



**HAL**  
open science

# Stress-Test Exercises and the Pricing of Very Long-Term Bonds

Simon Dubecq

► **To cite this version:**

Simon Dubecq. Stress-Test Exercises and the Pricing of Very Long-Term Bonds. General Mathematics [math.GM]. Université Paris Dauphine - Paris IX, 2013. English. NNT : 2013PA090001 . tel-00871760

**HAL Id: tel-00871760**

**<https://theses.hal.science/tel-00871760>**

Submitted on 10 Oct 2013

**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.

Université Paris-Dauphine  
Centre De Recherche en Mathématiques de la Décision UMR CNRS 7534

# **Stress-Test Exercises and the Pricing of Very Long-Term Bonds**

*Three Essays in Financial Econometrics*

Thèse présentée et soutenue publiquement le

**28 Janvier 2012**

en vue de l'obtention du

**Doctorat en Mathématiques Appliquées**

par

**Simon DUBECQ**

Directeur de Recherche : **Monsieur Christian GOURIEROUX**  
Professeur des Universités  
INSEE, Université Paris-Dauphine, Université de Toronto

Rapporteurs : **Monsieur Jean-Paul LAURENT**  
Professeur des Universités  
Université Paris 1 Panthéon-Sorbonne

**Monsieur Fabio TROJANI**  
Professeur des Universités  
Université de Lugano

Suffragants : **Monsieur Alain MONFORT**  
Professeur des Universités  
CREST, Banque de France, Université de Maastricht

**Monsieur Fulvio PEGORARO**  
Économiste Chargé de Recherche  
Banque de France, CREST

**Monsieur Olivier WINTENBERGER**  
Maître de Conférences  
Université Paris-Dauphine

L'Université Paris-Dauphine n'entend donner aucune approbation ni improbation  
aux opinions émises dans les thèses; ces opinions doivent être considérées comme  
propres à leur auteur.

# Remerciements

---

La réalisation de cette thèse a été un long voyage, qui, bien que solitaire et sans escale, n'aurait pu être mené à bien sans l'incalculable assistance dont j'ai bénéficié.

Mes premières pensées sont naturellement pour mon directeur de thèse, le Professeur Christian Gourieroux, dont les enseignements, le soutien, la patience mais aussi les critiques, ont été déterminant dans l'écriture de cet ouvrage. Être l'étudiant de Christian est une chance, une fierté, et un grand honneur, dont j'espère me montrer digne à l'avenir.

Je souhaite aussi rendre hommage à la Banque de France, où j'ai pu, dans un environnement formateur et passionnant, réaliser cette thèse et débiter ma carrière professionnelle. Au sein de la Direction des Etudes Monétaires et Financières, j'ai bénéficié de conditions de travail que beaucoup de thésards m'envieraient. Une grande part du mérite en revient à mes responsables hiérarchiques successifs Laurent Clerc, Caroline Jardet, Jean-Stéphane Mésonnier, Benoit Mojon et Anne-Marie Rieu, ainsi qu'à tous mes collègues de RECFIN: Alejandro Bernales,

Vladimir Borgy, Valère Fourel, Imen Ghattassi, Julien Idier, Fulvio Pegoraro, Jean-Paul Renne, Francesca Rinaldi, Guillaume Roussellet, Dilyara Salakhova et Luca Tiozzo-Pezzoli. Parmi ceux-ci, j'ai une dette particulière envers Jean-Paul Renne et Fulvio Pegoraro, dont les conseils, la créativité et la disponibilité m'ont été plus que précieux. Je souhaite aussi souligner l'aide qui m'a été apportée dans mes recherches bibliographiques, administratives, et dans le traitement des bases de données par l'équipe du Labolog, les secrétaires de la DEMFI Brigitte Arnaudo, Magali Rebagliato et Catherine Estrella, ainsi que les assistantes de recherche Béatrice Saes-Escorbiac et Aurélie Touchais.

Je suis par ailleurs reconnaissant envers toutes les personnes qui ont pris de leur temps pour lire les versions préliminaires des articles composant cette thèse. Je pense en particulier aux Professeurs Alain Monfort du CREST et de la Banque de France, et Andrew Siegel de l'Université de Washington à Seattle, dont les conseils m'ont énormément apporté.

J'aimerais aussi remercier infiniment les Professeurs Jean-Paul Laurent et Fabio Trojani d'avoir accepté d'être rapporteurs de cette thèse, ainsi que les autres membres du jury: le Professeur Alain Monfort, Fulvio Pegoraro, et Olivier Wintemberger.

La réalisation d'une thèse est un travail à plein temps, qui repose aussi bien souvent sur les encouragements et l'entrain de nombreux amis. Je souhaite profiter de cette opportunité pour les en remercier, du fond du cœur. Merci donc à Vincent Mansuy, Jean Beuve, Martin Koning, Baptiste Françon, Nicolas Sallée, Diego Castillo, Suzanne George, Catherine Haggioannou, Josephine et Julia Bastard, Cécile Peillon, Kevin Estel, Maxime Housard de la Potterie, Fabrice Rusig, Miguel Amaral, Hélène-Charlotte Pisarska, Agnès Redonnet, Arnaud Batu, Mathieu Murzyn,

Claire Willemont, Raphaëlle Hartout, Grégory Duflos et bien d'autres...

Je souhaite aussi faire un clin d'œil à Anabelle Canal, sans qui ce projet, comme bien d'autres qui m'ont fait grandir, n'aurait pas vu le jour. Je lui en suis infiniment reconnaissant.

Je veux ici exprimer toute ma gratitude à mes parents, Monique et Philippe Dubecq, ainsi qu'à mes frères et sœurs Marie, Etienne et Louise Dubecq. Leurs conseils me sont essentiels, leur confiance est mon plus puissant moteur. A l'heure où j'écris ces lignes, j'espère faire leur fierté. Ce travail leur est dédié.

Enfin, je souhaite conclure par mes profonds remerciements à Małgorzata Błażewska, dont j'ai la chance de partager la vie. Son amour et son soutien m'apportent un équilibre, une sérénité qui dépasse le cadre de ce travail. Cette thèse lui est aussi dédiée.



