})$. These two probabilities are inconsistent with each other, that is because $P(\text{black in urn B}) + P(\text{white in urn B})$ would not sum to 1. As the decision maker does not know the probability of each ball in urn B, the agent tends to prefer the known urn. This is called ambiguity aversion, which demonstrates that it is impossible to use Savage's model to find the optimal decision in this scenario.

In recent decades, there has been much theoretical interest in decision models that consider ambiguity and ambiguity aversion in order to resolve the Ellsberg Paradox, such as Schmeidler (1989), Gilboa and Schmeidler (1989), Klibanoff et al. (2005) among others. A more detailed discussion is provided below.

Home Bias and Decision Theory

Decision theory has been widely applied in optimal portfolio allocation problems trying in order to explain some asset allocation puzzles. Standard finance theory predicts that investors should hold a diversified portfolio of equities worldwide, assuming that capital is fully mobile across borders. Because foreign equities provide great diversification opportunities, falling barriers to international trade in financial assets should have led investors worldwide to re-balance their portfolio away from national assets toward foreign assets (Coeurdacier and Rey, 2013). The fact that people hold underdiversified portfolios can be linked to the *home bias*. Home bias is an investment behavior where investors tend to overweight their home country's market compared to its share in the overall market portfolio.

Coeurdacier and Rey (2013) show that home bias has decreased over the 1988-2008 period in developed countries across regions of the world but remains high in most countries. Furthermore, emerging markets have less diversified equity portfolios than developed countries and do not exhibit any clear downward trend in home bias. Table 4, taken from Coeurdacier and Rey (2013), shows the percentage of domestic assets that an investor should have in her portfolio and the percentage she has, which demonstrates that home bias still exists on the market.

Source country	Domestic market in % of world market capitalization	Share of portfolio in domestic equity in %
Australia	1.8	76.1
Euro Area	13.5	56.7
Sweden	0.7	43.6
Switzerland	2.3	50.9
United States	32.6	77.2

Table 4: *Home bias in equities in 2008 for selected countries. Source: Coeurdacier and Rey (2013)*

The home bias has been confirmed by different empirical studies such as Fidora et al. (2007), Lippi (2016), Lütje and Menkhoff (2007), Mishra (2008), Lin and Viswanathan (2016), among others.

Many researchers tried to detect reasons why investors show a home bias, and these reasons can be classified into three main categories:

- Institutional reasons such as taxes, transaction costs, and barriers
- Information asymmetries between investors

General Introduction

- Behavioral reasons include familiarity, patriotism, optimism, ambiguity aversion, etc.

The reader should refer to Gaar et al. (2020) for a more exhaustive literature review.

There have been different studies that examine the optimal portfolio asset allocation in order to study the elements that could generate home bias.

A lot of different decision theory models are used to study cognitive bias that an investor has when making portfolio allocation. Decision theory models can be in a static setting, such as Von Neumann and Morgenstern (1944) that use objective probability distributions, Schmeidler (1989) whose probability measure is not necessarily additive, Gilboa and Schmeidler (1989) where it maximizes the worst-case expected utility over a set of probabilities, Klibanoff et al. (2005) that identify a smooth ambiguity preference models. These models were also developed in a dynamic setting, such as Epstein and Schneider (2003) that extended the Gilboa and Schmeidler (1989), Klibanoff et al. (2005) and Hayashi and Miao (2011) that extended the Klibanoff et al. (2005). These are just a few examples of different decision theory models developed to study optimal asset allocation.

Researchers applied some of these models to study home bias by analyzing the optimal asset allocation when two risky stocks are in the market. We only discuss some of the recent studies on this topic. Guidolin and Liu (2016) studies the home bias using a model that considers uncertain beliefs over an asset pricing model, parameter uncertainty, and ambiguity aversion. These authors used a static smooth ambiguity model proposed by Klibanoff et al. (2005) to characterize ambiguity attitudes and include a Bayesian framework. Using their approach, they find that ambiguity aversion can lead to solid home bias in equity holding, regardless of an investor's belief in the domestic CAPM. Peijnenburg (2018) studies the evolution of home bias in a dynamic setting using a max-min expected utility model. She also finds that ambiguity helps explain under diversification and home bias. However, the model adopted does not separate ambiguity and ambiguity attitudes. A smooth ambiguity model in a dynamic setting is also studied in Yu et al. (2022), where they only concentrate on the sensitivity analysis of risk and ambiguity parameters. They found that as an individual becomes more ambiguous averse, she will consume less, buy more life insurance and invest less in risky assets.

Risk Management and Decision Theory

Risk and decision-making are two interrelated factors in organizational management, and they are both related to the presence of uncertainties in the organization (Lu et al., 2012). A risk manager should use the instruments

provided by decision theory to conduct choices that can make the organization profitable and avoid bankruptcy.

Risk management is defined as a series of activities aiming to direct and control an organization taking into consideration risk. It is necessary to identify and assess potential risks and then manage them by making appropriate decisions. The risk management process can be divided into four different steps:

1. Identification of the risks by using brainstorming, structure, interviews, questionnaire and surveys, loss data analysis, hypothetical what-if analysis, etc;
2. Assess the vulnerability of critical assets to specific threats;
3. Analysis and measure of the risk;
4. Management of the risks, which can be (i) avoid the risk, (ii) retain the risk through a mechanism such as risk capital allocation, (iii) mitigate the risk by reducing exposure, frequency, and severity, or (iv) transfer the risk to a third party.

Making appropriate decisions in real-world scenarios is tricky due to risks and uncertainties. A risk manager should be able to analyze problems using decision-making criteria to identify the outcomes in each scenario, the utility of each outcome, and to assign the probability of each outcome. Borgonovo et al. (2018) study the link between risk analysis with decision theory with the purpose of allowing a close integration of the risk management part of decision analysis to the output of risk assessment. Most researchers define risk analysis as quantifying the risk, identifying risk management options, and communicating the results to the risk manager, i.e., steps 1 to 3 of the risk management process. Even though risk analysis and decision analysis have some similarities and are often complementary, risk analysis is often performed without a known decision maker to assess probabilities and to choose risk management options based on his or her preferences and risk attitudes (Paté-Cornell and Dillon, 2006). Then, the outcome of risk analysis is communicated to the risk manager, who decides how to manage the overall risk, i.e., step 4 of the risk management process. The risk manager should choose the option that maximizes his or her utility function according to his or her preferences and properties.

Borgonovo et al. (2018) connect operational risk management with the theoretical foundations of decision theory. They show a direct link between risk analysis and decision theory when using objective or subjective probabilities. A different reformulation is necessary in case of uncertainty. In general, they

proved that risk managers can access the decision theory toolbox to define their strategies.

Regulators and Risk Measures

To protect a country or economic system from too high risks taken by financial institutions to obtain high profits, in the 1990s, risk management was considered necessary, as well as the introduction of regulation in the market. Regulators are committees that define different rules that a financial institution must respect to enter and stay in the market. To avoid bankruptcy, regulators usually define a risk measure, i.e., a capital requirement the company should retain to face significant losses.

In 1990 J.P. Morgan developed the Value-at-Risk (VaR), a global risk metric that measures the worst expected loss over a given horizon under normal market conditions at a given confidence level (Jorion, 2007). For instance, suppose that the 95% VaR at one year equals 2€ million. It means that, under normal market conditions, there is a 5% possibility that losses will be higher than 2€ million in the following year. A more detailed analysis of VaR can be found in Klugman et al. (2012), Jorion (2007), Linsmeier and Pearson (2000), among others. However, the main problem of VaR is that it fails to quantify how much risk is in the tails.

Another famous risk measure is the Tail Conditional Expectation (TCE), also called Expected Shortfall, Conditional Tail Expectation, or Tail-VaR, defined as the average of all losses higher than a specific threshold, usually the VaR at quantile p . By construction, TCE will always be at least equal to the VaR computed as the same quantile.

Even though all regulators have the same purpose, they disagree on which risk measures they should use, and not everybody uses the two risk measures highlighted before. The Canadian regulation of insurance and the United States prefer models requiring higher regulatory involvement in the supervision. In contrast, Solvency II adopts a VaR at 99.5% computed on a horizon of 1 year, and the Swiss Solvency Test uses a TCE at 99% computed on the same horizon. A comparison of different insurance regulatory regimes can be found in Comité Européen des Assurances and Mercer Oliver Wyman Limited (2005), or a comparison between an insurance regime (Solvency II) and a banking one (Basel III) can be found in Gatzert and Wesker (2011). It does not exist a general risk measure that is adopted worldwide; regulators decide which risk measures to adopt according to their level of risk aversion and by following some political process.

Both risk measures adopted by Solvency II and the Swiss Solvency Test are based on loss distribution. They use statistical quantities to describe the portfolio's conditional or unconditional loss distribution over some predetermined

horizon T at a specific confidence interval p .

In this thesis, we assume that losses follow a Pareto type I distribution or a Generalized Pareto distribution which is a distribution that can consider a lot of extreme values and whose probability density function presents fat tails. Vilfredo Pareto first introduced Pareto distribution in 1895 to study revenues. He identifies the so-called “Pareto principle”, which states that a smaller percentage of people in the society owns a large share of the society’s wealth. The Pareto Type I distribution is characterized by α , which identifies the riskiness in the market, and x_0 , which is the minimum possible value of the random variable studied. On the other hand, the generalized Pareto distribution has three parameters: ξ , which identifies the riskiness in the market and can be interpreted as the inverse of α , the location parameter μ , and the scale parameter σ . We model loss claims using this specific distribution as they are consistent with what is observed for claims with heavily distributed tails, and it allows us to derive readable results.

Coherent risk measures

Artzner et al. (1999) identify a series of properties that a risk measure should have in order to be defined coherent. They identify that a coherent risk measure should satisfy properties of monotonicity, subadditivity, positive homogeneity, and translation invariance.

The property of *monotonicity* stands that if the losses of one random variable are always higher than the losses of another random variable in each state of the world, then it makes sense for the risk measure of the former to be greater than the risk measure of the latter.

Subadditivity reflects that combining risks leads to diversification and, thus, a reduction of the total overall risk. It states that the risk measure related to an aggregated portfolio should be equal or lower than the sum of the risk measures computed for each item.

The property of *positive homogeneity* specifies that the risk measure of a constant multiple of the unexpected loss should be the constant multiple of the risk measure, whereas *translation invariance* says that the risk measure of combining an unexpected loss and a fixed loss should be the risk measure of the unexpected loss plus this fixed loss.

It has been proved by Artzner et al. (1999) that the Value-at-Risk is not a coherent risk measure, whereas the TCE is a coherent risk measure. However, there are different critics on the merit of the four properties, see for example Rau-Bredow (2019), Daniélsson et al. (2005), Cont et al. (2010), among others.

The importance of high-order moments

A useful instrument that the risk manager should use to assess the market risk is the computation of high-order moments of the returns' distribution that can capture asymmetry and tail-fatness of returns. Moreover, the more moments are known, the better the estimation of the distribution that can help the decision maker to identify the optimal strategy to avoid bankruptcy. For example, according to Dai et al. (2021), the dynamics of higher order moments is considered a measure of the level of risk being shared among asset classes. Furthermore, De Clerk and Savel'ev (2022) indicates that we can use the higher-order moment analysis as a tool to indicate changes in a financial asset's risk.

Many studies argue that in addition to mean and volatility, high-order moments can be included as risk factors in asset pricing, such as Sihem and Slaheddine (2014), Jurczenko and Maillet (2006), Ahmed and Al Mafrachi (2021), Chen et al. (2021), among others.

There is also a vast literature that reveals that high order distribution moments are used to predict market returns, such as Jondeau et al. (2019), Doan et al. (2010), Amaya et al. (2015), Mei et al. (2017), among others.

This different stream of papers identifies that high-order moments are a helpful tool for the risk manager to identify the most efficient strategy.

Thesis structure

This thesis concentrates on the study of the decision-making process and provides some tools that can help decision-making in the risk management field as well as the study of investors' cognitive bias by using models adopted in decisions under uncertainty. This thesis can be divided into two parts: one concentrates on risk management and decision under risk, and the second involves decisions under uncertainty and investors.

The first part is composed of chapters 1, 2, and 3. It discusses decision theory from a risk management viewpoint.

Chapter 1 studies the equivalence between different risk measures. Much research concentrates on the merits of VaR and TCE, the two most classic risk indicators financial institutions use. However, we decided to focus our research on the equivalence between such indicators to identify an implicit utility function that helps us understand how regulators choose such or such risk indicators for solvency computations. We know that regulators may make a decision following a political process. However, making the implicit functioning of actual regulations explicit should be a valuable contribution to assuring that regulations function in an effective and efficient manner. Further, TCE may not be the most accurate indicator to consider the nature of probability distribution

tails. For this reason, we introduce a new risk indicator that extends TCE, called high-order TCE or extended TCE, that can consider higher-order risks. Our study assumes that claims follow a simple Pareto framework and then follow a generalized Pareto framework. We also examine equivalence results between the quantiles of high-order TCEs. This chapter is based on the article published by Faroni et al. (2022).

Then, we identify the four properties that a risk measure should have to be defined as a coherent risk measure in chapter 2. This concept has been very influential in the further development of risk measures and in defining new instruments in risk management. We discuss that VaR is not a coherent risk measure, whereas TCE is. We also highlight some critics of this concept. Moreover, we discuss that the extended TCE introduced in the previous chapter is not a coherent risk measure. We also introduce a variation of the extended TCE that can respect the four properties identified by Artzner et al. (1999).

In chapter 3, we perform an empirical study on the stability of annual high-order moments in equity indexes among markets. We extend our study to partial annual high-order moments, i.e., moments computed using only values higher or lower than a specified threshold. We use seventeen equity indices to consider the world's most representative regions as well as the 3-month Treasury Bill Rate, and we collect daily closing prices from 2000 to 2022. This study aims to help risk managers identify which moment is more stable over time, which will lead them to a more reliable assessment of future market risks.

The second part of this thesis comprises chapter 4 in which we use decision models under uncertainty. We study the dynamic optimal asset allocation and the optimal consumption that the agent should adopt when there are two kinds of stocks in the market: local and foreign assets. We assume that the two stocks are identical on the market. However, the agent perceives them differently. She believes in knowing the local stock better and is more uncertain about the real return of the foreign asset. The aim is to study the home bias and its evolution over time when we assume that the decision maker is both risk-averse and ambiguity averse. We use the generalized recursive smooth ambiguity model axiomatized by Hayashi and Miao (2011) that can distinguish between risk aversion, ambiguity aversion, and intertemporal substitution. This separability allows us to understand better which parameters drive the home bias of the agent. We create a numerical implementation that includes ambiguity in the market and the fact that the decision-maker learns over time each year to reduce her uncertainty. Our implementation is explicitly computed for two risky and ambiguous stocks on the market. However, it can be generalized to n risky and ambiguous assets.

Chapter 1

Equivalent Risk Indicators: VaR, TCE, and Beyond

Due to a period that started in the early 1980s of less regulation and high volatility, regulators started to define risk-based capital requirements with the purpose of making capital available for absorbing losses when extreme events occur. Even though there is no consensus on the best risk measure to adopt, they all share a common feature: they are all related to the behavior of tails of the probability distribution of a firm's financial results.

A risk measure is defined as a mapping from the random variable representing a risk exposure to a set of real numbers. In Section 1.2, we introduce two classic risk measures, Value-at-Risk (VaR) and Tail Conditional Expectation (TCE), as well as a new risk measure that considers high-order risks, which we call high-order TCE or extended TCE. To give the reader a comprehensive view of the topic, we also provide an overview of the world's regulation system and their differences and similarities in Section 1.3.

We do not focus on the merits of the risk measure. However, we study the equivalence between the three risk measures presented here to understand how regulators choose a specific risk measure and its parameters. In Section 1.4, we study the equivalence between VaR, TCE, and the high-order TCE quantiles.

To obtain our results, we assume that claims follow the Pareto distribution and the generalized Pareto distribution for two main reasons: they allow us to derive readable results and are consistent with what is observed for claims with heavily distributed tails. Section 1.1 gives a brief overview of the Pareto distribution.

Finally, Section 1.5 concludes with a brief illustration based on real data and the suggestion of theoretical extensions.

The results described in this chapter are published in Faroni et al. (2022).

1.1 The Pareto distribution

In the XIX century, one of the most common ways to use statistics to describe real-world phenomena was to show that some variables are concentrated nearby a value (mean, mode, or median), and the big variables far from this value were considered rare and, consequently, they were not studied. However, this representation is not accurate for many situations where the big values, called “extreme values”, are not negligible and need to be included in the study.

In 1895, Vilfredo Pareto studied if revenues were distributed randomly or if they followed a specific rule yet to be discovered. Pareto (1895) (cited in Barbut (2003)) shows that there is a tendency to follow a specific curve characterized by asymmetry and extremely positive values. He shows that a smaller percentage of the people in that society (around 20%) own a larger share of the wealth (around 80%) of the society. This is called the “Pareto Principle”.

Pareto studies the cumulative function instead of the traditional density function to find the mathematical formula of this curve. Given a revenue x_i , he identifies $N(x_i)$ as the number of revenues equal to or higher than x_i . These coordinates are then represented in a logarithmic paper, $(\text{Log}(x_i), \text{Log}(N(x_i)))$. Pareto discovers that all the different series observed follow

$$\text{Log}N(x_i) = -\alpha \text{Log}(x_i) + C$$

where C is a positive constant value that depends on the distribution used. The previous equation can be written as:

$$N(x) = \left(\frac{x_0}{x}\right)^\alpha,$$

where α is the Pareto Index, and x_0 is the minimum possible value of the distribution. This is the expression for the tail function of the *Pareto Type I distribution*, discussed in Section 1.1.1. A hierarchy of Pareto distributions is known as Pareto Types I, II, III, and IV. The reader should refer to statistics books such as Arnold (2008) for a more comprehensive understanding of this topic.

In the context of extreme value theory that deals with extreme deviations from the median of probability distributions, Pickands III (1975) shows that, at a large threshold, the limiting distribution of values above this threshold follows a generalized Pareto distribution (GPD). We discuss this probability distribution in Section 1.1.2. The advantage of GPD is that it contains a lot of different sub-cases that can be agglomerated into three groups:

1. Gumbel distribution, whose tails decrease exponentially, such as the

Gaussian distribution;

2. Weibull distribution, whose tails are finite;
3. Fréchet distribution, which has heavy tails.

We will only concentrate on the last case as we are interested in the probability loss distributions, and it has been empirically demonstrated that they have fat tails.

The interested reader should refer to Coles et al. (2001), Arnold (2008), and Embrechts et al. (2013), among others, for a detailed discussion of these models.

We decide to model claims either as a Pareto distribution or a generalized Pareto distribution, as they are consistent with what is observed for claims with heavily distributed tails. We refer the reader to Le Courtois (2018), or Le Courtois and Walter (2014), among others, to explore situations where claims could be associated with semi-heavy tails.

1.1.1 Pareto Type I Distribution

Definition 1. Let X be a random variable that follows a *Pareto Type I distribution*, then the probability density function is

$$f_X(x) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & x \geq x_0 \\ 0 & x < x_0 \end{cases}, \quad (1.1)$$

and the cumulative distribution function is

$$F_X(x) = \begin{cases} 1 - \left(\frac{x_0}{x}\right)^\alpha & x \geq x_0 \\ 0 & x < x_0 \end{cases}, \quad (1.2)$$

where x_0 is the necessarily positive minimum possible value of X , and α is a positive parameter.

The Pareto type I distribution models only the values higher than a threshold, x_0 , which is the starting point of the observations. The shape parameter α , also called the ‘‘Pareto Index’’ or tail index, identifies the size of the tail distribution.

To give an interpretation of α , we introduce the hierarchical relationships in the random variable X . We classify each realization of X based on its size, from the largest to the smallest. Given n realization of X , the ordered sample is written as $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$. For each value x_k , the $\text{Nb}(X \geq x_k) = k$

which is the inverse of the cumulative function, $1 - F_X(x_k)$. The hierarchical relation of rang k is defined as¹:

$$\frac{X_{(k)}}{X_{(k+1)}} = \left(\frac{1 - F_X(X_{(k+1)})}{1 - F_X(X_{(k)})} \right)^{\frac{1}{\alpha}} = \left(\frac{k+1}{k} \right)^{\frac{1}{\alpha}} \quad \text{a.s.}$$

The hierarchical relation between the biggest value and second-biggest value is equal to $2^{\frac{1}{\alpha}}$ where α determines the size of the gap between the two. For example, if $\alpha = 1.4$, the ratio is equal to $2^{\frac{1}{1.4}} \approx 1.64$ whereas when $\alpha = 4$, the ratio is equal to $2^{\frac{1}{4}} \approx 1.18$ which means that the biggest value is 1.18 times the second-biggest value. We show that when α increases, the ratio diminishes, leading to having “less extreme” values. In other words, the inequality of the distribution increases when the shape parameter decreases, and vice versa.

Graphically, we can see the effect of α in Figure 1.1. We show that when α is big, a high probability is assigned to values close to x_0 , which leads the distribution to have a slim tail (compared to the one with a more minor α).

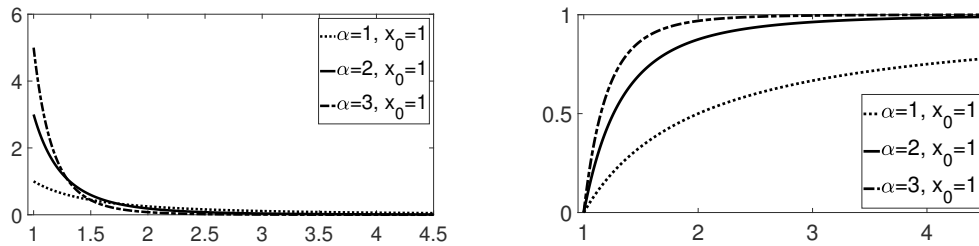


Figure 1.1: Probability density function (left panel) and cumulative distribution function (right panel) of Pareto Type I distribution with $x_0 = 1$ and $\alpha = [1, 2, 3]$.

Figure 1.2 shows that x_0 shifts the distribution as it is the minimum value of the distribution. This means that x_0 can be seen as a location parameter that does not influence the riskiness of the market, which is identified by the parameter α .

The Pareto Type I distribution has two primary limits.

First of all, it is impossible to use this distribution to represent the distribution of all values but only of all the values in the tail function, i.e., that are higher than a specific threshold. When using the Pareto distribution to represent the empirical data, we need to divide the empirical data into two sub-classes: (i) extreme values and (ii) all other values. Then, we should assign a distribution to each sub-class and combine the two regimes. In our case,

¹We consider that $1 - F_X(x_k) = \bar{F}_X(x_k) = \left(\frac{x_0}{x}\right)^\alpha$ that can be rewritten as $x_k = \frac{x_0}{(\bar{F}_X(x_k))^{\frac{1}{\alpha}}}$.

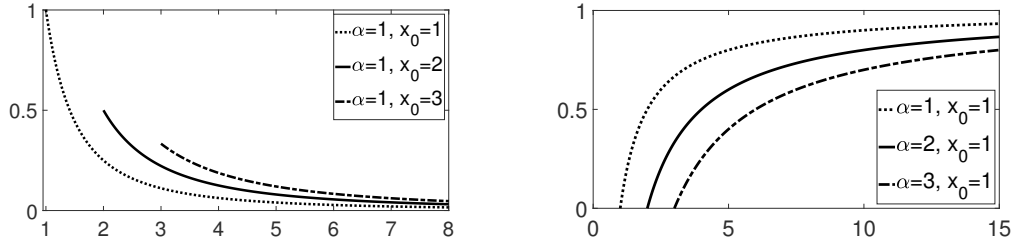


Figure 1.2: Probability density function (left panel) and cumulative distribution function (right panel) of Pareto Type I distribution with $x_0 = [1, 2, 3]$ and $\alpha = 1$.

this is not a limit as we study only the tail behavior of the distribution, which can be distributed as a Pareto Type I distribution.

The second problem is that it is not always possible to compute the moments of a distribution. Moments summarize the basic information of empirical data and allow us to understand the distribution better. From a statistical point of view, it is always possible to compute empirical moments. However, they are not always significant as they do not converge to the theoretical moments. From a mathematical viewpoint, it is possible to show that a Pareto distribution of a variable X has moments of order k only if k is lower than α . For example, if α is lower than 2, the empirical moment of order 2 will diverge, and the theoretical variance will be equal to infinity. If moments exist, they are as follows:

$$\text{Mean} = \frac{\alpha x_0}{\alpha - 1} \quad \text{for } \alpha > 1,$$

$$\text{Variance} = \frac{\alpha x_0^2}{(\alpha - 1)^2 (\alpha - 2)} \quad \text{for } \alpha > 2,$$

$$\text{Skewness} = \frac{2(1 + \alpha)}{\alpha - 3} \sqrt{\frac{\alpha - 2}{\alpha}} \quad \text{for } \alpha > 3,$$

$$\text{Kurtosis} = 3 + \frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)} \quad \text{for } \alpha > 4.$$

1.1.2 Generalized Pareto distribution

The generalized Pareto distribution (GPD) was introduced by Pickands III (1975), and it is usually used to model extreme data in the sense that they

exceed a particular designated high threshold.

Definition 2. Let X be a random variable that follows a *generalized Pareto distribution*, then the probability density function admits the Jenkinson-von Mises representation, which can be expressed as follows:

$$f_X(x) = \begin{cases} \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi} - 1} & \text{for } \xi \neq 0 \\ \frac{1}{\sigma} \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{for } \xi = 0 \end{cases} \quad (1.3)$$

and the cumulative distribution function is

$$F_X(x) = \begin{cases} 1 - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ 1 - \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{for } \xi = 0 \end{cases} \quad (1.4)$$

where μ identifies the location, σ the scale, and ξ the shape of the tail function.

The GPD is generalized in the sense that the parametric form subsumes three types of distributions according to the value of ξ :

- $\xi > 0$ is the Fréchet distribution,
- $\xi = 0$ is the Gumbel distribution,
- $\xi < 0$ is the Weibull distribution.

Figure 1.3 shows that the Weibull distribution is a short-tailed distribution with a finite right endpoint. The Gumbel and Fréchet distributions have infinite right tails, but the tail of the Fréchet distribution decreases slower than that of the Gumbel distribution.

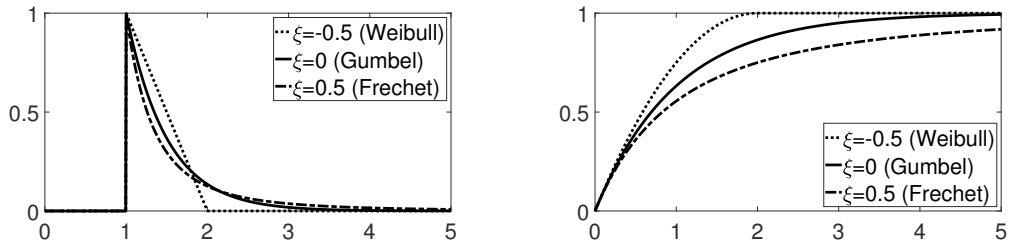


Figure 1.3: Probability density function (left panel) and cumulative distribution function (right panel) of GPD with $x_0 = [-0.5, 0, 0.5]$, $\mu = 0$, and $\sigma = 1$.

For our study, we use $\xi > 0$ as it allows us to consider a fat tail distribution which means that it considers extreme events. Moreover, it is essential to notice that when $\xi > 0$ and $\mu = \frac{\sigma}{\xi}$, the GPD is equivalent to the Pareto Type I distribution with $x_0 = \frac{\sigma}{\xi}$ and $\alpha = \frac{1}{\xi}$ which means that we can apply the same interpretation to ξ as in Section 1.1.1. In this case, an increase of ξ (i.e., a decrease of α) will lead to an increase in extreme values and an increase in the inequality of the distribution.

The moment of rank m of a GPD only exists if $\xi < \frac{1}{m}$. Otherwise, the moment is equal to infinity. If moments exist, they can be computed as follows:

$$\text{Mean} = \mu + \frac{\sigma}{1 - \xi} \quad \text{for } \xi < 1,$$

$$\text{Variance} = \frac{\sigma^2}{(1 - 2\xi)(1 - \xi)^2} \quad \text{for } \xi < \frac{1}{2},$$

$$\text{Skewness} = \frac{2(1 + \xi)\sqrt{1 - 2\xi}}{1 - 3\xi} \quad \text{for } \xi < \frac{1}{3},$$

$$\text{Kurtosis} = 3 + \frac{6(1 + \xi - 6\xi^2 - 2\xi^3)}{(1 - 3\xi)(1 - 4\xi)} \quad \text{for } \xi < \frac{1}{4}.$$

We can see that the third and fourth moment (skewness and kurtosis) only depends on ξ , i.e., the riskiness of the market. For information on the estimation of tail parameters, see, for instance, Hill (1975), or Hosking and Wallis (1987).

1.2 Risk Measures

The primary purpose of risk management is to identify the amount of capital a financial institution needs to hold to avoid unexpected losses that can lead to insolvency and bankruptcy. Following McNeil et al. (2015), we identify four different categories to measure the risk of a financial position.

The *notional-amount approach* consists of the weighted sum of the notional values of the individual securities in the portfolio, where the weight is a factor that represents an assessment of the riskiness of the class. This method cannot consider the benefits of diversification on the overall risk of the company and does not deal well with portfolios of derivatives.

Factor-sensitivity measures use portfolio value changes for a predetermined

change in one of the underlying risk factors. It is computed using mathematical derivatives. However, it is impossible to aggregate sensitivities computed using different risk factors, and it is not possible to aggregate this measure across markets.

Risk measures based on loss distribution uses statistical quantities to describe the portfolio's conditional or unconditional loss distribution over some predetermined horizon T . In this thesis, we use measures that fall into this category, such as Value-at-Risk, Tail Conditional Expectation, and high-order Tail Conditional Expectation discussed in detail in Sections 1.2.1, 1.2.2, and 1.2.3, respectively. Furthermore, in Section 1.2.4, we provide some examples to interpret the distribution parameters using the risk measures described here. Regulators tend to use this kind of risk measure and have to choose two parameters: the time horizon and the confidence level. The time horizon depends on the application of the company regulated. For example, it makes sense to use T equal to one trading day if the positions are liquid and actively managed, whereas to use T equals one month or year if the portfolio is traded less actively and some instruments of the portfolio are less liquid. The choice of the confidence level is tricky, and we leave this discussion for Section 1.4. The disadvantages of using a loss distribution is that it relies on past data that cannot be significant to predict the future.

Scenario-based risk measures rely on several possible future scenarios where the risk measure is computed as the maximum portfolio loss under all scenarios, where some extreme cases can be down-weighted to mitigate their effect. The main problem, in this case, is how to determine an appropriate set of scenarios and weighting factors. Furthermore, comparing scenario-based risk measures across portfolios is challenging if they are affected by different risk factors.

1.2.1 Value-at-Risk

The most widely used risk measure in a financial institution is the Value-at-Risk (VaR) (see, for instance, Klugman et al. (2012)), first introduced in the late 1980s to measure the risk of their trading portfolio by J.P. Morgan. VaR is a measure of the expected worst loss over a given horizon where losses greater than the VaR are suffered with just a small probability under normal market circumstances. Linsmeier and Pearson (2000) defines VaR as the loss expected to be exceeded with a probability of only $(1 - p)$ percent during the next T -day holding period.

Definition 3. Let $F_{X_T}(x)$ be the distribution function of outcomes over a fixed period T of a portfolio of risks where a loss is identified as a positive value of the random variable X_T . The VaR of the random variable X_T is the

p percentile of the distribution, denoted by:

$$\text{VaR}_p(X_T) = F_{X_T}^{-1}(p).$$

From a probabilistic viewpoint, VaR is just a quantile of the loss distribution. We should note that the VaR does not give any information about the severity of losses higher than the level of confidence chosen.

Following the approach of Denuit et al. (2006), we demonstrate that the VaR is an optimal capital requirement for insurance companies. Consider a portfolio where losses follow a distribution of X_T and suppose that a regulator wants a solvency capital requirement $Q[X_T]$ to be large enough to ensure that the shortfall risk, $\mathbb{E}[(X_T - Q[X_T])_+]$, remains small. However, regulators need to consider that retaining capital has a cost that diminishes the company's competitiveness. This cost can be expressed as a percentage ϵ of the capital requirement. Finally, the solution to the following minimization problem gives us the optimal capital requirement:

$$\min_{Q[X_T]} \{ \eta \cdot \mathbb{E}[(X_T - Q[X_T])_+] + (1 - \eta) Q[X_T] \epsilon \}, \quad (1.5)$$

where η is the weight that identifies the importance given to one element or the other by the regulator. For example, if the regulator wants to assign the same importance to both elements, it sets η equal to 0.5. The capital $Q[X_T]$ that is able to solve Eq. 1.5 is:

$$Q[X_T] = \text{VaR}_{(1 - \frac{1-\eta}{\eta}\epsilon)}(X_T).$$

Proof. See Appendix. □

Note that the quantile p of the optimal VaR just computed is a function of the weight η and the cost of capital parameter ϵ , i.e., $p = 1 - \frac{1-\eta}{\eta}\epsilon$. As the quantile is a number between 0 and 1, $\epsilon > \frac{p}{1-p}$ is a condition that has to be respected.

Figure 1.4 shows the relation between the quantile p , the weight η , and the cost of capital ϵ . The quantile p strongly increases with η for fixed values of ϵ . Similarly, an increase of η is traduced in an increase of p for fixed values of ϵ . Indeed, η shows the importance that the regulator gives to minimizing the residual risk, whereas $(1 - \eta)$ is the weight given to the cost of capital. In this case, if the regulator assigns a specific value to the VaR's quantile (which is usually close to 1), it gives much importance to minimizing the risk, and the cost of capital becomes less important aspect.

The Value-at-Risk has its limits, mostly because it is not a coherent risk measure. We leave this discussion for Chapter 2.

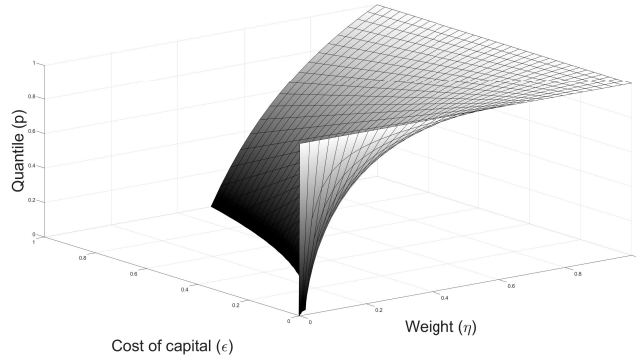


Figure 1.4: *VaR quantile w.r.t. optimization parameters.*

Pareto Type I Losses

Let X_T be a random variable that follows a Pareto Type I distribution with parameters α and x_0 . Given the cumulative function in Eq. 1.2 and the fact that $F_{X_T}(\text{VaR}_p(X_T)) = p$, it is possible to obtain that:

$$\text{VaR}_p(X_T) = x_0 (1 - p)^{-\frac{1}{\alpha}} \quad (1.6)$$

Then, we can compute the hierarchical relation between two Value-at-Risks at two different quantiles:

$$\begin{aligned} \frac{\text{VaR}_p(X_T)}{\text{VaR}_q(X_T)} &= \frac{x_0 (1 - p)^{-\frac{1}{\alpha}}}{x_0 (1 - q)^{-\frac{1}{\alpha}}} \\ &= \left(\frac{1 - p}{1 - q} \right)^{-\frac{1}{\alpha}}. \end{aligned}$$

If we assume $p > q$, we know that the Value-at-Risk at quantile p will be $\left(\frac{1 - p}{1 - q} \right)^{-\frac{1}{\alpha}}$ bigger than the Value-at-Risk at quantile q .

Example 1

Assume that the loss distribution has $x_0 = 20$ and $\alpha = 2$, compute the Value-at-Risk at 99% and the hierarchical relation between this quantile and the

quantile at 99.5%.

$$\begin{aligned}\text{VaR}_{0.99}(X_T) &= 20(1 - 0.99)^{-\frac{1}{2}} = 200, \\ \frac{\text{VaR}_{0.995}(X_T)}{\text{VaR}_{0.99}(X_T)} &= \left(\frac{1 - 0.995}{1 - 0.99}\right)^{-\frac{1}{2}} \approx 1.41.\end{aligned}$$

The Value-at-Risk computed at quantile 99% is equal to 200 Euros, whereas Value-at-Risk computed at 99.5% is approximately 1.41 times bigger than the one computed at quantile 99%. ■

Generalized Pareto distribution

Let X_T be a random variable that follows a Generalized Pareto distribution with three parameters: location μ , scale σ , and shape ξ . Given the cumulative distribution function in Eq. 1.4 and the fact that $F_{X_T}(\text{VaR}_p(X_T)) = p$, we can obtain after few algebraic passages:

$$\text{VaR}_p(X_T) = \mu + \frac{\sigma}{\xi} \left((1 - p)^{-\xi} - 1 \right). \quad (1.7)$$

Example 2

Assume that the loss distribution has $\mu = 100$, $\sigma = 3$, and $\xi = 0.2$, compute the Value-at-Risk at 99%.

$$\text{VaR}_{0.99}(X_T) = 100 + \frac{3}{0.2} \left((1 - 0.99)^{-0.2} - 1 \right) = 122.68.$$

The Value-at-Risk computed at the quantile 99% is equal to 122.68 Euros. ■

1.2.2 Tail Conditional Expectation

A risk measure that is more and more popular is the Tail Conditional Expectation (TCE), also called Conditional Tail Expectation, Tail-VaR, or Expected Shortfall, see for example Acerbi et al. (2001).