## Foreword

---

This Ph.D. dissertation, entitled "Stress-Test Exercises and the Pricing of Very Long-Term Bonds, Three Essays in Financial Econometrics", brings together three essays in the field of financial econometrics and asset pricing. Each essay corresponds to one chapter, and the three chapters are grouped in two different parts: the first one focuses on the methodology of stress-tests exercises, while the second part, which collects the two last chapters, addresses the issue of the pricing of very long-term bonds. The structure of the whole dissertation is explained in more details in the General Introduction in which we also provide a review of the related literature and define the questions of research we address (for the comfort of French readers, the General Introduction has been translated in French). Nevertheless, since each chapter corresponds to an independent essay, chapters can be read separately. This implies the presence of redundant information across chapters.





## Résumé

---

Cette thèse est constituée de deux parties distinctes, toutes deux d'intérêt pour la gestion du risque dit "de modèle", i.e. le risque lié à l'utilisation de modèles quantitatifs pour la valorisation d'actifs financiers, et la mesure du risque d'un investissement. Une des préconisations du comité de Bâle consiste à tester fréquemment la robustesse des institutions financières. A cette fin, les institutions financières se doivent de réaliser des tests de résistance, ou stress-tests, en évaluant les conséquences d'un environnement dégradé sur le risque de leurs positions. En pratique, ces exercices évaluent les conséquences d'un scénario donné, ou de manière équivalente, d'un choc déterministe sur les prix des actifs financiers. La première partie de cette thèse introduit une nouvelle méthodologie pour la réalisation d'exercices de stress-tests. Elle décrit les limites des pratiques actuelles, et introduit une nouvelle approche, où les chocs étudiés sont définis de façon stochastique. Nous y dérivons explicitement le lien entre les caractéristiques des chocs et leurs conséquences pour le portefeuille de l'institution considérée. Par ailleurs, alors que la plupart des tests de résistance sont réalisés pour une composition du portefeuille donnée (i.e. le portefeuille est dit cristallisé), nous relierons explicitement la réaction optimale du détenteur du portefeuille aux caractéristiques des chocs, et ceci pour plusieurs critères traditionnels d'investissement. Notre approche per-

met ainsi de considérer des scénarios de stress beaucoup plus riches, qui évaluent l'impact d'une modification de la distribution statistique des facteurs influençant les prix d'actifs, pas uniquement les conséquences d'une réalisation particulière de ces facteurs, et prennent en compte la réaction du gestionnaire de portefeuille au choc.

La deuxième partie de la thèse est consacrée au risque de mauvaise spécification des modèles de valorisation des actifs financiers. Il est nécessaire que les modèles de valorisation ne supposent pas l'existence d'opportunités illimitées d'arbitrage, ce qui risquerait de pervertir les stratégies d'investissement de l'institution financière utilisatrice, ou de la mettre en risque vis-à-vis de ses contreparties (ce serait par exemple le cas si un teneur de marché proposait des cotations potentiellement "arbitrables").

Nous nous intéressons dans cette thèse à la valorisation des obligations à maturité très longues (supérieure à 10 ans). Le segment très long des courbes de taux reste méconnu de la littérature, malgré l'accélération des émissions, due aux besoins de financement croissant des états (le gouvernement français émet par exemple des obligations de maturité 50 ans depuis 2005), et à l'allongement général de l'espérance de vie, qui accroît la demande pour les placements de longue durée. Pourtant, la modélisation des taux de très long terme est un défi, notamment parce que l'hypothèse d'absence d'opportunités d'arbitrage contraint plus fortement les taux très longs que très courts. Cette particularité explique pourquoi la plupart des modèles de taux d'intérêt impliquent un taux limite (de maturité infinie) constant en absence d'opportunités d'arbitrage. Ces modèles ont du mal à reproduire la dynamique des taux très long, parfois plus volatiles que les taux plus courts. De plus ils contraignent fortement les taux extrapolés des courbes de taux observées, extrapolation qui est parfois nécessaire pour la valorisation des flux monétaires de très long-terme [par exemple lors de l'évaluation du financement d'un système de retraite ou d'un projet d'infrastructure]. Les modèles à facteurs de la courbe des taux d'intérêt, dont les taux limites sont stochastiques, sont rares. La plupart incorporent un facteur dénommé "niveau", dont les variations ont un impact

uniforme sur l'ensemble de la courbe de taux modélisée, *a fortiori* sur les taux les plus longs. Parmi ceux-ci, le modèle de Nelson et Siegel, qui décompose la courbe de taux d'intérêt en facteurs "niveau", "pente", et "courbure" (du fait de leur impact respectif sur la courbe de taux) est très populaire, si populaire que la littérature interprète parfois ces facteurs comme des variables macroéconomiques. Le facteur "niveau" est ainsi associé aux anticipations d'inflation, et le facteur "pente" à l'activité économique dans le pays émetteur. Ce modèle n'est pourtant pas compatible avec l'hypothèse d'absence d'opportunités d'arbitrage. Une version non-arbitrable a récemment été introduite, mais celle-ci suppose une explosion des taux limites.

Le deuxième chapitre de cette thèse s'attache à étudier la compatibilité du facteur "niveau" avec l'absence d'opportunités d'arbitrage. Il démontre qu'un modèle de taux d'intérêt à facteur niveau requiert une dynamique très particulière de ce facteur pour être non-arbitrable. Nous introduisons dans le troisième chapitre de la thèse une nouvelle classe de modèle à facteurs, sans opportunités d'arbitrage, où le taux limite est stochastique. Nous présentons les propriétés empiriques de ce modèle sur une base de données de prix d'obligations du Trésor des Etats-Unis. Ce modèle peut aussi s'interpréter comme un modèle de Nelson Siegel non-arbitrable, bien que la dynamique des facteurs (et donc leur interprétation économique) soit profondément modifiée. Nous montrons que ce modèle à taux limite stochastique, pourtant très contraint, présente des performances empiriques comparables aux autres modèles Nelson-Siegel utilisés dans la littérature.

**Mots Clés:** Choc, Copule, Risque Extrême, Tests de Résistance, Modèle à Facteur, Risque Systémique, Gestion de Portefeuille, Obligations Souveraines, Taux d'intérêt, Structure par Terme, Modèle Affine, Facteur Niveau, Facteur Pente, Distribution Stable, Taux de Long-Terme Stochastique.



# Abstract

---

This thesis is composed of two different parts, though both related to the management of "model risk", i.e. the risk associated with the use of quantitative models for the pricing of financial assets and the measure of the risk of an investment.