According to the report of the National Association of Insurance Commissioners (2007), TCE measures the amount of risk within the tail of a distribution of outcomes, expressed as the probability-weighted average of the outcomes beyond a chosen point in the distribution. In other words, given a threshold (usually the VaR at level p), the TCE measures the conditional mean

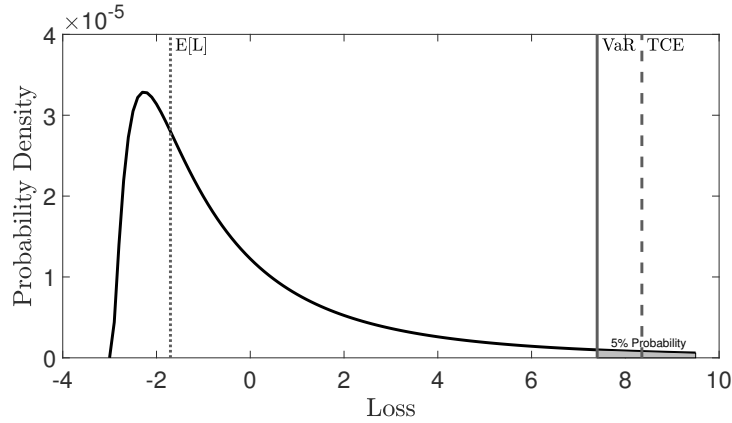


Figure 1.5: Example of loss distribution with 95% VaR and Tail Conditional Expectation.

value (Comité Européen des Assurances and Groupe Consultatif Actuariel Européen , 2007). A detailed description can be found in Society of Actuaries (2000), and the axiomatic foundation can be found in Wang and Zitikis (2021).

Definition 4. Let $F_{X_T}(x)$ be the distribution function of losses over a fixed period of time T , the TCE of the random variable X_T is defined as the expected value of all values higher than the threshold at percentile p of the distribution, denoted by:

$$\text{TCE}_p(X_T) = \mathbb{E}[X_T | X_T \geq \text{VaR}_p(X_T)].$$

If the random variable is continuous, we can compute the TCE as follows:

$$\begin{aligned} \text{TCE}_p(X_T) &= \frac{1}{1-p} \int_{\text{VaR}_p(X_T)}^{\infty} x f_{X_T}(x) dx \\ &= \frac{1}{1-p} \int_p^1 \text{VaR}_u(X_T) du. \end{aligned} \quad (1.8)$$

Figure 1.5 shows the notion of VaR and TCE together. Given a probability density function of a loss, we show with a vertical line the value of VaR at 95%, equal to 7.39€. This indicates a 5% chance that we will lose at least this amount. Moreover, it shows that the TCE computed at level p will always be at least equal to the VaR computed at the same level. In this example, the TCE is equal to 8.35€.

Pareto Type I losses

Let X_T be a random variable that follows a Pareto Type I distribution with parameters α and x_0 . Using Eq. 1.6 and Eq. 1.8, we can write that:

$$\begin{aligned} \text{TCE}_p(X_T) &= \frac{1}{1-p} \int_p^1 x_0 (1-u)^{-\frac{1}{\alpha}} du. \\ &= \frac{x_0}{1-p} \left[\frac{\alpha}{1-\alpha} (1-u)^{-\frac{1}{\alpha}+1} \right]_{u=p}^{u=1} \\ &= \frac{\alpha}{\alpha-1} \text{VaR}_p(X_T). \end{aligned} \tag{1.9}$$

We show that the TCE of a random variable that follows a Pareto Type I distribution is a multiplier of the VaR, and the size is given by the shape parameter α . We also notice that the hierarchical relation between two TCEs at different orders is the same as the hierarchical relation between two VaR at different orders.

Example 3

Assume that the loss distribution has $x_0 = 20$ and $\alpha = 2$, as in example 1. Compute the TCE at 99%.

$$\text{TCE}_{0.99}(X_T) = \frac{2}{2-1} \cdot 200 = 400$$

The Value-at-Risk computed at the quantile 99% is equal to 200 Euros as showed in example 1 and the TCE at the same level is 2 times the Value-at-Risk, i.e. 400 Euros. ■

Generalized Pareto losses

Let X_T be a random variable that follows a Generalized Pareto distribution with three parameters: location μ , scale σ , and shape ξ . Given Eq. 1.7 that

computes the VaR and Eq. 1.8 that gives the definition of TCE, we can write:

$$\begin{aligned}
\text{TCE}_p(X_T) &= \frac{1}{1-p} \int_p^1 \left(\mu + \frac{\sigma}{\xi} \left((1-u)^{-\xi} - 1 \right) \right) du. \\
&= \frac{1}{1-p} \left(\mu(1-p) + \frac{\sigma}{\xi} \left[\frac{1}{\xi-1} (1-u)^{-\xi+1} \right]_{u=p}^{u=1} - \frac{\sigma}{\xi} (1-p) \right) \\
&= \mu + \frac{\sigma}{\xi} \left(\frac{(1-p)^{-\xi}}{1-\xi} - 1 \right). \tag{1.10}
\end{aligned}$$

Example 4

Assume that the loss distribution has $\mu = 100$, $\sigma = 3$, and $\xi = 0.2$, compute the TCE at 99%.

$$\text{TCE}_{0.99}(X_T) = 100 + \frac{3}{0.2} \left(\frac{(1-0.99)^{-0.2}}{1-0.2} - 1 \right) = 132.10$$

The Value-at-Risk computed at the quantile 99% equals 122.68 Euros, as shown in example 2, and the TCE at the same level is equal to 132.10 Euros. ■

1.2.3 High-order Tail Conditional Expectation

The high-order Tail Conditional Expectation Indicator $\text{TCE}_p^{(m)}(X_T)$, also called extended TCE, is a conditional higher-order moment of the probability distribution under study X_T . This new indicator is introduced in Faroni et al. (2022).

Definition 5. Let X_T be the distribution of the outcomes over a fixed period T where losses are positive, and gains are negative values. The $\text{TCE}_p^{(m)}(X_T)$, also called TCE at order m , is defined as:

$$\text{TCE}_p^{(m)}(X_T) = \mathbb{E}[X_t^m | X_T \geq \text{VaR}_p(X_T)]. \tag{1.11}$$

As an illustration, $\text{TCE}_p^{(2)}(X_T)$ is a conditional non-central second-order moment, where the condition is that losses exceed the $1-p$ quantile. The $\text{TCE}_p^{(3)}(X_T)$ is a conditional non-central third-order moment, and so on. On another hand, when $m = 1$, $\text{TCE}_p^{(1)}(X_T)$ is the standard tail conditional expectation indicator described in Section 1.2.2.

Because

$$\mathbb{E}[X_T^m | X_T \geq VaR_c(X_T)] = \frac{\mathbb{E}[X_T^m \mathbf{1}_{X_T \geq VaR_c(X_T)}]}{Pr(X_T \geq VaR_c(X_T))},$$

the extended TCE indicator can be rewritten as follows:

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \mathbb{E}[X_T^m \mathbf{1}_{X_T \geq VaR_c(X_T)}],$$

so that

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \int_{VaR_c(X_T)}^{+\infty} x^m dF(x), \quad (1.12)$$

where $c = F(VaR_c(X_T))$.

Let us change variables as follows: $F(x) = s$, $x = F^{-1}(s) = VaR_s(X_T)$, and $ds = dF(x)$. We readily obtain a third equivalent representation of the extended TCE indicator:

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \int_c^1 (VaR_s(X_T))^m ds. \quad (1.13)$$

Note that another extended TCE indicator $\Xi^{(m)}$ can be found in the risk management literature (see for instance Barczy et al. (2022)). This indicator is defined by:

$$\Xi^{(m)} = \frac{m}{1-c} \int_c^1 \left(\frac{s-c}{1-c}\right)^{m-1} VaR_s(X_T) ds = \frac{m}{1-c} \int_c^1 \left(\frac{s-c}{1-c}\right)^{m-1} F^{-1}(s) ds.$$

If we again change variables as follows: $F(x) = s$, $x = F^{-1}(s) = VaR_s(X_T)$, and $ds = dF(x)$, we obtain:

$$\Xi^{(m)} = \frac{m}{1-c} \int_{VaR_c(X_T)}^{+\infty} \left(\frac{F(x)-c}{1-c}\right)^{m-1} x dF(x).$$

All of these expressions are distinct from equations 1.11 to 1.13 and confirm that $\Xi^{(m)}$ cannot be interpreted as a partial higher order moment, contrary to the indicator examined here.

Pareto Type I losses

Let X_T be the losses distribution that follows a Type I Pareto distribution whose parameters are α and x_0 . To compute $\text{TCE}_p^{(m)}(X_T)$, we apply the definition in Eq. 1.12 using the probability density function in Eq. 1.1 which leads us to write:

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \int_{\text{VaR}_p(X_T)}^{+\infty} x^m \frac{\alpha x_0^\alpha}{x^{\alpha+1}} dx.$$

Theorem 1. *When losses are Pareto Type I distributed, the extended TCE indicator admits the following expression:*

$$\text{TCE}_p^{(m)}(X_T) = \frac{\alpha}{\alpha - m} (\text{VaR}_p(X_T))^m \quad (1.14)$$

when $\alpha > m$.

Proof. See Appendix. □

The hierarchical ratio between two $\text{TCE}_p^{(m)}$ computed at two different quantiles is equal to the hierarchical ratio between two Value-at-Risk at different quantiles at power m , i.e.,

$$\frac{\text{TCE}_p^{(m)}(X_T)}{\text{TCE}_q^{(m)}(X_T)} = \left(\frac{(\text{VaR}_p(X_T))^m}{(\text{VaR}_q(X_T))^m} \right) = \left(\frac{1-p}{1-q} \right)^{-\frac{m}{\alpha}}.$$

Example 5

Assume that the loss distribution has $x_0 = 20$ and $\alpha = 2$, as in example 1, compute the TCE at 99% at order 2, 3, and 4.

$$\text{TCE}_{0.99}^{(2)}(X_T) = \frac{2}{2-1} \cdot 200^2 = 80,000$$

$$\text{TCE}_{0.99}^{(3)}(X_T) = \frac{2}{2-1} \cdot 200^3 = 16,000,000$$

$$\text{TCE}_{0.99}^{(4)}(X_T) = \frac{2}{2-1} \cdot 200^4 = 3,200,000,000$$

The Value-at-Risk computed at the quantile 99% equals 200 Euros, as shown in example 1, and the TCE at the same level is 400 Euros, computed in example 3. Here, we show that the second non-central high-order moment is 80,000 Euros², the third non-central high-order moment is 16,000,000 Euros³ and then the fourth non-central high-order moment is 3,200,000,000 Euros⁴. We cannot

compare their different values as they are expressed in different units. ■

Generalized Pareto Losses

Let X_T follows a generalized Pareto distribution with parameters ξ , μ , and σ , we can obtain a quasi-closed formula for the extended TCE indicator using the definition of the extended TCE at order m in Eq. 1.12 and the probability density function in Eq. 1.3.

Theorem 2. *In the GPD case, the extended TCE indicator can be computed as follows:*

$$\begin{aligned}
TCE_p^{(m)}(X_T) &= \frac{1}{1-p} \cdot \int_{VaR_p(X_T)}^{+\infty} x^m \frac{1}{\sigma} \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)} dx \\
&= \frac{1}{1-p} \left[-(-1)^{-(m-\frac{1}{\xi}+1)} \frac{1}{\xi} \frac{\Gamma\left(-m + \frac{1}{\xi}\right) \Gamma(m+1)}{\Gamma\left(1 + \frac{1}{\xi}\right)} \left(\frac{\mu\xi - \sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi}\right)^m \right. \\
&\quad \left. - (1-p) \frac{(VaR_p(X_T))^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} {}_2F_1\left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{VaR_p(X_T) \xi}{\mu\xi - \sigma}\right) \right], \tag{1.15}
\end{aligned}$$

where ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the hypergeometric function, $\Gamma(\cdot)$ is the gamma function, $0 < \xi < \frac{1}{m}$, and $VaR_p(X_T) = \mu + \frac{\sigma}{\xi} \left((1-p)^{-\xi} - 1\right)$.

Proof. See Appendix. □

To numerically solve Eq. 1.15, the following three conditions must be met:

- $0 < \xi < \frac{1}{m}$, to ensure the convergence of the integral in Eq. 1.15 and to avoid the appearance of complex numbers in Eq. 1.15.
- $(-1)^{-m-\frac{1}{\xi}+1} \cdot \left(\frac{\mu\xi-\sigma}{\sigma}\right)^{-\frac{1}{\xi}}$ has to be a real number to avoid the appearance of complex numbers in Eq. 1.15.
- $\left| \frac{VaR_p \xi}{\mu\xi - \sigma} \right| < 1$, which is a necessary property of the fourth parameter of the hypergeometric function.

Example 6

Assume that the loss distribution has $\mu = 100$, $\sigma = 3$ and $\xi = 0.2$, compute

the TCE at order 2, 3, and 4 at a percentile of 99%.

$$\text{TCE}_{0.99}^{(2)}(X_T) = 17,597$$

$$\text{TCE}_{0.99}^{(3)}(X_T) = 2,372,049$$

$$\text{TCE}_{0.99}^{(4)}(X_T) = 326,010,258$$

In this example, we see that the TCE at order 2 is equal to 17,597 Euros², the TCE at order 3 is equal to 2,372,049 Euros³ and the TCE at order 4 is equal to 326,010,258 Euros⁴. We cannot compare their values as they are expressed in different units. ■

1.2.4 Examples

In this section, we compute Value-at-Risk, TCE, and high-order TCE using the formulas obtained from the previous section in the case of Pareto Type I distribution and Generalized Pareto distribution using different parameters of either the Pareto Type I distribution or Generalized Pareto distribution. This section aims to analyze the effect of distribution parameters on the risk measures computation.

Pareto Type I Losses

Let X_T be the losses distribution that follows a Pareto type I distribution. Table 1.1 shows the computation of risk measures with different α (cases I and II), different x_0 (cases I and III), and different quantiles (cases I and IV). This confrontation allows us to interpret the Pareto Type I parameters more exhaustively.

We show in Section 1.1.1 that α is inversely proportional to extreme risks. This is also confirmed here as an increase of α is translated into a decrease in the risk measure. In this specific example, when the shape parameter halves, the TCE is 4.74 times bigger, and the extended TCE at order 4 is 150 times bigger.

Furthermore, an increase of the minimum possible value, x_0 , increases the risk measure. When x_0 is multiplied by 2, the VaR and TCE are multiplied by 2, whereas the higher-order TCE is multiplied by 2^m .

Finally, we notice that an increase in the percentile will increase the risk measures because we are considering more extreme risks in our computation.

Case	α	x_0	p	VaR_p	TCE_p	$\text{TCE}_p^{(2)}$	$\text{TCE}_p^{(3)}$	$\text{TCE}_p^{(4)}$
I	2	20	0.99	200	400	80,000	16,000	3,200,000,000
II	4	20	0.99	63	84	5,333	337,309	21,333,333
III	2	10	0.99	100	200	20,000	2,000,000	200,000,000
IV	2	20	0.995	282	565.69	160,000	45,254,834	12,800,000,000

Table 1.1: Risk Measures computed with different parameters assuming that losses follow a Pareto type I distribution. Numbers are truncated.

Case	μ	σ	ξ	p	VaR_p	TCE_p	$\text{TCE}_p^{(2)}$	$\text{TCE}_p^{(3)}$	$\text{TCE}_p^{(4)}$
I	100	3	0.2	0.99	122	132	17,597	2,372,049	326,010,258
II	100	3	0.1	0.99	117	122	15,122	1,866,586	231,086,590
III	100	6	0.2	0.99	145	164	27,551	4,784,955	892,276,963
IV	200	3	0.2	0.99	222	232	54,017	12,614,304	2,959,085,291
V	100	3	0.2	0.995	128	139	19,544	2,785,588	406,902,010

Table 1.2: Risk Measures computed with different parameters assuming that losses follow a generalized Pareto distribution. Numbers are truncated.

Generalized Pareto Losses

Let X_T follow a Generalized Pareto distribution with three parameters: location μ , scale σ , and shape ξ . Table 1.2 shows the computation of risk measures with different ξ (cases I and II), different σ (cases I and III), different μ (cases I and IV), and different quantiles (cases I and V).

We demonstrate that when ξ increases, the risk measure increases as the losses modeled are more unequal and there are more extreme losses. The parameter ξ is used to model the shape parameter, i.e., the riskiness of the market. This interpretation goes together with the parameter α of the Pareto Type I distribution.

Next, an increase in the location parameter μ and the scale parameter σ leads to an increase in the risk measures. The risk measures are more sensitive to μ than σ .

Finally, when the percentile used increases, the risk measure increases as we are more interested in extreme risks.

1.3 Risk Measures and Regulation

No unique risk measure is identified as the “best” risk measure, leading to different opinions among regulators. Even though the regulatory purpose is the same for each regulator, i.e., to make capital available for absorbing losses

occurring in events of significant financial losses that could bring insolvency, regulators only agree on using a risk measure related to the behavior of tails of the probability distribution of a firm's financial results.

Some regulators impose the use of quantiles of the distribution (Value-at-Risk), while others impose the use of partial moments (TCE). Furthermore, regulators also differ in their time horizon and confidence level choice.

A comparison of different regulatory regimes for insurance companies can be found in Comité Européen des Assurances and Mercer Oliver Wyman Limited (2005) and Comité Européen des Assurances and Groupe Consultatif Actuariel Européen (2007). The European Union's regulation is given by Solvency II, which imposes risk-based capital requirements computed using the VaR over a one-year period and with a confidence level of 99.5%. When Solvency II was designed, the banking regulation Basel II decided to adopt VaR as a risk measure for capital requirement purposes. However, the credit crisis of 2008 exposed the weaknesses of VaR, and a new system of capital regulation for banking has been developed, Basel III (see Basel Committee on Banking Supervision (2019) and Basel Committee on Banking Supervision (2022)). We refer the reader to Gatzert and Wesker (2011) for comparing Solvency II and Basel regulations.

Solvency II has become a model for risk-based capital requirements worldwide in the insurance world. However, the United States is one major exception to this trend primarily because the U.S. model was developed before the introduction of Solvency II. The U.S. model is designed around a formula provided by the regulatory body, the National Association of Insurance Commissioners (NAIC) (see also National Association of Insurance Commissioners (2007) on ORSA perspectives).

The Canadian regulation of insurance capital differs from that of the European Union, and it is based on the firm's risk assessment in the context of specific extreme events. Canada's Office of Supervision of Financial Institutions (OSFI) risk assessment process begins with an evaluation of the inherent risk within each significant activity of an insurer, and the quality of risk management applied to mitigate these risks (see Canada Office of Supervision of Financial Institutions (2022)). After considering this information, OSFI identifies the level of net risk and the rating trend for each significant activity to see if it is increasing, stable, or decreasing. The combination of the net risk of each significant activity will compute the overall net risk of the insurer. Moreover, OSFI provides additional capital requirement guidelines that must be included in the insurer's risk and solvency assessment.

In the cases of both the United States and Canada, there is significant regulatory involvement in the supervision of risk-based capital. In contrast, in the case of the European Union and Switzerland, the regulation is more

principle-based. The Swiss regulation includes the use of a capital standard together with stress-testing of specific extreme scenarios. The capital standard is based on the Tail Conditional Expectation at 99% over a period of 1 year.

Over the years, new alternatives to evaluate risks are arising due to the weaknesses of VaR and its lack of coherence (discussed in Chapter 2). Rostek (2010) defines a model of preferences in which, given beliefs about uncertain outcomes, an individual evaluates an action by a quantile of the induced distribution. Fadina et al. (2021) designed a unified axiomatic framework for risk evaluation principles that jointly quantify a random loss variable and a set of plausible probabilities. Faroni et al. (2022) provides the extended TCE described in Section 1.2.3.

Fuchs et al. (2017) show that a notion of a quantile risk measure is a natural generalization of that of a spectral risk measure and provides another view of the distortion risk measures generated by a distribution function on the unit interval. They prove several results on quantile risk measures in this setting. Finally, Denuit et al. (2006) provide a comprehensive review or modeling risk in incomplete markets, emphasizing insurance risks, expanding on and combining in a comprehensive review the existing literature on quantitative risk management.

1.4 Relation between Risk Measures

As it does not exist a general risk measure adopted worldwide, we want to understand how regulators choose such or such risk indicators for solvency computations.

Li and Wang (2022) noted that the Basel Committee on Banking Supervision proposed the shift from the 99% VaR to the 97.5% TCE for internal models in market risk assessment (see Basel Committee on Banking Supervision (2022)). Using this idea, Li and Wang introduce a new distributional index, the *Probability Equivalent Level of VaR and ES* (PELVE), which identifies the balancing point for the equivalence between VaR and TCE. This new indicator has interesting theoretical properties and can distinguish empirically heavy-tailed distributions from light-tailed ones. A generalization of the PELVE measure and the application of a high-order TCE is given by Barczy et al. (2022). We show that the extended TCE introduced in Section 1.2.3 differs from the high-order TCE presented in Barczy et al. (2022). Another indicator related to monotone risk measure is given in Fiori and Rosazza Gianin (2021).

The results described in the following sections and presented in the paper by Faroni et al. (2022) are developed independently from the previous stream of papers. We did not introduce an intermediate indicator to compare the

different risk measure quantiles.

The goal of this section is to study the relationship between different risk measure confidence intervals in order to understand the implicit utility function of insurance regulators and to understand how equivalent risk valuation systems can be put in place, but we do so without introducing any utility function. We know that the regulations put in place are not full results of a particular intent but rather of a political process, so what we determine may not be the output of an actual utility function of a specific regulator. However, making the implicit functioning of existing regulations explicit should be a valuable contribution to assuring that regulations function in an effective and efficient manner.

First, we study the relation between VaR and TCE in Section 1.4.1. Then, we study the relationship between the new high-order indicator and the VaR in Section 1.4.2. Finally, Section 1.4.3 examines equivalence results between the quantiles of extended TCE, with a specific analysis of equivalence between TCE and extended TCE at order m .

1.4.1 Relation between VaR and TCE

In this subsection, we focus on the comparison between VaR and TCE. This can be seen by comparing two regulation systems: Solvency II and the Swiss Solvency Test (SST). The SST was implemented in 2004 and preceded Solvency II, but in 2015 the European Union recognized the SST as the first regime to be fully equivalent to Solvency II. From a risk measure point of view, Solvency II imposes a capital requirement computed using VaR over a 1-year period and a confidence level of 99.5%. In contrast, SST uses TCE with a confidence level of 99% over a 1-year period. As both risk measures use a 1-year period, we can concentrate our analysis on the confidence level adopted for both risk measures.

If we compute VaR and TCE using the same quantile, TCE will always be higher or equal than VaR by construction. However, the quantiles for VaR and TCE are usually chosen to be different by regulators (see Solvency II and SST, for example). We study which relation should exist between these two quantiles. Specifically, we examine how it is possible to find c and q such that $\text{VaR}_q = \text{TCE}_c$, where $q > c$.

Let X_T be a random variable that follows a generalized Pareto distribution with three parameters as described in Section 1.1.2. $\text{VaR}_q(X_T)$ and $\text{TCE}_c(X_T)$ can be rewritten as a function of parameters as shown in Eq. 1.7 and Eq. 1.10, respectively. Thus, $\text{TCE}_c(X_T) = \text{VaR}_q(X_T)$ is equivalent to

$$\mu + \frac{\sigma}{\xi} \left(\frac{(1-c)^{-\xi}}{1-\xi} - 1 \right) = \mu + \frac{\sigma}{\xi} \left((1-q)^{-\xi} - 1 \right).$$

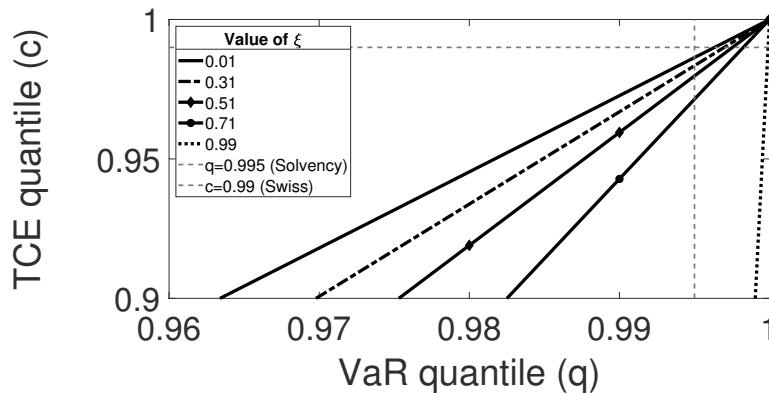


Figure 1.6: *TCE quantile as a function of VaR quantile.*

Using the above equality, we can relate the quantiles c and q as follows.

Theorem 3. *In the generalized Pareto framework, the quantile of TCE and the quantile of VaR obey the following relationship when the two risk indicators are equal:*

$$c = 1 - (1 - \xi)^{-\frac{1}{\xi}} \cdot (1 - q) \quad (1.16)$$

where $0 < \xi < 1$.

Note that Eq. 1.16 depends on ξ , but not on μ or σ . Also note that if $\xi = \frac{1}{\alpha}$, the result obtained using the generalized Pareto distribution boils down to an identical result in the subcase of Pareto Type I distribution².

Next, Figure 1.6 illustrates Theorem 3 where we plot the TCE quantile c as a function of its equivalent VaR quantile q . We plot this relation for different values of the market risk parameter ξ where ξ takes values between 0.01 and 0.99. We recall that a higher value of ξ is equivalent to an increased presence of extreme risks in the distribution used. We only concentrate on high values of VaR and TCE quantiles as we study the equivalence from a risk management point of view.

When there is more extreme risk in the market, i.e., when ξ tends to 1, the VaR quantile has to be a number close to 1, whereas the TCE quantile can take a broad range of values between 0.9 and 1.

This means that, in the presence of extreme risk, there is a large variability in the choice of the TCE quantile, making it a difficult choice for the regulator. Conversely, the VaR quantile is easy to set as the regulator only needs to choose a sufficiently high value.

²Knowing that $\text{VaR}_q(X_T)$ and $\text{TCE}_c(X_T)$ can be rewritten as Eq. 1.6 and Eq. 1.9 respectively, we can rewrite the relation between the two quantiles as $\frac{\alpha}{\alpha-1}x_0(1-c)^{-\frac{1}{\alpha}} = x_0(1-q)^{-\frac{1}{\alpha}}$ that can easily conduct to Eq. 1.16

However, we do not observe this feature when ξ is small, where many different values are considered admissible by both VaR and TCE quantiles, and where these two quantiles vary linearly.

Figure 1.6 also illustrates that many of the solutions to Theorem 3 are consistent with Solvency II and the Swiss Insurance regulations. However, the two regulations are inconsistent. For example, given the VaR quantile equal to 99.5% (Solvency II), the corresponding TCE quantile should be lower than the quantile chosen by the Swiss regulation. This would lead us to say that the implicit utility function of the Swiss regulator is more risk-averse than the one of Solvency II.

1.4.2 Relation between VaR and high-order TCE

This subsection studies the equivalence relation of high-order TCE with Value-at-Risk. This study is done using two different approaches. First, we study the quantiles that allow us to have the two risk measures equal, and then, we study the relation that allows us to have the risk measures equal, assuming that they are expressed in the same unit. This is called a homogeneous comparison, as we compare two measures with the same unit, i.e., euro with euro.

Using risk measure equivalence

We examine how it is possible to find a relation between c and q such that

$$\text{VaR}_q(X_T) = \text{TCE}_c^{(m)}(X_T).$$

Pareto Distributed Losses Suppose that X_T follows a Type I Pareto distribution where VaR is given in Eq. 1.6 and the high-order TCE is given in Eq. 1.14. We deduce that $\text{VaR}_q(X_T) = \text{TCE}_c^{(m)}(X_T)$ is equivalent to

$$\frac{\alpha}{\alpha - m} (\text{VaR}_c(X_T))^m - \text{VaR}_q(X_T) = 0. \quad (1.17)$$

This equation allows us to find the relation between the extended TCE and the VaR quantile.

Theorem 4. *In the Pareto framework, the extended TCE quantile (c) and the VaR quantile (q) obey the following relationship when the two risk indicators are equal:*

$$c = 1 - \left(\frac{(\alpha - m) x_0^{1-m}}{\alpha} \right)^{-\frac{\alpha}{m}} (1 - q)^{\frac{1}{m}}, \quad (1.18)$$

where $\alpha > m$.

Proof. See Appendix. □

For consistency with the generalized Pareto approach, we rewrite the previous equation as:

$$c = 1 - \left((1 - m\xi) x_0^{1-m} \right)^{-\frac{1}{m\xi}} (1 - q)^{\frac{1}{m}}$$

where $0 < \xi < \frac{1}{m}$ and knowing that $\xi = \frac{1}{\alpha}$.

It appears that Eq. 1.18 generalized the result of Theorem 3 in the case of Pareto Type I distribution.

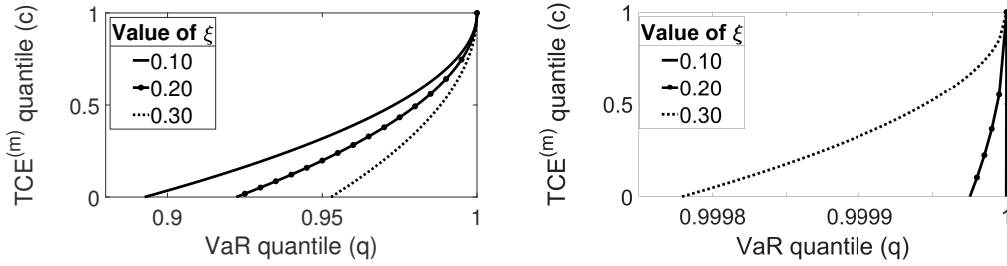


Figure 1.7: *Extended TCE quantile as a function of VaR quantile. Left panel: $m = 2$ and $x_0 = 1$. Right panel: $m = 2$ and $x_0 = 5$.*

Figure 1.7 shows the high-order TCE quantile as a function of the VaR quantile for different values of the risk parameter ξ . In the left panel, the minimum value of the distribution is equal to $x_0 = 1$, while in the right panel, the value is equal to $x_0 = 5$. Both panels assume $m = 2$.

We see that a higher value of ξ leads to more complex situations as there is more extreme risk in the market to be considered. In this case, the VaR quantile takes a value closer to one, while the high-order TCE quantile can take a broad range of values. Moreover, Figure 1.7 shows that a higher value of x_0 leads to an even more complex situation where the VaR quantile would lead to the worst-case scenario. Thus, when the Value-at-Risk cannot distinguish between extreme risk situations, a more sophisticated indicator, such as the extended TCE indicator, can produce such a distinction.

Figure 1.8 is constructed similarly to Figure 1.7, but both panels are now computed using $m = 3$. When we increase m , we are giving more importance to extreme risks, and we show that the curves are pushed to the right for higher values of m by comparing Figures 1.7 and 1.8. Thus, varying the value of m allows us to construct risk indicators that are more or less sensitive to extreme risks.

GPD Losses Suppose X_T follows a generalized Pareto distribution with its three parameters where the VaR equation is given in Eq. 1.7 and the extended

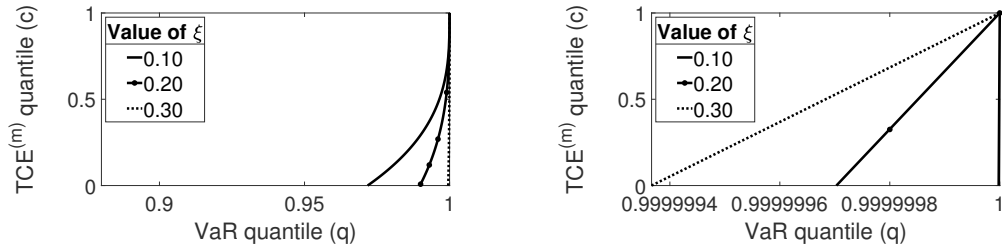


Figure 1.8: *Extended TCE quantile as a function of VaR quantile. Left panel: $m = 3$ and $x_0 = 1$. Right panel: $m = 3$ and $x_0 = 5$.*

TCE equation is given in Eq. 1.15.

Although Eq. 1.15 does not admit a closed-form solution, we numerically solve the equation:

$$\text{VaR}_q(X_T) = \text{TCE}_c^{(m)}(X_T).$$

to show the relation between the $\text{TCE}^{(m)}$ quantile and the VaR quantile, as a function of order m and of the three generalized Pareto distribution parameters ξ , μ , and σ .

Figure 1.9 shows the relation between the extended TCE and the VaR quantiles when $m = 2$ and $\sigma = 0.1$. The left panel of the figure presents the situation where the location parameter is equal to $\mu = -0.05$, while the right panel presents $\mu = 0.05$.

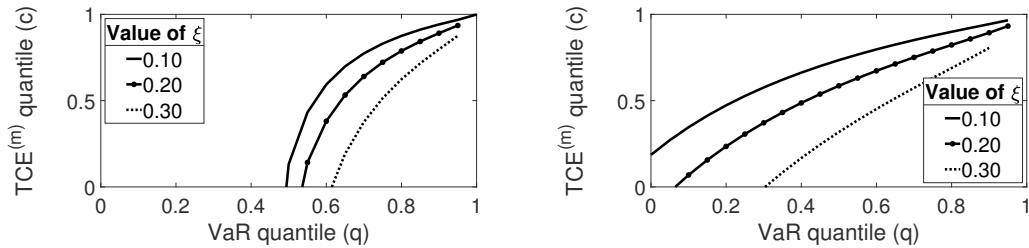


Figure 1.9: *Extended TCE quantile as a function of VaR quantile. Left panel: $m = 2$, $\sigma = 0.1$, and $\mu = -0.05$. Right panel: $m = 2$, $\sigma = 0.1$, and $\mu = +0.05$.*

Figure 1.9 shows that μ substantially impacts quantile dependences. When μ is high, the relation between indicator quantiles becomes nearly linear, which is an ideal situation from a risk management viewpoint.

Next, we plot in Figure 1.10 the relation between the high-order TCE and the VaR quantiles when $m = 2$ and $\mu = 0$. The left panel of the figure presents the situation where $\sigma = 0.1$, while the right panel presents the situation where $\sigma = 0.4$. We see that the relation is quasi-linear for small-scale parameter values, which is the ideal situation.

Finally, we plot the relation between $\text{TCE}^{(m)}$ and the VaR quantiles when

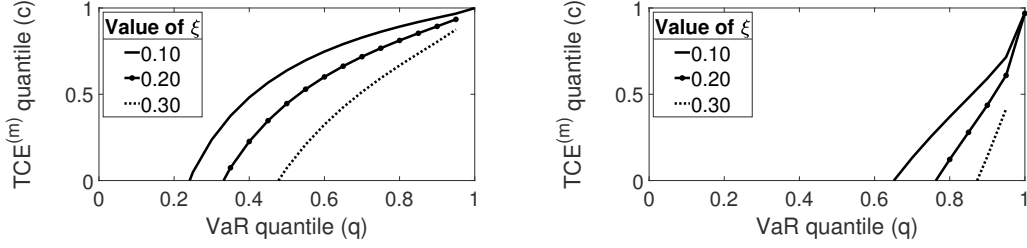


Figure 1.10: *Extended TCE quantile as a function of VaR quantile. Left panel: $m = 2$, $\sigma = 0.1$, and $\mu = 0$. Right panel: $m = 2$, $\sigma = 0.4$, and $\mu = 0$.*

$m = 3$ and $\mu = 0$ in Figure 1.11. We set $\sigma = 0.1$ in the left panel and $\sigma = 0.4$ in the right panel. Figures 1.10 and 1.11 change only in the order of the parameter m .

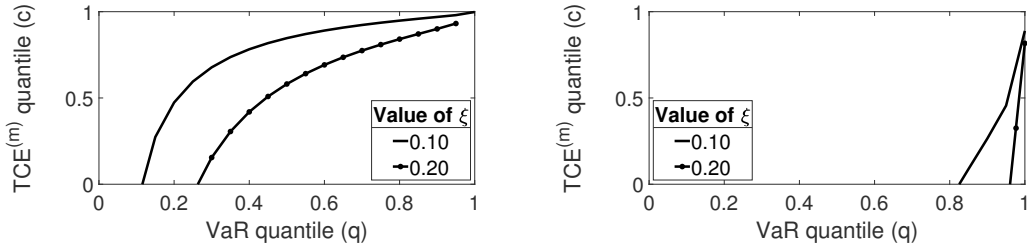


Figure 1.11: *Extended TCE quantile as a function of VaR quantile. Left panel: $m = 3$, $\sigma = 0.1$, and $\mu = 0$. Right panel: $m = 3$, $\sigma = 0.4$, and $\mu = 0$.*

From the comparison of Figures 1.10 and 1.11, we show that higher values of the order parameter m lead to a more complex situation from a risk manager point of view, that is mainly because we are giving more importance to extreme events. Moreover, the presence of extreme risk in the system makes risk management more complicated in the sense that choosing an indicator quantile becomes a more critical and sensitive decision.