The Basel Committee recommends frequently assess the robustness of financial institutions, by performing stress-test exercises, which try to anticipate the consequences of a deteriorated environment for a financial institution. In practice, these exercises approximate such "deteriorated environment" by a given scenario, or equivalently, by a deterministic shock on the price of financial assets. In the first part of this thesis, we introduce a new methodology for stress-test exercises. We describe the limitations of the current practices, and propose as a new approach to consider stochastic rather than deterministic shocks in the implementation of stress-test exercises. We explicitly derive the relationship between the characteristics of the shock, which hit the factors driving the financial assets' prices, and their consequences for the asset portfolio of the considered institution. Moreover, we also explicitly link the optimal reaction of the portfolio manager to the characteristics of the shock, while the common practice perform stress-test exercises for a given portfolio allocation (the portfolio is said "crystallized"), for three classic investment criteria. Our approach allows to consider richer stress-test exercises,

which assess the impact of a modification of the whole distribution of asset prices' factors, rather than focusing on a single realization of these factors, and take into account the potential reaction to the shock of the portfolio manager.

The second part of the thesis is devoted to the risk of misspecified models for the pricing of financial assets. It is indeed essential for asset pricing models to not assume the existence of unlimited arbitrage opportunities, which would corrupt the investment strategies of the model users, and put the financial institution at risk vis-à-vis its counterparts (for instance, it would be the case if the quotes announced by a market maker were not internally consistent).

In this thesis, we focus on the pricing of bonds with very long-term time-to-maturity (more than ten years). The very long segment of the yield curve remains relatively unknown from the literature, in spite of the expansion of the market, fed by the increasing funding needs of public borrowers (the French government issues for instance fifty-years bonds since 2005), and the stronger appetite for very long-term investment (related to the general increase in life expectancy). However, the modeling of the very long-term rates is a challenge, in particular because of the no-arbitrage assumption, which constraints more the long-term rather the short-term rates. This peculiarity explains why most of the no-arbitrage term structure models assume a constant limiting rate (of infinite maturity). These models badly reproduce the dynamics of the very long-term rates, which are sometimes more volatile than the shorter ones. Besides, they have strong implications for the rates extrapolated from observed term structures [this extrapolation is sometimes required for the pricing of very long-term cash-flows, such as the ones generated by some infrastructure projects, or by retirement systems].

Term structure models, whose limiting rates are stochastic, are rare in the literature. Most of them incorporate a so-called "level" factor, whose variations have a uniform impact on the modeled yield curve. Among them, the Nelson-Siegel model, which decompose the yield curve in factors "level", "slope", and "curvature" (due to the respective impact of each factor on the yield curve) is very popular, so popular that it is now common in the literature to relate the Nelson-Siegel fac-

tors to macroeconomic variables. The "level" and "slope" factors are thus usually related to the expected inflation, and the economic activity in the issuer country, respectively. The Nelson-Siegel model is however not arbitrage-free. An arbitrage-free version has recently been introduced, which assume explosive limiting rates. The second chapter of this thesis investigates the compatibility of the so-called "level" factor, with the no-arbitrage assumptions. It shows that, to be arbitrage-free, a term structure model incorporating a "level" factor shall imply a very peculiar dynamics of this factor. We introduce in the third chapter a new class of arbitrage-free term structure factor models, which allows the limiting rate to be stochastic. We present the empirical properties of this model on a dataset of prices of bonds issued by the US Treasury. This new model belongs to the Nelson-Siegel family, even though the dynamics of the factors (and thus their economic interpretation) is dramatically modified. We show that this model with stochastic limiting rate, though much more constrained, feature similar empirical properties than other Nelson-Siegel models used in the literature.

**Keywords:** Shock, Copula, Extreme Risk, Stress-Tests, Factor Model, Systemic Risk, Portfolio Management, Sovereign Bonds, Interest Rate, Term Structure, Affine Model, No Arbitrage, Level Factor, Slope Factor, Stable Distribution, Stochastic Long-Term Rate.





# Contents

---

<b>Remerciements</b>	<b>1</b>
<b>Foreword</b>	<b>4</b>
<b>Résumé</b>	<b>6</b>
<b>Abstract</b>	<b>10</b>
<b>Introduction Générale</b>	<b>17</b>
<b>General Introduction</b>	<b>46</b>
<b>I The Methodology of Stress-Test Exercises</b>	<b>73</b>
<b>1 Shock on Variable or Shock on Distribution with Application to Stress-Tests</b>	<b>75</b>
1.1 Introduction . . . . .	75
1.2 Family of Distributions or Family of Variables . . . . .	80
1.2.1 Copula . . . . .	80
1.2.2 Extension to families . . . . .	82
1.2.3 Shocks . . . . .	82
1.3 Local Analysis . . . . .	84
1.3.1 Expansion of the distribution . . . . .	84
1.3.2 Expansion in terms of variable . . . . .	85
1.4 Shock on Risk Management . . . . .	89
1.4.1 Risk management . . . . .	89

1.4.2	Expansion in terms of distribution . . . . .	91
1.4.3	Expansion in terms of variable . . . . .	94
1.5	Systematic Shock . . . . .	101
1.5.1	Shocks on tails . . . . .	101
1.5.2	Shocks on a systematic factor . . . . .	106
1.6	Stress-testing the European Sovereign Bond Market . . . . .	111
1.6.1	Stress-test . . . . .	111
1.6.2	Excess gains on the European sovereign bond market . . . . .	112
1.6.3	Contamination on the systematic factor . . . . .	113
1.6.4	Impact of the systematic shock on portfolio characteristics . . . . .	114
1.7	Concluding Remarks . . . . .	120

## **II Pricing Very Long-Term Bonds under No-Arbitrage Assumptions** **121**

<b>2</b>	<b>Term Structure Models with Level Factor, a Long-Term Rates Perspective</b>	<b>125</b>
2.1	Introduction . . . . .	125
2.2	No-Arbitrage Condition for Buy-and-Hold Strategies Based on Two Zero-Coupon Bonds. . . . .	128
2.3	Behavior of the Long-Term Interest Rate . . . . .	132
2.4	Risk-Neutral Factor Dynamic . . . . .	134
2.5	Term Structure Model with Level and Slope Factors . . . . .	138
2.5.1	The affine term structure . . . . .	138
2.5.2	The risk-neutral factor dynamics . . . . .	141
2.5.3	Constraints on the affine factor dynamics . . . . .	142
2.5.4	The absence of solution with affine level and slope factors . . . . .	144
2.5.5	Non-affine level and slope factors . . . . .	145
2.6	Concluding Remarks . . . . .	146
<b>3</b>	<b>Stable Term Structure Models with Stochastic Long-Term Rates.</b>	<b>147</b>
3.1	Introduction . . . . .	147
3.2	Stable Positive Process . . . . .	150
3.2.1	Positive stable distribution . . . . .	150
3.2.2	Moving average process with positive stable innovation . . . . .	151
3.3	Stable Term Structure Models . . . . .	157
3.3.1	Stable AutoRegressive Term Structure Model (SARTSM) . . . . .	158
3.3.2	A stochastic and almost flat term structure . . . . .	160
3.4	Discrete Time Nelson-Siegel Term Structure Model . . . . .	162

*Table of Contents*

3.4.1	The triangular Gaussian ATSM . . . . .	162
3.4.2	The extensions of the Nelson-Siegel-Svensson term structure model . . . . .	165
3.5	Application . . . . .	168
3.5.1	Methodology . . . . .	168
3.5.2	Results . . . . .	169
3.6	Concluding Remarks . . . . .	173
 <b>General Conclusion</b>		<b>174</b>
 <b>References</b>		<b>179</b>
 <b>III Appendix</b>		<b>193</b>
 <b>List of Tables</b>		<b>234</b>
 <b>List of Figures</b>		<b>235</b>



# Introduction Générale

---

Le risque financier a de multiples dimensions: les régulateurs financiers distinguent ainsi le risque de crédit (risque inhérent au défaut d'un emprunteur), le risque de marché (le risque de pertes sur les positions au bilan et hors-bilan d'une institution financière, suite aux variations des prix de marché), ou encore le risque opérationnel (risque de pertes résultant d'une défaillance attribuable à des agents, de procédures ou modèles internes inadéquats, ou d'évènements physiques extérieurs). Cette thèse s'intéresse au risque dit "de modèle", i.e. le risque de mauvaise spécification des modèles internes des institutions financières pour la valorisation d'actifs financiers et la mesure du risque de leurs investissements. La récente crise financière a dramatiquement souligné combien une mauvaise maîtrise de ces modèles pouvait mettre en danger n'importe quelle institution financière, l'amenant à mal apprécier le rendement et le risque de ses positions. Cette thèse est constituée de deux parties relativement distinctes, bien que toutes les deux d'importance pour la gestion du risque financier.

Un modèle financier se doit d'être robuste et cohérent. Robuste, au sens où il n'est pas dramatiquement affecté par une légère modification de ses hypothèses sous-jacentes. Cohérent, au sens où il n'identifie pas des opportunités illimitées d'investissement au couple rendement/risque irréaliste (ce que la littérature économique

dénomme un arbitrage). Cette thèse s'intéresse à ces deux dimensions : la première partie porte sur la méthodologie des test de résistance, ou stress-tests, réalisés par les banques pour évaluer la robustesse de leurs modèles financiers, tandis que la seconde partie s'intéresse à la valorisation des actifs financiers de très longues maturités en absence d'opportunités d'arbitrage.

Cette introduction s'organise de la façon suivante : tout d'abord, nous présentons les produits financiers dont la valorisation occupe une large part de ce travail, en particulier les obligations dites "zéro-coupon". Nous introduisons à cette occasion plusieurs notations nécessaires à la lecture de ce manuscrit. Nous rappelons ensuite les principes généraux de valorisation des actifs financiers, en nous concentrant particulièrement sur la modélisation de la structure par terme des obligations zéro-coupon et de leur taux d'intérêt, sous hypothèse d'absence d'opportunités d'arbitrage<sup>1</sup>. Ce rappel nous permet par la suite de préciser le concept de chocs dans les modèles financiers, et son enjeu pour la gestion du risque financier. Nous concluons finalement cette introduction par une présentation rapide des différents chapitres qui composent cette thèse.

## PRÉSENTATION DES ACTIFS FINANCIERS, LE CAS DES OBLIGATIONS

Un actif financier est un contrat qui donne à son détenteur le droit (ou l'obligation) de recevoir (de payer) des flux monétaires futurs. Une obligation par exemple, est un contrat selon lequel l'émetteur de l'obligation s'engage à verser au détenteur du titre, sur une période prédéterminée, des flux monétaires dont le mode de calcul est défini à l'émission du titre. Une obligation se distingue ainsi d'une action, pour laquelle chaque versement des flux financiers (les dividendes) relève d'une décision de l'assemblée générale des actionnaires. Une obligation zéro-coupon nominale est un titre financier élémentaire, qui promet à son détenteur un unique flux d'une

---

<sup>1</sup> Nous nous inspirons librement des livres de Cochrane (2001) et de Gourieroux et Jasiak (2001) pour la réalisation de cette partie.

unité monétaire à la date de maturité  $T$  de l'obligation<sup>2</sup>. Nous dénotons  $B(t, h)$  le prix à la date  $t$  d'une obligation zéro-coupon de maturité résiduelle<sup>3</sup>  $h = T - t$ . Les schémas 1 et 2 présentent les flux monétaires associés à ce titre en l'absence de défaut de l'émetteur, pour un investisseur détenant l'obligation jusqu'à maturité (schéma 1), et pour un investisseur revendant l'obligation avant maturité à une date intermédiaire  $t + k < t + h$  (schéma 2) :

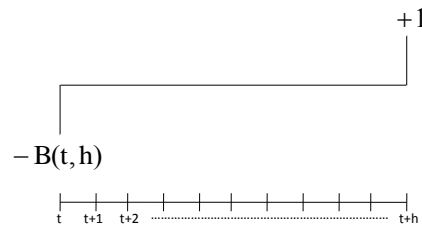


Schéma 1 : Flux monétaires associés à la détention jusqu'à maturité d'une obligation zéro-coupon de maturité résiduelle  $h$  périodes, et de prix  $B(t, h)$  à la date  $t$ .

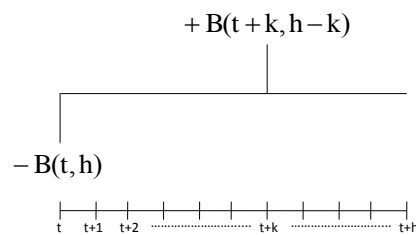


Schéma 2 : Flux monétaires associés à une obligation zéro-coupon revendue avant maturité, acquise à la date  $t$  et revendue  $k$  périodes après.

<sup>2</sup>Cette thèse se concentre sur les obligations nominales. Pour ne pas alourdir les notations, le terme "d'obligation" fait référence à une obligation nominale.