Using risk measure equivalence and homogeneous comparison

In the previous subsection, we studied the quantile equivalence between two risk measures; however, VaR is expressed in a unit of money, whereas the high-order TCE is expressed in unit ^{m} of money. We study the equivalence between these two risk measures when both of them are expressed in the same unit of money, i.e.

$$\text{VaR}_q(X_T) = \left(\text{TCE}_c^{(m)}(X_T) \right)^{\frac{1}{m}}.$$

Pareto Distributed Losses Let X_T follow a Pareto Type I distribution with parameter x_0 and α . where VaR is given in Eq. 1.6 and the high-order

TCE is given in Eq. 1.14. We can write the previous relation as

$$\left(\frac{\alpha}{\alpha - m} (\text{VaR}_c(X_T))^m\right)^{\frac{1}{m}} - \text{VaR}_q(X_T) = 0. \quad (1.19)$$

Theorem 5. *In the Pareto framework, the extended TCE quantile (c) and the VaR quantile (q) obey the following relationship when the two risk indicators are equal and expressed in the same unit of money:*

$$c = 1 - \left(\frac{\alpha}{\alpha - m}\right)^{\frac{\alpha}{m}} (1 - q) \quad (1.20)$$

where $\alpha > m$.

Proof. See Appendix. □

For consistency with the generalized Pareto approach, we rewrite the previous equation as:

$$c = 1 - \left(\frac{1}{1 - m\xi}\right)^{\frac{1}{m\xi}} (1 - q)$$

where $0 < \xi < \frac{1}{m}$ and knowing that $\xi = \frac{1}{\alpha}$.

Figure 1.12 shows the relation described in Eq. 1.20, which is the relation between the high-order TCE scaled at the unit and the Value-at-Risk. The relation between the two is linear, and we note that the parameter x_0 does not play a role in this relation. We see that a higher order of m tends to push the line on the right, which means that the risk manager should consider a higher VaR quantile given the same high-order TCE quantile. As before, an increase of ξ will translate into a complex situation where the VaR quantile takes a value close to one.

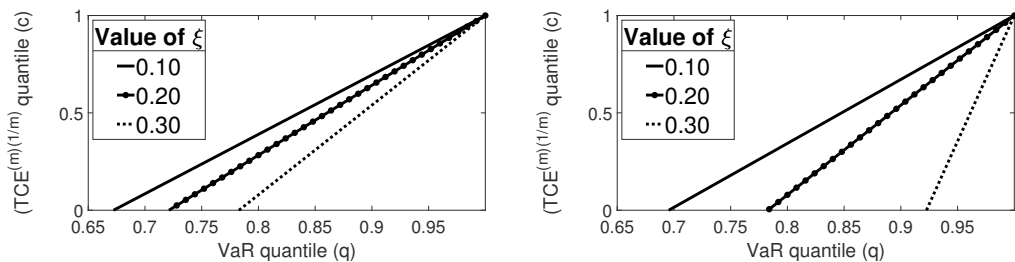


Figure 1.12: *Extended TCE quantile at power $1/m$ as a function of VaR quantile. Left panel: $m = 2$. Right panel: $m = 3$.*

GPD Losses Suppose X_T follows a generalized Pareto distribution with its three parameters where the VaR equation is given in Eq. 1.7 and the extended TCE equation is given in Eq. 1.15.

Although Eq. 1.15 does not admit closed-form solutions, we numerically solve the equation:

$$\text{VaR}_q(X_T) = \left(\text{TCE}_c^{(m)}(X_T) \right)^{\frac{1}{m}}.$$

Figure 1.13 plot the relation between the extended TCE quantile at power $1/m$ as a function of the VaR quantile. The location parameter shifts the curves to the right when it increases.

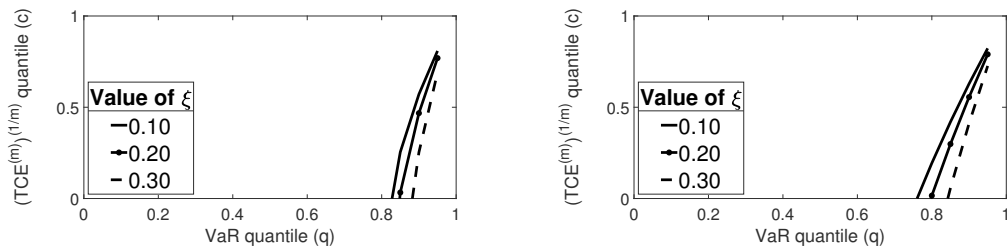


Figure 1.13: *Extended TCE quantile at power $1/m$ as a function of VaR quantile. Left panel: $m = 2$, $\sigma = 0.1$, and $\mu = -0.05$. Right panel: $m = 2$, $\sigma = 0.1$, and $\mu = +0.05$.*

Next, we plot in Figure 1.14 the same relation when $m = 2$ and $\mu = 0$. The left panel of the figure presents the situation where $\sigma = 0.1$, while the right panel presents the situation where $\sigma = 0.4$. The figure shows that, even though it is not clear from the equation, the parameter σ does not influence the relation between the extended TCE quantile at power $1/m$ and the VaR quantile.

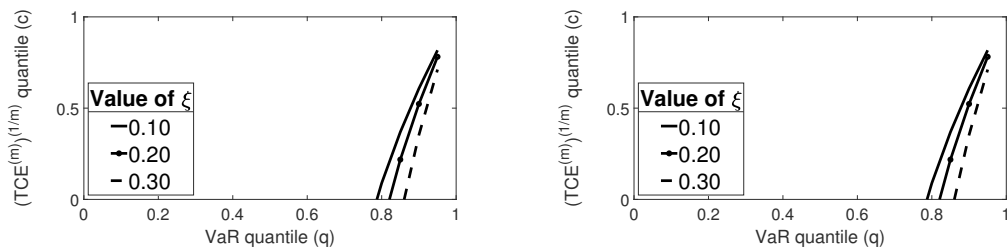


Figure 1.14: *Extended TCE quantile power $1/m$ as a function of VaR quantile. Left panel: $m = 2$, $\sigma = 0.1$, and $\mu = 0$. Right panel: $m = 2$, $\sigma = 0.4$, and $\mu = 0$.*

Finally, Figure 1.15 shows the relationship when m is bigger compared to Figure 1.14. We show that when m increases, the VaR quantile is closer to 1, and the decision of the quantile is tedious for the risk manager.

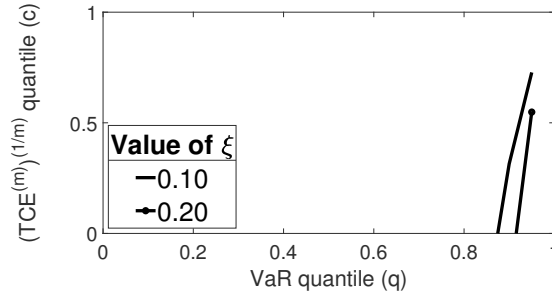


Figure 1.15: Extended TCE quantile power $1/m$ as a function of VaR quantile with $m = 3$, $\sigma = 0.1$, and $\mu = 0$.

1.4.3 Relation between high-order TCEs

In this subsection, we study the relation between the quantiles $q^{(m)}$ and $q^{(n)}$ of distinct extended tail conditional expectation indicators, where each indicator is associated with a different order m or n . This study is done using two different approaches. First, we study the quantiles that allow us to have the risk measures equal, and then we study the relation that allows us to have the two risk measures equal and expressed with the same unit. This second approach is identified as a homogeneous comparison.

For both situations we will study the sub-case where TCE is compared with a high-order TCE.

Using risk measure equivalence

We study

$$\text{TCE}_{q^{(m)}}^{(m)}(X_T) = \text{TCE}_{q^{(n)}}^{(n)}(X_T).$$

Pareto Distributed Losses Suppose we model losses as a Type I Pareto distribution, we can compute $\text{TCE}_{q^{(m)}}^{(m)}$ using Eq. 1.14 where $\alpha > m$. Then, we can compare two different extended TCE as follows:

$$\frac{\alpha}{\alpha - m} \left(\text{VaR}_{q^{(m)}}(X_T) \right)^m = \frac{\alpha}{\alpha - n} \left(\text{VaR}_{q^{(n)}}(X_T) \right)^n,$$

where

$$\text{VaR}_{q^{(m)}}(X_T) = x_0 \left(1 - q^{(m)} \right)^{-\frac{1}{\alpha}}.$$

We obtain:

$$\frac{\alpha}{\alpha - m} x_0^m \left(1 - q^{(m)} \right)^{-\frac{m}{\alpha}} = \frac{\alpha}{\alpha - n} x_0^n \left(1 - q^{(n)} \right)^{-\frac{n}{\alpha}},$$

which leads us to:

$$q^{(n)} = 1 - \left(\frac{\alpha - n}{\alpha - m} x_0^{m-n} \right)^{-\frac{\alpha}{n}} \left(1 - q^{(m)} \right)^{-\frac{m}{n}} \quad (1.21)$$

where $\alpha > m$, $\alpha > n$, and $x_0 > 0$.

Then, we denote $\xi = \frac{1}{\alpha}$, we can rewrite the previous equation as in the following theorem.

Theorem 6. *When losses follow a classic Pareto distribution, the quantiles of high order TCEs that solve Eq. 1.4.3 can be related as follows:*

$$q^{(n)} = 1 - \left(\frac{\frac{1}{\xi} - m}{\frac{1}{\xi} - n} \right)^{\frac{1}{\xi n}} x_0^{\frac{n-m}{\xi n}} \left(1 - q^{(m)} \right)^{-\frac{m}{n}}, \quad (1.22)$$

where $0 < \xi < \frac{1}{m}$, $0 < \xi < \frac{1}{n}$, and $x_0 > 0$.

We now illustrate this theorem.

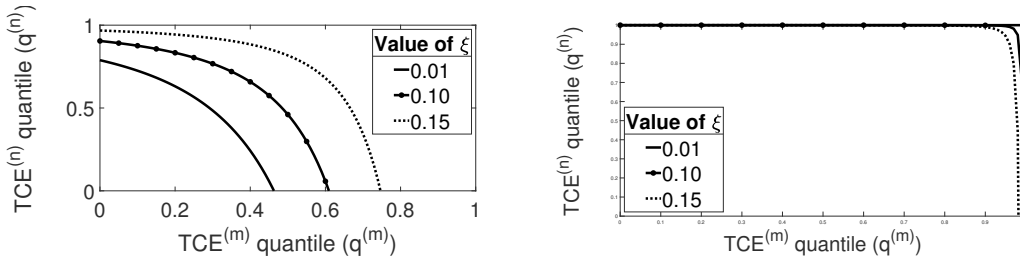


Figure 1.16: *Extended TCE quantile at order $n = 2$ as a function of extended TCE quantile at order $m = 5$. Left panel: $x_0 = 1$. Right panel: $x_0 = 2$.*

Figure 1.16 plots the relation between the quantiles $q^{(m)}$ and $q^{(n)}$ when $m = 5$ and $n = 2$. The left panel of the figure shows the situation where $x_0 = 1$, while the right panel shows the situation where $x_0 = 2$.

The left panel of Figure 1.16 can be interpreted as follows. For instance, the relation between the high-order quantiles is countermonotonic, contrary to the relation between the TCE and VaR quantiles. This means that a high value of $q^{(m)}$ corresponds to a small value of $q^{(n)}$, and conversely.

This feature is a consequence of high-order TCEs concentrating on different parts of probability tails. Thus, the figure shows us that a manager that reduces high-order extreme risks at a given order, say m , is not simultaneously reducing high-order extreme risks at another order, say n . The right panel of the figure tells us that this aspect is even more pronounced for higher values of x_0 .

We now come to the specific case where $n = 1$, that is to the study of the relation between TCE and a higher order TCE:

$$\frac{\alpha}{\alpha - 1} (\text{VaR}_q(X_T))^1 = \frac{\alpha}{\alpha - m} (\text{VaR}_{q^{(m)}}(X_T))^m .$$

Eq. 1.21 becomes

$$q = 1 - \left(\frac{\alpha - m}{(\alpha - 1) x_0^{m-1}} \right)^\alpha (1 - q^{(m)})^{-m} ,$$

when $\alpha > m$ and $x_0 > 0$. Similarly, Eq. 1.22 becomes

$$q = 1 - \left(\frac{\frac{1}{\xi} - m}{\left(\frac{1}{\xi} - 1\right) x_0^{m-1}} \right)^{\frac{1}{\xi}} (1 - q^{(m)})^{-m}$$

when $0 < \xi < \frac{1}{m}$ and $x_0 > 0$.

Figure 1.17 shows the relation between the TCE and high-order TCE quantile when $m = 2$. The left panel of the figure shows the situation where $x_0 = 1$, while the right panel shows the situation where $x_0 = 2$.

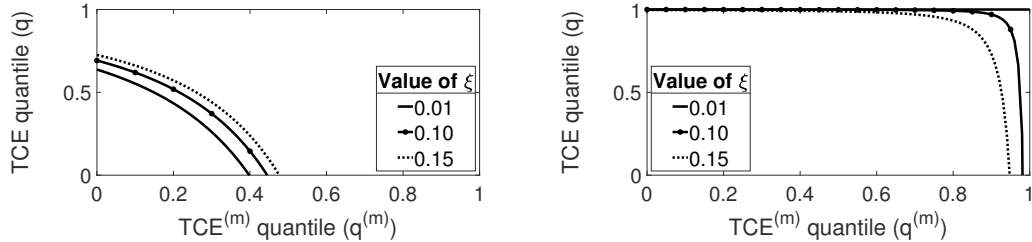


Figure 1.17: *TCE quantile as a function of extended TCE quantile at order 2. Left panel: $x_0 = 1$. Right panel: $x_0 = 2$.*

Figure 1.17 confirms the results of Figure 1.16. Reducing risks using TCE does not necessarily reduce risks as measured by a high order TCE, and conversely. Again, this effect is more pronounced for higher values of x_0 .

GPD Losses Let us now consider to the more general situation where losses are modeled using a generalized Pareto distribution. Our goal is to solve Eq. 1.4.3 when the extended TCE indicator $\text{TCE}_{q^{(m)}}^{(m)}$ is given by Eq. 1.15. Thus,

to derive the relation between $q^{(m)}$ and $q^{(n)}$, we numerically solve:

$$\begin{aligned}
& \frac{1}{1 - q^{(m)}} \left[-(-1)^{-(m - \frac{1}{\xi} + 1)} \frac{1}{\xi} \frac{\Gamma(-m + \frac{1}{\xi}) \Gamma(m + 1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi}\right)^m \right. \\
& \quad \left. - (1 - q^{(m)}) \frac{(\text{VaR}_{q^{(m)}}(X_T))^{m+1}}{(m + 1) \cdot (\sigma - \mu\xi)} {}_2F_1\left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{\text{VaR}_{q^{(m)}}(X_T) \xi}{\mu\xi - \sigma}\right) \right] \\
& = \frac{1}{1 - q^{(n)}} \left[-(-1)^{-(n - \frac{1}{\xi} + 1)} \frac{1}{\xi} \frac{\Gamma(-n + \frac{1}{\xi}) \Gamma(n + 1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi}\right)^n \right. \\
& \quad \left. - (1 - q^{(n)}) \frac{(\text{VaR}_{q^{(n)}}(X_T))^{n+1}}{(n + 1) \cdot (\sigma - \mu\xi)} {}_2F_1\left(1, n - \frac{1}{\xi} + 1; n + 2; \frac{\text{VaR}_{q^{(n)}}(X_T) \xi}{\mu\xi - \sigma}\right) \right] \tag{1.23}
\end{aligned}$$

Figure 1.18 plots the relation between the quantiles $q^{(m)}$ and $q^{(n)}$ when $m = 5$, $n = 2$ and $\sigma = 0.1$. The left panel of the figure shows the situation where $\mu = -0.05$, while the right panel of the figure shows the situation where $\mu = 0.05$.

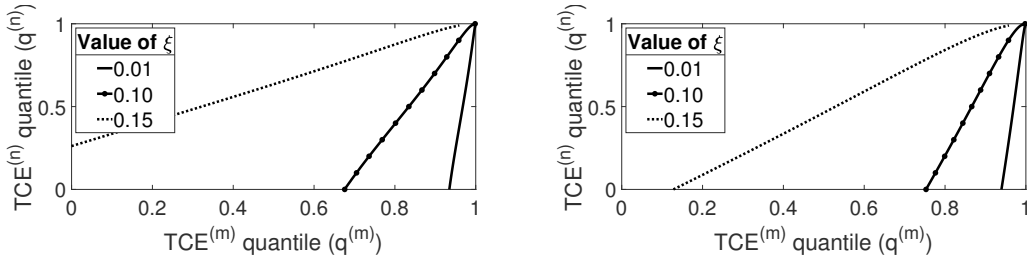


Figure 1.18: Extended TCE quantile at order $n = 2$ as a function of extended TCE quantile at order $m = 5$ with $\sigma = 0.1$. Left panel: $\mu = -0.05$. Right panel: $\mu = +0.05$.

From Figure 1.18, we deduce that the link between the high-order TCE quantiles is linear when $\sigma = 0.1$ so that this parameter of the GPD distribution is not problematic. By comparing the two panels of the figure, we see that the parameter μ has little effect on the curves linking the high-order TCE quantiles.

Figure 1.19 plots the relation between the quantiles $q^{(m)}$ and $q^{(n)}$ when $m = 5$, $n = 2$, and $\mu = 0$. The left panel of the figure shows the situation where $\sigma = 0.1$, while the right panel shows the situation where $\sigma = 0.4$.

Figure 1.19 shows us that high values of σ can yield problematic links between the high-order TCE quantiles, hinting at probability tails that are

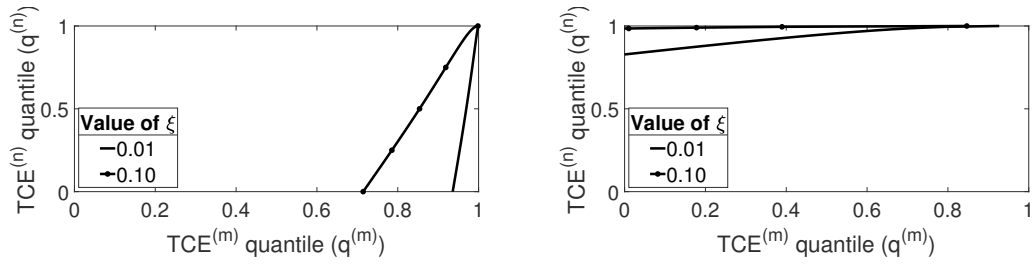


Figure 1.19: *Extended TCE quantile at order $n = 2$ as a function of extended TCE quantile at order $m = 5$ with $\mu = 0$. Left panel: $\sigma = 0.1$. Right panel: $\sigma = 0.4$.*

quantified differently by different high-order TCE indicators.

We now come to the specific case where $n = 1$ is the study of the relation between TCE and a higher-order TCE. In that case, the solutions are also numerically obtained by solving Eq. 1.23.

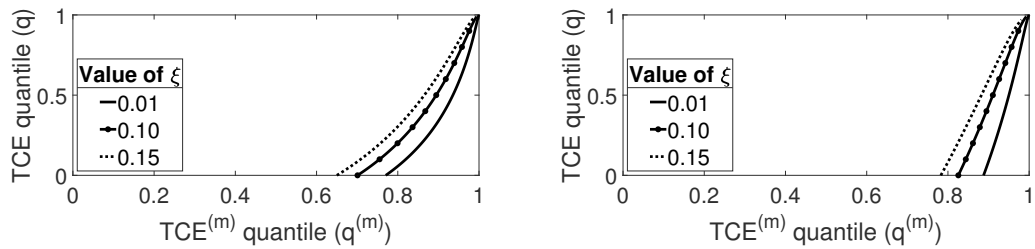


Figure 1.20: *TCE quantile as a function of extended TCE quantile at order $m = 2$ with $\sigma = 0.1$. Left panel: $\mu = -0.05$. Right panel: $\mu = +0.05$.*

Figure 1.20 plots the relation between the quantiles $q^{(m)}$ and q when $m = 2$ and $\sigma = 0.1$. The left panel of the figure shows the situation where $\mu = -0.05$, while the right panel shows the situation where $\mu = 0.05$.

Figure 1.20 tells us that the link between the second-order TCE quantile and the TCE quantile is close-to-linear when $\sigma = 0.1$, so that, again, this parameter of the GPD distribution is not problematic when it is not set to a high value.

By comparing the two panels of Figure 1.20, we see, as in Figure 1.18, that large variations of the parameter μ have a pretty limited impact on the position of the curves relating a high order TCE quantile to the TCE quantile.

Figure 1.21 plots the relation between the quantiles $q^{(m)}$ and q when $m = 2$ and $\mu = 0$. The left panel of the figure shows the situation where $\sigma = 0.1$, while the right panel shows the situation where $\sigma = 0.4$.

Figure 1.21 confirms the conclusion of Figure 1.19. Specifically, high values of σ can yield problematic links between a high-order TCE quantile and the TCE quantile.

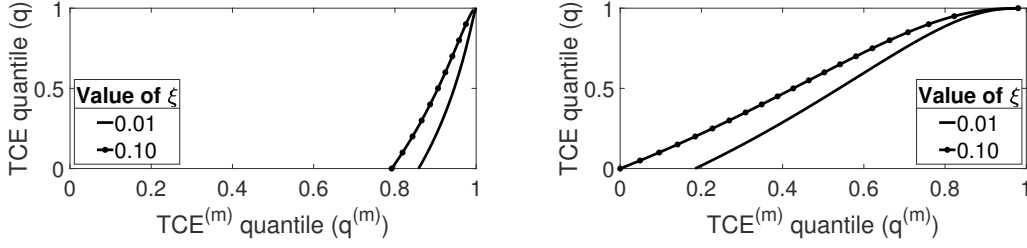


Figure 1.21: *TCE quantile as a function of extended TCE quantile at order $m = 2$ with $\mu = 0$. Left panel: $\sigma = 0.1$. Right panel: $\sigma = 0.4$.*

Using risk measure equivalence and homogeneous comparison

We also study the situation where we compare the two risk measure using the same unit, i.e.

$$\left(\text{TCE}_{q^{(m)}}^{(m)}(X_T)\right)^{\frac{1}{m}} = \left(\text{TCE}_{q^{(n)}}^{(n)}(X_T)\right)^{\frac{1}{n}}. \quad (1.24)$$

Pareto Distributed Losses Suppose we model losses as a Type I Pareto distribution, we can compute $\text{TCE}_{q^{(m)}}^{(m)}$ using Eq. 1.14 where $\alpha > m$. Then, we can compute Eq. 1.24 as follows:

$$\left(\frac{\alpha}{\alpha - m} \left(\text{VaR}_{q^{(m)}}(X_T)\right)^m\right)^{\frac{1}{m}} = \left(\frac{\alpha}{\alpha - n} \left(\text{VaR}_{q^{(n)}}(X_T)\right)^n\right)^{\frac{1}{n}},$$

where

$$\text{VaR}_{q^{(m)}}(X_T) = x_0 \left(1 - q^{(m)}\right)^{-\frac{1}{\alpha}}.$$

We obtain:

$$\left(\frac{\alpha}{\alpha - m}\right)^{\frac{1}{m}} \left(1 - q^{(m)}\right)^{-\frac{1}{\alpha}} = \left(\frac{\alpha}{\alpha - n}\right)^{\frac{1}{n}} \left(1 - q^{(n)}\right)^{-\frac{1}{\alpha}},$$

which leads us to:

$$q^{(n)} = 1 - \left(\frac{\alpha}{\alpha - n}\right)^{\frac{\alpha}{n}} \left(\frac{\alpha}{\alpha - m}\right)^{-\frac{\alpha}{m}} \left(1 - q^{(m)}\right) \quad (1.25)$$

where $\alpha > m$ and $\alpha > n$ and we note that the equation does not depends on x_0 .

Then, we denote $\xi = \frac{1}{\alpha}$, we can rewrite the previous equation as in the following theorem.

Theorem 7. *When losses follow a classic Pareto distribution, the quantiles of*

high order TCEs that solve Eq. 1.24 can be related as follows:

$$q^{(n)} = 1 - \frac{1}{1 - n\xi} (1 - m\xi)^{\frac{1}{m\xi}} (1 - q^{(m)}) \quad (1.26)$$

where $0 < \xi < \frac{1}{m}$ and $0 < \xi < \frac{1}{n}$.

Figure 1.22 shows Eq. 1.26. In the left panel, we plot the extended TCE quantile at order $n = 2$ as a function of extended TCE quantile at order $m = 5$, while in the right panel, the TCE quantile as a function of extended TCE quantile at order $m = 2$. In the Pareto Type I distribution case, the relation is linear and does not depend on the minimum value of the distribution, x_0 . This is an ideal situation for the regulator as the two quantiles are linearly related, and they only depend on the riskiness of the market, i.e., ξ .

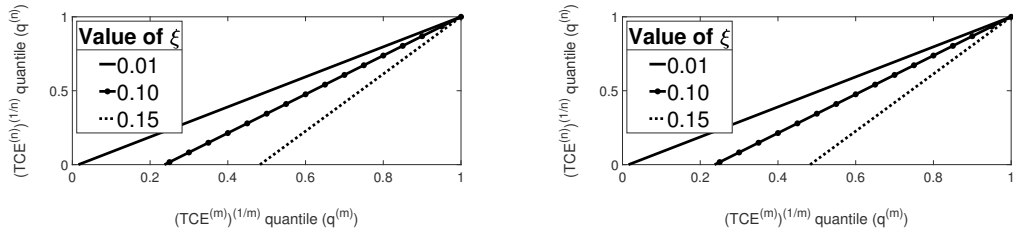


Figure 1.22: Left panel: Extended TCE quantile at order $n = 2$ as a function of extended TCE quantile at order $m = 5$. Right panel: TCE quantile as a function of extended TCE quantile at order $m = 2$.

GPD Losses In a more general situation, we assume losses follow a generalized Pareto distribution. The aim is to solve Eq. 1.24 when the extended TCE indicator $\text{TCE}_{q^{(m)}}^{(m)}$ is given by Eq. 1.15. For this reason, to derive the relation between $q^{(m)}$ and $q^{(n)}$, we need to numerically solve:

$$\begin{aligned}
& \left(\frac{1}{1 - q^{(m)}} \left[- (-1)^{-(m - \frac{1}{\xi} + 1)} \frac{1}{\xi} \frac{\Gamma(-m + \frac{1}{\xi}) \Gamma(m + 1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma} \right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi} \right)^m \right. \right. \\
& \quad \left. \left. - (1 - q^{(m)}) \frac{(\text{VaR}_{q^{(m)}}(X_T))^{m+1}}{(m + 1) \cdot (\sigma - \mu\xi)} {}_2F_1 \left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{\text{VaR}_{q^{(m)}}(X_T) \xi}{\mu\xi - \sigma} \right) \right] \right)^{\frac{1}{m}} \\
& = \left(\frac{1}{1 - q^{(n)}} \left[- (-1)^{-(n - \frac{1}{\xi} + 1)} \frac{1}{\xi} \frac{\Gamma(-n + \frac{1}{\xi}) \Gamma(n + 1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma} \right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi} \right)^n \right. \right. \\
& \quad \left. \left. - (1 - q^{(n)}) \frac{(\text{VaR}_{q^{(n)}}(X_T))^{n+1}}{(n + 1) \cdot (\sigma - \mu\xi)} {}_2F_1 \left(1, n - \frac{1}{\xi} + 1; n + 2; \frac{\text{VaR}_{q^{(n)}}(X_T) \xi}{\mu\xi - \sigma} \right) \right] \right)^{\frac{1}{n}} \\
& \hspace{15em} (1.27)
\end{aligned}$$

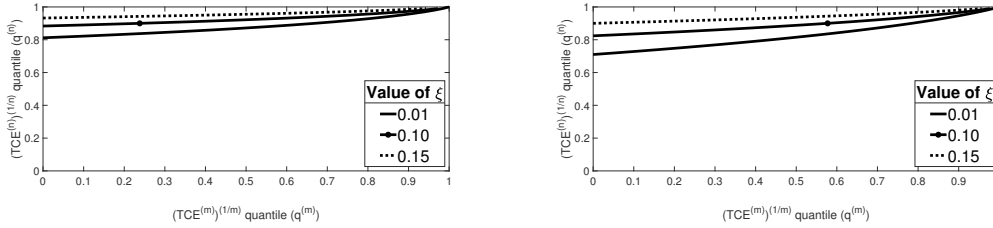


Figure 1.23: Extended TCE quantile at power $1/n$ at order $n = 2$ as a function of extended TCE quantile at power $1/m$ at order $m = 5$ with $\sigma = 0.1$. Left panel: $\mu = -0.05$. Right panel: $\mu = +0.05$.

Figure 1.23 plots the relation between the quantiles when $m = 5$, $n = 2$, and $\sigma = 0.1$. The left panel shows the situation where $\mu = -0.05$, while the right panel of the figure shows the situation where $\mu = 0.05$. We can deduce from this figure that the parameter μ has little effect on the relation linking the high-order TCE quantiles.

Then, Figure 1.24 shows the relationship in Eq. 1.27 when $m = 5$, $n = 2$, and $\mu = 0$. In the left panel, $\sigma = 0.1$, while in the right panel, we plot the situation where $\sigma = 0.4$. Even though it is not immediate to see from the equation, the parameter σ has no effect on this relation.

Finally, Figure 1.25 shows the study of the relation between TCE and a higher-order TCE at power $1/m$ when $m = 2$. The figure shows the situation where $\sigma = 0.1$ and $\mu = -0.05$ in the left panel, while the right panel shows $\mu = 0.05$. As noticed before, a variation of the parameter μ has a pretty limited

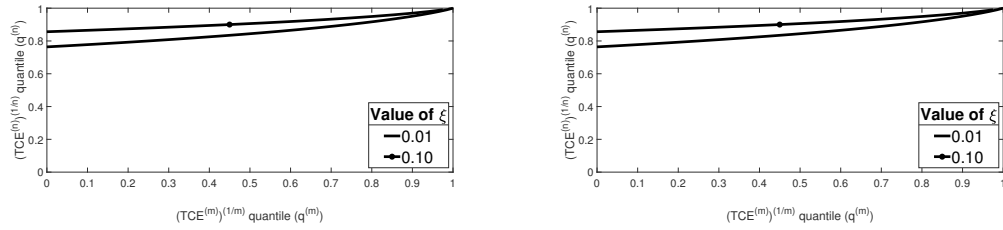


Figure 1.24: *Extended TCE quantile at power $1/n$ at order $n = 2$ as a function of extended TCE quantile at power $1/m$ at order $m = 5$ with $\mu = 0$. Left panel: $\sigma = 0.1$. Right panel: $\sigma = 0.4$.*

impact on the position of the curves relating a high-order TCE quantile to the TCE quantile.

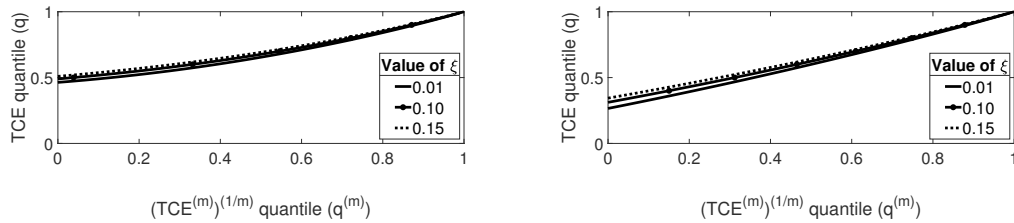


Figure 1.25: *TCE quantile as a function of extended TCE quantile at power $1/m$ at order $m = 2$ with $\sigma = 0.1$. Left panel: $\mu = -0.05$. Right panel: $\mu = 0.05$.*

1.5 Conclusion

We end this chapter with a brief illustration based on actual data from a fire insurance claim data set, labeled “beaonre” within the R package CASdatasets. This dataset includes 1823 observations of fire insurance claims from the year 1997. We transform this dataset of claim costs into a dataset of reimbursements from which we can compute VaR, TCE, and $\text{TCE}^{(m)}$. In order to compare the extended TCE with the VaR, we need to adjust the former quantity in terms of scale. For instance, we may want to solve $\text{VaR}_q(X_T) = \left(\text{TCE}_c^{(m)}(X_T)\right)^{\frac{1}{m}}$. While exact solutions to this equation do not exist, we can still deduce a relation between VaR and $\text{TCE}^{(m)}$.

We show in Figure 1.26 the link between VaR and the high-order TCE indicator (computed with $m = 2$) in the case of fire data. This figure is consistent with the theoretical results shown, for instance, in Figure 1.12. Thus, Figure 1.26 confirms, in passing, the relevance of the GPD assumption.

Further illustration with actual data is out of the present chapter's scope but could be an interesting extension.

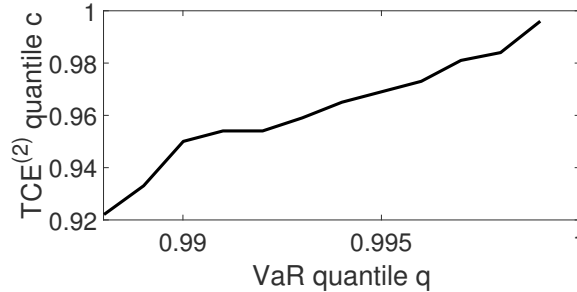


Figure 1.26: *Extended TCE quantile, with $m = 2$, as a function of VaR.*

To conclude, after giving a significant overview of the distribution used in this chapter and the definition of the classic risk measures, we introduce our risk indicator called high order TCE risk measure. We compare the quantiles of this indicator to the quantiles of VaR, TCE, and high-order TCEs to illustrate the interplay between regulators' implicit choices of risk measures and the characteristics of probability distribution tails.

Among the possible extensions of the results described in this chapter is the analysis of coherence of the high-order indicator presented here, which is detailed in Chapter 2. Another possible extension could be examining the relation between high-order TCEs when the probability distribution admits tails that are not modeled using the generalized Pareto distribution but using, for instance, the semi-heavy tails of infinitely divisible probability distributions. Finally, it could be interesting to examine the stability of high-order TCE indicators (see, for instance, the discussion in Le Courtois et al. (2020) on the cross-stability of second and fourth-order moments).

Appendix

Solution of Minimization Problem 1.5

To solve Eq. 1.5, we need to satisfy two conditions:

- First order condition:

$$\frac{\partial f(Q[X_T])}{\partial Q[X_T]} = 0,$$

- Second order condition:

$$\frac{\partial^2 f(Q[X_T])}{\partial Q[X_T]^2} > 0,$$

where $f(Q[X_T]) = \eta \cdot \mathbb{E}[(X_T - Q[X_T])_+] + (1 - \eta) Q[X_T] \epsilon$.

Thus, we solve the FOC as follows:

$$\begin{aligned} \frac{\partial}{\partial Q[X_T]} \{ \eta \cdot \mathbb{E}[(X_T - Q[X_T])_+] + (1 - \eta) Q[X_T] \epsilon \} &= 0, \\ \eta \cdot \frac{\partial}{\partial Q[X_T]} \left\{ \int_{Q[X_T]}^{\infty} (1 - F_{X_T}(x)) \right\} + (1 - \eta) \epsilon &= 0. \end{aligned}$$

We obtain:

$$\eta \cdot (-1 + F_{X_T}(Q[X_T])) + (1 - \eta) \epsilon = 0$$

or, equivalently,

$$Q[X_T] = F_{X_T}^{-1} \left(1 - \frac{1 - \eta}{\eta} \epsilon \right).$$

Finally, we have:

$$Q[X_T] = \text{VaR}_{(1 - \frac{1 - \eta}{\eta} \epsilon)}(X_T),$$

which is our result.

To make sure that the result is a local minimum, the need to check the second order condition. We compute the second derivative of the objective function,

$$\frac{\partial^2 f(Q[X_T])}{\partial Q[X_T]^2} = \frac{\partial}{\partial Q[X_T]} \eta \cdot (-1 + F_{X_T}(Q[X_T])) + (1 - \eta) \epsilon,$$

so that

$$\frac{\partial^2 f(Q[X_T])}{\partial Q[X_T]^2} = \eta \frac{\partial}{\partial Q[X_T]} (-1 + F_{X_T}(Q[X_T])),$$

or

$$\frac{\partial^2 f(Q[X_T])}{\partial Q[X_T]^2} = \eta \left[\frac{\partial}{\partial Q[X_T]} \int_{-\infty}^{Q[X_T]} f_{X_T}(x) dx \right].$$

Finally, we can check that

$$\frac{\partial^2 f(Q[X_T])}{\partial Q[X_T]^2} = \eta [f_{X_T}(Q[X_T])] > 0,$$

where the second order condition is true because $\eta > 0$ and $f_{X_T}(x) > 0$ for all $x \in \mathcal{R}$.

Proof of Theorem 1

Our goal is to solve

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \cdot \int_{\text{VaR}_p(X_T)}^{+\infty} x^m f_{X_T}(x) dx,$$

where $f_{X_T}(x) = \frac{\alpha x_0^\alpha}{x^{\alpha+1}}$.

We have:

$$\begin{aligned} \text{TCE}_p^{(m)}(X_T) &= \frac{1}{1-p} \cdot \int_{\text{VaR}_p(X_T)}^{+\infty} x^m \frac{\alpha x_0^\alpha}{x^{\alpha+1}} dx, \\ &= \frac{1}{1-p} \left[-\frac{\alpha x_0^\alpha}{\alpha-m} \cdot x^{m-\alpha} \right]_{\text{VaR}_p(X_T)}^{+\infty}. \end{aligned}$$

The quantity $\lim_{x \rightarrow +\infty} \left(-\frac{\alpha x_0^\alpha}{\alpha-m} \cdot x^{m-\alpha} \right)$ converges to zero only if $m - \alpha < 0$. Assuming that $\alpha > m$, we obtain:

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \frac{\alpha x_0^\alpha}{\alpha-m} \cdot (\text{VaR}_p(X_T))^{m-\alpha}.$$

Next, we write:

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \frac{\alpha x_0^\alpha}{\alpha-m} (\text{VaR}_p(X_T))^{-\alpha} \cdot (\text{VaR}_p(X_T))^m,$$

and we replace VaR with its expression:

$$\text{VaR}_p(X_T) = x_0 (1-p)^{-\frac{1}{\alpha}},$$

to obtain:

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \frac{\alpha x_0^\alpha}{\alpha - m} \left(x_0 \cdot (1-p)^{-\frac{1}{\alpha}} \right)^{-\alpha} \cdot (\text{VaR}_p(X_T))^m.$$

Finally, we have:

$$\text{TCE}_p^{(m)}(X_T) = \frac{\alpha}{\alpha - m} \cdot (\text{VaR}_p(X_T))^m,$$

which is our result.

Proof of Theorem 2

The aim of this appendix is to demonstrate that the integral:

$$\text{TCE}_p^{(m)}(X_T) = \frac{1}{1-p} \int_{\text{VaR}_p(X_T)}^{+\infty} x^m \frac{1}{\sigma} \left(1 + \frac{\xi(x-\mu)}{\sigma} \right)^{-\frac{1}{\xi}-1} dx$$

can be computed to provide the result in Eq. 1.15.

Using classic results on special functions (see for instance Lebedev (1972)), we rewrite the integral as follows:

$$\begin{aligned} \text{TCE}_p^{(m)}(X_T) &= \frac{1}{1-p} \left[\frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(x-\mu)}{\sigma} \right)^{-\frac{1}{\xi}} \right. \\ &\quad \left. \times {}_2F_1 \left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{x\xi}{\mu\xi - \sigma} \right) \right]_{\text{VaR}_p(X_T)}^{+\infty}. \end{aligned} \quad (1.28)$$

To compute the limit when x tends to infinity of the quantity J defined by:

$$J = \frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(x-\mu)}{\sigma} \right)^{-\frac{1}{\xi}} {}_2F_1 \left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{x\xi}{\mu\xi - \sigma} \right),$$

we rewrite the hypergeometric function using a linear transformation:

$$\begin{aligned}
& {}_2F_1\left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{x\xi}{\mu\xi - \sigma}\right) \\
&= \frac{\Gamma\left(m - \frac{1}{\xi}\right)\Gamma(m + 2)}{\Gamma\left(m - \frac{1}{\xi} + 1\right)\Gamma(m + 1)} \left(-\frac{x\xi}{\mu\xi - \sigma}\right)^{-1} {}_2F_1\left(1, -m; -m + \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right) \\
&+ \frac{\Gamma\left(-m + \frac{1}{\xi}\right)\Gamma(m + 2)}{\Gamma(1)\Gamma\left(1 + \frac{1}{\xi}\right)} \left(-\frac{x\xi}{\mu\xi - \sigma}\right)^{-(m - \frac{1}{\xi} + 1)} \\
&\times {}_2F_1\left(m - \frac{1}{\xi} + 1, -\frac{1}{\xi}; m - \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right). \tag{1.29}
\end{aligned}$$

Thus, J can be rewritten as follows:

$$\begin{aligned}
J &= K x^m \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} {}_2F_1\left(1, -m; -m + \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right) \\
&+ L x^{\frac{1}{\xi}} \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} {}_2F_1\left(m - \frac{1}{\xi} + 1, -\frac{1}{\xi}; m - \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right),
\end{aligned}$$

where K and L are functions of the parameters that are independent of x . Specifically,

$$K = \frac{\Gamma\left(m - \frac{1}{\xi}\right)}{\Gamma\left(m - \frac{1}{\xi} + 1\right) \xi}$$

and

$$L = \frac{\Gamma\left(-m + \frac{1}{\xi}\right)\Gamma(m + 1)}{\Gamma\left(1 + \frac{1}{\xi}\right)} \frac{1}{(\sigma - \mu\xi)} \left(\frac{\xi}{\sigma - \mu\xi}\right)^{-(m - \frac{1}{\xi} + 1)}.$$

Then, we use the fact that

$$\begin{aligned}
& \lim_{x \rightarrow +\infty} {}_2F_1\left(1, -m; -m + \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right) = \\
& \lim_{x \rightarrow \infty} \left[\left(\frac{\mu\xi - \sigma}{x\xi}\right)^0 + \frac{-m}{-m + \frac{1}{\xi} + 1} \frac{\mu\xi - \sigma}{x\xi} + \dots \right] = 1
\end{aligned}$$

and

$$\begin{aligned} & \lim_{x \rightarrow +\infty} {}_2F_1 \left(m - \frac{1}{\xi} + 1, -\frac{1}{\xi}; m - \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi} \right) = \\ & = \lim_{x \rightarrow \infty} \left[\left(\frac{\mu\xi - \sigma}{x\xi} \right)^0 + \frac{-\frac{1}{\xi} \left(m - \frac{1}{\xi} + 1 \right) \mu\xi - \sigma}{m - \frac{1}{\xi} + 1} \frac{\mu\xi - \sigma}{x\xi} + \dots \right] = 1, \end{aligned}$$

and also

$$\lim_{x \rightarrow +\infty} x^m \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} = 0$$

and

$$\lim_{x \rightarrow +\infty} x^{\frac{1}{\xi}} \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} = \left(\frac{\xi}{\sigma} \right)^{-\frac{1}{\xi}},$$

to show that

$$\lim_{x \rightarrow \infty} J = L \left(\frac{\xi}{\sigma} \right)^{-\frac{1}{\xi}},$$

when $\xi < \frac{1}{m}$.

A few elementary operations allow us to write that

$$\begin{aligned} & \lim_{x \rightarrow +\infty} \frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(x - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} {}_2F_1 \left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{x\xi}{\mu\xi - \sigma} \right) \\ & = -(-1)^{-(m-\frac{1}{\xi}+1)} \frac{1}{\xi} \frac{\Gamma(-m + \frac{1}{\xi}) \Gamma(m+1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma} \right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi} \right)^m \quad (1.30) \end{aligned}$$

when $\xi < \frac{1}{m}$.

Finally, we compute the value of the primitive in Eq. 1.28 when x is equal to $\text{VaR}_p(X_T)$. This quantity is equal to

$$-\frac{(\text{VaR}_p(X_T))^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(\text{VaR}_p(X_T) - \mu)}{\sigma} \right)^{-\frac{1}{\xi}} {}_2F_1 \left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{\text{VaR}_p(X_T) \xi}{\mu\xi - \sigma} \right). \quad (1.31)$$

We input Eq. 1.30 and 1.31 into Eq. 1.28 and we derive Eq. 1.15, which is our result.

Proof of Theorem 4

Our goal is to solve Eq. 1.17 when $\text{VaR}_c(X_T) = x_0(1-c)^{-\frac{1}{\alpha}}$, so when X_T follows a Pareto Type I distribution. Replacing VaR with its expression, we can rewrite Eq. 1.17 as follows:

$$\frac{\alpha}{\alpha - m} \cdot \left(x_0(1-c)^{-\frac{1}{\alpha}}\right)^m - \left(x_0(1-q)^{-\frac{1}{\alpha}}\right) = 0.$$

Next, we write:

$$\frac{\alpha}{\alpha - m} \cdot x_0^m (1-c)^{-\frac{m}{\alpha}} - x_0(1-q)^{-\frac{1}{\alpha}} = 0,$$

so that

$$c = 1 - \left(\frac{\alpha - m}{\alpha x_0^m} x_0(1-q)^{-\frac{1}{\alpha}}\right)^{-\frac{\alpha}{m}},$$

which can readily be rewritten as Eq. 1.18.

Proof of Theorem 5

Our goal is to solve Eq. 1.19 when $\text{VaR}_c(X_T) = x_0(1-c)^{-\frac{1}{\alpha}}$, so when X_T follows a Pareto Type I distribution. Replacing VaR with its expression, we can rewrite Eq. 1.19 as follows:

$$\left(\frac{\alpha}{\alpha - m} \cdot \left(x_0(1-c)^{-\frac{1}{\alpha}}\right)^m\right)^{\frac{1}{m}} - \left(x_0(1-q)^{-\frac{1}{\alpha}}\right) = 0.$$

Next, we write:

$$\left(\frac{\alpha}{\alpha - m}\right)^{\frac{1}{m}} \cdot \left(x_0(1-c)^{-\frac{1}{\alpha}}\right) - \left(x_0(1-q)^{-\frac{1}{\alpha}}\right) = 0.$$

so that

$$c = 1 - \left(\left(\frac{\alpha}{\alpha - m}\right)^{-\frac{1}{m}} \left(x_0(1-q)^{-\frac{1}{\alpha}}\right)\right)^{-\alpha}$$

which can readily be rewritten as Eq. 1.20.

Chapter 2

Coherent Extended TCE

In the previous chapter, we define a risk measure as an instrument that can define the optimal capital the company should retain to avoid bankruptcy. In this chapter, we introduce the concept of coherence of risk measures introduced in the paper by Artzner et al. (1999). This concept has been very influential in the further development of risk measures and in defining new instruments in risk management.

A coherent risk measure is defined by the following four properties: sub-additivity, monotonicity, positive homogeneity, and translation invariance. A mathematical definition, as well as interpretation, can be found in Section 2.1.

Then, we study whether the most popular risks measures respect these properties. In Section 2.2, we show that the VaR is not subadditive, whereas in Section 2.3, we show that the TCE can be defined as a coherent risk measure. In both sections, we highlight that using a coherent risk measure is not always the best possible scenario.

We study here that the high-order TCE introduced by Faroni et al. (2022) is not a coherent risk measure as discussed in Section 2.4. However, a new risk measure strictly related to the high-order TCE is introduced here, and we study its coherence in Section 2.5.

Finally, Section 2.6 concludes.

2.1 Coherent Measures of Risk

Artzner et al. (1999) propose a list of properties that a good risk measure should have. In their seminal paper, they identify a series of axioms that a risk measure should satisfy to be called a coherent risk measure.

We use the approach of McNeil et al. (2015) where the financial risks are interpreted as portfolio losses over some time horizon left unspecified and losses expressed as positive values, instead of the approach of Artzner et al. (1999)

that identifies the random variable as the future value of the position currently held. This leads to some changes in the discussion of the axiom.

Definition 6. Given a probability space (Ω, \mathcal{F}, P) , financial risks are represented by a set M of random variables. A risk measure is a real-valued function $Q(\cdot) : M \rightarrow \mathbb{R}$ is called a coherent risk measure if it satisfies the following axioms:

- **Axiom 1: Translation Invariance.** For all $L \in M$ and every $c \in \mathbb{R}$, we have $Q(L + c) = Q(L) + c$.
- **Axiom 2: Subadditivity.** For all $L_1, L_2 \in M$ we have $Q(L_1 + L_2) \leq Q(L_1) + Q(L_2)$.
- **Axiom 3: Positive Homogeneity.** For all $L \in M$ and every $\lambda > 0$, we have $Q(\lambda L) = \lambda Q(L)$.
- **Axiom 4: Monotonicity.** For all $L_1, L_2 \in M$ such that $L_1 \geq L_2$ almost surely we have $Q(L_1) \geq Q(L_2)$.

Given $L \in M$, $Q(L)$ can be interpreted as the amount of capital that should be added to a position (or a portfolio) with a loss given by L so that the position (or a portfolio) is acceptable to an external or internal risk controller. Position with $Q(L) \leq 0$ are acceptable without capital injection.

The *translation invariance* (axiom 1) property states that by adding or subtracting a deterministic quantity c to a loss L , the capital requirements will be changed by the same amount c . From a financial risk management viewpoint, it implies that adding a certain amount of capital (negative c) reduces the risk by the same amount. This axiom is necessary to add sense to the capital requirement definition. Consider a position L and its capital requirement $Q(L) > 0$. Then, we add the amount equal to the capital requirement to our position, i.e., $\hat{L} = L - Q(L)$, and compute the new risk measure. If axiom one is respected, $Q(\hat{L}) = Q(L) - Q(L) = 0$, which means that now the position \hat{L} is acceptable to a risk controller viewpoint.

The *subadditivity* (axiom 2) principle can be explained by the fact that the risk of two (or more) portfolios together cannot be higher than the sum of the risks. This is the idea that risks can be reduced by using diversification. Suppose a regulator uses a non-subadditive risk measure to determine a financial institution's regulatory capital. In that case, that institution is incentivized to legally break up into various subsidiaries to reduce its regulatory capital requirements (McNeil et al., 2015). Moreover, subadditivity makes decentralization of risk-management systems possible (McNeil et al., 2015). Suppose that there are two trading desks with possible future losses equal to L_1 and L_2 and that the risk manager wants to ensure that $Q(L) = Q(L_1 + L_2)$

is lower than a threshold T . In order to achieve that, if the risk manager uses a subadditive risk measure, he needs to choose bounds M_1 and M_2 such that $M_1 + M_2 \leq M$ and he imposes on each of the desks the constraint that $Q(L_i) = M_i$. This allows the risk manager to manage the two different positions effectively, and subadditivity will ensure automatically that $Q(L) \leq M$. The principle of subadditivity is also seen as problematic. For example, Rauh-Bredow (2019) discussed that mergers of two banks can create additional risk using two banks, A and B, where bank B has borrowed heavily in the inter-bank market. Suppose that there is a significant loss for bank A. This vast loss could trigger the default of the newly merged group and banks from which bank B had taken loans before the merger. However, if the merger had not occurred, only Bank A would default. Another example of the problematic nature of the subadditivity principle is given by Dhaene et al. (2008), where they study the fact that mergers can increase the probability of systemic failure.

Axiom 3, *Positive Homogeneity*, implies that the risk of a position is proportional to its size, and it can be justified by using subadditivity that, if it holds, it requires:

$$Q(\lambda L) = Q(L + L + \dots + L) \leq nQ(L).$$

Because there is no diversification between these losses, it is natural to require that $Q(L + L + \dots + L) = nQ(L)$, which is the positive homogeneity property. For $\lambda = 0$, this axiom implies $Q(0) = 0$, the *normalization principle* that is the risk of holding no position is equal to zero. The positive homogeneity axiom has been highly criticized, and it has been suggested that for large values of λ , we should have $Q(\lambda L) > \lambda Q(L)$ to penalize the concentration of risk and the ensuing liquidity problems. The alternative is to use a class of convex risk measures (see Föllmer and Schied (2002) and Frittelli and Gianin (2002)) where the conditions of subadditivity and positive homogeneity have been relaxed. They only require a weaker property of convexity. However, this chapter will consider positive homogeneity in its standard form. The reader should note that subadditivity and positive homogeneity imply that the risk measure Q is convex on M .

Finally, the *monotonicity* axiom can be interpreted as positions that lead to higher losses in every state of the world requiring more risk capital.

In this chapter, we refer to a coherent risk measure as identified in definition 6. We study the coherence of Value-at-Risk in Section 2.2 and the coherence of Tail Conditional Expectation in Section 2.3. Then, we check whether the Extended TCE introduced in Section 1.2.3 is coherent in Section 2.4. Once we found that this new risk measure is not coherent, we introduced a new coherent risk extended TCE whose coherence is studied in Section 2.5. Table 2.1 summarizes the results obtained in the following sections.

Risk Measures	Axioms			
	Translation Invariance	Subadditivity	Positive Homogeneity	Monotonicity
Value-at-Risk	✓		✓	✓
Tail Conditional Expectation	✓	✓	✓	✓
Extended Tail Conditional Expectation				✓
Weak Coherent Extended Tail Conditional Expectation	✓	✓	✓	(✓)

Table 2.1: Summary of the coherent axiom is respected for each risk measure studied to define a coherent risk measure. If the risk measure respects all 4 axiom, it is coherent. The monotonicity property of the weak coherent extended TCE is demonstrated by adding an additional condition.

2.2 (Not) Coherent VaR

The Value-at-Risk (VaR) presented in Section 1.2.1 is not a coherent risk measure as it is not able to respect the subadditivity property as argued by Artzner et al. (1999). However, it is possible to demonstrate that VaR enjoys the properties of monotonicity, translation invariance, and positive homogeneity, as shown in the appendix. Moreover, VaR is a coherent risk measure when the losses are Normally distributed.

To demonstrate that VaR is not coherent, we provide a standard example that shows that the VaR portfolio is higher than the sum of VaR computed for both risks, i.e., the risk increases when diversification increases, which means that VaR does not respect the subadditivity property.

Consider Project A and Project B, where each has a probability of 0.02 of a loss of 10,000€ and a probability of 0.98 of a loss of 1,000€ during a one-year period. The one-year $\text{VaR}_{0.975}$ is equal to 1,000€ for each project. Suppose the two independent projects are put together in the same portfolio whose losses and probability are schematized in the following table.

Losses	Probability
20,000€	$0.02 \times 0.02 = 0.0004$
11,000€	$2 \times 0.02 \times 0.98 = 0.0392$
2,000€	$0.98 \times 0.98 = 0.9604$

This will lead to computing the $\text{VaR}_{0.975}$ of the portfolio equals 11,000€, whereas the sum of both VaR at the same quantile is 2,000€. This violates the subadditivity condition.

Even though VaR does not respect subadditivity, it is used mainly for its simplicity of implementing it operationally and its simplicity in understanding

it. In order to justify this use, Daniélsson et al. (2005) argues that the non-subadditivity of the VaR should not be considered a problem as they demonstrate that VaR is subadditive in the tails for all fat-tailed distributed, provided that tails are not super fat (i.e., assets whose first moment is not defined) at probabilities that are most relevant for practical applications. However, Daniélsson et al. (2013) note that VaR estimated from historical simulations may lead to violations of subadditivity, and they showed that a reduction of these violations happens when VaR is computed using semi-parametric extreme value techniques.

2.3 *Coherent TCE*

As VaR has been proved to be a risk measure that is not coherent by Artzner et al. (1999), they introduced a new risk measure called Tail Conditional Expectation (TCE), presented in section 1.2.2, which is a coherent risk measure.

We show in the appendix the demonstration that the TCE is able to respect all four properties of a coherent risk measure. Inui and Kijima (2005) show that a convex combination of Tail Conditional Expectation gives any coherent risk measure. Moreover, Acerbi and Tasche (2002) compares different risk measure definitions. They demonstrate that the definition proposed here is robust in yielding a coherent risk measure regardless of the underlying distributions.

Even though the coherence principle has been widely accepted, Dhaene et al. (2008) demonstrate that the subadditivity condition can lead to some undesirable situation where the shortfall risk increases by a merger, moreover, we should consider that merging or splitting portfolios may change management, business strategy, cost structure, and so on, this could lead to changes in the loss under consideration. In their paper, they also suggest additional conditions to avoid these undesirable situations. Cont et al. (2010) shows that using the TCE can lead to a less robust risk measurement procedure than using historical Value-at-Risk as there is a conflict between the subadditivity property and the robustness of risk measurement procedures.

2.4 *(Not) Coherent Extended TCE*

We discuss the non-coherence of the Extended TCE, or high-order TCE, presented in Section 1.2.3.

As discussed above, to be a coherent risk measure, the extended high-order TCE should have the following properties: translation invariance, subadditivity, positive homogeneity, and monotonicity. We show here that the high-order

TCE violates three of the four properties.

We should note that if $m = 1$, the $\text{TCE}^{(m)}$ is reduced to the TCE formula, which means that it respects the coherence properties.

Translation Invariance

Given a loss distribution X with two possible values: $X_1 = 1,000\text{€}$ with probability 0.4 and $X_2 = 500\text{€}$ with probability 0.6. The high-order TCE of order 2 is equal to:

$$\text{TCE}_{0.99}^{(2)}(X) = 1,000^2 = 1,000,000.$$

Let's assume now that there is a loss, c , of 100€ in both scenarios, the high-order TCE of order 2 is equal to:

$$\text{TCE}_{0.99}^{(2)}(X + c) = 1,100^2 = 1,210,000.$$

The risk measure of $\text{TCE}_{0.99}^{(2)}(X + c)$ is not equal to $\text{TCE}_{0.99}^{(2)}(X) + c$. This demonstrates that the high-order TCE does not respect the translation invariance.

Subadditivity

Consider Project A and Project B, where each project has a probability of 0.40 of a loss of $1,000\text{€}$ and a probability of 0.60 of a loss of 500€ . The high-order TCE for both projects at order 2 and a quantile of 0.99 is equal to $1,000,000\text{€}$. Suppose that these two independent projects are put together in the same portfolio whose losses and probability are schematized in the following table:

Losses	Probability
$\text{€ } 2,000$	$0.40 \times 0.40 = 0.16$
$\text{€ } 1,500$	$2 \times 0.40 \times 0.60 = 0.48$
$\text{€ } 500$	$0.60 \times 0.60 = 0.36$

If we compute the extended TCE for the portfolio that includes both projects, the result is $\text{TCE}_{0.99}^{(2)}(X) = 4,000,000\text{€}$ which is higher than the sum of the risk measures computed for each stand-alone project. This violates the subadditivity properties.

Positive Homogeneity

Let us consider the same loss distribution as in the translation invariance example, and given a value $\lambda = 2$, the high-order TCE can be computed as

follows:

$$ETCE_{0.99}^{(2)}(\lambda X) = (2 \cdot 1,000)^2 = 4,000,000.$$

The $TCE_{0.99}^{(2)}(\lambda X)$ is not equal to $\lambda TCE_{0.99}^{(2)}(X)$ which leads us to show that this risk measure does not respect positive homogeneity.

However, we should notice that this property is respected if we rewrite this risk measure as $(TCE_{\alpha}^{(m)}(X))^{\frac{1}{m}}$.

Monotonicity

As m is a positive number, it is possible to demonstrate using the same approach used for the TCE in the appendix that the high-order TCE respects the monotonicity property.

2.5 Weakly Coherent Extended TCE

In the previous section, we show that the high-order TCE presented in Section 1.2.3 and firstly introduced by Faroni et al. (2022) is not a coherent risk measure. Here, we introduce a new extended TCE related to a high-order moment (as the standard high-order TCE). It is a weakly coherent measure of risk given an additional hypothesis used to demonstrate the monotonicity property.

The weakly coherent extended TCE writes as follows:

$$ETCE_p^{(m)}[X] = \mathbb{E}[X] + (\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_p[X]])^{\frac{1}{m}}.$$

The risk measure proposed here is different from the one proposed by Barbosa and Ferreira (2004) as they propose a risk measure whose threshold is an arbitrary t , whereas our measure is widely general, and the threshold depends on the quantile of the loss distribution. Fischer (2003) presented a similar coherent risk measure whose threshold is the expected value of the loss distribution but is not linked to the risk considered. The same approach has been studied by Krokmal (2007), where the threshold is not defined in the same way as ours.

In the following paragraph, we prove that this modified version of the high-order TCE can respect translation invariance, subadditivity, positive homogeneity, and weak monotonicity.

As the high-order risk measure, when $m = 1$, this new risk measure reduces to the standard TCE.

Translation Invariance

Next, proving that ETCE respects translation invariance amounts to proving that

$$ETCE_\alpha [X + c] = ETCE_\alpha [X] + c.$$

We start by computing $ETCE_\alpha [X + c]$ as follows:

$$ETCE_\alpha [X + c] = \mathbb{E}[X+c] + (\mathbb{E}[(X + c - \mathbb{E}[X + c])^m | X + c > \text{VaR}_\alpha[X + c]])^{\frac{1}{m}}.$$

Because VaR satisfies the translation invariance property, we can rewrite the previous equation as follows:

$$ETCE_\alpha [X + c] = \mathbb{E}[X+c] + (\mathbb{E}[(X + c - \mathbb{E}[X + c])^m | X + c > \text{VaR}_\alpha[X] + c])^{\frac{1}{m}}.$$

We note that the event $\{X + c > \text{VaR}_\alpha[X] + c\}$ is identical to the event $\{X > \text{VaR}_\alpha[X]\}$. Thus, we can write:

$$ETCE_\alpha [X + c] = \mathbb{E}[X + c] + (\mathbb{E}[(X + c - \mathbb{E}[X + c])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}},$$

or, after elementary computations,

$$ETCE_\alpha [X + c] = \mathbb{E}[X] + c + (\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}}.$$

We recognize

$$ETCE_\alpha [X + c] = ETCE_\alpha [X] + c,$$

which is the translation invariance property applied to the ETCE indicator.

Subadditivity

To show that ETCE follows the subadditivity property, we need to prove that

$$ETCE_\alpha [X + Y] \leq ETCE_\alpha [X] + ETCE_\alpha [Y].$$

First of all, we see that the inequality:

$$\begin{aligned} & \mathbb{E}[X + Y] + (\mathbb{E}[(X + Y - \mathbb{E}[X + Y])^m | (X + Y) > \text{VaR}_\alpha[X + Y]])^{\frac{1}{m}} \leq \\ & \mathbb{E}[X] + (\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}} + \mathbb{E}[Y] + (\mathbb{E}[(Y - \mathbb{E}[Y])^m | Y > \text{VaR}_\alpha[Y]])^{\frac{1}{m}} \end{aligned}$$

is equivalent to

$$\begin{aligned} & (\mathbb{E}[(X + Y - \mathbb{E}[X + Y])^m | (X + Y) > \text{VaR}_\alpha[X + Y]])^{\frac{1}{m}} \leq \\ & (\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}} + (\mathbb{E}[(Y - \mathbb{E}[Y])^m | Y > \text{VaR}_\alpha[Y]])^{\frac{1}{m}}, \end{aligned}$$

which can in turn be rewritten as follows:

$$\begin{aligned} & \left(\int_{\text{VaR}_\alpha[X+Y]}^{+\infty} (X+Y - \mathbb{E}[X+Y])^m f(x+y) d(x+y) \right)^{\frac{1}{m}} \leq \\ & \left(\int_{\text{VaR}_\alpha[X]}^{+\infty} (X - \mathbb{E}[X])^m f(x) dx \right)^{\frac{1}{m}} + \left(\int_{\text{VaR}_\alpha[Y]}^{+\infty} (Y - \mathbb{E}[Y])^m f(y) dy \right)^{\frac{1}{m}}. \quad (2.1) \end{aligned}$$

Recall that $X_{[i,n]}$ denotes the i -th largest value of a sequence of random variables $\{X_1, \dots, X_n\}$. We use the following result, originally proved by Van Zwet (1980):

$$\int_{\text{VaR}_\alpha[X]}^{+\infty} x f(x) dx = \lim_{n \rightarrow +\infty} \frac{1}{n(1-\alpha)} \sum_{i=[n(1-\alpha)]}^n X_{[i,n]},$$

where $[n(1-\alpha)]$ is the smallest integer greater than or equal to $n(1-\alpha)$.

First, we write that

$$\int_{\text{VaR}_\alpha[X]}^{+\infty} (x - \mathbb{E}[X])^m f(x) dx = \lim_{n \rightarrow +\infty} \frac{1}{n(1-\alpha)} \sum_{i=[n(1-\alpha)]}^n \left(X_{[i,n]} - \frac{X_1 + \dots + X_n}{n} \right)^m.$$

Then, we define $\hat{X}_i = X_i - \frac{X_1 + \dots + X_n}{n}$, which yields the following simplification:

$$\int_{\text{VaR}_\alpha[X]}^{+\infty} (x - \mathbb{E}[X])^m f(x) dx = \lim_{n \rightarrow +\infty} \frac{1}{n(1-\alpha)} \sum_{i=[n(1-\alpha)]}^n (\hat{X}_{[i,n]})^m,$$

We can now rewrite Eq. 2.1 as follows:

$$\begin{aligned} & \left(\lim_{n \rightarrow +\infty} \sum_{i=[n(1-\alpha)]}^n ((\widehat{X+Y})_{[i,n]})^m \right)^{\frac{1}{m}} \leq \\ & \left(\lim_{n \rightarrow +\infty} \sum_{i=[n(1-\alpha)]}^n (\hat{X}_{[i,n]})^m \right)^{\frac{1}{m}} + \left(\lim_{n \rightarrow +\infty} \sum_{i=[n(1-\alpha)]}^n (\hat{Y}_{[i,n]})^m \right)^{\frac{1}{m}}. \end{aligned}$$

Equivalently, we should prove that

$$\lim_{n \rightarrow +\infty} \left(\sum_{i=\lceil n(1-\alpha) \rceil}^n ((\widehat{X+Y})_{[i,n]})^m \right)^{\frac{1}{m}} \leq \lim_{n \rightarrow +\infty} \left(\sum_{i=\lceil n(1-\alpha) \rceil}^n (\widehat{X}_{[i,n]})^m \right)^{\frac{1}{m}} + \lim_{n \rightarrow +\infty} \left(\sum_{i=\lceil n(1-\alpha) \rceil}^n (\widehat{Y}_{[i,n]})^m \right)^{\frac{1}{m}}.$$

This inequality holds true provided that

$$\left(\sum_{i=\lceil n(1-\alpha) \rceil}^n ((\widehat{X+Y})_{[i,n]})^m \right)^{\frac{1}{m}} \leq \left(\sum_{i=\lceil n(1-\alpha) \rceil}^n (\widehat{X}_{[i,n]})^m \right)^{\frac{1}{m}} + \left(\sum_{i=\lceil n(1-\alpha) \rceil}^n (\widehat{Y}_{[i,n]})^m \right)^{\frac{1}{m}}$$

for all n beyond a certain range.

Because $((\widehat{X+Y})_{[i,n]})^m = (\widehat{X}_{[i,n]} + \widehat{Y}_{[i,n]})^m$, we can readily apply a Minkowski inequality to prove the above result, which ultimately confirms the validity of the subadditivity property.

Positive Homogeneity

To prove the positive homogeneity of the extended TCE indicator, we need to prove that

$$ETCE_\alpha[\lambda X] = \lambda ETCE_\alpha[X].$$

We start by computing $ETCE_\alpha[\lambda X]$:

$$ETCE_\alpha[\lambda X] = \mathbb{E}[\lambda X] + (\mathbb{E}[(\lambda X - \mathbb{E}[\lambda X])^m | \lambda X > \text{VaR}_\alpha[\lambda X]]).$$

Because VaR respects the positive homogeneity property, we can rewrite the previous equation as follows:

$$ETCE_\alpha[\lambda X] = \mathbb{E}[\lambda X] + (\mathbb{E}[(\lambda X - \mathbb{E}[\lambda X])^m | \lambda X > \lambda \text{VaR}_\alpha[X]])^{\frac{1}{m}}.$$

The event $\{\lambda X > \lambda \text{VaR}_\alpha[X]\}$ is identical to the event $\{X > \text{VaR}_\alpha[X]\}$. This feature allows us to rewrite the previous equation as follows:

$$ETCE_\alpha[\lambda X] = \mathbb{E}[\lambda X] + (\mathbb{E}[(\lambda X - \mathbb{E}[\lambda X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}}.$$

Then, elementary computations allow us to confirm that

$$ETCE_\alpha[\lambda X] = \lambda \mathbb{E}[X] + \lambda (\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}} = \lambda ETCE_\alpha[X].$$

We have thus shown that the ETCE respects the positive homogeneity property.

(Weak) Monotonicity

To prove the monotonicity property, we slightly change the axiom and use a weak definition.

Weak Monotonicity Axiom. A risk measure respect $Q(\cdot)$ the weak monotonicity property if for all loss distributions, X and Y , such that $X \leq Y$ and $X - \mathbb{E}[X] \leq Y - \mathbb{E}[Y]$, the following inequality holds:

$$Q(X) \leq Q(Y).$$

As $X \leq Y$, it is easily showed that $\mathbb{E}[X] \leq \mathbb{E}[Y]$ which means that to demonstrate that:

$$ETCE_\alpha[X] \leq ETCE_\alpha[Y],$$

we can just demonstrate that:

$$(\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]])^{\frac{1}{m}} \leq (\mathbb{E}[(Y - \mathbb{E}[Y])^m | Y > \text{VaR}_\alpha[Y]])^{\frac{1}{m}},$$

or equivalently

$$\mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]] \leq \mathbb{E}[(Y - \mathbb{E}[Y])^m | Y > \text{VaR}_\alpha[Y]].$$

In order to demonstrate the last inequality, we use the same procedure applied in the case of TCE:

$$\begin{aligned} \mathbb{E}[(X - \mathbb{E}[X])^m | X > \text{VaR}_\alpha[X]] &< \mathbb{E}[(Y - \mathbb{E}[Y])^m | X > \text{VaR}_\alpha[X]] \\ &\leq \mathbb{E}[(Y - \mathbb{E}[Y])^m | Y > \text{VaR}_\alpha[Y]]. \end{aligned}$$

The first inequality is obtained using the weak monotonicity axiom, whereas the second inequality relates to the fact that VaR respects the monotonicity property.

We have demonstrated that the ETCE respects the weak monotonicity property.

2.6 Conclusion

In this chapter, we identify the different properties a risk measure should have to be considered a coherent instrument. Moreover, we include some standard results related to the no coherence of VaR and the coherence of TCE.

We study that the high-order tail conditional expectation presented in Faroni et al. (2022) is not a coherent risk measure. Moreover, as demonstrated in this chapter, we create a weak coherent risk measure based on the high-order TCE that can respect the properties of positive homogeneity, translation invariance, subadditivity, and a weak version of the monotonicity property.

Appendix

VaR Properties

This subsection demonstrates that VaR respects three out of four coherent properties.

Translation Invariance Given a loss distribution, X , and $c \in \mathbb{R}$, we define $Y = X + c$, we can write the Value-at-Risk for Y as:

$$\begin{aligned}\text{VaR}_p(Y) &= \min \{y : P(X + c \leq y) \geq p\} \\ &= \min \{y : P(X \leq y - c) \geq p\} \\ &= c + \min \{y - c : P(X \leq y - c) \geq p\} \\ &= c + \text{VaR}_p(X).\end{aligned}$$

This demonstrates that the translation invariance holds.

Positive Homogeneity Given a loss distribution X and a positive quantity $\lambda > 0$, we can compute the Value-at-Risk for $Y = \lambda X$ as:

$$\begin{aligned}\text{VaR}_p(Y) &= \min \{y : P(\lambda X \leq y) \geq p\} \\ &= \min \left\{ y : P\left(X \leq \frac{y}{\lambda}\right) \geq p \right\} \\ &= \lambda \min \left\{ \frac{y}{\lambda} : P\left(X \leq \frac{y}{\lambda}\right) \geq p \right\} \\ &= \lambda \text{VaR}_p(X),\end{aligned}$$

which demonstrates that VaR respects positive homogeneity.

Monotonicity Given two loss distribution, X and Y , such that $X \leq Y$. The fact that

$$p \leq P(Y \leq \text{VaR}_p(Y)) \geq P(X \leq \text{VaR}_p(Y))$$

can lead to $\text{VaR}_p(X) \leq \text{VaR}_p(Y)$.

Coherent TCE

We provide mathematical proof that TCE respects the four properties that identify a coherent risk measure.

Translation Invariance Given a loss distribution, X , and $c \in \mathbb{R}$, we define $Y = X + c$, the TCE can be written as:

$$\begin{aligned}\text{TCE}_p(Y) &= \mathbb{E}[Y|y > \text{VaR}_p(Y)] \\ &= \mathbb{E}[X + c|X + c > \text{VaR}_p(X + c)]\end{aligned}$$

As VaR satisfies the translation invariance property, it is possible to rewrite the previous equation as:

$$\mathbb{E}[X + c|X + c > \text{VaR}_p(X) + c]$$

and we note that the event $\{X + c > \text{VaR}_p(X) + c\}$ is identical to the event $\{X > \text{VaR}_p(X)\}$. Thus, we can write:

$$\begin{aligned}\text{TCE}_p(Y) &= \mathbb{E}[X + c|X > \text{VaR}_p(X)] \\ &= \mathbb{E}[X|X > \text{VaR}_p(X)] + c,\end{aligned}$$

where the last passage is possible as c is a quantity known. This proves that TCE follows the translation invariance property.

Subadditivity To have complete proof of the subadditivity property for the TCE, we refer the interested reader to consult Embrechts and Wang (2015), where it is possible to find seven ways to prove this property by using different approaches.

Positive Homogeneity Given a loss distribution X and a positive quantity $\lambda > 0$, we can compute the TCE for $Y = \lambda X$ as:

$$\begin{aligned}\text{TCE}_p(\lambda X) &= \mathbb{E}[\lambda X|\lambda X > \text{VaR}_p(\lambda X)] \\ &= \mathbb{E}[\lambda X|\lambda X > \lambda \text{VaR}_p(X)]\end{aligned}$$

by using the positive homogeneity property of the Value-at-Risk.

Then, as the event $\{\lambda X > \lambda \text{VaR}_p(X)\}$ is identical to the event $\{X >$

$\text{VaR}_p(X)$ }, we can rewrite the previous equation as

$$\begin{aligned}\text{TCE}_p(\lambda X) &= \mathbb{E}[\lambda X | X > \text{VaR}_p(X)] \\ &= \lambda \mathbb{E}[\lambda X | X > \text{VaR}_p(X)],\end{aligned}$$

which demonstrates that TCE respects the positive homogeneity property.

Monotonicity Given two loss distribution, X and Y , such that $X \leq Y$. It is possible to prove that TCE follows the monotonicity property as follows:

$$\begin{aligned}\text{TCE}_p(X) &= \frac{1}{1-p} \int_{\text{VaR}_p(X)}^{+\infty} x f_X(x) dx \\ &< \frac{1}{1-p} \int_{\text{VaR}_p(X)}^{+\infty} y f_Y(y) dy \\ &\leq \frac{1}{1-p} \int_{\text{VaR}_p(Y)}^{+\infty} y f_Y(y) dy \\ &= \text{TCE}_p(Y).\end{aligned}$$

The first strict inequality above is related to the strict monotonicity property of the integral and second inequality is related to the fact that VaR respects the monotonicity property.