<sup>3</sup>La plupart des résultats de cette thèse se concentrent sur la valorisation d'actifs financiers en temps discret, i.e. pour  $h \in \mathbb{N}^+$ .



L'ensemble des taux d'intérêt, dénotés  $r(t, h)$  à la date  $t$  pour une obligation de maturité résiduelle  $h$ , des obligations zéro-coupon d'un même emprunteur constitue une structure par terme des taux d'intérêt (ou courbe de taux)<sup>4</sup>, où :

$$r(t, h) = -\frac{1}{h} \log B(t, h), \forall h \in \mathbb{N}^+. \quad (0.0.1)$$

Enfin, nous notons  $B^f(t, h_1, h_2)$  le prix de l'obligation zéro-coupon à terme (ou *forward*) par lequel le détenteur du contrat s'engage à la date  $t$  à payer  $B^f(t, h_1, h_2)$  à l'émetteur à la date  $t + h_1$  contre la promesse de recevoir 1 unité monétaire à la date  $t + h_2$ , tel que présenté par le schéma 3 ci-dessous. Le taux à terme (ou *forward*) associé  $r^f(t, h_1, h_2)$  est alors :

$$r^f(t, h_1, h_2) = -\frac{1}{h_2 - h_1} \log B^f(t, h_1, h_2). \quad (0.0.2)$$

Par simplicité, nous notons  $r^f(t, h) = -\log B^f(t, h, h + 1)$  le taux court *forward*. L'ensemble des taux courts *forward*  $r^f(t, h)$  de différentes dates d'engagements  $h \in \mathbb{N}^+$  constitue une courbe de taux à terme (ou *forward*).

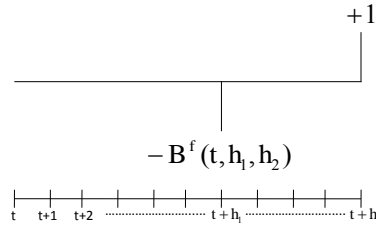


Schéma 3 : Flux monétaires associés à un contrat à terme portant sur l'achat d'une obligation zéro-coupon à une date future  $t + h_1$ , de maturité résiduelle  $h_2$ , à un prix prédéterminé à la date  $t$ .

<sup>4</sup>Le taux d'intérêt d'une obligation zéro-coupon de prix  $B(t, h)$  peut être défini de deux façons, selon qu'il est capitalisé périodiquement ou de manière continue :

- taux d'intérêt capitalisé périodiquement :  $r(t, h) = B(t, h)^{-1/h} - 1$ ,
- taux d'intérêt capitalisé continument :  $r(t, h) = -(1/h) \log B(t, h)B(t, h)$ ,

Sans perte de généralité, les résultats de ce manuscrit sont formulés en terme de capitalisation continue.

La valorisation des obligations zéro-coupon sans-risque (ou de la courbe de taux associée) est fondamentale en finance : nombres d'actifs financiers plus complexes peuvent être définis (et valorisés) comme des portefeuilles d'obligations zéro-coupon. Par exemple, une obligation nominale de maturité 30 ans, qui paye semi-annuellement des coupons fixes et rembourse le principal *in fine* peut-être décomposée en 61 obligations zéro-coupon différentes (30x2 associées aux flux de coupons + 1 pour le paiement du principal). Ces obligations zéro-coupon sont parfois échangeables directement sur le marché sous le nom de STRIPS, c'est à dire (actifs) démembrés (c'est notamment le cas pour celles issues de la décomposition des obligations à coupon fixe émise par le Trésor des Etats-Unis<sup>5</sup>).

## PRINCIPES GÉNÉRAUX DE VALORISATION DES ACTIFS FINANCIERS

D'un point de vue économique, un titre financier permet à un investisseur de lisser sa consommation au cours du temps (l'investissement a un motif d'épargne), et entre les futurs états du monde possibles (l'investissement se fait alors pour motif d'assurance). La valeur que l'investisseur accorde au titre dépend donc de la date des flux monétaires associés au titre, et de l'incertitude entourant ces flux. L'équation de valorisation fondamentale exprime la valeur de tout actif, pour un investisseur donné, en termes d'espérance conditionnelle :

$$P_{i,t} = \mathbb{E}_t \left[ M_{i,t+1}^j g_{i,t+1} \right], \quad (0.0.3)$$

où  $P_{i,t}$  désigne le prix de l'actif  $i$  à la date  $t$ ,  $\mathbb{E}_t$  est l'opérateur mathématique d'espérance conditionnelle à l'information disponible à la date  $t$ ,  $g_{i,t+1}$  est le flux monétaire associé à la détention de l'actif à la période future  $t + 1$  (qui peut correspondre à la valeur de revente de l'actif en  $t + 1$ ), et  $M_{i,t+1}^j$  désigne le facteur d'escompte stochastique de l'investisseur  $j$ . Le facteur d'escompte stochastique est central dans l'équation (0.0.24), car il pondère les flux futurs en fonction de

---

<sup>5</sup>Voir Sack (2000) pour une présentation détaillée du marché des STRIPS aux Etats-Unis.

l'impatience et de l'appétit pour le risque de l'investisseur. Par exemple, considérons une économie à la Lucas (1978), où l'investisseur reçoit à chaque date  $t$  un revenu externe dénoté  $e_t$ , qu'il utilise pour consommer une quantité unique de bien  $c_t$ , de prix  $q_t$ , ou pour investir dans une quantité  $\omega_t$  de titres de prix  $P_t$  (où  $\omega_t$  et  $P_t$  sont des vecteurs)<sup>6</sup>. Par ailleurs, les actifs financiers disponibles ne versent pas de flux intermédiaires : la valeur de revente du titre est le seul flux attendu par l'investisseur. L'objectif de l'investisseur est de choisir son profil de consommation ( $c_t$ ), son portefeuille financier ( $\omega_t$ ) pour toute date  $t \in \mathbb{N}^+$ , afin de maximiser l'espérance de sa fonction d'utilité intertemporelle sous contrainte de budget<sup>7</sup> :

$$\begin{aligned} \max_{c_t, \omega_t} \quad & \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \delta^k u(c_{t+k}) \right] \\ \text{s.c.} \quad & q_t c_t + \omega'_t P_t = e_t + \omega'_{t-1} P_t, \forall t \in \mathbb{N}^+, \end{aligned} \quad (0.0.4)$$

où  $u(c_{t+k})$  désigne l'utilité que l'investisseur retire de sa consommation en  $t+k$ , et  $\delta$  correspond à son facteur d'escompte subjectif.