Chapter 3

Stability of High Order Moments: a Risk Management Approach for Assessing the Risk of Insurance Companies

Risk management wants to reduce financial risks in an uncertain future. In order to avoid the normality assumption when assessing the market risk, a risk manager should take into account the high-order moments of the returns' distribution where the extant literature emphasizes the importance of heavy tail (kurtosis) and left-tailed events (skewness) to explain the stock price behavior. We give an overview of the literature related to the high-order moments and their use in asset pricing and risk management in Section 3.1.

In this work, we study the stability of annual high-order moments in equity indexes among markets. Section 3.2 describes the idea that we want to test, i.e., how we define the stability of moments. We extend our study to partial annual high-order moments, i.e., those computed using only values higher (lower) than a specified threshold.

The data used for this study are presented in Section 3.3, whereas computation and results are shown in Section 3.4. We demonstrate that raw non-centered moments are more stable than standardized moments, and third-order moments are more stable than fourth-order moments. However, that situation is the opposite if we consider partial moments.

This empirical study aims to help risk managers identify which moment is more stable over time, which will lead them to a more reliable assessment of future market risks.

Finally, Section 3.5 concludes.

3.1 The importance of high order moments

Following the traditional mean-variance CAPM of Sharpe (1964) and Lintner (1965), returns follow a Gaussian distribution, which would lead asset pricing models and investment strategies should only be based on the first two moments of the return distribution, mean and standard deviation. Under this assumption, probabilities of upside and downside outcomes are equal as the distribution is assumed symmetrical. However, in reality, returns tend to depart from normality, as shown in the seminal paper by Fama (1965). For this reason, it is essential to model the asymmetry and tail-fatness of returns captured by high-order moments. On this aspect, Robinson and Shaver (1973) shows that investors care about high-order moments when constructing their optimal asset allocation.

High-order moments are used to identify the shape of the stock returns' distribution. A raw non-centered high-order moment of order m of the distribution X is computed as:

$$\int_{-\infty}^{+\infty} x^m f_X(x) dx,$$

where $f_X(x)$ is the probability density function of the random variable X .

The lack of symmetry around the mean can be verified by the third-order moment. A zero value implies symmetry, as the right and left sides of the distributions are mirror images of each other. Positive (negative) skewness is rendered by a positive (negative) third-order moment, which suggests that the right (left) side of the distribution contains more extreme values than the other side. The fourth-order moment reflects a distribution's degree of flatness or peakedness around its mean. A distribution that has a fourth-order moment higher (lower) than three is said to be leptokurtic (platykurtic), which means that it has a thinner (flatter) peak and heavier (lighter) tails than the Normal distribution. Moreover, a high value of the fourth-order moment indicates the presence of extreme values and outliers in the data. We should remember that raw moments reflect both the return distribution shape and the variance level. In contrast, standardized moments are only tied to the shape of the return distribution. In this chapter, we study both raw non-central¹ and standardized moments.

Many studies argue that high-order moments can be included as risk factors in asset pricing, in addition, to mean and volatility. Sihem and Slaheddine (2014) test the reliability for the first four moments to describe the extreme values distribution on the French financial market. They find this new risk

¹Similar results could have been obtained using raw central moments instead of raw non-central moments because the return of expected value is too small to show any differences in the study.

measure suitable for avoiding significant crisis losses. However, it is insufficient to accurately predict the probability of crisis occurrences. This study is based on the multi-moment asset allocation and pricing models presented by Jurczenko and Maillet (2006). Among others, Ahmed and Al Mafrachi (2021) underscore the significance of considering higher-order realized moments of the return distribution for asset pricing and the prediction in cryptocurrency returns. Then, Chen et al. (2021) find that realized skewness positively relates to U.S. industry returns. However, they did not find a significant relationship between kurtosis and industry returns. Grigoletto and Lisi (2011) analyze through empirical studies the practical effects of modeling high-order moments on risk management. These authors find that models accounting for skewness and kurtosis outperform symmetric models when making VaR predictions. In general, vast literature reveals that high-order distributional moments are helpful in predicting market returns, such as Jondeau et al. (2019), which studies the ability of the average asymmetry to predict returns in the U.S. stock markets. Doan et al. (2010) show that high-order moments are essential in pricing stocks in the U.S. and Australian markets, but the degree of importance depends on the stocks' distribution. Amaya et al. (2015) demonstrate a reliable and significant negative relation between realized skewness and future stock returns, whereas Mei et al. (2017) show that realized skewness could help to forecast stock volatility. In contrast, realized kurtosis does not improve the models' performance. Even though there are many papers on the importance of high-order moments for stock market forecasting and optimal asset allocation, there is no study on the stability of such moments, which is the aim of the empirical study presented here.

To conduct an exhaustive analysis, the chapter also focuses on partial moments, i.e., moments only computed from the values higher or lower than a given threshold. This is because, from the investors' viewpoint, the uncertainty on the left-hand side of the return distribution represents risks. In contrast, the uncertainty on the right-hand side means better investment possibilities (Deng and Chen, 2022). Partial moment functions are proposed as a flexible way to characterize and test asymmetric effects as deviations from a reference value. Raw non-central lower (upper) partial moments, so LPMs (UPMs), measure the moments for values that are lower (higher) than a specific threshold, K . In the case of a continuous distribution and given $x \in [-\infty, +\infty]$, these partial moments can be computed as:

$$\text{LPM}_{K,t}^{(m)}(x) = \mathbb{E}[X^m | X < K] = \int_{-\infty}^K x^m f_X(x) dx,$$

and

$$\text{UPM}_{K,t}^{(m)}(x) = \mathbb{E}[X^m | X > K] = \int_K^{+\infty} x^m f_X(x) dx,$$

where $f(x)$ is the probability density distribution. These moments can also be standardized as shown in Section 3.4. LPMs and UPMs can capture high-order moments for values that are in the tails.

LPMs can be used as a measure of downside risk in financial decision-making, asset pricing, and so on. For example, Ling et al. (2020) propose a multi-period portfolio selection approach based on a variation of the LPMs defined here. Moreover, Nesaz et al. (2020) show that the method using LPMs is more efficient than the standard method in terms of portfolio performance.

In this chapter, we study the stability of high-order moments and high-order partial moments over time. The aim is to find which moment is the most stable over time and whether stability is affected by the market characteristics, region, industry, etc. This research could be used by risk managers to identify the main elements they should use to assess future market risk.

3.2 Key concept

A moment is considered stable over time if it does not change too much between a given year and the next. To study how much a moment changes from the previous year, we compute the quantity:

$$\left| f_t^{(m)} - f_{t-1}^{(m)} \right|,$$

where $f_t^{(m)}$ represents the moment at order m computed with the data available in year t .

In our study, we are also interested in studying the partial high-order moments, i.e., high-order moments computed only for values higher or lower than a threshold defined as a quantile α of the distribution X_t . We call $f_{\alpha,t}^{(m)}$ such a partial high-order moment. Section 3.4 details the computation of partial upper and lower high-order moments in different scenarios.

Furthermore, we apply a transformation to express high order moments in the same unit and make them comparable with high order moments computed at different orders:

$$\left| \left(f_{\alpha,t}^{(m)} \right)^{\frac{1}{m}} - \left(f_{\alpha,t-1}^{(m)} \right)^{\frac{1}{m}} \right|.$$

Finally, to study the stability of a moment over time, we compute

$$\sum_{t=2}^T \left| \left(f_{\alpha,t}^{(m)} \right)^{\frac{1}{m}} - \left(f_{\alpha,t-1}^{(m)} \right)^{\frac{1}{m}} \right|, \quad (3.1)$$

where the result is a non-negative number. When the result is equal to zero, the moment does not vary over time. However, this is only a hypothetical situation that never occurs in practice. When the result is close to zero, it can be interpreted as a stable moment that does not vary too much. A higher result can be interpreted as a moment that varies a lot over time, which means that it cannot be trusted to use to predict the future.

We aim to study the stability over time among different areas of the world, including developed and developing equity markets and the 3-month Treasury Bill Rate. Our study concentrates on the last 20 years, from 1st January 2000 until 31st December 2022. Section 3.3 gives a detailed description of the selected indices.

3.3 Data

In this chapter, we use seventeen equity indices to consider the world's most representative regions (including American, Asian, and European developed and developing markets) as well as the 3-month Treasury Bill Rate.

The equity index values are downloaded from a Bloomberg terminal, whereas the 3-month Treasury Bill is derived from the Federal Reserve Economic Data database. We use daily closing prices from 2000 to 2022 to compute daily log returns for each index i on each day for each year t . We have 23 years of observation for all the seventeen equity indices except for the Shanghai Stock Exchange, for which data are available only from 2005, which leads us to use 18 years of observation in the case of that index. We compute the standard return for the 3-month T-bill because negative yields in the years 2015 and 2020 do not allow us to compute the log returns.

Table 3.1 lists the indices used in this study and the country of origin. By looking at the country of origin, the reader should note that the study does not concentrate on one market but covers different areas.

Figure 3.2 shows the daily skewness and kurtosis computed on the overall period from 2000 to 2022, except for the Shanghai Stock Exchange, computed from 2005 to 2022. We see that all log-returns of the equity indices have negative skewness. We see that the Argentina Index has a higher kurtosis than all the other indices and a lower skewness. This table shows that the indices used in this study are pretty heterogeneous.

	Ticker	Country
3 Month Treasury Bill Rate	DTB3	U.S.
CAC 40	CAC40	France
DAX Performance Index	DAX	Germany
Dow Jones Industrial Average	DJI	U.S.
EURO STOXX 50	SX5E	Eurozone
FTSE 100 Index	UKX	U.K.
FTSE MIB Index	FTSEMIB	Italy
FTSE/JSE Top 40 Index	TOP40	South Africa
Indice Bovespa	IBOV	Brazil
KOSPI 50 Index	KOSPI 50	South Korea
Merval Index	Merval	Argentina
NIKKEI 225	NKI225	Japan
RTS Index	RTSI	Russia
S&P 100 Index	OEX	U.S.
S&P BSE SENSEX	SENSEX	India
S&P/ASX 50	AS31	Australia
S&P/TSX 60	SPTSX60	Canada
Shanghai Stock Exchange 50	SSE50	China

Table 3.1: *List of the indices used in the study together with their country of origin. The ticker is the one displayed by Bloomberg.*

	Skewness	Kurtosis
3 Month Treasury Bill Rate	6.8415	178.5248
CAC 40	-0.1674	10.1406
DAX Performance Index	-0.1745	8.8584
Dow Jones Industrial Average	-0.3595	15.3770
EURO STOXX 50	-0.2011	8.9622
FTSE 100 Index	-0.3314	13.3773
FTSE MIB Index	-0.5101	11.7640
FTSE/JSE Top 40 Index	-0.3721	8.1896
Indice Bovespa	-0.4539	9.3357
KOSPI 50 Index	-0.3035	12.6449
Merval Index	-5.3056	124.9266
NIKKEI 225	-0.2614	7.2284
RTS Index	-1.9815	47.4169
S&P 100 Index	-0.2979	12.4403
S&P BSE SENSEX	-0.2807	11.9711
S&P/ASX 50	-0.8034	12.6260
S&P/TSX 60	-0.7916	15.7892
Shanghai Stock Exchange 50	-0.2880	7.2216

Table 3.2: *Daily skewness and kurtosis computed on daily log-returns from 2000 to 2022 with two exceptions: Shanghai Stock Exchange 50 is computed from 2005 to 2022, and the 3-month T-Bill is computed in standard returns.*

3.4 Empirical Study

In this section, we perform an empirical study to check whether moments are stable over time and whether some are always more stable than others.

This research aims to compute Eq. 3.1 for all the difference indices defined in Section 3.3. Because the previous equation is defined for a general moment, $f_{\alpha,t}^{(m)}$ computed in year t , we specify in Section 3.4.1 the computation that we adopt for the different moments.

Then, Section 3.4.2 shows the results of our empirical study and comments on the findings we obtain.

3.4.1 Computation

Given the daily log-return of day i at year t for an equity index or the daily return of day i at year t of the 3-month T-bill, $x_{i,t}$, and n_t , the number of observations present in the dataset for the index in year t , we compute the m -th non-central moment for a year t as

$$\mathbb{E}[X_t^m] = \frac{1}{n_t} \sum_{i=1}^{n_t} x_{i,t}^m,$$

whereas the m -th standardized central moment for year t is computed as

$$\mathbb{E}\left[\frac{(X_t - \mathbb{E}[X_t])^m}{\sigma_t^m}\right] = \frac{1}{n_t} \sum_{i=1}^{n_t} \frac{(x_{i,t} - \mathbb{E}[X_t])^m}{\sigma_t^m},$$

where $\mathbb{E}[X_t]$ is the daily return average computed in year t and σ_t is the daily standard deviation of year t . When $m = 3$, this high-order moment is called skewness; when $m = 4$, it is called kurtosis.

To compute the m -th partial non-central moment, we need to sort the daily returns in ascending order for each year. We call $\hat{x}_{i,t}$ the sorted daily return where $\hat{x}_{1,t}$ is the smallest daily return and $\hat{x}_{n_t,t}$ is the largest ones in year t . Then, we can compute the partial m -th non-central moments as follows:

$$\mathbb{E}\left[X_t^m | X_t \leq \hat{x}_{[\alpha n_t],t}\right] = \frac{1}{[\alpha n_t]} \sum_{i=1}^{[\alpha n_t]} \hat{x}_{i,t}^m,$$

and

$$\mathbb{E}\left[X_t^m | X_t \geq \hat{x}_{[\alpha n_t],t}\right] = \frac{1}{[\alpha n_t]} \sum_{i=[\alpha n_t]}^{N_t} \hat{x}_{i,t}^m.$$

Where $\alpha \in (0, 1]$ defines the quantiles and $[\alpha n_t]$ is the least integer greater

than or equal to αn_t , whereas $\lfloor \alpha n_t \rfloor$ is the greatest integer less than or equal to αn_t . The first equation computes the partial $m - th$ non-central moment when the returns are lower than a specific threshold defined by $\hat{x}_{\lfloor \alpha n_t \rfloor, t}$, whereas the second equation computes the $m - th$ non-central moment conditional on the fact that returns are higher than the threshold, $\hat{x}_{\lfloor \alpha n_t \rfloor, t}$.

In the case of partial standardized central moments, we use the following formulas:

$$\mathbb{E} \left[\frac{(X_t - \mathbb{E}[X_t])^m}{\sigma_t^m} \middle| X_t \leq \hat{x}_{\lfloor \alpha n_t \rfloor, t} \right] = \frac{1}{\lfloor \alpha n_t \rfloor} \sum_{i=1}^{\lfloor \alpha n_t \rfloor} \frac{(\hat{x}_{i,t} - \mathbb{E}[X_t])^m}{\sigma_t^m}$$

$$\mathbb{E} \left[\frac{(X_t - \mathbb{E}[X_t])^m}{\sigma_t^m} \middle| X_t \geq \hat{x}_{\lfloor \alpha n_t \rfloor, t} \right] = \frac{1}{\lfloor \alpha n_t \rfloor} \sum_{i=\lfloor \alpha n_t \rfloor}^{n_t} \frac{(\hat{x}_{i,t} - \mathbb{E}[X_t])^m}{\sigma_t^m},$$

where $\mathbb{E}[X_t]$ and σ_t are the non-conditional daily average and daily standard deviations in year t .

Each moment is computed using around 250 observations, i.e., the daily returns available in one year. To be sure that the conditional moments are computed using more than a single value, α cannot be lower than 0.01 when the conditioning is expressed as $X_t \leq \hat{x}_{\lfloor \alpha n_t \rfloor, t}$ whereas it cannot be higher than 0.99 in case the moment is conditioned on $X_t \geq \hat{x}_{\lfloor \alpha n_t \rfloor, t}$.

3.4.2 Results

First of all, we remind the reader that 3-month T-Bill rate moments are computed using arithmetic returns, whereas equity moments are computed using log returns. Moreover, the Shanghai Stock Exchange 50 moments are computed from 2005 to 2022, whereas the others are computed from 2000 to 2022.

Table 3.3 shows the sum of the variation of raw non-central moments from 2000 to 2022 for each index in our database, where the definition of each moment is defined at the top row of the table. By analyzing the first two columns, we notice that in most of our indices (excluding 3-month T-Bill, DJIA, and S&P 100 Index), the third-order raw moment is higher than the fourth-order raw moment, which means that the manager should use the fourth-order moment to take decisions on the overall historical returns. This relation exists even though our data are heterogeneous with different levels of skewness and kurtosis, as shown in Table 3.2. The exceptions, highlighted in bold in the table, are all related to the U.S. market. In the other columns, we compute the lower raw moments when $\alpha = 0.05$ and $\alpha = 0.01$ and the upper raw moments when $\alpha = 0.95$ and $\alpha = 0.99$. In both cases, we see that the third-order par-

tial raw moment is lower than the fourth-order partial raw moment, without exception. From table 3.3, we see that fourth-order raw moments are more stable than third-order raw moments, whereas for partial moments the opposite is true: third-order partial raw moments are more stable than fourth-order partial raw moments. The risk manager is usually interested in the behavior of tails and in the avoidance of risks, which will lead us to advise him or her to use fourth-order partial raw moments. We also see that raw moments tend to be stable over time in most markets. For example, the third-order raw moment of CAC 40 moves in 23 years by only 0.2484 euros. Partial moments move a bit more because we consider extreme events, which vary more over time. Moreover, partial lower moments are less stable than partial upper moments, which leads us to say that extreme adverse events' distribution varies more than extreme positive events, which justifies paying greater attention from a risk management viewpoint.

Table 3.4 shows Eq. 3.1 computed using standardized moments, i.e., raw moments adjusted by the mean and standard deviation of the returns distribution. The previous table shows that skewness is less stable than kurtosis when computed from the overall distributions. However, in terms of partial standardized moments, the third-order moment is more stable than the fourth-order moment, without exceptions. Moreover, lower partial moments are less stable than upper partial moments. The table shows that the risk manager should use the more stable indicator because the historical value is more able to predict future values.

By comparing this table with Table 3.3 we notice that raw non-central moments are more stable than standardized ones where this information should be taken into consideration from the risk managers viewpoint ².

²A similar result can be obtained if we compare raw central moments with standardized ones.

	$\mathbb{E}[X_t^m]$		$\mathbb{E}\left[X_t^m X_t \leq \hat{x}_{[a_N, 1, t]}\right]$							
			$\alpha = 0.05$				$\alpha = 0.01$			
	$m = 3$	$m = 4$	$m = 3$	$m = 4$	$m = 3$	$m = 4$	$m = 3$	$m = 4$	$m = 3$	$m = 4$
3 Month Treasury Bill Rate	4.6341	5.1162	8.6629	9.7403	12.0488	12.6941	4.65	5.12	8.66	9.74
CAC 40	0.2484	0.1949	0.3532	0.3881	0.5179	0.5374	0.27	0.28	0.35	0.39
DAX Performance Index	0.2883	0.2074	0.3749	0.4048	0.5341	0.5491	0.30	0.31	0.37	0.39
Dow Jones Industrial Average	0.1768	0.1850	0.3157	0.34007	0.4641	0.4723	0.29	0.30	0.31	0.32
EURO STOXX 50	0.2311	0.1744	0.3192	0.3431	0.4423	0.4558	0.29	0.30	0.31	0.32
FTSE 100 Index	0.2240	0.1968	0.3493	0.3814	0.5089	0.5285	0.29	0.30	0.31	0.32
FTSE MIB Index	0.3512	0.2282	0.4302	0.4881	0.6468	0.6816	0.29	0.30	0.31	0.32
FTSE/JSE Top 40 Index	0.3313	0.1923	0.3816	0.4061	0.5468	0.5552	0.29	0.30	0.31	0.32
Indice Bovespa	0.4306	0.2679	0.4970	0.5586	0.7720	0.7999	0.33	0.34	0.35	0.36
KOSPI 50 Index	0.2853	0.2070	0.3566	0.3763	0.4969	0.5042	0.33	0.34	0.35	0.36
Merval Index	0.7640	0.7340	1.3088	1.6249	2.1530	2.3601	0.36	0.37	0.38	0.39
NIKKEI 225	0.2530	0.1746	0.3047	0.3365	0.4712	0.4844	0.29	0.30	0.31	0.32
RTS Index	0.4766	0.4351	0.7837	0.8890	1.2063	1.2536	0.44	0.45	0.46	0.47
S&P 100 Index	0.1844	0.1864	0.3164	0.3356	0.4532	0.4576	0.29	0.30	0.31	0.32
S&P BSE SENSEX	0.4017	0.2256	0.4140	0.4640	0.6360	0.6643	0.29	0.30	0.31	0.32
S&P/ASX 50	0.2417	0.1741	0.3400	0.3689	0.5082	0.5186	0.29	0.30	0.31	0.32
S&P/TSX 60	0.2135	0.2041	0.3719	0.3983	0.5201	0.5287	0.29	0.30	0.31	0.32
Shanghai Stock Exchange 50	0.2350	0.1222	0.2762	0.2863	0.3673	0.3720	0.14	0.15	0.16	0.17

Table 3.3: Stability Coefficients computed as in Eq. 3.1 where $f_{\alpha,t}^{(m)}$ is defined as the top r should remember that we use log returns for equity indices, whereas arithmetic returns for the

	$\mathbb{E} \left[\frac{(X_t - \mathbb{E}[X_t])^m}{\sigma_t^m} \right]$		$\mathbb{E} \left[\frac{(X_t - \mathbb{E}[X_t])^m}{\sigma_t^m} \right]$		$\mathbb{P} \left[X_t \leq \hat{x}_{[\alpha N_t], t} \right]$		α
	$m = 3$	$m = 4$	$\alpha = 0.05$	$\alpha = 0.01$	$m = 3$	$m = 4$	
3 Month Treasury Bill Rate	17.8703	8.3823	18.5003	24.2333	32.7122	36.4709	11.820
CAC 40	18.7204	4.2111	8.7574	11.1441	18.7535	19.9870	9.3197
DAX Performance Index	18.0811	3.7496	9.0009	11.0834	17.8275	18.6856	8.0109
Dow Jones Industrial Average	10.1340	2.0122	5.4032	6.3310	10.1573	10.7252	4.2707
EURO STOXX 50	16.5004	3.1439	6.4316	8.3546	14.1361	15.1887	7.3533
FTSE 100 Index	13.6814	3.3476	6.5963	8.6012	13.7887	15.2332	5.2055
FTSE MIB Index	19.2055	5.1652	10.0190	12.7463	19.0090	20.5158	8.5106
FTSE/JSE Top 40 Index	14.6136	2.2434	6.9807	8.0322	12.7312	13.2829	4.9340
Indice Bovespa	12.1685	4.4961	10.4664	13.2412	20.4044	21.8173	5.8213
KOSPI 50 Index	9.8163	2.4621	4.7501	5.8166	9.6821	10.3260	4.9255
Merval Index	16.4122	10.2687	19.7594	27.0851	38.2957	42.9115	7.3303
NIKKEI 225	14.8011	3.7130	7.9913	10.7197	17.2674	18.5624	5.3785
RTS Index	12.4006	6.0765	12.2795	15.8006	24.4519	25.8950	6.7028
S&P 100 Index	9.2912	1.8042	4.7536	5.8260	11.1942	11.4995	3.6791
S&P BSE SENSEX	16.0621	4.8086	9.9969	12.2843	18.0385	19.4538	7.4381
S&P/ASX 50	8.2932	2.0934	6.1750	7.4505	12.2660	12.8354	3.9497
S&P/TSX 60	9.7927	3.4928	7.9644	9.8007	14.9959	15.8336	5.1006
Shanghai Stock Exchange 50	15.8204	1.9937	7.5407	8.5799	12.7894	13.1346	6.0744

Table 3.4: Stability Coefficients computed as in Eq. 3.1 where $f_{\alpha t}^{(m)}$ is defined as the top r should remember that we use log returns for equity indices, whereas arithmetic returns for the

3.5 Conclusion

A risk manager should take into account the overall returns distribution in order to reduce financial risks in the future. We discuss in this chapter the stability of annual high-order moments in equity indexes among markets. We discover skewness (third-order raw moments) is less stable than kurtosis (fourth-order raw moments). In contrast, partial third-order raw (standardized) moments are more stable than partial fourth-order raw (standardized) moments. This leads the risk manager to rely on the third moment if he or she is interested in the overall returns distribution or on the fourth moment if he or she wants to define the tail return distribution.

Moreover, we show that partial upper moments are more stable than partial lower moments, which means that extreme positive events are more stable than extreme adverse events.

Finally, raw moments are more stable than standardized moments over time.

The goal of this empirical study is to help the risk manager to identify which moment is more stable over time, which leads him or her to a more reliable assessment of future market risks according to his or her objectives.

Chapter 4

Home Bias and Learning in a Dynamic Portfolio Choice under Smooth Ambiguity

In this chapter, we study the dynamic optimal asset allocation and the optimal consumption level that the decision-maker should choose when there are two kinds of stocks in the market: local and foreign. We want to study the evolution of the home bias over time from the viewpoint of a single investor. Home bias is the phenomenon where investors tend to invest more in their home country's market than its share in the overall market. This phenomenon leads to underdiversified portfolios as agents invest more in domestic assets rather than foreign ones.

We analyze the optimal investor's portfolio through a specific period of time, assuming that the investor re-balances her portfolio at the end of each year where she aims to maximize her utility function modeled using the generalized recursive smooth ambiguity model axiomatized by Hayashi and Miao (2011). We use this model as it is one of the most general models in dynamic decision theory, and it includes different decision theory models as special cases. Moreover, the generalized recursive smooth ambiguity model is able to distinguish between risk aversion, ambiguity aversion, and intertemporal substitution, which will allow us to understand better which parameters drive the home bias in this model.

Section 4.1 gives an overview of the existing literature on home bias and model decision theory in order to understand the importance of our findings better.

We model a financial market with one risk-free stock and two kinds of risky and ambiguous stocks. Stock returns follow a geometric Brownian motion where the investor ignores the real value of the drift. However, she has a Normally distributed prior over it, and the decision maker learns over time

and updates the prior at each time using the new information available on the market. In order to model foreign and local stocks, we assume that the decision maker has a short period of time where she does not invest but studies the local stock, which will lead her to be more confident about this asset. Then, when she starts to invest, she discovers that there is another type of asset (foreign stocks) on the market, which will lead her to be more ambiguous about the foreign asset at the start as she discovers them later. She has little information about it compared to the information available for domestic stocks. The market is described as ambiguous as the investor does not know the real mean of the geometric Brownian motion that drives the stock returns. Then, the decision maker will be considered risk averse and ambiguity averse. Section 4.2 describes the theoretical framework and preference model used.

In this framework, the investor needs to choose the optimal consumption level and the optimal portfolio allocation at each time. Section 4.3 gives a detailed description of this investor's problem and highlights the Bellman equation and some theoretical results.

To solve the investor's problem theoretically described in 4.3, we use a numerical approach using a backward approach and solve the expected value using a double Gauss-Hermite quadrature. This model is explicitly computed for two risky and ambiguous stocks on the market. However, it can be generalized to n risky and ambiguous stocks. The numerical approach used is detailed in Section 4.4, and the results are shown in Section 4.5.

Finally, Section 4.6 concludes, and we suggest possible expansions on our findings.

4.1 Literature Review

In this section, we provide a short literature review on the home bias (Section 4.1.1) and decision theory (Section 4.1.2). Moreover, Section 4.1.3 discuss some articles that studied home bias and portfolio choice under ambiguity that can be related to the results obtained in this chapter and how our findings are different from the existing literature.

The literature presented here is meant to be partial, but we want to identify where our topic can be placed in the literature. We advise the reader to refer to the bibliography cited in this section to master the subject.

4.1.1 Home Bias

Home bias is defined as an investment behavior where investors tend to overweight their home country's market compared to its share in the overall market portfolio. It was defined as one of the officially recognized six major

puzzles in international macroeconomics by Obstfeld and Rogoff (2001). The seminal paper of French and Poterba (1991) shows that most corporate equity is held by domestic investors. For example, in 1989, U.S. residents held 92% of the United States stock market, and the same pattern can be seen for Japanese, U.K., Germany, and France investors. Fidora et al. (2007), among others, confirmed the patterns of the home bias for these countries. Lippi (2016) proves the existence of the home bias specifically for Italian professional occupational pension fund managers, whereas Lütje and Menkhoff (2007) prove it for German investors, Mishra (2008) shows the home bias in the Australian investors' portfolio. Additional evidence can be found worldwide and even in the online market. See, for example, Lin and Viswanathan (2016) that study home bias in a debt-based crowdfunding platform. We refer the interested reader to Gaar et al. (2020) for a more exhaustive literature review.

Even though, the first empirical study was done in a market that was far from being friction-less, we see that home bias is still alive nowadays even though there has been an increase in cross-border asset trade thanks to electronic trading, exchanges of information across borders and falling transaction costs (Lane and Milesi-Ferretti, 2003). Over the 1988-2008 period, Coeurdacier and Rey (2013) shows that home bias has decreased in developed countries but remains high. This means that the share of foreign equities in investors' portfolios is roughly a third of what it should be. Furthermore, emerging markets have less diversified equity portfolios and do not exhibit any clear downward trend in home bias. According to the International CAPM model, investors in these countries hold one-tenth of the amount of foreign equities they should be holding.

Similar to home bias, *local bias* describes the behavior where investors are more inclined to invest in assets of firms located close to them, independent from country borders. Baltzer et al. (2013) bridge the home and local biases by studying the investment strategies from Germany and neighborhood countries. They discover that investors living close to a foreign country display less home bias towards investment opportunities in that country. With the same approach, they argue that the overweight of regionally close stocks extends beyond domestic borders.

In this chapter, we use two kinds of assets: domestic stock and foreign stock. However, domestic and foreign definitions are left to the decision maker. For example, an investor in Geneva could define domestic stocks as an asset issued by a company in France or Switzerland. In contrast, an investor living in Zurich could define it as a stock issued by a firm located in Switzerland or Germany.

Why home bias exists?

The vast literature on home and local bias has not provided a generally accepted explanation. However, researchers agree that the reason could be that home bias is caused by a combination of factors that can be institutional, information-related, or individual investor-related. We now discuss each one of them. However, we refer the reader to Ardalan (2019) and Gaar et al. (2020) for finding references used in this subsection.

Institutional reasons relates to the costs associated with international investing. These costs can induce the home bias because the net return on domestic equities is higher than on foreign equities. A list of them follows.

- *Capital controls* where governments created barriers to international investment. This was true many years ago when governments imposed taxes through the 1980s and 1990s. However, all developed markets and many emerging ones are now open to foreign investors. For this, capital control is considered an unrealistic reason to use to explain home bias.
- *Transaction costs* such as international taxes, informational costs, management fees, and other barriers to trade equity. However, it does not provide a reasonable explanation as it has been shown that the capital flows on foreign equity transactions tend to be higher than capital flows on domestic equity transactions. Moreover, in recent years, transaction costs have been declining over time, and it has become more challenging to argue that they are a cause of under-diversification.
- *Information costs*, i.e., the cost of acquiring information about foreign accounting practices, foreign corporate relationships, and legal environment.

Nowadays, information costs and government restrictions cannot be used as a stand-alone reason to justify home bias in developed countries, mostly because of the Internet and the fact that governments in developed countries do not impose such restrictions.

Information asymmetries can also play a role in home bias and can be placed into the following categories.

- *Home advantage*, i.e., the fact that domestic investors have better information about domestic stocks leads them to invest in them. This can be modeled, for example, using Bayesian learning, where domestic investors begin with less accurate prior beliefs about foreign countries' fundamentals.

-
- *Distance*. Empirical evidence shows that geographical distance between two countries is a good proxy for information asymmetries.
 - *Cultural differences* also play a role in international portfolio choices. For example, Finnish investors whose native language is Swedish are more likely to invest in stocks of Finnish companies that publish their reports in Swedish than Finnish investors whose native language is Finnish (Grinblatt and Keloharju, 2001).
 - *Accounting standard*, in the sense that using different accounting standards would act as information barriers.

Behavioral biases Information asymmetries and institutional reasons rely on the rationality of investors. However, psychologists and experimental economists noted that individuals act with behavioral biases, which is the third factor that can be used to explain home bias. A list of some behavioral factors and how they affect home bias follows.

- *Familiarity* as people tend to root for their home team, eat local food, and invest in a company's stocks that are visible to them.
- *Confidence*, which leads home bias to occur when domestic investors perceive that they have an information advantage relative to foreign investors and overestimate their ability to forecast the performance of familiar assets.
- *Competence* in terms that investors believe they are more competent in investing in their home country than abroad, and investors treat foreign stock differently for fear of showing incompetence.
- *Optimism* as investors tend to have higher expectations about the rates of return on domestic stocks than their expectations on the rates of return on foreign stocks.
- *Cultural differences and patriotism* can increase home bias. Patriotism is also associated with investors' loyalty toward a specific company.
- *Investor characteristics* such as the level of experience. Literature reports that less experienced investors tend to be more home-biased than more sophisticated ones.
- *Loss aversion* means that some investors experience losses asymmetrically, i.e., more severely than equivalent gains. For this reason, they tend to avoid losses. However, as in international stock markets returns

correlation is higher in market downturns than in upturns. Investors will avoid international stock markets, which will create a home bias.

- *Beliefs* as local investors consider more reliable information on local stocks relative to signals on foreign stocks.
- *Ambiguity Aversion* where ambiguity refers to investment with unknown probabilities of future outcomes. Empirical evidence, such as Dimmock et al. (2016), shows that ambiguity aversion has a negative relation with stock market participation and portfolio allocations to equity and helps explain home bias.

In this chapter, we study the presence of home bias over time. We justify the home bias using the investor's ambiguity aversion and the fact that the investor has (or perceives to have) better information about domestic stocks at start. This last point will be corrected as the investor learns over time about the foreign stocks, which will decrease the home bias in the investor's portfolio.

4.1.2 Decision Theory

Decision theory started as the study of the games of chance in the XVI and XVII century and nowadays continues to study how to explain the decision-maker process from a normative and descriptive viewpoint.

Classic theory

One of the first decision theory models considered a cornerstone of this theory is the **objective expected utility** or expected utility. This theory was first introduced by Bernoulli (1738) as the solution to the Saint Petersburg Paradox that was created by his cousin Nicholas Bernoulli. The St. Petersburg Paradox states as follows:

Given a game of chance in which a fair coin is tossed at each stage. The initial stake begins at one ducat and is doubled every time a head appears. The first time a tail appears, the game ends, and the player wins whatever is the current stake. How much will the player be willing to pay to enter the game?

Suppose the price of the game is computed using just the expected value. In that case, the player should be willing to play at any price if offered the opportunity (that is because the expected value of this game is infinite). However, Nicholas Bernoulli states that most players would not pay a price higher

than twenty ducats, and the discrepancy between the price and the expected value creates this paradox. Daniel Bernoulli suggested that the price of the game should be computed not using the expected value but the expected utility, which means that each ducat should be considered not as a ducat itself but as the utility that the item will be given to the player if won:

$$\mathbb{E}[U(x)] = \sum_{i=1}^n U(x_i)p_i.$$

He described the concept of utility using the idea that one ducat has a different meaning for a poor person or a rich one. The objective expected utility has been axiomatized by Von Neumann and Morgenstern (1944) where one of the essential axioms is the independence one that states that a decision maker should not consider the states where the lotteries have the same outcomes. The objective expected utility has different properties, such as the risk aversion parameters studied by different researchers as Pratt (1964) and others. However, literature shows violations of the independence axiom, creating different types of paradoxes such as those of Allais (1953). For this reason, new non-expected utility models of preferences over objective lotteries have been developed. To see a non-exhaustive discussion of them, we refer the reader to Machina (1992).

The Objective Expected Utility described below is only limited to gambles whose probability can be defined objectively, such as the toss of a fair coin, the roll of a dice, and so on. However, in the real world, the major of decisions are made under uncertainty where probabilities are not subjectively defined, which led to the development of the **Subjective Expected Utility** (SEU) by Savage (1972). This new axiomatization considers that the expected utility is computed using the utility function and subjective probability measures that change among decision-makers. This theory is based on two main principles: (i) the sure-thing principle, similar to the independent principle, states that if two acts yield the same outcome for each state in the event, it should not matter what those states common outcomes are, and (ii) weak comparative probability which states the decision maker has a well-defined comparative likelihood ranking over events. However, the SEU has been highly criticized primarily due to the difficulty of determining such subjective probabilities. For a description and interpretation of Savage's result, we refer the reader to Kreps (1988), Gilboa (2009), Fishburn (1970), and others.

To create a link between subjective and objective probabilities, Anscombe et al. (1963) defines a joint objective-subjective approach that allows for an axiomatic derivative of subjective probability simpler than that of Savage (1972). The **horse-roulette theory** is based on the fact that agents have expected utility preferences over primitive objective lotteries. Put in different words,

they assume that an agent has subjective probabilities over a horse race. However, instead of receiving the winnings on his bet in cash, the agent will receive a ticket for a lottery with objective probabilities.