La résolution du programme de l'investisseur (0.0.25) donne la condition d'Euler à toute date  $t$  pour chaque actif  $i$  :

$$P_{i,t} = \mathbb{E}_t \left[ \frac{q_t}{q_{t+1}} \delta \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)} P_{i,t+1} \right]. \quad (0.0.5)$$

L'investisseur achète (épargne) ou vend (emprunte) l'actif financier tant que la condition d'Euler (0.0.26) n'est pas vérifiée.

Dans cet exemple, le facteur d'escompte stochastique est :

$$M_{t,t+1} = \frac{q_t}{q_{t+1}} \delta \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}. \quad (0.0.6)$$

Il est fonction de l'inflation entre  $t$  et  $t+1$  ( $\frac{q_t}{q_{t+1}}$ ), de l'impatience de l'investisseur ( $\delta$ ), et du rapport des utilités marginales entre  $t$  et  $t+1$  ( $\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}$ ).

Toutes choses égales par ailleurs, la valeur de l'actif sera plus élevée si :

- i) l'investisseur est patient,
- ii) l'inflation est faible,

<sup>6</sup>Pour ne pas alourdir les notations, nous abandonnons l'indice  $j$  désignant l'investisseur. Les résultats s'entendent pour un investisseur donné.

<sup>7</sup>L'utilité des investisseurs est ici supposée s'additionner au cours du temps.

iii) la valeur future de l'actif est élevée lorsque le rapport des utilités marginales est élevé.

Il est généralement supposé que l'utilité d'un agent économique augmente avec sa consommation, mais de moins en moins vite à mesure que son niveau de consommation s'accroît (l'agent est peu à peu "rassasié"). Un rapport élevé des utilités marginales  $\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}$  correspond dans ce cas à une forte baisse de la consommation entre  $t$  et  $t + 1$ . iii) illustre ici la valeur que l'investisseur accorde aux propriétés d'assurance du titre: un titre dont la valeur future est élevée dans les états du monde où la consommation de l'agent baisse (i.e.  $\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}$  élevé) permettra à l'investisseur de "lisser" sa consommation dans le temps, et sera ainsi plus désirable.

Remarques : Le modèle ci-dessus est présenté afin d'illustrer intuitivement le concept de facteur d'escompte stochastique. Plusieurs remarques sont nécessaires à ce stade.

- Lorsque la fonction d'utilité de l'investisseur est linéaire, le rapport des utilités marginales est constant, et l'investisseur est dit neutre au risque. Dans ce cas, la valeur de l'actif est égale à l'espérance de sa valeur future actualisée (au taux d'escompte subjectif) :

$$P_{i,t+1} = (\delta r)^{-1} \mathbb{E}_t(P_{i,t+1}), \text{ avec } r = \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}, \forall c_t, c_{t+1}. \quad (0.0.7)$$

- La fonction objectif de l'investisseur est ici écrite en termes de son niveau de consommation. Le modèle n'est donc pas directement transposable à un gestionnaire de fond dont la fonction objectif dépend de la valeur future de son portefeuille.
- Les intuitions données en i), ii) et iii) s'entendent pour un investisseur donné. Le prix de marché observé dépend des offres et demandes de tous les investisseurs, dont l'agrégation est complexe, même dans l'exemple simple ci-dessus. Une large part de la littérature contourne le problème en supposant

l'existence d'un agent représentatif. L'équation de valorisation fondamentale (0.0.24) est alors utilisée pour relier l'ensemble des prix de marché à un unique facteur d'escompte stochastique, celui de l'investisseur représentatif. Sous cette hypothèse additionnelle, le modèle ci-dessus est communément appelé le CCAPM (pour *Consumption based Capital Asset Pricing Model*). Dans les paragraphes suivants, les prix des titres peuvent s'entendre comme les prix de marché sous hypothèses d'agent représentatif, ou comme les hypothétiques "justes" prix d'actifs pour un investisseur donné.

## PRINCIPES DE VALORISATION DES OBLIGATIONS ZÉRO-COUPON

Intéressons nous maintenant au cas particulier des obligations zéro-coupon. L'équation de valorisation fondamentale (0.0.24) donne :

$$B(t, 1) = \mathbb{E}_t [M_{t,t+1}], \quad (0.0.8)$$

pour l'obligation courte, et pour les obligations de maturité résiduelle supérieure à une période :

$$\begin{aligned} B(t, h) &= \mathbb{E}_t [M_{t,t+1} B(t+1, h-1)] \\ &= \mathbb{E}_t (M_{t,t+1} \mathbb{E}_{t+1} [M_{t+1,t+2} B(t+2, h-2)]) \\ &= \mathbb{E}_t [M_{t,t+1} \dots M_{t+h-1,t+h}], \end{aligned} \quad (0.0.9)$$

par le théorème des espérances itérées. Le prix de l'obligation zéro-coupon correspond à l'espérance conditionnelle du produit des facteurs d'escompte stochastique de court-terme sur la durée de vie résiduelle du titre. En combinant (0.0.30) avec (0.0.27) dans l'exemple ci-dessus, la valeur de l'obligation serait fonction du facteur d'actualisation subjectif, de l'inflation et du rapport des utilités marginales sur la durée de vie résiduelle de l'obligation :

$$B(t, h) = \mathbb{E}_t \left[ \delta^h \frac{q_t}{q_{t+h}} \frac{\frac{du}{dc}(c_{t+h})}{\frac{du}{dc}(c_t)} \right]. \quad (0.0.10)$$

La valorisation de la structure par terme des obligations (i.e. de l'ensemble des obligations zéro-coupon pour un continuum de maturités résiduelles) est donc liée à la dépendance des facteurs d'escompte stochastique  $M_{t,t+1}, M_{t+1,t+2}, \dots, M_{t+h-1,t+h}$ . Le juste prix de la structure par terme des obligations sera notamment très différent selon que les facteurs d'escompte stochastique sont indépendants entre eux (au sens probabiliste du terme) ou non. La structure par terme des obligations peut ainsi être utilisée pour identifier les propriétés du facteur d'escompte stochastique de l'investisseur représentatif [Alvarez et Jermann (2005), Bachus, Chernov et Zin (2011)].