Ellsberg's Paradox and ambiguity

These last two models were a significant improvement in the decision-making theory. However, they do not capture the degree of confidence the decision maker has when deciding the subjective probabilities to use. Ellsberg (1961) shows that in a situation in which probabilities are not known, Savage's axioms can be violated. Specifically, one of **Ellsberg's experiment** involves two urns, A and B, that contain 100 balls each. Urn A has 50 red and 50 black balls, whereas Urn B has 100 balls that can be either red or black, but the decision maker does not have any information on the proportion of red and black. Most decision-makers are indifferent between betting on either color within each urn. However, they all prefer to bet on Urn A instead of Urn B. No probability measure can justify this betting behavior, creating the Ellsberg Paradox. This experiment revealed the phenomenon of ambiguity aversions, i.e., agents prefer gambles with known probabilities to unknown ones.

On this note, we want to remember that in the literature, uncertainty, and ambiguity are not always clearly defined and distinguished. We use these terms equivalently, and we define an ambiguous event as a situation where the agent does not know the probabilistic distribution of the outcomes as in Guidolin and Rinaldi (2013). We refer the reader to Etner et al. (2012) for a discussion on the definition of ambiguity attitudes in different decision theory models.

Alternative models

A series of alternative models that are all been axiomatized was developed to answer the Ellsberg paradox. We will cite the most popular ones.

Schmeidler (1989) introduces the **Choquet Expected Utility**, CEU, whose probability measure is not necessarily additive, i.e., given two events A, B and the capacity v represents the measure of credence, we may have:

$$v(A) + v(B) < v(A \cup B).$$

To resolve the problem of the computation of an expected utility using a non-additive probability measure, he adopted the Choquet Integral, and we refer the reader to Gilboa (2009) for a more detailed explanation of the Choquet Integral. This new model allows us to relax the independence axiom and use the definition of comotonic independence, which is able to distinguish between mixing operations that reduce uncertainty (via hedging) and mixing operations that do not.

Gilboa and Schmeidler (1989) introduces the **Max-Min preferences**, MEU, also called multiple prior preferences, where it maximizes the worst-case expected utility over a set of probabilities. The idea is that the agent has several probability measures as potential beliefs, and each belief creates an expected utility index for each act. This will lead each act to have as many expected utility values as the potential belief. The agent should use the worst-case expected utility to define the act's desirability. They should consider the most unfavorable prior as the decision maker is ambiguity averse.

The **smooth ambiguity preferences model** (KMM) has been axiomatized by Klibanoff et al. (2005). Given an act $f : S \rightarrow X$, the expected utility is computed as:

$$\int_{\Delta(S)} \phi \left(\int U(f(\cdot)) d\mu \right) dM(\mu), \quad (4.1)$$

where $U(\cdot)$ is the von Neumann-Morgenstern utility function, M is the individual's second order prior over the set $\Delta(S)$ of measure μ on the state space S , and $\phi(\cdot)$ is the individual's second order utility function. The idea is that the individual is uncertain about which probability measure μ describes the likelihood of state realizations, and the second-order uncertainty is modeled using M . Using other words, the agent distorts the expected utilities computed as $\int U(f(\cdot)) d\mu$ using the function $\phi(\cdot)$ whose concavity corresponds to the ambiguity aversion. The key feature of this model is that it achieves a separation between ambiguity, identified as a characteristic of the decision maker's subjective beliefs, and ambiguity attitude, a characteristic of the decision maker's tastes. This model will be further discussed in Section 4.2.1.

There are many different alternative models, such as the α -MEU, which is a generalization of the max-min preferences, the vector expected utility by Siniscalchi (2009), and the variational preferences by Maccheroni et al. (2006), among others. We refer the interested reader to Gilboa and Marinacci (2016), and Machina and Siniscalchi (2014) for a more exhaustive overview of alternative models on decision theory.

Dynamic theory

In the axiomatization described till now, time is not explicitly considered, which leads to one-period decision problems: the individual chooses an act, then the state is realized, and the prize is determined. It is possible to extend the model decision theory studied till now to a dynamic decision problem in which the individual may acquire partial information about the state of nature over time and take action at several decision points. In order to pass from a static setting to a dynamic one, we need to consider two key ingredients.

- Updating a prior probability measure μ to reflect new information ac-

quired. We discuss the updating process in Section 4.2.3.

- Dynamic consistency or consequentialism links conditional and unconditional preferences. Dynamic consistency ensures that if a multi-period choice problem is solved by backward induction, one should obtain an optimal ex-ante solution. In contrast, consequentialism is the fact that only the states whose outcomes are still possible can matter for updated preferences. Both cannot coexist in a dynamic model, and one must be relaxed. Even though there is no consensus on whether dynamic consistency or consequentialism is the more plausible assumption. Dominiak et al. (2012) shows that most subjects violate dynamic consistency rather than consequentialism by using the 3-color experiment Ellsberg urn.

The seminal paper on dynamic decision modeling is by Kreps and Porteus (1978). The decision maker uses a von Neumann-Morgenstern utility function computed at each node of the decision tree that expresses all the actions the decision-maker can take and the outcomes at each state. The application of the Kreps-Porteus model to non-expected utility models is made by Epstein and Zin (1989). The **Epstein-Zin model** develops a class of recursive preferences over intertemporal consumption lotteries that are empirically studied in Epstein and Zin (1991). The agents' preferences can be represented recursively by:

$$\forall t, U_t = W \left(c_t, V^{-1} \left(\mathbb{E}_t \left[V(\tilde{U}_{t+1}) \right] \right) \right),$$

for some increasing aggregator function W , \tilde{U}_{t+1} is the future stochastic utility, V is an increasing function, and \mathbb{E}_t is the mathematical expectation conditional on information available at time t . The reader should note that $V^{-1} \left(\mathbb{E}_t \left[V(\tilde{U}_{t+1}) \right] \right)$ represents the certainty equivalent of future utility where V represents risk preferences and the aggregator W defines the intertemporal substitution preferences between the consumption at time t and the future utility. This was the first model to separate the role of risk aversion and intertemporal substitution; however, Etner (2006) shows a relation between these two concepts via the notion of the timing of resolution of uncertainty.

Following the same approach, different models have been proposed that also include ambiguity, such as, for example, the **recursive MEU** by Epstein and Schneider (2003) in the discrete-time version and by Chen and Epstein (2002).

In this chapter, we study the **generalized recursive smooth ambiguity preferences model** axiomatized by Hayashi and Miao (2011) that allows for a three-way separation among risk aversion, ambiguity aversion, and intertemporal substitution. This model is a dynamic model of the KMM model. It nests some popular utility models as special cases, including Epstein-Zin model, smooth ambiguity preferences (Klibanoff et al., 2009), and risk-sensitive pref-

erences (Hansen and Sargent, 2001). We discuss this model in detail in Section 4.2.1.

4.1.3 Home Bias and Decision theory

The findings presented in this chapter are related to a series of papers that studies portfolio choice and home bias using decision theory models under ambiguity. We will comment here on why our findings differ from those in the literature.

Epstein and Miao (2003) studies the international portfolio choice in a two-agent equilibrium setting in which agents differ in their ambiguity about returns and use a recursive MEU model. Uppal and Wang (2003) focuses on asset allocation with multiple risky assets in an i.i.d. setting where different levels of ambiguity are attached to the return distributions of domestic and foreign assets. Garlappi et al. (2007) studies dynamic portfolio choice in the multiple priors framework, and Peijnenburg (2018) uses a max-min expected utility model over the life cycle in a dynamic setting. She finds that ambiguity helps explain under diversification and home bias, but the model does not allow a separation between ambiguity and ambiguity attitudes. It only concentrates on the worst-case scenario.

Among the few papers applying the smooth ambiguity model, Chen et al. (2014) consider an optimal consumption and portfolio choice problem where the stock returns can be modeled as i.i.d. returns or using return predictability. However, they focus on a simple application using one single risky asset. Indeed, few articles apply the model with more than one risky or ambiguous stock. Similar to our article, Yu et al. (2022) studies the portfolio management problem under the smooth ambiguity model by deriving the Hamilton-Jacobi-Bellman equation for the equilibrium value function. The analysis is done in a life cycle setting, and they focus on the sensitivity analysis of risk and ambiguity parameters. They found that as an individual becomes more ambiguous (or risk) averse, he will consume less, buy more life insurance and invest less in risky assets. Another article related to ours, but in a static setting, is by Guidolin and Liu (2016), where they study the asset allocation under smooth ambiguity aversion where the investor relies on the conditional distribution of them.

Our research differs from the previous papers as we use a smooth ambiguity model and not the max-min model that considers only the worst-case scenario. We use a dynamic model that can consider the future continuation value, and we concentrate on the optimal asset allocation, assuming that there are two kinds of stock on the market: domestic and foreign. Moreover, we provide an in-depth analysis of the home bias and its evolution. To our knowledge, it does not exist a numerical approach that is able to solve the optimal asset

allocation for more than one stock using the generalized smooth ambiguity model adopted here.

4.2 The model

This section describes the theoretical model used to study the investor's optimization problem.

Section 4.2.1 described the generalized smooth ambiguity modeled and its properties in detail. Here, we discuss why it can be considered a generalized model as well as its importance of it as it can distinguish between risk aversion, ambiguity aversion, and intertemporal substitution.

The decision maker that follows the decision model in Section 4.2.1 has to decide at each time $t \in \{0, \dots, T\}$ the consumption level as well as the allocation of her portfolio. The optimal decision problem will be discussed in Section 4.3.

Before going further, Section 4.2.2 shows the composition of the financial market and how it is perceived from the decision-maker's viewpoint. Finally, in Section 4.2.3, we discuss how the investor updates her beliefs when new information is available in the financial market.

4.2.1 Generalized Recursive smooth ambiguity preferences

In this section, we introduce the generalized recursive smooth ambiguity modeled axiomatized in Hayashi and Miao (2011) that has been built on the static smooth ambiguity modeled developed by Klibanoff et al. (2005).

In the static setting, a decision maker's ambiguity preferences over consumption are represented by the utility function in terms of two certainties equivalent as:

$$v^{-1} \left(\int_{\Pi} v \left(u^{-1} \left(\int_S u(C) d\pi \right) \right) d\mu(\pi) \right),$$

where u and v are increasing functions and μ represents a subjective prior over set Π of probability measures on S that the decision maker thinks possible. When we define $\phi = v(u^{-1})$, the previous utility function is ordinally equivalent to the one computed by Klibanoff et al. (2005):

$$\int_{\Pi} \phi \left(\int_S u(C) d\pi \right) d\mu(\pi). \quad (4.2)$$

As explained in the previous section, this model separates ambiguity, characterized by properties of the subjective set of measures Π , and ambiguity attitudes, characterized by the shape of ϕ or v . Thanks to the separability of

this concept, it is possible to create comparative statics of ambiguity attitudes while holding ambiguity fixed. Moreover, the attitudes toward pure risk are characterized by the shape of u . In particular, the decision maker displays risk aversion if u is concave, while she displays ambiguity aversion if ϕ is concave or, equivalently, if v is a concave transformation of u . In addition, this model can include the multiple prior model of Gilboa and Schmeidler (1989) as a special case when ambiguity aversion goes to infinity (Klibanoff et al., 2005). On another note, if ϕ is linear, the decision maker is ambiguity neutral, reducing this model to the standard expected model. This model can also be re-interpreted as a model of robustness in which the decision maker is concerned about model misspecification where each distribution $\pi \in \Pi$ describes an economic model. However, the decision maker is ambiguous regarding the correct distribution, but she has a prior μ over alternative models. Moreover, she is averse to model uncertainty, which is why the function ϕ is concave. While the preference model just described achieves the task of separating ambiguity and ambiguity attitudes, and it is able to obtain smooth indifference curves (and not kinked ones), it is limited by the fact that it is a timeless framework.

The first attempt to extend Klibanoff et al. (2005) to an intertemporal framework was done by Klibanoff et al. (2009); however, this dynamic model suffers from a limitation that intertemporal substitution and attitudes toward risk or uncertainty are intertwined. This makes comparative statistics hard to interpret. That is why we use the model proposed by Hayashi and Miao (2011) that contains the model by Klibanoff et al. (2009) as a specific case.

We consider a dynamic setting with finite horizon T , time is denoted by $t = \{0, 1, 2, \dots, T\}$, and the state space for each period is denoted by S . At time t , the decision maker's information consists of history $s^t = \{s_0, s_1, \dots, s_t\}$. The investor has an initial wealth W_0 and maximizes her lifetime utility by controlling her consumption stream $C = \{C_t\}_{0 \leq t \leq T}$ and investment decision, that is her portfolio composition $\theta = \{\theta_t\}_{0 \leq t < T}$ where each C_t, θ_t is a function of the observed states up to time t . The decision maker is ambiguous concerning the true distribution over the full state space. The uncertainty is described by an unobservable random state z in the space Z , which drives a probability distribution π_z over the full state space. The investor has a starting prior μ_0 over the hidden state z , which can be interpreted as an unknown parameter, and the posterior μ_t and the conditional likelihood $\pi_{z,t}$ can be obtained by Bayes' rule as described in Section 4.2.3. Finally, let u and v be increasing functions with the same interpretation as the static model, let \mathcal{A} be a time aggregator, and \mathcal{R}_t be a certainty equivalent functional that maps an s^{t+1} -measurable random variable ξ to an s^t -measurable random function.

The recursive smooth ambiguity function is given by:

$$\begin{cases} V_t(C) = \mathcal{A}(C_t, \mathcal{R}_t(V_{t+1}(C))) \\ \mathcal{R}_t(\xi) = v^{-1} \left(\mathbb{E}_{\mu_t} \left[v \circ u^{-1} \mathbb{E}_{\pi_{z,t}} [u(\xi)] \right] \right) \end{cases} \quad (4.3)$$

where V_{t+1} is the continuation value at date t . This equation can be rewritten equivalently as:

$$V_t(C) = \mathcal{A} \left(C_t, v^{-1} \left(\int_Z v \circ u^{-1} \left(\int_S u(V_{t+1}(C)) d\pi_{z,t} \right) d\mu_t(z) \right) \right).$$

The generalized recursive smooth ambiguity model in Eq. 4.3 permits a three-way separation among risk aversion, ambiguity aversion, and intertemporal substitution. This representation is based on the following axioms: (i) weak order, continuity, and sensitivity, (ii) preference over acts for the future is independent of current consumption, (iii) preference over risky consumption is independent of history, (iv) independence for timeless lotteries à la von Neumann and Morgenstern, (v) dynamic consistency, (vi) the preference over second-order acts falls in the SEU theory of Savage (1972) and (vii) the preference over the subdomain of one step-ahead acts and the preference over the subdomain of the corresponding second-order acts are consistent with each other. We refer the reader to Hayashi and Miao (2011) for a more detailed discussion on the axioms and more mathematical details on this model. When $v \circ u^{-1}$ is linear, for example, $v = u$, this model reduces to the recursive utility model of Epstein and Zin (1989) described in Section 4.1.2 as the decision maker is ambiguity neutral.

Ju and Miao (2012) consider the following elasticity of intertemporal substitution aggregator and function of the CRRA family:

$$\begin{aligned} \mathcal{A}(c, y) &= \left[(1 - \beta)c^{1-\rho} + \beta y^{1-\rho} \right]^{\frac{1}{1-\rho}} & \rho > 0, \rho \neq 1 \\ u(x) &= \frac{x^{1-\gamma}}{1-\gamma} & \gamma > 0, \gamma \neq 1 \\ v(x) &= \frac{x^{1-\eta}}{1-\eta} & \eta > 0, \eta \neq 1. \end{aligned}$$

We use the ordinally equivalent aggregator $\mathcal{A}(c, y) = [c^{1-\rho} + \beta y^{1-\rho}]^{\frac{1}{1-\rho}}$, so that utility is given by:

$$V_t(C) = \left(C_t^{1-\rho} + \beta \left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_{\pi_{z,t}} [V_{t+1}^{1-\gamma}] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right)^{\frac{1-\rho}{1-\eta}} \right)^{\frac{1}{1-\rho}}, \quad (4.4)$$

where $\beta = (0, 1)$ is the subjective discount factor, $1/\rho$ represents the elasticity of intertemporal substitution, i.e., the willingness on the part of the consumer to substitute future consumption for present consumption. Also, γ is the risk aversion parameter, and η is the ambiguity aversion parameter. The decision maker displays ambiguity aversion when $\eta > \gamma$ and ambiguity neutrality when $\eta = \gamma$.

We demonstrate here, as in Ju and Miao (2012), that the generalized recursive smooth ambiguity model nests other decision theory models as special cases.

In the limiting case with $\rho = 1$ and $U_t = \ln V_t$, Eq. 4.4 reduces to:

$$U_t = \ln C_t + \frac{\beta}{1 - \eta} \ln \left\{ \mathbb{E}_{\mu_t} \left[\exp \left(\frac{1 - \eta}{1 - \gamma} \ln \left(\mathbb{E}_{\pi_{z,t}} \left(\exp \left((1 - \gamma) U_{t+1} \right) \right) \right) \right) \right] \right\}. \quad (4.5)$$

This specification reduces to the multiplier model with hidden states studied by Hansen and Sargent (2010) as there are two risk-sensitive adjustments: for the distribution $\pi_{z,t}$ that reflects the agent's concerns about the misspecification in the underlying Markov law given a hidden state z and for the distribution μ_t that reflects the agent's concerns about the misspecification of the probabilities assigned to the hidden states. More generally, the recursive smooth model presented in Eq. 4.3 nests the recursive multiplier model with hidden states in Hansen and Sargent (2006) as shown by Hayashi and Miao (2011).

If we take the limit in Eq. 4.5 when $\gamma \rightarrow 1$ and $\eta \rightarrow \infty$, Eq. 4.5 becomes

$$U_t = \ln C_t + \beta \min_z \mathbb{E}_{\pi_{z,t}} [U_{t+1}],$$

which is a utility function that belongs to the class of the recursive multiple-prior model of Epstein and Schneider (2003). In this case, the agent is extremely ambiguity averse by choosing the worst continuation value each period.

Moreover, Hayashi and Miao (2011) shows that the general model in Eq. 4.3 nests the model of Klivanoff et al. (2009) as a special case, i.e.:

$$V_t(C) = u(C_t) + \beta \phi^{-1} \left(\mathbb{E}_{\mu_t} \left[\phi \left(\mathbb{E}_{\pi_{z,t}} [V_{t+1}(C)] \right) \right] \right),$$

where risk aversion and intertemporal substitution are confounded.

In this chapter, we use the representation in Eq. 4.4 to solve our optimal dynamic portfolio choice problem.

4.2.2 Financial Assets

We consider a universe where the decision maker can trade N risky assets and a risk-free asset.

The *gross risk free rate* from time t to time $t + 1$ is defined by $R_{f,t+1} = 1 + r_{f,t+1}$.

The vector of *log-returns* of the risky assets from time t to time $t + 1$ is denoted by $r_{t+1} = (r_{1,t+1}, \dots, r_{N,t+1})'$. The log-returns of the risky asset are assumed to be jointly normally distributed with mean vector $m = (m_1, \dots, m_N)'$ and covariance matrix $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq N}$. While this covariance matrix is supposed to be known, the investor does not know the real value of m . However, she has a normal probability distribution over m characterized by a mean vector $\mu_t = (\mu_{1,t}, \dots, \mu_{N,t})'$ and a covariance matrix $\Lambda_t = (\lambda_{ij,t})_{1 \leq i, j \leq N}$. The decision maker has as starting prior at time $t = 0$ the mean vector $\mu_0 = (\mu_1, \dots, \mu_N)'$ and the diagonal covariance matrix $\Lambda_0 = \text{diag}(\lambda_1^2, \dots, \lambda_N^2)$.

At each time t , the investor chooses the level of consumption C_t , assuming that there is only one consumption good, and the composition of her portfolio, described by the weights $\theta_t = (\theta_{1,t}, \dots, \theta_{N,t})'$ so that the weight invested in the risk-free asset is $\theta_{0,t} = 1 - \sum_{n=1}^N \theta_{n,t}$.

The gross return on her wealth between time t and $t + 1$, given a portfolio θ_t , is

$$R_{t+1}(\theta_t, r_{t+1}) = R_{f,t+1} + \theta_t'(e^{r_{t+1}} - R_{f,t+1}\mathbf{1})$$

where $\mathbf{1}$ is the N -vector of ones and the exponential is computed pointwise.

4.2.3 Learning

In the above utility function specified in Eq. 4.4, the unknown parameter z is the mean of the log-returns of the risky asset, i.e., m . Moreover, the conditional likelihood $p(r|m)$ of observing some returns r is independent of time, which leads us to denote expectations concerning this likelihood by $\mathbb{E}_m[\cdot]$.

Even though the real value of m is unknown to the investor, she has a prior over it. At $t = 0$, the decision maker has a starting prior, and then she will use the observed returns r_t to learn about the distribution of this mean vector. After t observations of the asset returns $\{r_s\}_{s=1}^t$, let \bar{r} be the mean vector. The resulting posterior has a Normal multivariate distribution¹ with mean

$$\mu_t = (\Lambda_0^{-1} + t\Sigma^{-1})^{-1} (\Lambda_0^{-1}\mu_0 + t\Sigma^{-1}\bar{r})$$

¹See section 3.5, multivariate normal models with known variance, in Gelman et al. (2015).

and covariance matrix

$$\Lambda_t^{-1} = \Lambda_0^{-1} + t\Sigma^{-1}.$$

The update rule at each time t is then

$$\begin{aligned}\Lambda_{t+1}^{-1} &= \Lambda_t^{-1} + \Sigma^{-1} \\ \mu_{t+1} &= \Lambda_{t+1}^{-1} \left(\Lambda_t^{-1} \mu_t + \Sigma^{-1} r_{t+1} \right).\end{aligned}\tag{4.6}$$

It is possible to see the update rule for the mean as a weighted average between the previous mean μ_t and the real value on the market r_{t+1} where the ratio between the covariance matrices gives the weights. When the uncertainty becomes null, $\lambda_{t+1} \rightarrow 0$, the Bayesian update gives a higher weight to μ_t and a quasi-null weight to r_{t+1} that is because the mean tends to its real value (no uncertainty) which means that there is no need to use new information on the market. On the other hand, when there is a high uncertainty, λ_{t+1} is significant, and the Bayesian update gives importance to both the previous value and the new information on the market.

We denote the prior and posterior distributions by μ_t to avoid cumbersome notations. Hence, to highlight the dependencies, we write

$$\mu_{t+1} = \mathcal{B}(\mu_t, r_{t+1}),\tag{4.7}$$

where \mathcal{B} is the belief updating function. In this case, we mean the updates of the mean vector and the covariance matrix. Accordingly, expectations with respect to these posteriors should be denoted by $\mathbb{E}_{\mu_t} [\cdot]$.

Now, we show the evolution of Bayesian learning over time by using Eq. 4.6. We assume $T = 100$ and two risky and ambiguous assets on the market. Both assets are equal on the financial market whose log returns follow a normal distribution with mean $m_1 = m_2 = 8\%$ and standard deviation $\sigma_1 = \sigma_2 = 20\%$ and correlation equals 0.5. However, we assume that the decision-maker does not know the real value of the log returns' mean, and the investor has a prior over m with means equal to $\mu_1 = \mu_2 = 11\%$. In order to distinguish between these two stocks, we assume that the decision maker is less ambiguous over asset 1, which means that the starting variance of her prior is lower in asset 1 than asset 2, i.e., $\Lambda_1^2 = 0.004 < \Lambda_2^2 = 0.01$.

Figure 4.1 shows the Bayesian learning evolution of the mean and variance of the posterior probability distribution. The right panel shows that the variance of the belief distribution decreases unevenly over time. At the start, the variance tends to decrease faster compared to the last years. We can justify this behavior as the decision maker gives more weight to the first information she receives as she does not know anything about the assets. In contrast, in the last few years, she already knows a lot from the past, meaning new in-

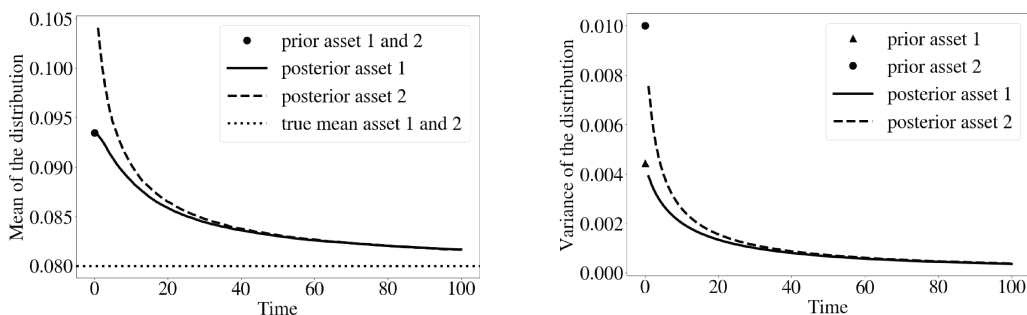


Figure 4.1: Bayesian update of the mean and variance of the posterior probability distribution. The stocks returns' mean are $m_1 = m_2 = 8\%$ and the variance matrix is $\Sigma = \begin{bmatrix} 0.04 & 0.02 \\ 0.02 & 0.04 \end{bmatrix}$, the starting prior means are $\mu_1 = \mu_2 = 11\%$ and variance matrix is $\Lambda_0 = \text{diag}(0.004, 0.01)$.

formation has a smaller weight. The left panel shows that the return of both assets converges to the true value. We see that $\mu_{2,t}$ is closer to the asset's real value even though the decision maker is more ambiguous over this stock. This happens as the update rule for μ_t uses the variance value, which in the case of asset 2 is more significant than asset 1. We should never forget that, even if, on average, $\mu_{2,t}$ is closer to the real value, the second asset prior has a higher variance.

This chapter studies the bayesian learning evolution of two kinds of stocks: foreign and domestic assets. We do not use the previous starting prior as the decision maker would have been equally wrong on both assets and believed to be less wrong over the foreign assets. This seems controversial, so we developed the starting prior as described below.

Starting prior for domestic and foreign stocks

We study the optimal asset allocation in a market with two kinds of stocks: one domestic asset and one foreign asset. We assume that both stocks are equal on the market, i.e., $m_1 = m_2 = 8\%$ and the variance matrix equal to

$$\Sigma = \begin{bmatrix} 0.04 & 0.02 \\ 0.02 & 0.04 \end{bmatrix}.$$

To avoid using an arbitrary value for the decision-maker starting prior, we assume that the decision-maker studies the market for a short time before investing in it. We identify the studying period as $t = -t_{\text{study}}, \dots, 0$ where at $t = 0$ the decision maker starts to invest. In the studying period, the investor only knows about the domestic stock. The investor has a prior over

this domestic stock with mean $\hat{\mu}_t$ and standard deviation $\hat{\lambda}_t$. The update of the prior is done using just the information related to the domestic stock as follows:

$$\hat{\lambda}_{t+1}^2 = \left(\frac{1}{\hat{\lambda}_t^2} + \frac{1}{\sigma_{11}} \right)^{-1}$$

$$\hat{\mu}_{t+1} = \hat{\lambda}_{t+1}^2 \left(\frac{\hat{\mu}_t}{\hat{\lambda}_t^2} + \frac{r_{1,t+1}}{\sigma_{11}} \right).$$

After the first study period, the decision maker decides to invest from time $t = 0$ to T . When she starts to invest at time 0, the agent discovers that another kind of asset exists on the market, i.e., the foreign asset and she updates her posterior distribution using Eq. 4.6. We assume that whenever the decision maker discovers a new stock, she always has the same starting prior for each stock, however in the case of the domestic stock, this prior has been updated for t_{study} years which will lead to having a prior closer to the real value and with less uncertainty. Kilka and Weber (2000) empirically show that subjective probability distributions of stock returns are significantly less dispersed for stocks associated with high competence levels than for stocks associated with low competence levels.

Throughout this chapter, we use two examples to study the different results obtained in the optimization problem. In both cases, we assume a five-year study period without investing.

In the first example, we assume that the decision-maker **overestimates** the stock returns. At time $t = -5$, the starting prior for the domestic asset is $(\hat{\mu}_{-5} = 11\%, \hat{\lambda}_{-5}^2 = 0.01)$ that is updated using the previous rule. At the time $t = 0$, the decision maker starts to invest, and the starting prior for both assets are:

$$\mu_0 = \begin{bmatrix} 9.34\% \\ 11\% \end{bmatrix}$$

with variance:

$$\Lambda_0 = \begin{bmatrix} 0.004444 & 0 \\ 0 & 0.01 \end{bmatrix},$$

where asset 1 is the domestic asset and asset 2 is the foreign one. We represent the posterior in Figure 4.2. In this case, asset 1 is always closer to the true value, and the variance is always lower compared to the posterior distribution of the foreign asset. However, we notice that after a few times, the two distributions tend to converge, and we see that there is little or no difference between them when t is higher than 50 years.

The second example is when the decision maker **underestimates** the stock returns of both stocks. A time $t = -5$, the starting prior for the domestic asset

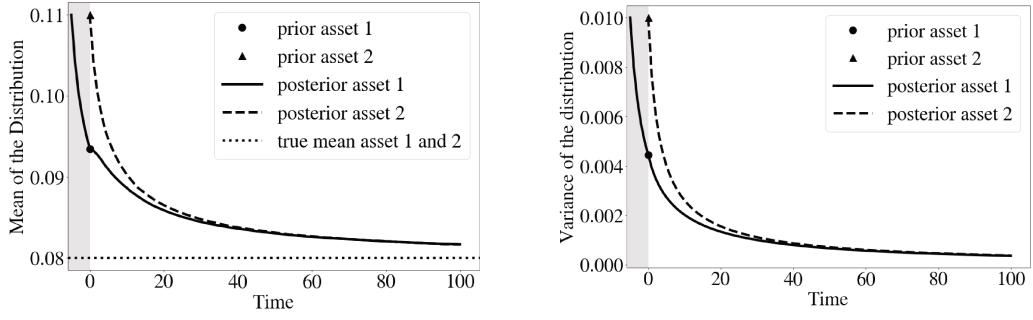


Figure 4.2: Mean and variance of the probability distribution assuming a study period of 5 years and the decision maker overprices the asset. Asset 1 is the domestic asset and asset 2 is the foreign stock.

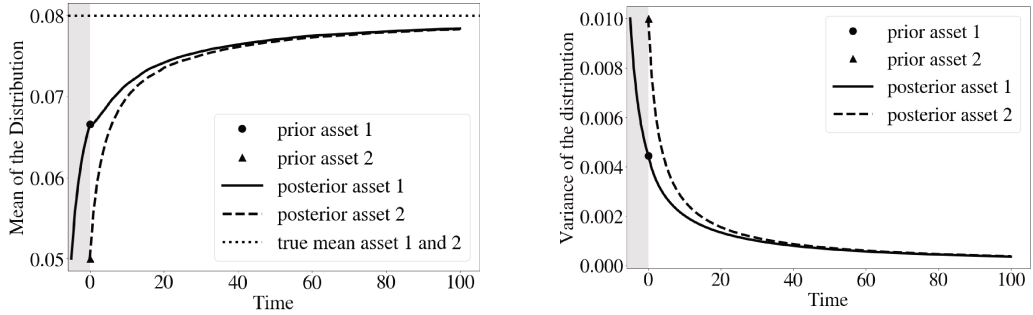


Figure 4.3: Mean and variance of the probability distribution assuming a study period of 5 years and the decision maker underprices the asset. Asset 1 is the domestic asset and asset 2 is the foreign stock.

is $(\hat{\mu}_{-5} = 5\%, \hat{\lambda}_{-5}^2 = 0.01)$. Then, after 5 years of studying period, the decision maker starts to invest, and the starting prior means are:

$$\mu_0 = \begin{bmatrix} 6.65\% \\ 5\% \end{bmatrix}$$

with variance:

$$\Lambda_0 = \begin{bmatrix} 0.004444 & 0 \\ 0 & 0.01 \end{bmatrix}.$$

We see that the variance in the underprice and overprice examples are equal. That is because the variance decreases over time as new information is available on the market. However, it only cares about what the information says. Figure 4.3 shows the Bayesian updating in this example. We see that, even in this case, the mean of the domestic asset is always a little bit closer to the real value of the mean of the distribution compared to the foreign asset.

4.3 The investor's problem

We now describe the investor's problem considering that she has a finite horizon T at which it is optimal for her to consume all her remaining wealth. She has an initial wealth called W_0 and, at each date t , she decides to consume an amount C_t of her available wealth W_t and to invest the remaining ($W_t - C_t$) into the available financial assets. Her intertemporal budget constraint then writes:

$$W_{t+1} = (W_t - C_t) R_{t+1}(\theta_t, r_{t+1}). \quad (4.8)$$

We additionally impose standard borrowing and short sales constraints. Let $W := \{W_t\}_{0 \leq t \leq T}$ be the wealth process, the consumption process C must be in $\mathcal{C}(W) := \{C: 0 \leq C_t \leq W_t, \forall t\}$ and the portfolio process $\theta := \{\theta_t\}_{0 \leq t \leq T}$ must be in Θ^T where $\Theta := \{x \in \mathbb{R}_+^N: \mathbf{1}'x \leq 1\}$. This rules out the possibility that a leveraged investor faces bankruptcy before the horizon T .

The investor must define the consumption level and optimal asset allocation at each time to maximize her utility function. Formally, she has to find the optimal controls $\{C_t^*\}_{t=0}^T$ and $\{\theta_t^*\}_{t=0}^T$ which maximize V_0 :

$$\max_{C \in \mathcal{C}, \theta \in \Theta} V_0(W, \mu, C, \theta)$$

where $W := \{W_t\}_{0 \leq t \leq T}$ and $\mu := \{\mu_t\}_{0 \leq t \leq T}$ are the state variables driven by Eq. 4.8 and 4.7. We recall that μ_t denotes the whole posterior distribution and not only its mean.

Due to the recursive structure of the preference, this problem is amenable to dynamic programming methods. Define the value function

$$J_t(W_t, \mu_t) = \max_{\{C_s\}_t^T \in \mathcal{C}, \{\theta_s\}_t^T \in \prod_t^T \Theta} V_t(\{W_s\}_t^T, \{\mu_s\}_t^T, \{C_s\}_t^T, \{\theta_s\}_t^T)$$

which satisfies the Bellman equation

$$J_t(W_t, \mu_t) = \max_{C_t \in \mathcal{C}(W_t), \theta_t \in \Theta} \left(C_t^{1-\rho} + \beta \left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_m \left[J_{t+1}^{1-\gamma}(W_{t+1}, \mu_{t+1}) \right] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right)^{\frac{1-\rho}{1-\eta}} \right)^{\frac{1}{1-\rho}} \quad (4.9)$$

with terminal value $J_T(W_T, \mu_T) = W_T$ as $C_T = W_T$.

Solving the Bellman equation

The homogeneity to the value function leads to the conjecture $J_t(W_t, \mu_t) = W_t G_t(\mu_t)$. Substituting into the Bellman equation in Equation 4.9 with the

budget constraint, Eq. 4.8, leads to

$$G_t(\mu_t) = \max_{C_t \in \mathcal{C}(W_t), \theta_t \in \Theta} \left(\left(\frac{C_t}{W_t} \right)^{1-\rho} + \beta \left(1 - \frac{C_t}{W_t} \right)^{1-\rho} H_t^{1-\rho}(\mu_t, \theta_t) \right)^{\frac{1}{1-\rho}} \quad (4.10)$$

where

$$H_t(\mu_t, \theta_t) := \left(\mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_m \left[(G_{t+1}(\mathcal{B}(\mu_t, r)) R_{t+1}(\theta_t, r))^{1-\gamma} \right] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right)^{\frac{1}{1-\eta}}. \quad (4.11)$$

Note that we used the belief updating rules $\mu_{t+1} = \mathcal{B}(\mu_t, r)$ to list the relevant variables for the function H_t and note that the inner expectation is over values of r , see appendix for more information.

FOC with respect to consumption: Assuming the solution is an interior point, the first-order condition for C_t leads to

$$\left(\frac{C_t^*/W_t}{1 - C_t^*/W_t} \right)^{-\rho} = \beta H_t^{1-\rho}(\mu_t, \theta_t).$$

As the r.h.s. is independent of W_t , the optimal consumption must be proportional to wealth, that is $C_t/W_t = c_t(\mu_t, \theta_t)$ where

$$c_t(\mu_t, \theta_t) = \left(1 + \beta^{\frac{1}{\rho}} H_t^{\frac{1-\rho}{\rho}}(\mu_t, \theta_t) \right)^{-1}. \quad (4.12)$$

Moreover, as $G_T(\mu_T) = 1 > 0$ and $R_{t+1} > 0$, we obtain by backward recursion that H_t is always positive, hence that $c_t(\mu_t, \theta_t) < 1$ for all t except $t = T$ at which we have equality. Hence the constraint $C_t \in \mathcal{C}(W_t)$ on the consumption the process is not binding, given the prohibition of leverage in the portfolio composition.

Inject this optimal consumption path into the Bellman Eq. 4.10 and use Eq. 4.12 to obtain:

$$G_t(\mu_t) = \max_{\theta_t \in \Theta} \left\{ c_t^{-\frac{\rho}{1-\rho}}(\mu_t, \theta_t) \right\} = \max_{\theta_t \in \Theta} \left\{ \left(1 + \beta^{\frac{1}{\rho}} H_t^{\frac{1-\rho}{\rho}}(\mu_t, \theta_t) \right)^{\frac{\rho}{1-\rho}} \right\}. \quad (4.13)$$

As the maximand is a monotonically increasing function of H_t , we conclude that the optimal portfolio is obtained by maximizing H_t :

$$\theta_t^*(\mu_t) = \operatorname{argmax}_{\theta_t \in \Theta} H_t(\mu_t, \theta_t) \quad (4.14)$$

from which follows the optimal consumption path $C_t^* = W_t c_t(\mu_t, \theta_t^*)$ and the

value function

$$G_t(\mu_t) = c_t^{-\frac{\rho}{1-\rho}}(\mu_t, \theta_t^*). \quad (4.15)$$

FOC with respect to the portfolio composition: The Lagrangian for the optimal portfolio problem in Eq. 4.14 at time t is

$$\mathcal{L}(\theta_t, \boldsymbol{\lambda}_1, \lambda_2) = H_t(\mu_t, \theta_t) + \boldsymbol{\lambda}_1' \theta_t + \lambda_2(1 - \mathbf{1}' \theta_t)$$

where $\boldsymbol{\lambda}_1 \in \mathbb{R}_+^N$ and $\lambda_2 \in \mathbb{R}_+$ are the Lagrange multipliers.