Il est aussi possible de combiner (0.0.29) et (0.0.24), tel que :

$$P_{i,t} = \mathbb{E}_t \left[ \frac{M_{t,t+1}}{\mathbb{E}_t(M_{t,t+1})} B(t, 1) P_{i,t+1} \right], \quad (0.0.11)$$

où  $\frac{M_{t,t+1}}{\mathbb{E}_t(M_{t,t+1})}$  définit une densité d'une fonction de probabilité conditionnelle à l'information disponible en  $t$ . Par conséquent, l'équation de valorisation fondamentale (0.0.24) peut se réécrire sous la forme d'une espérance sous une mesure de probabilité modifiée, dénotée  $\mathbb{Q}$  (par opposition à la mesure de probabilité "physique" que nous dénotons  $\mathbb{P}$ ) :

$$P_{i,t+1} = \mathbb{E}_t^{\mathbb{Q}} [B(t, 1) P_{i,t+1}]. \quad (0.0.12)$$

Sous  $\mathbb{Q}$ , les prix d'actifs s'expriment comme l'espérance des prix futurs actualisés. Pour cette raison, la probabilité modifiée  $\mathbb{Q}$  est souvent appelée probabilité risque-neutre.

Le changement de probabilité correspond à l'idée suivante : un investisseur averse au risque peut de manière équivalente être vu comme un investisseur neutre au risque, qui modifie les probabilités d'occurrence des états du monde en surpondérant les états défavorables (lorsque le facteur d'escompte stochastique  $M_{t,t+1}$  est élevé, i.e. lorsque l'investisseur est sensible à une augmentation marginale de la valeur future de l'actif), et en sous-pondérant les états favorables (lorsque le facteur d'escompte stochastique  $M_{t,t+1}$  est faible, l'investisseur est peu sensible à la valeur future de l'actif).

Dans le cas précis des obligations zéro-coupon, le prix d'une obligation correspond

à l'espérance sous  $\mathbb{Q}$  du prix futur actualisé de l'obligation, ou à l'espérance "risque-neutre" du produit des obligations courtes jusqu'à maturité de l'obligation :

$$\begin{aligned} B(t, h) &= \mathbb{E}_t^{\mathbb{Q}} [B(t, 1)B(t + 1, h - 1)] \\ &= \mathbb{E}_t^{\mathbb{Q}} [B(t, 1) \dots B(t + h - 1, 1)], \end{aligned}$$

à ne pas confondre avec le prix de l'obligation pour un individu neutre au risque, noté  $\tilde{B}(t, h)$  :

$$\tilde{B}(t, h) = \mathbb{E}_t^{\mathbb{P}} [B(t, 1) \dots B(t + h - 1, 1)]. \quad (0.0.13)$$

La prime de risque sur les obligations, ou prime de terme, se définit comme l'écart entre le prix de l'obligation  $B(t, h)$ , et celui pour un individu neutre au risque  $\tilde{B}(t, h)$ , ou plus fréquemment comme l'écart entre le taux de l'obligation  $r(t, h) = -\frac{1}{h} \log B(t, h)$  et le taux "neutre au risque"  $\tilde{r}(t, h) = -\frac{1}{h} \log \tilde{B}(t, h)$ <sup>8</sup>.

## LES MODÈLES À FORME RÉDUITE DE VALORISATION DES ACTIFS FINANCIERS EN ABSENCE D'OPPORTUNITÉS D'ARBITRAGE

L'exemple du CCAPM ci-dessus détermine avec précision la spécification du facteur d'escompte stochastique (0.0.27). Cette spécification est fortement dépendante des hypothèses de départ du modèle, et d'autres formulations peuvent être obtenues à partir d'hypothèses différentes (sur la fonction objectif de l'investisseur par exemple).

---

<sup>8</sup>Parfois, la prime de risque se définit comme un écart à la théorie des anticipations rationnelles, qui postule de façon équivalente que [Cochrane (2001), p. 355] :

- i) le taux zéro-coupon de maturité résiduelle  $h$  est égale à l'espérance "physique" de la moyenne des taux courts futurs,
- ii) le taux *forward* court  $r^f(t, h)$  est égal à l'espérance "physique" du taux court futur  $\mathbb{E}^{\mathbb{P}} [r(t + h, 1)]$ ,
- iii) l'espérance "historique" du rendement des obligations  $\log \left( \frac{B(t+1, h-1)}{B(t, h)} \right)$  est la même pour toutes les obligations.

La théorie des anticipations rationnelles et la neutralité au risque sont proches, mais pas équivalentes puisque la prime de risque selon la théorie des anticipations rationnelles ignore le terme dit "de convexité" qui apparaît lorsque l'on passe des prix, fonctions exponentielles des taux d'intérêt, aux taux eux-mêmes.

Il n'est en fait pas indispensable de spécifier le facteur d'escompte stochastique pour valoriser les actifs financiers. Plusieurs travaux ont ainsi démontré que l'existence du facteur d'escompte stochastique découle de l'absence d'opportunités d'arbitrage. Un arbitrage revient à constituer un portefeuille autofinancé (i.e. sans apport externe de richesse), qui génère un rendement au minimum nul, parfois positif. L'absence d'opportunités d'arbitrage est communément admise sur les marchés à la liquidité suffisante (au sens où il y est possible d'acheter ou vendre une grande quantité de titres sans impact significatif sur les prix de marché), et constitue aujourd'hui une hypothèse de base des modèles modernes de valorisation d'actifs<sup>9</sup>. L'absence d'opportunité d'arbitrage est une condition nécessaire et suffisante pour l'existence d'au moins un facteur d'escompte stochastique strictement positif, i.e.  $M_{t,t+1} > 0 \forall t$  [Harrison et Kreps (1979), Hansen et Richard (1987)]. De manière équivalente, en absence d'opportunités d'arbitrage, il est toujours possible de valoriser les actifs financiers comme une espérance de leur valeur future actualisée, sous une mesure de probabilité modifiée  $\mathbb{Q}$ , équivalente à la probabilité physique  $\mathbb{P}$  [Harrison et Kreps (1979)]. Cette probabilité modifiée est faiblement contrainte : elle doit surtout être telle que les évènements de probabilité physique nulle ont également une probabilité nulle sous  $\mathbb{Q}$ , et inversement. Autrement dit, tout évènement irréalisable dans le monde physique doit avoir un prix nul, et inversement, tout évènement de prix nul ne peut avoir une probabilité strictement positive de réalisation dans le monde physique.