For each asset $1 \leq n \leq N$, the first-order conditions are,

$$\frac{\partial H_t}{\partial \theta_{n,t}} + \boldsymbol{\lambda}_{1,n} - \lambda_2 = 0,$$

where

$$\frac{\partial H_t}{\partial \theta_{n,t}} = H_t^\eta \mathbb{E}_{\mu_t} \left[\left(\mathbb{E}_m \left[G_{t+1}^{1-\gamma} R_{t+1}^{1-\gamma} \right] \right)^{\frac{\gamma-\eta}{1-\gamma}} \mathbb{E}_m \left[G_{t+1}^{1-\gamma} R_{t+1}^{-\gamma} e^{r_{n,t+1}} \right] \right].$$

In the following numerical implementation, this problem is solved numerically and not through this system of equations.

4.4 Numerical Implementation

In this section, we discuss the numerical implementation of the Bellman equation derived in the section before.

We implement the Bellman equation with two ambiguous stocks, i.e., $N = 2$. However, this approach can be extended to a generic N . Firstly, we solve the Bellman equation without learning and then compute it with learning.

The Python code is available upon request.

4.4.1 Solving without learning

Without learning $\mu_t = \mu_0$ and $\Lambda_t = \Lambda_0$ for all t and Eq. 4.11 can be rewritten as $H_t(\mu_t, \theta_t) = G_{t+1}(\mu_0) \tilde{H}(\mu_0, \theta_t)$ where

$$\tilde{H}(\mu_0, \theta_t) := \left(\mathbb{E}_{\mu_0} \left[\left(\mathbb{E}_m \left[(R_{t+1}(\theta_t, r))^{1-\gamma} \right] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right)^{\frac{1}{1-\eta}}. \quad (4.16)$$

At each time step we have

$$\theta^* = \arg \max_{\theta_t \in \Theta} \tilde{H}(\mu_0, \theta_t) \quad (4.17)$$

from which we see that θ^* is independent of t .

The double expectations in \tilde{H} are computed using Gauss-Hermite quadrature (see appendix for details) and `scipy.optimize.minimize()` is applied to obtain θ^* with the constraints and the optimal value $\tilde{H}^*(\mu_0) := \tilde{H}(\mu_0, \theta^*)$.

Once identified the optimal consumption, θ^* , the optimal level of consumption is obtained by using Eq. 4.12 and the value function using Eq. 4.15 using a backward approach.

4.4.2 Approximating the value function

This subsection allows us to obtain some results helpful in solving the optimization problem with learning.

From Eq. 4.15, we know that:

$$G_t^{\frac{1-\rho}{\rho}}(\mu_t) = c_t^{*-1}(\mu_t, \theta_t^*).$$

We identify that it is possible to rewrite $c_t^{-1}(\mu_t, \theta_t^*)$ as a polynomial function of order 2 of the parameters that characterized the posterior belief distribution:

$$c_t^{-1}(\mu_t, \theta_t^*) = \text{Poly}(\mu_t, \Lambda_t \mid \alpha_t).$$

We estimate the parameter α_t using all sample paths available at time t . We obtain an adjusted R^2 close to 1 and a mean squared error that tends to 0. Moreover, the value estimated using the polynomial approximation is close to the real values.

This representation is helpful as we need to express $G_t^{\frac{1-\rho}{\rho}}(\mu_t)$ as a function of μ_t, Λ_t in order to compute the expected value in the case solving with learning.

4.4.3 Solving with learning

1. Generate B samples paths $\left\{ \left\{ r_t^{(b)} \right\}_{t=1}^T : 1 \leq b \leq B \right\}$ and compute the corresponding Bayesian updates $\left\{ \left\{ \mu_t^{(b)} \right\}_{t=0}^T : 1 \leq b \leq B \right\}$ and $\left\{ \left\{ \Lambda_t \right\}_{t=0}^T \right\}$ of the prior.
2. At T : We have $C_T^* = W_T$, equivalently, $c_T^* = 1$.
Therefore $J_T(W_T, \mu_T) = W_T$ and $G_T(\mu_T) = 1$ is constant.
3. At $T - 1$
 - (a) For each $1 \leq b \leq B$ compute the optimal portfolios $\theta_{T-1}^{*(b)}$: given that G_T is independent of the posterior μ_T , we solve for $\theta_{T-1}^{*(b)} =$

$\arg \max_{\theta \in \Theta} \tilde{H}(\mu_{T-1}^{(b)}, \Lambda_{T-1}, \theta_{T-1})$ and we store the optimized value

$$H(\mu_{T-1}^{(b)}, \theta_{T-1}^{*(b)}) = \tilde{H}(\mu_{T-1}^{(b)}, \theta_{T-1}^{*(b)}).$$

- (b) Compute $c_{T-1}(\mu_{T-1}^{(b)}, \theta_{T-1}^{*(b)})$ with Eq. 4.12
- (c) Compute $G_{T-1}(\mu_{T-1}^{(b)})$ with Eq. 4.15.
- (d) Regress the inverse of $c_{T-1}^{(b)}$ on a polynomial basis of degree 2 of $(\mu_{T-1}^{(b)}, \Lambda_{T-1})$, as described in Section 4.4.2, in order to find the polynomial coefficients α_{T-1} :

$$G_{T-1}^{\frac{1-\rho}{\rho}}(\mu_{T-1}^{(b)}) = c_{T-1}^{-1}(\mu_{T-1}^{(b)}, \theta_{T-1}^{*(b)}) = \text{Poly}(\mu_{T-1}^{(b)}, \Lambda_{T-1} \mid \alpha_{T-1}).$$

4. For each $t = T - 2$ to $t = 0$

- (a) For each $1 \leq b \leq B$ compute the optimal portfolios $\theta_t^{*(b)}$ using Eq. 4.11:

$$\theta_t^{*(b)}(\mu_t^{(b)}) = \arg \max_{\theta_t \in \Theta} \left(\mathbb{E}_{\mu_t^{(b)}} \left[\left(\mathbb{E}_m \left[\left(\text{Poly}(\mathcal{B}(\mu_t^{(b)}, r_{t+1}^{(b)}), \Lambda_{t+1} \mid \alpha_{t+1}) \right)^{\frac{1-\rho}{\rho}} \times R_{t+1}(\theta_t^{(b)}, r_{t+1}^{(b)})^{1-\gamma} \right] \right)^{\frac{1-\eta}{1-\gamma}} \right] \right)^{\frac{1}{1-\eta}}$$

where G_{t+1} is rewritten using the polynomial function described in Section 4.4.2 and we remember that $\mu_{t+1} = \mathcal{B}(\mu_t, r_{t+1})$ as in Eq. 4.7.

- (b) Retrieve $H_t(\mu_t^{(b)}, \theta_t^{*(b)})$ from the optimizer, then compute $c_t(\mu_t^{(b)}, \theta_t^{*(b)})$ and $G_t(\mu_t^{(b)})$ using Eq. 4.12 and Eq. 4.15 respectively.
- (c) If $t > 0$, regress the inverse of $c_t(\mu_t^{(b)}, \theta_t^{*(b)})$ as a function of $(\mu_t^{(b)}, \Lambda_t)$ using a polynomial basis of order 2 as as described in Section 4.4.2 in order to obtain the polynomial coefficients α_t :

$$G_t^{\frac{1-\rho}{\rho}}(\mu_t^{(b)}) = c_t^{-1}(\mu_t^{(b)}, \theta_t^{*(b)}) = \text{Poly}(\mu_t^{(b)}, \Lambda_t \mid \alpha_t)$$

The double expectations in H_t are computed using Gauss-Hermite quadrature (see appendix for details) and `scipy.optimize.minimize()` is applied to obtain θ_t^* with the constraints. The Python code is available upon request.

4.5 Results

In this section, we present some numerical examples obtained using 10,000 Monte Carlo simulations, allowing us to obtain smooth results.

In our standard scenario, we assume one risk-free asset and two risky stocks. The risk-free rate is 2% whereas the two risky stocks are equal on the market whose mean is $m_1 = m_2 = 8\%$, and variance is $\sigma_{11} = \sigma_{22} = 0.04$, and they are correlated with a coefficient equal to $\rho = 0.5$.

As described in our theoretical model, the decision maker is ambiguous about the actual value of the mean of the return's distribution. We assume that for asset 1, considered the domestic stock, the decision maker believes that the mean of the distribution is equal to $m_1 + c$ at time t_{-study} and she learns for the period $\{t_{-study}, \dots, 0\}$. More specifically, we will assume two cases: case I (overvalue of the mean of the distribution) where $c = 3\%$ which will lead to having μ_1 at time $t = 0$ equal to 9.38% and case II (undervalue of the mean of the distribution) where $c = -3\%$ which will lead to having μ_1 at time $t = 0$ equal to 6.63%. For the second stock, the foreign asset, the decision maker believes that the mean of the distribution at time 0 is equal to $m_2 + c$, i.e., $\mu_2 = 11\%$ in case I and $\mu_2 = 5\%$ in case II.

The variance of the decision maker's prior belief for the domestic stock at time t_{-study} is equal to 1%. However, it arrives at $\Lambda_1^2 = 0.44\%$ at time 0. For the foreign asset, the variance equals 1% at time 0. The covariance matrix at time $t = 0$ is then equal to

$$\begin{bmatrix} 0.44\% & 0\% \\ 0\% & 1\% \end{bmatrix}.$$

We refer the reader to Section 4.2.3 for a more comprehensive understanding of the bayesian updating rule and the starting prior used.

We see that in this standard scenario, we create a financial market where two stocks are entirely equal in distribution. However, the decision maker has a different probability belief over them, allowing us to identify domestic and foreign stocks.

We define the parameters of the generalized recursive smooth ambiguity model as in Bansal and Yaron (2004): $\gamma = 2$, $\eta = 8.864$, $\beta = 0.975$, and $\rho = 1/1.5$. Moreover, the decision maker starts to invest at time 0 till $T = 75$, as in Peijnenburg (2018).

We perform a sensitivity analysis of the decision maker's home bias when the generalized recursive smooth ambiguity model parameter changes in Section 4.5.1. In order to better comment on the sensitivity, we assume that the decision-maker invests at $t = 0$ by using the prior distribution she has; however, she does not rebalance her portfolio each year, i.e., one-shot investing.

Section 4.5.2 studies the optimal asset allocation and optimal consumption in the case of a dynamic setting where the decision maker rebalances her portfolio each year by using the new posterior belief distribution. We study how the portfolio composition changes if the decision maker is not ambiguous-averse. We also do a comparative analysis with the static model and one-shot investing.

4.5.1 No Rebalancing portfolio

In this section, we assume that the decision-maker does not rebalance her portfolio using the new information available on the market. She invests at time $t = 0$ using the starting prior she has, but she does not adjust her portfolio using the posterior belief distribution. We will use the results obtained in this section as a benchmark for the case where the decision maker rebalances her portfolio each time.

As she does not rebalance her portfolio, the percentage invested in each stock does not change over time. For this reason, we use this scenario to perform a sensitivity analysis of the model parameters.

Different results can be obtained using different priors and market configurations. However, the interpretation of the parameters should be the same.

Risk Parameter

We describe here how the risk parameter can affect the portfolio composition.

Figure 4.4 shows the percentage invested in each stock as a percentage of the total invested (left panel) as well as the market participation (right panel) in the risky market when the risk parameter γ changes. This figure shows the case where the decision-maker overprices the stock returns. When the risk parameter increases, market participation decreases as the investor will try as much as possible to avoid risky situations. However, the quantity invested in stock 2 (i.e., more ambiguous stock) increases compared to the quantity invested in stock 1. This behavior can be justified because the decision-maker believes that asset 2 has a higher return. Moreover, as the decision maker becomes hugely risk averse, she cares less about the uncertainty in the market. Finally, we should consider that the investor is considered ambiguity averse only if $\eta > \gamma$ and, in this case, $\eta = 8.864$.

The portfolio composition and the market participation are also shown in Figure 4.5 in the case of undervaluation of the stock returns. As in the previous Figure, an increase in the risk parameter will lead to a decrease in market participation. In this case, the percentage invested in each stock is constant over time. That is because the more ambiguous stock is believed to

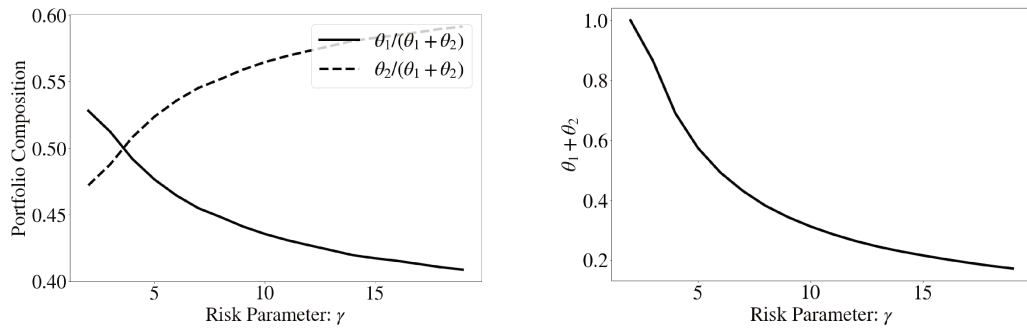


Figure 4.4: *Portfolio composition and market participation when the decision maker does not rebalance her portfolio and when she overestimates the market stock returns.*

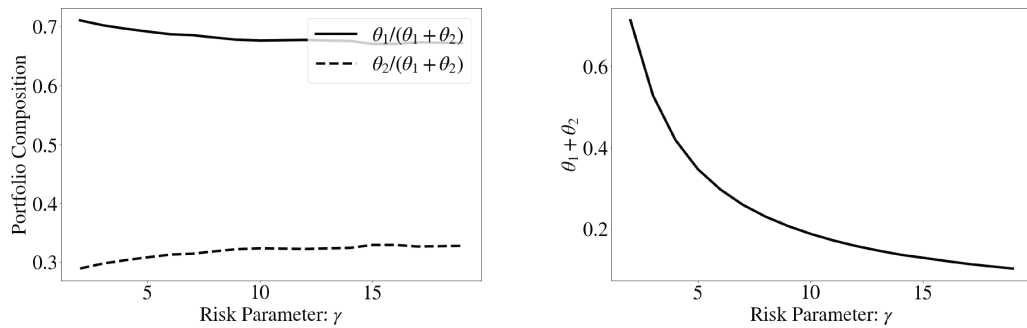


Figure 4.5: *Portfolio composition and market participation when the risk parameter changes in the case without rebalancing and when the decision maker underestimate the market stock returns.*

have less return which means that even though the decision maker would prefer to have a more ambiguous stock in her portfolio, the return is not tempting enough.

In both cases, the home bias decreases when the risk parameter increases. In the first example, we see that the home bias decreases at a point to become negative primarily due to the fact that asset 2 is believed to have a higher return than asset 1.

Moreover, it is possible to show that the optimal level of consumption slightly increases when the risk parameter increases. This can be because the decision maker would prefer to consume a little bit more than invest all the remaining wealth in a risk-free asset.

Ambiguity Parameter

To see the interpretation of the ambiguity parameter, we refer to Figure 4.6 in the case of overpriced returns. We should notice that this is a partic-

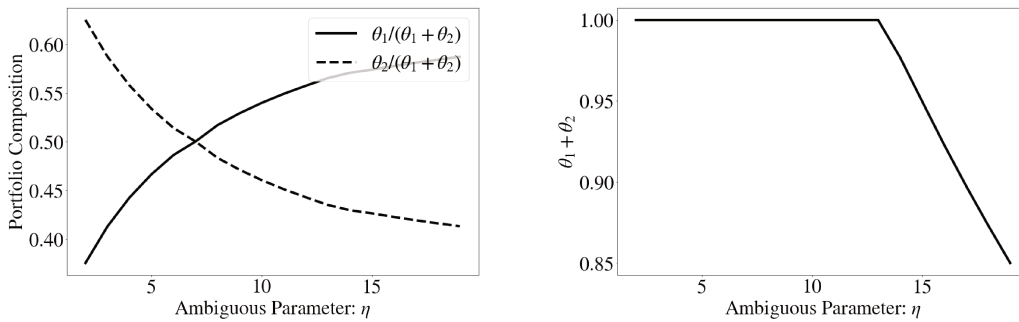


Figure 4.6: *Portfolio composition and market participation when the ambiguity parameter changes in the case without rebalancing and when the decision maker overestimate the market stock returns.*

ular case as using the parameters described at the start of this section and $\gamma = 2$, the optimal portfolio computed without ambiguity would have given 100% participation in the market. This can explain the fact that the market participation stays at 100% even if the ambiguous parameter increases until the ambiguity becomes too big, which will decrease participation. The left panel of the Figure allows us to see different behavior. First of all, at the start, the decision maker is ambiguity neutral, $\eta = \gamma = 2$, and the decision maker tends to invest more in asset 2 as she believes that the returns are higher without caring about the ambiguity of her belief. As she becomes more and more ambiguity averse, she tends to decrease the quantity invested in asset 2. She prefers to invest in asset 1 even though she believes the asset has a smaller return. This is due to the fact that asset 1 is less ambiguous than asset 2. We need this compensation as the investor market participation is equal to 100%.

In the case of overpricing, we are not in the same situation as before, and Figure 4.7 shows that when the ambiguity-averse parameter increases, market participation decreases. We do not have the same problems as in the previous case because the CRRA market portfolio is not equal to 100% without ambiguity. The left panel of the Figure shows that an increase of the ambiguity parameter will not change the portfolio composition much, even though we see that the decision-maker will tend to slightly increase the percentage invested in stock 1 by decreasing stock 2.

In both cases, an increase in the ambiguity aversion parameter will tend to increase the home bias. In Figure 4.6, we see that the home bias is negative at the start, primarily because the decision maker believes that asset 2 has higher returns than asset 1. However, when the ambiguity parameter increases, the home bias increases (becoming less negative and then positive).

An increase in the ambiguity aversion parameter provokes a slight increase in the optimal consumption level. As discussed for the risk parameter, the

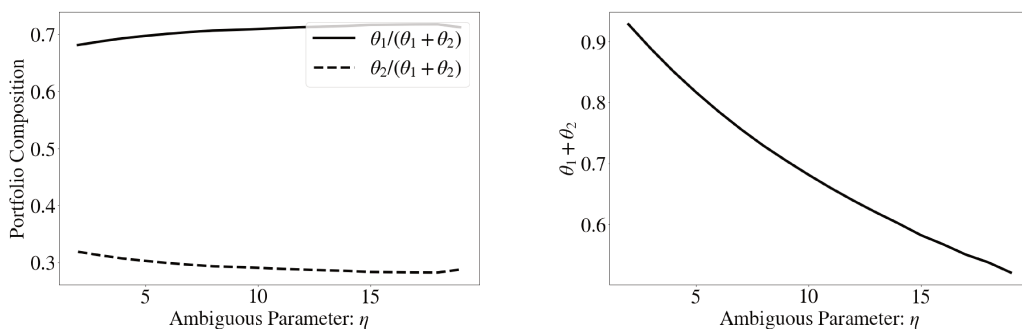


Figure 4.7: *Portfolio composition and market participation when the ambiguity parameter changes in the case without rebalancing and when the decision maker overestimate the market stock returns.*

increase is relatively small, and we can conclude that the ambiguity and risk parameters do not significantly influence the optimal consumption level.

Intertemporal Substitution Parameter

The elasticity of the intertemporal substitution (EIS) parameter, which is the inverse of ρ , influences the optimal consumption level. However, it does not influence the quantity invested in the risky assets or the quantity invested in each asset. The EIS is usually defined as the percentage change in consumption growth induced by a 1 percent increase in the rate of return on investment. When EIS is higher than 1, this effect pushes the agent to substitute consumption today for more continuation utility and vice versa. Flynn et al. (2023) proves that the previous simple intuition is fully general.

Figure 4.8 shows that the optimal consumption level is an exponential function that reaches 1 when $t = T$. In order to better show the effect of ρ , the results are also presented in Figure 4.9, which shows the logarithm of the optimal consumption level. Both Figures show the optimal consumption level when the decision-maker overvalues (left panel) and undervalues (right panel) the asset returns. As expected, the consumption level increases over time till it reaches 100%, which is a condition of our model. When ρ increases, the consumption level decreases, assuming the same t , and vice versa. This is because when EIS increases, we give more value to future consumption, which means that the actual consumption is small in order to have a higher amount of wealth in the future to consume. From another point of view, a decrease in the elasticity of intertemporal substitution, or an increase of ρ , will lead to prefer actual consumption at the expense of future consumption. The same behavior happens whether the decision-maker overvalues or undervalues asset returns.

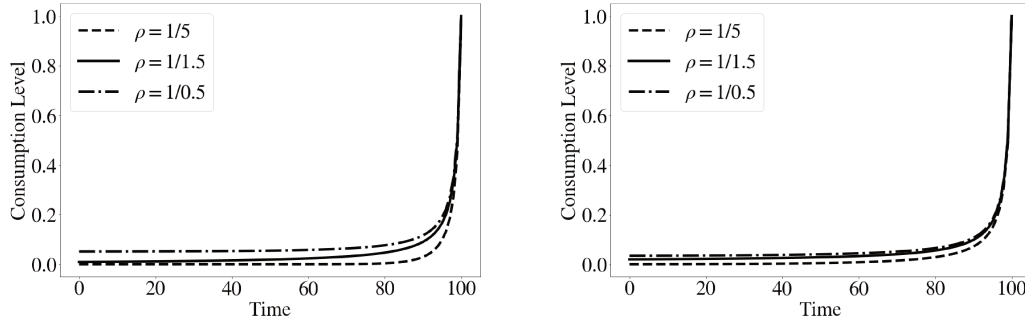


Figure 4.8: Optimal consumption level with different intertemporal substitution parameter in the case without rebalancing. Left Panel: the decision maker overprice the asset returns, Right panel: the decision maker underprice the asset returns.

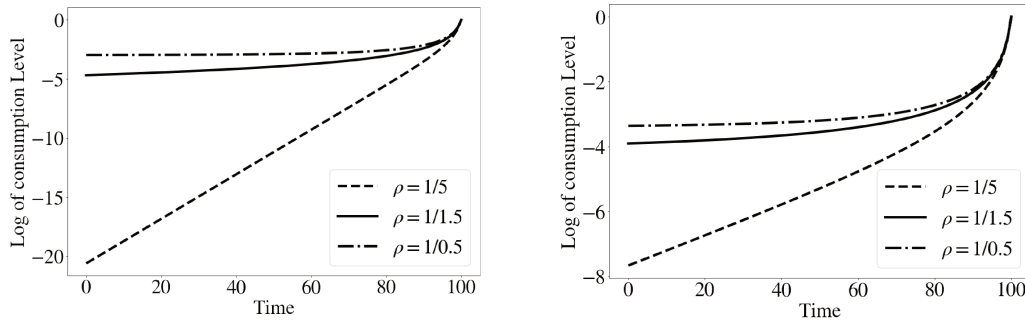


Figure 4.9: The logarithm of the optimal consumption level with different intertemporal substitution parameter in the case without rebalancing. Left Panel: the decision maker overprice the asset returns, Right panel: the decision maker underprice the asset returns.

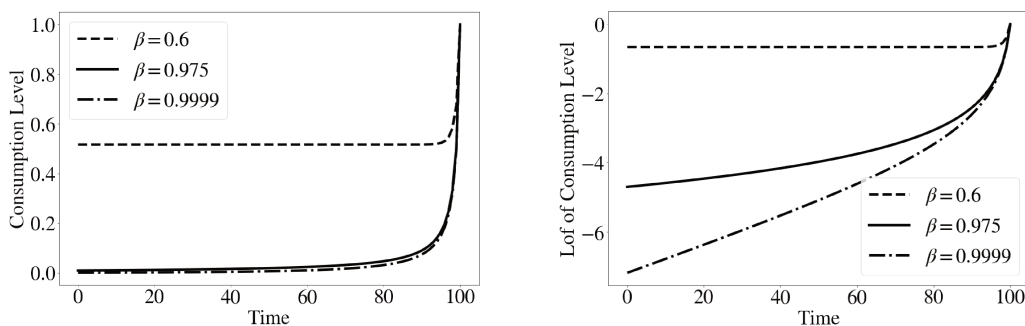


Figure 4.10: Optimal consumption level with different subjective discount parameters in the case without rebalancing.

Subjective discount parameter

In Figure 4.10, we show how β can include the optimal consumption level and the logarithm of the optimal consumption level. The results do not change for cases where the agent overvalues or undervalues the assets. We see that when the subjective discount parameter is far from 1, the agent tends to increase her consumption level as it gives less weight to the future utility, whereas in the case beta is closer to 1, the agent tends to give more weight to the future utility which will lead to consuming less at the start.

Changes in the decision maker's prior

In this part, we discuss how optimal asset allocation and the optimal consumption level change when the decision maker's prior changes. On the basic scenario, we assume that both stocks are equal on the market and the prior on both stocks are equal, $\mu_1 = \mu_2 = 8\%$ and $\sigma_{11} = \sigma_{22} = 4\%$ with a correlation value of 0.5. We also assume in this case that the decision maker has the same confidence level for both stocks, i.e., $\lambda_{11} = \lambda_{22} = 1\%$. We remind the reader that, in this case, the decision maker cannot rebalance her portfolio using the new information available on the market.

Figure 4.11 shows how the optimal asset allocation and market participation change when the decision maker's prior mean for asset 1 changes. First, we should note that this change does not influence consumption. Everything else equal, when the decision maker believes that the mean return of one stock is higher than another asset, she will tend to invest more in that asset and more in the market.

Then, we study what happens to the portfolio composition in case of changes in the standard deviation of the decision maker's prior. Figure 4.12 shows that when the decision maker believes to have more uncertainty over the mean's return, she will tend to invest less in that stock and compensate

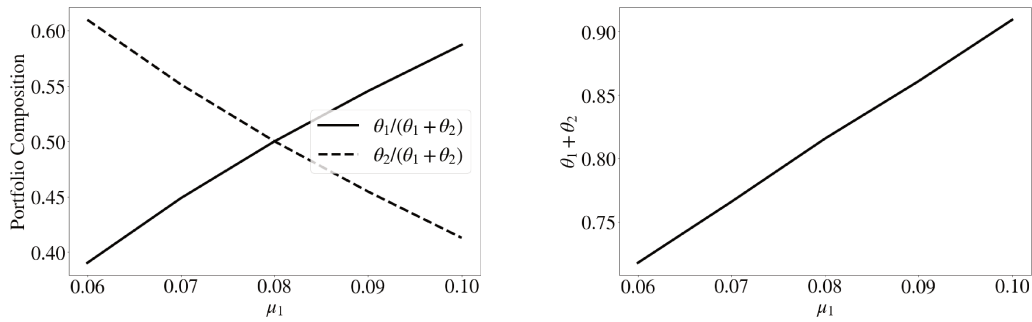


Figure 4.11: *Optimal asset allocation and market participation given different value of the decision maker prior's mean without rebalancing.*

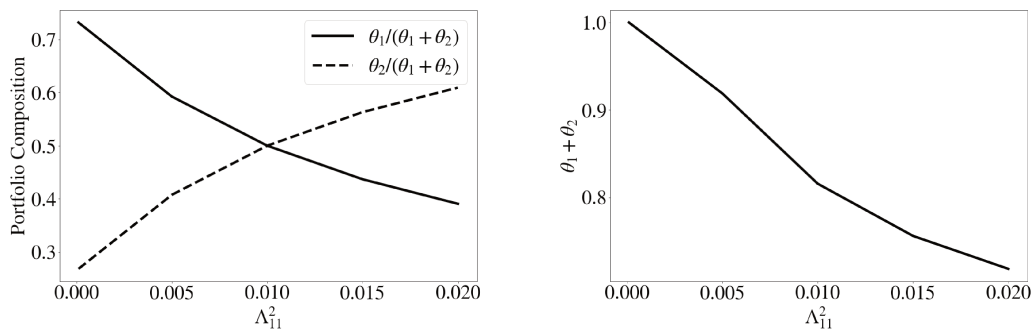


Figure 4.12: *Optimal asset allocation and market participation given different value of the decision maker prior's variance without rebalancing.*

by investing in the less ambiguous stock. However, as the compensation is insufficient, the agent will invest less in the risky market.

4.5.2 Solving with rebalancing

In this section, we study the evolution of home bias over time using the generalized recursive smooth ambiguity model. The results are obtained by applying the numerical approach described in Section 4.4 using 10,000 simulations, assuming that the decision-maker re-balances her portfolio each year.

To give an overview of the time that this numerical approach takes, an example with 1,000 simulations and $T = 75$ takes 44 minutes on a computer with the following characteristics: CPU INTEL roket lake I7-11700KF 3.6GHZ, and 2 RAM sticks DDR4 16GB 3600MHZ.

We study the effect of learning and dynamic setting on optimal portfolio choice and home bias. Moreover, we analyze the effect of ambiguity aversion on optimal portfolio choice, home bias, wealth, and consumption.

We show that the home bias diminish over time, as empirically shown by Bekaert et al. (2017).

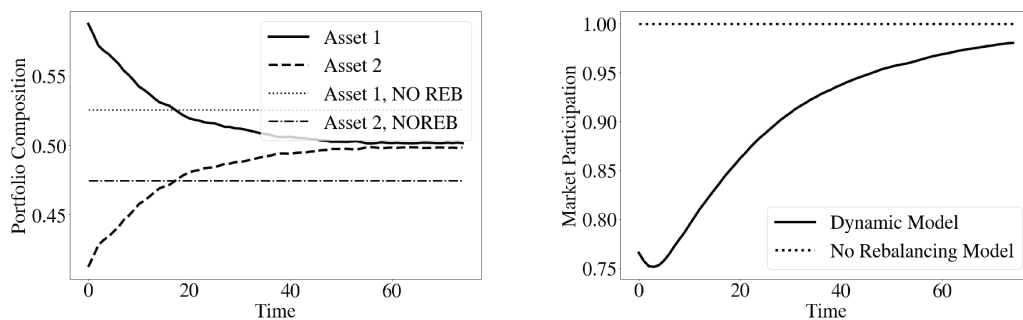


Figure 4.13: *Optimal asset allocation and market participation with and without rebalancing in a dynamic setting when the decision-maker overestimates the stocks' returns.*

Effect of learning on optimal portfolio choice and home bias

To study the effect of learning, we decide to compare two scenarios: the standard scenario, where the decision-maker updates her belief each year, and she rebalances her portfolio consequently, and the scenarios one-shot investing portfolio, where the decision-maker creates her portfolio according to her starting belief and then she does not update it.

Figure 4.13 shows the optimal portfolio allocation (as a percentage of the overall risky portfolio) and the risky market participation of the investors. We assume that the investor overestimates the stock's returns, and we compare a rebalancing situation with a no-rebalancing setting that gives a static portfolio composition. The right panel shows that the decision maker has higher participation in the case of no rebalancing. This is because she believes that the stock's returns are higher than the real ones, and she does not update her belief according to that. Moreover, without rebalancing, she does not have to consider the future value given by G_{t+1} . In the case of rebalancing, we see that market participation decreases at the start. This is justified by the fact that the impact tied to the diminishing uncertainty of the decision-maker over the stock's returns is not significant enough to compensate for the negative impact linked to the decision-maker, which learns that the real mean of the stocks is lower than expected. We use the left panel to discuss the portfolio composition and the home bias. We see that in the case of rebalancing, the home bias decreases over time mainly because the decision-maker learns over time that the two stocks are precisely equal. However, in the case without rebalancing, the gap is always constant. We see that in the early years, the home bias is more significant in the case of rebalancing, which could be justified by the fact that the decision-maker has to take into consideration the future value when optimizing her portfolio. She should also consider that she will not rebalance her portfolio.

Figure 4.14 shows the optimal portfolio composition and market participation when the decision-maker underestimates the stocks' returns. On the right panel, we see that the market participation without learning is around 70%. At the start, the decision maker tends to invest less, which could be justified by the fact that she considers that she can adjust her portfolio at each time according to her belief. Then, the market participation after 10 years in the case of the rebalancing model is higher than the no rebalancing model mainly because the investor knows that the stocks' returns have a higher mean and she has less uncertainty over the market. In the left panel, the home bias is exaggerated in the case without rebalancing: the decision maker believes that asset 1 has a higher return, and she is more confident about it. As she cannot rebalance her portfolio, she invests more in the asset she believes has a higher return and less uncertainty. Moreover, without rebalancing, the investor does not care about the future value. Optimal asset allocation and market participation with and without rebalancing in a dynamic setting when the decision maker overestimates the stocks' returns.

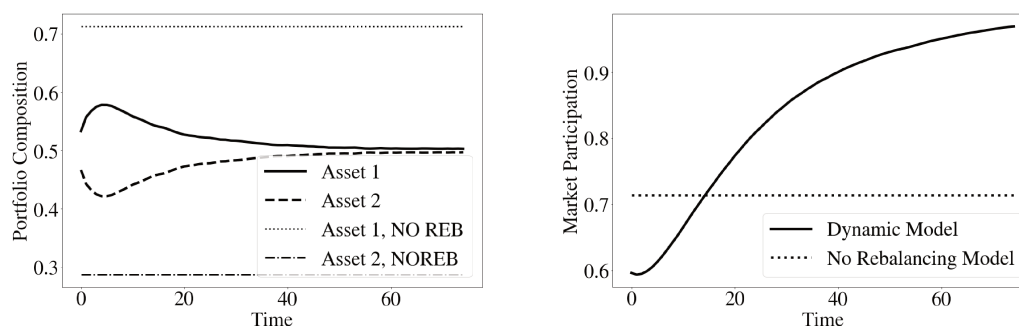


Figure 4.14: *Optimal asset allocation and market participation with and without rebalancing in a dynamic setting when the decision-maker underestimates the stocks' returns.*

Effect of continuation value function

In this subsection, we compare the standard dynamic model and the KMM model, which is the static version of the model used in this chapter. In the static model, the decision maker does not take into consideration the utility that she can obtain by future consumption, and this model re-computes at each time the optimal portfolio allocation by using the update beliefs, but without considering the continuation value function, i.e., $G_{t+1}(\cdot)$. The previous section compared the standard dynamic model with a model that does not rebalance. In this section, the decision maker can rebalance her portfolio. However, she does not consider the utility that she can obtain for future consumption as a parameter to use for solving her decision problem.

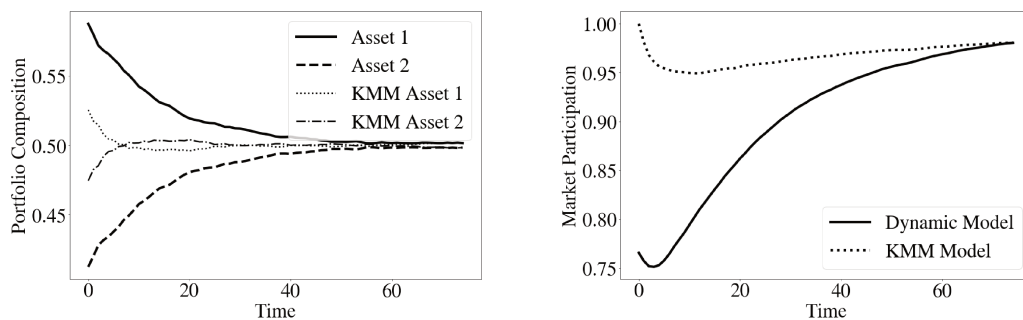


Figure 4.15: *Optimal asset allocation and market participation with learning in a dynamic and static setting when the decision-maker overestimates the stocks' returns.*

Figure 4.15 shows the portfolio composition and the market participation when the decision-maker overestimates the stocks' returns. In the static setting, market participation is higher as it does not consider future consumption. Moreover, we see that it decreases at the start (due to learning that the stocks' return is, in reality, lower than the one at the start), and then it slightly increases (due to the increase in confidence). We have the same behavior in the dynamic model. However, it is more accentuated as market participation at the start is lower than in the static setting, primarily because it should consider future consumption. In the left panel, we show market participation. In the case of a dynamic setting, the home bias seems to exist only at the start, whereas in the dynamic model it persists for a longer time. Moreover, the home bias is more pronounced in the dynamic setting. For this reason, we can say that the dynamic model exaggerates the home bias presented in the data.

A similar effect is seen in Figure 4.16 where the optimal portfolio composition and market participation is shown in case of undervalue of the stocks' return. In the right panel, we see that the KMM model has higher market participation than the dynamic model mostly due to the fact that the decision-maker does not care about the future value of the KMM model, which will lead her to invest more in the risky market. We also see that the market participation increases over time because the ambiguity decreases, and the investor learns that the stock returns are higher than expected at the start. Moreover, the presence of the continuation value slightly decrease the market participation in the first years. Using the information from the left panel, we can discuss the home bias. In both cases, we see that the home bias decreases over time as the ambiguity decreases and the decision maker discovers that both stocks are equal. We see that the KMM model has a higher home bias in the early years than the dynamic model. This is because they have higher market participation, and they need to create a portfolio that is also coherent with

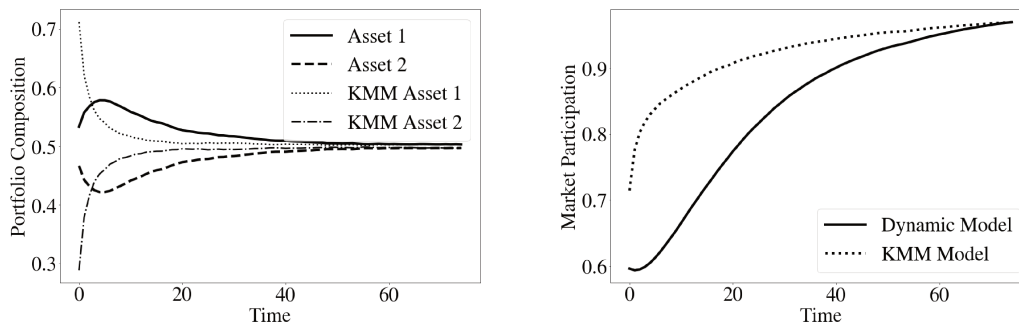


Figure 4.16: *Optimal asset allocation and market participation with learning in a dynamic and static setting when the decision-maker underestimates the stocks returns.*

the ambiguity aversion parameter. In absolute term, the quantity invested in asset 1 for the KMM and the dynamic model is the same. However, it is scaled differently inside the portfolio. Moreover, the presence of the continuation value creates a controversial result: at start the home bias is lower than in the following years. As both model takes in consideration the fact that the decision maker learns over time and the uncertainty decreases, this can be seen as the effect of the continuation value function.

Effect of horizon on optimal portfolio choice and home bias

We study the effect of the different horizons on portfolio composition and market participation. We compare the results obtained with $T = 75$ and $T = 35$.

In Figure 4.17 when the agent overestimates the stock's returns and Figure 4.18 when the agent underestimates the stock returns. We see that the home bias changes according to the horizon. If we have a shorter horizon, the home bias is more significant at the start but decreases faster. Moreover, we see that with a shorter horizon, the decision maker tends to invest less at the start compared to a longer horizon. This can be justified because the investor with a shorter horizon has less time to compensate for negative returns, which will lead him to avoid investing in the risky market and to have a more extensive home bias as she prefers to invest in the less uncertain asset. As she learns that the two stocks are equal and there is less uncertainty in the market, she tends to invest more than an investor with a longer horizon as she has less future to consider.

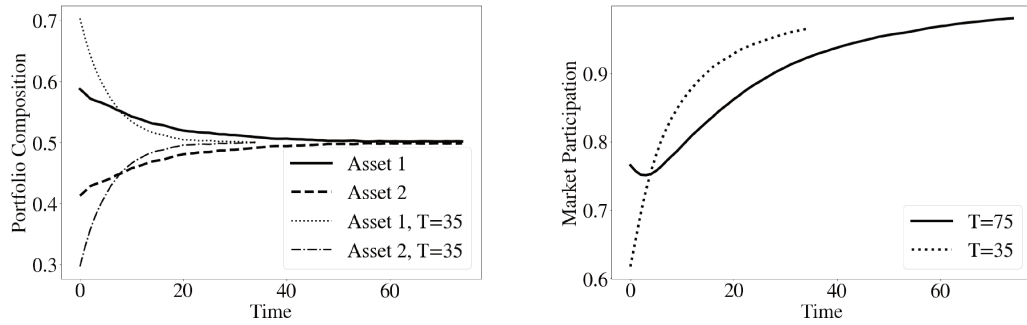


Figure 4.17: *Optimal asset allocation and market participation with learning in a dynamic setting with two different horizon when the decision-maker over-estimates the stocks returns.*

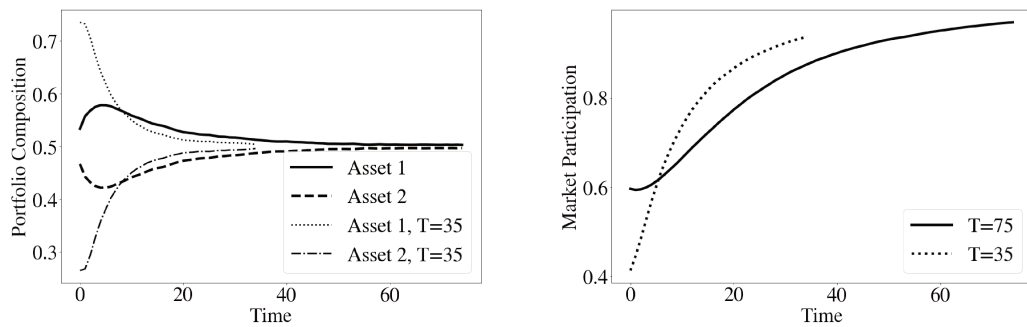


Figure 4.18: *Optimal asset allocation and market participation with learning in a dynamic setting with two different horizon when the decision-maker under-estimates the stocks returns.*

Effect of ambiguity aversion on optimal portfolio choice and home bias

In order to better understand the effect of ambiguity in our dynamic model, we compare two models: when the decision maker is ambiguity averse and when the decision maker is ambiguity neutral.

Figure 4.19 compares these two models. The right panel shows that an ambiguity-averse agent will tend to invest less than an ambiguity-neutral agent. Moreover, an ambiguity-neutral agent will have a more negligible home bias than an ambiguity-averse agent. Even though the agent is ambiguity neutral, she will slightly prefer asset 1 over asset 2 even though asset 1 has a lower return.

Figure 4.20 shows the optimal portfolio allocation and market participation when the investor is ambiguity averse and ambiguity neutral. The investor tends to invest less when she is ambiguity averse. However, an ambiguity-averse investor has a higher home bias at the early stages than an ambiguity-neutral investor. This could be due to the fact that asset 1 is considered to have a higher mean than asset 2 and the fact that the investor does not care about the ambiguity in the market.

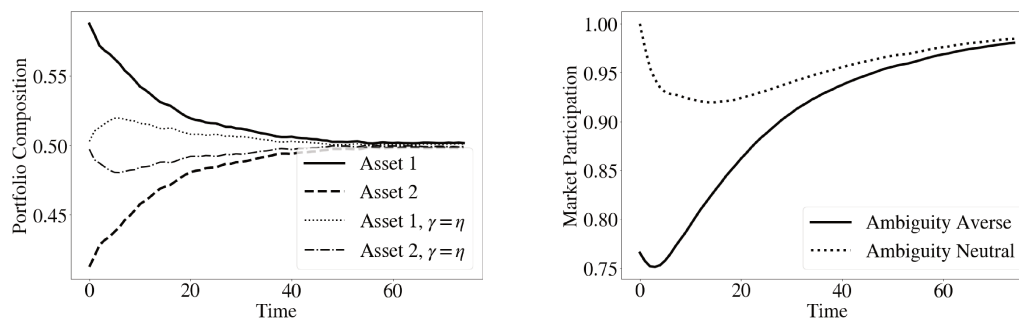


Figure 4.19: *Optimal asset allocation and market participation with learning in a dynamic setting when the decision-maker overestimates the stocks returns.*

Effect of ambiguity aversion and learning on wealth and consumption

Finally, we study the effect of ambiguity aversion and learning on wealth and consumption.

Figure 4.21 shows that the logarithm of the optimal level of consumption, c_t , does not depend on the investors' ambiguous belief or the portfolio composition rebalancing. We see that the consumption increases over time, reaching 100% at time T , which is our starting constraint. We see the same behavior in the case of overestimation of the stock returns as well as in the underestimation case.

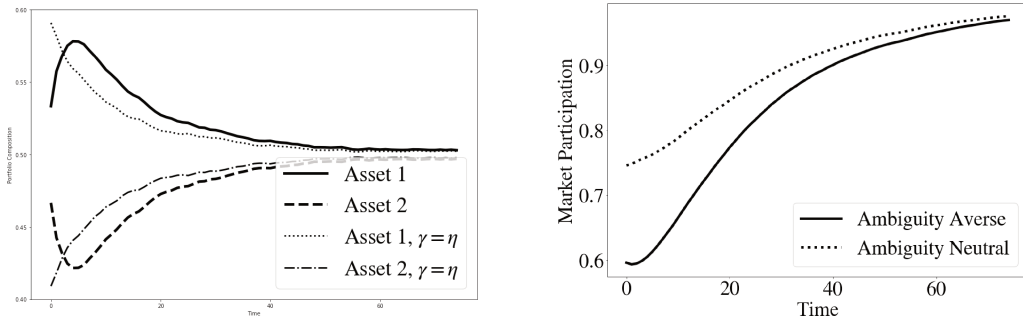


Figure 4.20: *Optimal asset allocation and market participation with learning in a dynamic setting when the decision-maker underestimates the stocks returns.*

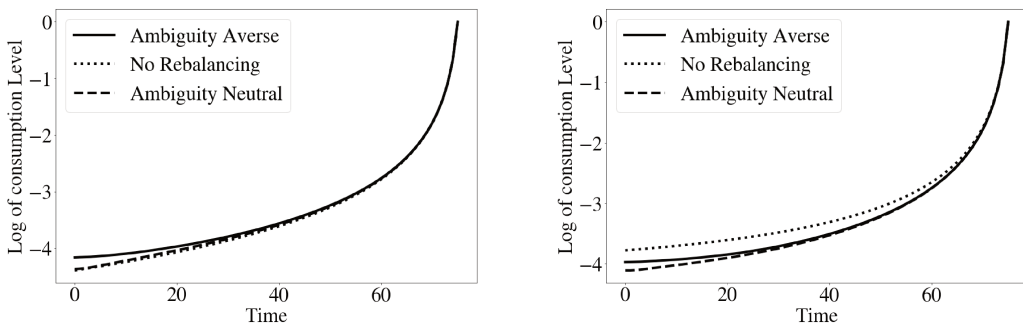


Figure 4.21: *Optimal consumption level in the case the decision-maker is ambiguity averse or ambiguity neutral. Right Panel: the decision-maker overestimates the stocks returns. Left Panel: the decision-maker underestimates the stocks returns.*

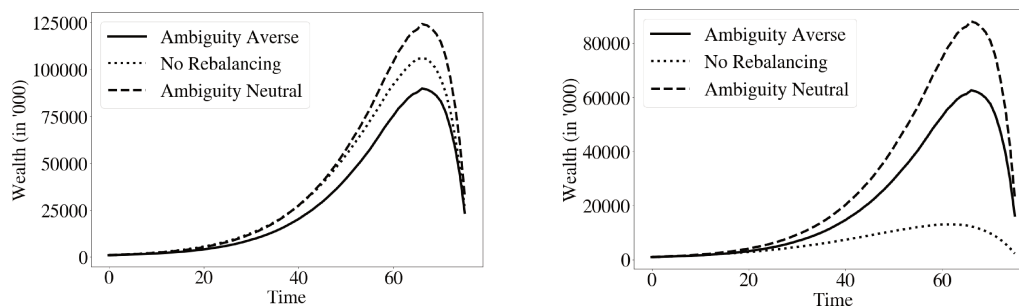


Figure 4.22: *Optimal wealth in the case the decision maker is ambiguity averse or ambiguity neutral. Right Panel: the decision-maker overestimates the stocks returns. Left Panel: the decision-maker underestimates the stocks returns.*

Figure 4.22 shows the wealth of our investors, assuming that the starting wealth is equal to 1,000€. In the left panel, we show the wealth in the case the decision maker overestimates the stock’s return, whereas, in the right panel, we show the wealth in the case the decision maker underestimates the stock’s return. We notice that in both panels, an ambiguity-averse investor will have lower wealth than an ambiguity-neutral one because an ambiguity-neutral investor tends to invest more than an ambiguity-averse investor. The case of no-rebalancing is linked to the starting prior. In the left panel, the starting prior is higher than the real return, which will lead the decision maker to invest more in the markets, i.e., higher wealth. However, in the case of bankruptcy, this could create extreme losses. In the case of the underestimation of stock returns, we see that the ambiguity-neutral investor will always have a higher wealth than ambiguity-neutral ones (as described before). However, we see that in the case of no rebalancing, the wealth is low primarily because the decision-maker underestimates the stocks’ return and does not adjust her portfolio. This will lead to less participation in the market and lower wealth.

4.6 Conclusion

In this chapter, we study the home bias in a dynamic setting and the portfolio choice problem where investors are ambiguous about stock returns but learn about their distribution. The investor’s preferences are modeled with the generalized recursive smooth ambiguity model of Hayashi and Miao (2011) that can separate among risk aversion, ambiguity aversion, and intertemporal substitution. To solve this problem, we develop a numerical approach that extends the existing ones to tackle the specificity of our framework by using a double Gauss Hermite quadrature and a backward recursive approach. Our re-

sults show that ambiguity-averse investors participate less in financial markets and have a higher home bias compared to ambiguity-neutral decision-makers. However, learning over time allows the investor to adjust her prior closer to the real mean of stock return and adjust her optimal portfolio composition.

This study could be extended by using a different specification for the utility function. We use a CRRA function that allows us to solve the optimal asset allocation independently from the level of wealth. Using a different utility function would allow computing the optimal portfolio composition in the function of the wealth level. Another path for extension is to use a different way to update the posterior belief of the decision-maker instead of Bayesian learning. Bayesian learning assumes that an investor weights the new information equally independently of the portfolio composition. However, an investor may not learn about the asset if it is not in her portfolio, and, on the other hand, she learns a lot about that asset if it is a significant share of her portfolio. For example, Abreu et al. (2011) shows that investors tend to learn by trading such as the more active they are in the domestic market, the earlier they start to enter in new markets given the new knowledge they have.

The solution of a model that uses wealth as a parameter and a model that includes the fact that learning depends on the previous portfolio composition is more complex than the model we propose here. This is due to the fact that our model solves the optimal portfolio composition backward, i.e., from T to 0. It could be possible to solve this problem using a grid search point, but it would become more complex and time-consuming.

Appendix

Numerical Implementation: Gauss-Hermite Quadrature

This procedure is used to find the optimal θ_t for each time $t \in [0, T - 2]$ and for each simulation $b \in B$. Gauss-Hermite (GH) quadrature

1. Create two 4th order N_{Gauss} -dimensional tensors which give the values of $r_{1,t+1}$ and $r_{2,t+1}$ for the GH quadrature using the Cholesky decomposition

$$\begin{aligned}\Xi_{p_1 p_2 n_1 n_2}^{1,(b)}(\mu_t, \Lambda_t) &= \mu_{1,t}^{(b)} + \sqrt{2(\lambda_{11})_t} \xi_{p_1} + \sqrt{2\sigma_{11}} \xi_{n_1} \\ \Xi_{p_1 p_2 n_1 n_2}^{2,(b)}(\mu_t, \Lambda_t) &= \mu_{2,t}^{(b)} + \varrho \sqrt{2(\lambda_{22})_t} \xi_{p_1} + \sqrt{2(\lambda_{22})_t(1-\varrho_t^2)} \xi_{p_2} \\ &\quad + \rho \sqrt{2\sigma_{22}} \xi_{n_1} + \sqrt{2\sigma_{22}(1-\rho^2)} \xi_{n_2}\end{aligned}$$

where

$$\begin{aligned}\varrho_t &= \frac{(\lambda_{12})_t}{\sqrt{(\lambda_{11})_t \cdot (\lambda_{22})_t}} \\ \rho &= \frac{\sigma_{12}}{\sqrt{\sigma_{11} \cdot \sigma_{22}}}\end{aligned}$$

- (a) At each GH node $(p_1 p_2 n_1 n_2)$

- i. Compute the return of the portfolio

$$\begin{aligned}R_{p_1 p_2 n_1 n_2}^{(b)}(\theta_t)_{t+1} &= R_{f,t+1} + \theta_{1,t} \left(\exp \left(\Xi_{p_1 p_2 n_1 n_2}^{1,(b)}(\mu_t, \Lambda_t) \right) - R_{f,t+1} \right) \\ &\quad + \theta_{2,t} \left(\exp \left(\Xi_{p_1 p_2 n_1 n_2}^{2,(b)}(\mu_t, \Lambda_t) \right) - R_{f,t+1} \right)\end{aligned}$$

- ii. Compute the Bayesian update of the prior

$$\begin{aligned}\Lambda_{t+1} &= \left(\Lambda_t^{-1} + \Sigma^{-1} \right)^{-1} \\ \mu_{p_1 p_2 n_1 n_2}^{(b)}(\mu_t^{(b)}, \Lambda_t)_{t+1} &= \Lambda_{t+1} \left(\Lambda_t^{-1} \mu_t^{(b)} + \Sigma^{-1} \left(\begin{array}{c} \Xi_{p_1 p_2 n_1 n_2}^{1,(b)}(\mu_t, \Lambda_t) \\ \Xi_{p_1 p_2 n_1 n_2}^{2,(b)}(\mu_t, \Lambda_t) \end{array} \right) \right)\end{aligned}$$

- iii. Compute the values of G_{t+1} knowing the α_{t+1}

$$\begin{aligned}G_{p_1 p_2 n_1 n_2}^{(b)} &\left(\mu_{p_1 p_2 n_1 n_2}^{(b)}(\mu_t^{(b)}, \Lambda_t)_{t+1}, \Lambda_{t+1} \mid a_{t+1} \right)_{t+1} \\ &= \text{Poly} \left(\mu_{p_1 p_2 n_1 n_2}^{(b)}(\mu_t^{(b)}, \Lambda_t)_{t+1}, \Lambda_{t+1} \mid \alpha_{t+1} \right)^{\frac{T-(t+1)}{1-\rho}}\end{aligned}$$

(b) Evaluate the double integration with Gauss Hermite quadrature

i. In the case without learning:

$$\tilde{H}(\mu_0, \Lambda_0, \theta^*) \approx \left(\pi^{-1} \sum_{p_1, p_2=1}^{N_{\text{Gauss}}} \omega_{p_1} \omega_{p_2} \left[\pi^{-1} \sum_{n_1, n_2=1}^{N_{\text{Gauss}}} \omega_{n_1} \omega_{n_2} \left(R_{p_1 p_2 n_1 n_2}(\theta_t)_{t+1} \right)^{1-\gamma} \right]^{\frac{1-\eta}{1-\gamma}} \right)^{\frac{1}{1-\eta}}$$

ii. In case with learning:

$$H_t(\mu_t^{(b)}, \Lambda_t, \theta_t^{*(b)}) \approx \left(\pi^{-1} \sum_{p_1, p_2=1}^{N_{\text{Gauss}}} \omega_{p_1} \omega_{p_2} \left[\pi^{-1} \sum_{n_1, n_2=1}^{N_{\text{Gauss}}} \omega_{n_1} \omega_{n_2} \left(G_{p_1 p_2 n_1 n_2}^{(b)}(\cdot | a_{t+1})_{t+1} R_{p_1 p_2 n_1 n_2}^{(b)}(\theta_t)_{t+1} \right)^{1-\gamma} \right]^{\frac{1-\eta}{1-\gamma}} \right)^{\frac{1}{1-\eta}}$$

iii. `scipy.optimize.minimize()` is applied to $-H_t^{(b)}$ to obtain $\theta_t^{*(b)}$ with the constraints.

General Conclusion

This thesis is at the frontier of economics and finance as it studies decision theory both from a risk management viewpoint and from an investor viewpoint.

Chapter 1 gives an overview of the two most classic risk measures adopted by risk managers in order to avoid bankruptcy: Value-at-Risk (VaR) and Tail Conditional Expectation (TCE). In order to study an implicit utility function between regulators, we compare the quantiles of VaR and TCE. We discover that the Swiss Solvency Test can be considered more risk-averse than Solvency II regulation if we assume that claims follow a generalized Pareto distribution. Moreover, we introduce a new risk measure, the high-order TCE, which can consider high-order risk in the distribution's tails. We compare the quantiles of this new risk measure with more classic risk measures to illustrate the interplay between implicit choices of risk measures by regulators and the characteristics of probability distribution tails. The study assumes that claims follow the Pareto type I distribution or the generalized Pareto distribution. However, a possible extension could examine the relationship between risk measures where semi-heavy tails are modeled using the infinitely divisible probability distribution.

In chapter 2, we study the four properties that a risk measure should have in order to be defined as coherent: subadditivity, monotonicity, positive homogeneity, and translation invariance. We discuss that VaR is not a coherent risk measure, whereas TCE is defined as a coherent risk measure. However, we see that it is not always possible to respect the four axioms. We highlight that the extended TCE introduced in the previous chapter is not a coherent risk measure. However, we introduce a variation of the extended TCE that is able to respect the coherence axioms, if not fully, at least in a weak sense.

To help risk managers to identify which annual moment is more stable over time, we perform an empirical study in chapter 3 by analyzing seventeen equity indices and the 3-month T-Bill rate. We discover that skewness (or third-order raw moment) is less stable than kurtosis (or fourth-order raw moment). In contrast, partial third-order raw (standardized) moments are more stable than partial fourth-order raw (standardized) moments. Moreover, we show that raw moments are more stable than standardized moments over time. This empirical

study aims to give risk managers enough information on which moments they should rely on when constructing their strategy.

Then, we study decision theory under uncertainty in chapter 4, using the generalized recursive smooth ambiguity model of Hayashi and Miao (2011) to study the home bias in a dynamic setting. We develop a numerical approach that extends the existing ones to tackle the specificity of our framework, and we perform sensitivity analysis to see how the home bias varies when our model parameters vary. Moreover, the fact that the decision-maker learns over time allows the agent to adjust her prior closer to the real mean of stock return and, therefore, to adjust her optimal portfolio composition. This study could be extended by using a different utility specification that is not independent of the level of wealth. However, this is a more complex situation where we should use a grid point research that would increase the problem's difficulty.

Bibliography

- Abreu, M., Mendes, V., and Santos, J. A. (2011). Home country bias: Does domestic experience help investors enter foreign markets? *Journal of Banking & Finance*, 35(9):2330–2340.
- Acerbi, C., Nordio, C., and Sirtori, C. (2001). Expected shortfall as a tool for financial risk management. *arXiv preprint cond-mat/0102304*.
- Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.
- Ahmed, W. M. and Al Mafrachi, M. (2021). Do higher-order realized moments matter for cryptocurrency returns? *International Review of Economics & Finance*, 72:483–499.
- Allais, M. (1953). Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école américaine. *Econometrica: Journal of the Econometric Society*, pages 503–546.
- Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118(1):135–167.
- Anscombe, F. J., Aumann, R. J., et al. (1963). A definition of subjective probability. *Annals of mathematical statistics*, 34(1):199–205.
- Ardalan, K. (2019). Equity home bias: a review essay. *Journal of Economic Surveys*, 33(3):949–967.
- Arnold, B. C. (2008). *Pareto and Generalized Pareto Distributions*, pages 119–145. Springer New York.

- Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.
- Baltzer, M., Stolper, O., and Walter, A. (2013). Is local bias a cross-border phenomenon? evidence from individual investors' international asset allocation. *Journal of Banking & Finance*, 37(8):2823–2835.
- Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, 59(4):1481–1509.
- Barbosa, A. M. and Ferreira, M. A. (2004). Beyond coherence and extreme losses: Root lower partial moment as a risk measure. *Available at SSRN 609221*.
- Barbut, M. (2003). Homme moyen ou homme extrême? de Vilfredo Pareto (1896) à Paul Lévy (1936) en passant par Maurice Fréchet et quelques autres. *Journal de la Société française de Statistique*, 144(1-2):113–133.
- Barczy, M., Nedényi, F. K., and Sütő, L. (2022). Probability equivalent level of value at risk and higher-order expected shortfalls. *arXiv preprint arXiv:2202.09770*.
- Basel Committee on Banking Supervision (2019). The market risk framework: In brief. *Available online*.
- Basel Committee on Banking Supervision (2022). Basel iii: international regulatory framework for banks. *Available online*.
- Bekaert, G., Hoyem, K., Hu, W.-Y., and Ravina, E. (2017). Who is internationally diversified? evidence from the 401 (k) plans of 296 firms. *Journal of Financial Economics*, 124(1):86–112.
- Bernoulli, D. (1738, 1954). Specimen theorize naval de mensura sortis (translated as "exposition of a new theory on the measurement of risk". *Econometrica*.
- Borgonovo, E., Cappelli, V., Maccheroni, F., and Marinacci, M. (2018). Risk analysis and decision theory: A bridge. *European Journal of Operational Research*, 264(1):280–293.

BIBLIOGRAPHY

- Canada Office of Supervision of Financial Institutions (2022). Regulatory capital and internal capital targets. *Available online*.
- Chen, H., Ju, N., and Miao, J. (2014). Dynamic asset allocation with ambiguous return predictability. *Review of Economic Dynamics*, 17(4):799–823.
- Chen, X., Li, B., and Worthington, A. C. (2021). Higher moments and us industry returns: realized skewness and kurtosis. *Review of Accounting and Finance*, 20(1):1–22.
- Chen, Z. and Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443.
- Coeurdacier, N. and Rey, H. (2013). Home bias in open economy financial macroeconomics. *Journal of Economic Literature*, 51(1):63–115.
- Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.
- Comité Européen des Assurances and Groupe Consultatif Actuariel Européen (2007). Solvency II Glossary. *Public Document*.
- Comité Européen des Assurances and Mercer Oliver Wyman Limited (2005). Solvency assessment models compared: : Essential groundwork for the solvency ii project. *Public Document*.
- Cont, R., Deguest, R., and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative finance*, 10(6):593–606.
- Dai, X., Xiao, L., Wang, Q., and Dhesi, G. (2021). Multiscale interplay of higher-order moments between the carbon and energy markets during phase iii of the eu ets. *Energy Policy*, 156:112428.
- Daniélsson, J., Jorgensen, B. N., Mandira, S., Samorodnitsky, G., and De Vries, C. G. (2005). Subadditivity re-examined: the case for value-at-risk.

- Daniélsson, J., Jorgensen, B. N., Samorodnitsky, G., Sarma, M., and de Vries, C. G. (2013). Fat tails, var and subadditivity. *Journal of econometrics*, 172(2):283–291.
- De Clerk, L. and Savel’ev, S. (2022). An investigation of higher order moments of empirical financial data and their implications to risk. *Heliyon*, 8(2):e08833.
- Debreu, G. (1954). Representation of a preference ordering by a numerical function. *Decision processes*, 3:159–165.
- Deng, X. and Chen, J. (2022). The lower partial moments risk measure in a novel fuzzy framework based on possibility density function. *Computers & Industrial Engineering*, 169:108309.
- Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2006). *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons.
- Dhaene, J., Laeven, R. J., Vanduffel, S., Darkiewicz, G., and Goovaerts, M. J. (2008). Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75(2):365–386.
- Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, 119(3):559–577.
- Doan, P., Lin, C.-T., and Zurbruegg, R. (2010). Pricing assets with higher moments: Evidence from the Australian and US stock markets. *Journal of International Financial Markets, Institutions and Money*, 20(1):51–67.
- Dominiak, A., Duersch, P., and Lefort, J.-P. (2012). A dynamic Ellsberg urn experiment. *Games and Economic Behavior*, 75(2):625–638.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics*, pages 643–669.
- Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

BIBLIOGRAPHY

- Embrechts, P. and Wang, R. (2015). Seven proofs for the subadditivity of expected shortfall. *Dependence Modeling*, 3(1).
- Epstein, L. G. and Miao, J. (2003). A two-person dynamic equilibrium under ambiguity. *Journal of Economic Dynamics and Control*, 27(7):1253–1288.
- Epstein, L. G. and Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31.
- Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.
- Epstein, L. G. and Zin, S. E. (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99(2):263–286.
- Etner, J. (2006). A note on the relation between risk aversion, intertemporal substitution and timing of the resolution of uncertainty. *Annals of Economics & Finance*, 7(2).
- Etner, J., Jeleva, M., and Tallon, J.-M. (2012). Decision theory under ambiguity. *Journal of Economic Surveys*, 26(2):234–270.
- Fadina, T., Liu, P., and Wang, R. (2021). One axiom to rule them all: An axiomatization of quantiles. *Available at SSRN at 3944312*.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105.
- Faroni, S., Le Courtois, O., and Ostaszewski, K. (2022). Equivalent risk indicators: Var, tce, and beyond. *Risks*, 10(8):142.
- Fidora, M., Fratzscher, M., and Thimann, C. (2007). Home bias in global bond and equity markets: the role of real exchange rate volatility. *Journal of International Money and Finance*, 26(4):631–655.
- Fiori, A. M. and Rosazza Gianin, E. (2021). Generalized p-entropy and applications to risk measures. *Available at SSRN 3949592*.

- Fischer, T. (2003). Risk capital allocation by coherent risk measures based on one-sided moments. *insurance: Mathematics and Economics*, 32(1):135–146.
- Fishburn, P. C. (1970). Utility theory for decision making. Technical report, Research analysis corp McLean VA.
- Flynn, J. P., Schmidt, L. D., and Toda, A. A. (2023). Robust comparative statistics for the elasticity of intertemporal substitution. *Theoretical Economics*, 18(1):231–265.
- Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447.
- French, K. R. and Poterba, J. M. (1991). Investor diversification and international equity markets. *The American Economic Review*, 81(2):222–226.
- Frittelli, M. and Gianin, E. R. (2002). Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486.
- Fuchs, S., Schlotter, R., and Schmidt, K. D. (2017). A review and some complements on quantile risk measures and their domain. *Risks*, 5(4):59.
- Gaar, E., Scherer, D., and Schiereck, D. (2020). The home bias and the local bias: A survey. *Management Review Quarterly*, pages 1–37.
- Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies*, 20(1):41–81.
- Gatzert, N. and Wesker, H. (2011). A comparative assessment of basel ii/iii and solvency ii. *Geneva Papers on Risk and Insurance: Issues and Practice*, 37:539–570.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2015). *Bayesian Data Analysis*. Chapman and Hall/CRC, New York, 3 edition.
- Gerber, H. U. and Pafumi, G. (1998). Utility functions: from risk theory to finance. *North American Actuarial Journal*, 2(3):74–91.

BIBLIOGRAPHY

- Gilboa, I. (2009). *Theory of decision under uncertainty*, volume 45. Cambridge university press.
- Gilboa, I. (2010). *Making better decisions: Decision theory in practice*. John Wiley & Sons.
- Gilboa, I. and Marinacci, M. (2016). Ambiguity and the bayesian paradigm. In *Readings in formal epistemology*, pages 385–439. Springer.
- Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:pp–141.
- Grigoletto, M. and Lisi, F. (2011). Practical implications of higher moments in risk management. *Statistical Methods & Applications*, 20:487–506.
- Grinblatt, M. and Keloharju, M. (2001). How distance, language, and culture influence stockholdings and trades. *The Journal of Finance*, 56(3):1053–1073.
- Guidolin, M. and Liu, H. (2016). Ambiguity aversion and underdiversification. *Journal of Financial and Quantitative Analysis*, 51(4):1297–1323.
- Guidolin, M. and Rinaldi, F. (2013). Ambiguity in asset pricing and portfolio choice: A review of the literature. *Theory and Decision*, 74(2):183–217.
- Hansen, L. and Sargent, T. (2006). Robust estimation and control without commitment. forthcoming in journal of economic theory.
- Hansen, L. P. and Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2):60–66.
- Hansen, L. P. and Sargent, T. J. (2010). Fragile beliefs and the price of uncertainty. *Quantitative Economics*, 1(1):129–162.
- Hayashi, T. and Miao, J. (2011). Intertemporal substitution and recursive smooth ambiguity preferences. *Theoretical Economics*, 6(3):423–472.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174.

- Hosking, J. and Wallis, J. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349.
- Inui, K. and Kijima, M. (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of banking & finance*, 29(4):853–864.
- Jondeau, E., Zhang, Q., and Zhu, X. (2019). Average skewness matters. *Journal of Financial Economics*, 134(1):29–47.
- Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. The McGraw-Hill Companies, Inc.
- Ju, N. and Miao, J. (2012). Ambiguity, learning, and asset returns. *Econometrica*, 80(2):559–591.
- Jurczenko, E. and Maillet, B. (2006). *Multi-moment asset allocation and pricing models*. John Wiley & Sons Hoboken, NJ.
- Kilka, M. and Weber, M. (2000). Home bias in international stock return expectations. *The Journal of Psychology and Financial Markets*, 1(3-4):176–192.
- Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.
- Klibanoff, P., Marinacci, M., and Mukerji, S. (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 144(3):930–976.
- Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss models: from data to decisions*, volume 715. John Wiley & Sons.
- Kreps, D. (1988). *Notes on the Theory of Choice*. Routledge.
- Kreps, D. M. and Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica: journal of the Econometric Society*, pages 185–200.
- Krokhmal, P. A. (2007). Higher moment coherent risk measures.

BIBLIOGRAPHY

- Lane, P. R. and Milesi-Ferretti, G. M. (2003). International financial integration. *IMF Staff Papers*, 50(1):82–113.
- Le Courtois, O. (2018). Some further results on the tempered multistable approach. *Asia-Pacific Financial Markets*, 25(2):87–109.
- Le Courtois, O., Lévy-Véhel, J., and Walter, C. (2020). Regulation risk. *North American Actuarial Journal*, 24(3):463–474.
- Le Courtois, O. and Walter, C. (2014). The computation of risk budgets under the lévy process assumption. *Finance*, 35(2):87–108.
- Lebedev, N. (1972). *Special Functions and their Applications*. Dover Publications.
- Li, H. and Wang, R. (2022). Pelve: Probability equivalent level of var and es (november 25, 2019). *Journal of Econometrics*, *Forthcoming*.
- Lin, M. and Viswanathan, S. (2016). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science*, 62(5):1393–1414.
- Ling, A., Sun, J., and Wang, M. (2020). Robust multi-period portfolio selection based on downside risk with asymmetrically distributed uncertainty set. *European Journal of Operational Research*, 285(1):81–95.
- Linsmeier, T. J. and Pearson, N. D. (2000). Value at risk. *Financial Analysts Journal*, 56(2):47–67.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, 20(4):587–615.
- Lippi, A. (2016). (country) home bias in italian occupational pension funds asset allocation choices. *The Quarterly Review of Economics and Finance*, 59:78–82.
- Lu, J., Jain, L. C., and Zhang, G. (2012). Risk management in decision making. *Handbook on Decision Making: Vol 2: Risk Management in Decision Making*, pages 3–7.

- Lütje, T. and Menkhoff, L. (2007). What drives home bias? evidence from fund managers' views. *International Journal of Finance & Economics*, 12(1):21–35.
- Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.
- Machina, M. J. (1992). Choice under uncertainty: Problems solved and unsolved. In *Foundations of Insurance economics*, pages 49–82. Springer.
- Machina, M. J. and Siniscalchi, M. (2014). Ambiguity and ambiguity aversion. In *Handbook of the Economics of Risk and Uncertainty*, volume 1, pages 729–807. Elsevier.
- McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press.
- Mei, D., Liu, J., Ma, F., and Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help? *Physica A: Statistical Mechanics and its Applications*, 481:153–159.
- Mishra, A. V. (2008). Australia's equity home bias. *Australian Economic Papers*, 47(1):53–73.
- National Association of Insurance Commissioners (2007). *Own risk and solvency assessment (ORSA) guidance manual*. NAIC.
- Nesaz, H. H., Jasemi, M., and Monplaisir, L. (2020). A new methodology for multi-period portfolio selection based on the risk measure of lower partial moments. *Expert Systems with Applications*, 144:113032.
- Obstfeld, M. and Rogoff, K. (2001). The six major puzzles in international macroeconomics: is there a common cause? *NBER macroeconomics annual*, 15:339–390.
- Pareto, V. (1965, [1895]). *Ecrits sur la courbe de la répartition de la richesse*. Librairie Droz.

BIBLIOGRAPHY

- Paté-Cornell, M. E. and Dillon, R. L. (2006). The respective roles of risk and decision analyses in decision support. *Decision Analysis*, 3(4):220–232.
- Peijnenburg, K. (2018). Life-cycle asset allocation with ambiguity aversion and learning. *Journal of Financial and Quantitative Analysis*, 53(5):1963–1994.
- Peterson, M. (2017). *An introduction to decision theory*. Cambridge University Press.
- Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136.
- Rau-Bredow, H. (2019). Bigger is not always safer: a critical analysis of the subadditivity assumption for coherent risk measures. *Risks*, 7(3):91.
- Robinson, J. P. and Shaver, P. R. (1973). Measures of social psychological attitudes.
- Rostek, M. (2010). Quantile maximization in decision theory. *The Review of Economic Studies*, 77(1):339–371.
- Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica: Journal of the Econometric Society*, pages 571–587.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.
- Sihem, M. and Slaheddine, H. (2014). The impact of higher order moments on market risk assessment. *Procedia Economics and Finance*, 13:143–153.
- Siniscalchi, M. (2009). Vector expected utility and attitudes toward variation. *Econometrica*, 77(3):801–855.
- Society of Actuaries (2000). Getting to know cte. *Risk Management Newsletter*.

- Uppal, R. and Wang, T. (2003). Model misspecification and underdiversification. *The Journal of Finance*, 58(6):2465–2486.
- Van Zwet, W. R. (1980). A strong law for linear functions of order statistics. *The Annals of Probability*, 8(5):986–990.
- Von Neumann, J. and Morgenstern, O. (1944). Theory of games and economic behavior.
- Wang, R. and Zitikis, R. (2021). An axiomatic foundation for the expected shortfall. *Management Science*, 67(3):1413–1429.
- Yu, L., Lin, L., Guan, G., and Liu, J. (2022). Time-consistent lifetime portfolio selection under smooth ambiguity. *Mathematical Control and Related Fields*.