La littérature académique s'intéressant à la valorisation des actifs financiers (et plus particulièrement à la valorisation des obligations, ou des taux d'intérêts) a donc suivi deux routes complémentaires. La première a établi des modèles, dits structurels, spécifiant précisément la forme du facteur d'escompte stochastique comme une fonction de variables d'intérêts choisies à partir d'un modèle économique sous-jacent, à la manière du modèle CCAPM ci-dessus. Les travaux de Piazzesi et Schneider (2007) s'inscrivent parmi de nombreux autres dans cette approche, utile

---

<sup>9</sup>Les modèles économiques en équilibre général font aussi implicitement l'hypothèse d'absence d'opportunités d'arbitrage, une opportunité d'arbitrage étant par définition une situation de déséquilibre, que les agents économiques rationnels devraient exploiter sans limite.



pour la compréhension économique du niveau absolu et relatif du prix des actifs financiers et de leur variation dans le temps.

Malheureusement, la confrontation de ces modèles aux données empiriques montre qu'ils sont encore perfectibles, et justifie le recours à la seconde approche présentée ci-dessous. Celle-ci, initiée par Ross (1976) pour les marchés d'actions, tire avantage des faibles contraintes pesant sur la forme du facteur d'escompte stochastique pour modéliser les prix des actifs, ou leurs flux monétaires associés, comme une fonction ad-hoc d'un nombre déduit de facteurs, tout en respectant l'hypothèse d'absence d'opportunités d'arbitrage. En comparaison des modèles structurels, les modèles de prix d'actifs à forme réduite sont plus flexibles et reproduisent mieux les prix de marché observés.

Nous avons présenté précédemment comment mesure de probabilité physique  $\mathbb{P}$ , facteur d'escompte stochastique  $M_{t,t+1}$ , et mesure de probabilité risque-neutre  $\mathbb{Q}$  étaient liés : la spécification de deux d'entre eux définissant automatiquement la troisième. Trois approches sont donc acceptables pour la valorisation en forme réduite des prix d'actifs [voir Bertholon, Monfort, Pegoraro (2008)] :

i) la modélisation **directe** de la distribution conditionnelle jointe du facteur d'escompte stochastique et des flux monétaires associés au titre, d'où l'on dérive la distribution risque-neutre des flux monétaires futurs (actualisés), et par conséquent, le prix de l'actif,

ii) la modélisation **risque-neutre "contrainte"**, qui spécifie les distributions conditionnelles physique et risque-neutre des flux monétaires (ou des facteurs déterminants ces flux). Le facteur d'escompte stochastique est alors déterminé par la confrontation des probabilités physique et risque-neutre,

iii) la modélisation **inverse**, qui détermine la distribution conditionnelle, sous  $\mathbb{Q}$ ,

des flux monétaires, et la forme du facteur d'escompte stochastique, dont on dérive la distribution physique des flux monétaires.

Chaque approche a ses avantages et inconvénients, et est mobilisée en fonction des objectifs de modélisation : le contrôle de la dynamique risque-neutre des facteurs, donc de la distribution risque-neutre des flux monétaires futurs [approches ii) et iii)] peut faciliter les formules de valorisation ; alternativement, il peut être souhaité de déterminer explicitement la dynamique historique des flux monétaires [approches i) et ii)] au regard de la dynamique des prix observée dans le passé ; enfin, spécifier formellement le facteur d'escompte stochastique [approches i) et iii)] permet de contrôler explicitement l'attitude de l'investisseur envers le risque.

## LES MODÈLES AFFINES DES TAUX D'INTÉRÊT D'OBLIGATIONS ZÉRO-COUPON EN ABSENCE D'OPPORTUNITÉS D'ARBITRAGE

Pour les obligations, les modèles de structure par terme des taux d'intérêt à forme réduite en absence d'opportunités d'arbitrage dits "affines" sont devenus particulièrement populaire depuis le travail original de Duffie et Kan (1996). Ces modèles à forme réduite contraignent la forme du facteur d'escompte stochastique ainsi que la dynamique physique et risque-neutre des facteurs de façon à ce que les taux d'intérêts modélisés soient des fonctions affines des facteurs [voir Gourieroux, Monfort, Polimenis (2006) pour une description complète des modèles affines de taux d'intérêt en temps discret<sup>10</sup>, Duffie, Filipovic, et Schachermayer (2003) pour les modèles à temps continu] de type :

$$r(t, h) = \alpha'(h)X_t + \beta(h), h \in \mathbb{N}^+, \quad (0.0.14)$$

---

<sup>10</sup>En particulier, Gourieroux, Monfort, Polimenis (2006) Section 4, pour la valorisation des obligations zéro-coupon sans risque de défaut.

où  $X_t$  désignent les facteurs. Les fonctions  $\alpha(h)$ ,  $\beta(h)$ , correspondent à des structures par terme de base fonction de la maturité des taux  $h \in \mathbb{N}^+$ , respectivement stochastique et déterministe, qui, pour une valeur des facteurs  $X_t$  donnée, se combinent pour former la courbe de taux modélisée à la date  $t$ .

Ces modèles spécifient généralement le facteur d'escompte stochastique comme une fonction exponentielle affine des facteurs<sup>11</sup> :

$$M_{t,t+1} = \exp[-\gamma_0(X_t) - \gamma_1'(X_t)X_t], \quad (0.0.15)$$

où les fonctions  $\gamma_0(X_t)$ ,  $\gamma_1(X_t)$  déterminent la sensibilité au risque de l'investisseur<sup>12</sup>. Cette spécification, qui assure la positivité du facteur d'escompte stochastique (nécessaire à l'absence d'opportunité d'arbitrage) quelque soit la valeur des facteurs, est très générale : la plupart des facteurs d'escompte stochastiques des modèles structurels peuvent être écrit sous une forme exponentielle affine [Gourieroux, Monfort (2007)].

La seconde hypothèse fondamentale des modèles affines de taux d'intérêt stipule que la dynamique physique ou risque-neutre des facteurs soit CAR, c'est-à-dire telle que la transformée de Laplace conditionnelle des facteurs soit une fonction exponentielle affine des valeurs présentes et passées des facteurs<sup>13</sup> [Darolles, Gourieroux, Jasiak (2006)]. Par exemple, si les  $K$  facteurs  $X_t$  sont CAR d'ordre 1 sous la dynamique physique  $\mathbb{P}$  :

$$\mathbb{E}^{\mathbb{P}}[\exp(-u'X_{t+1})|X_t] = \exp(-a[u]X_t - b[u]), u \in \mathbb{R}^K. \quad (0.0.16)$$

La classe des processus CAR est très flexible. Par exemple, les processus peuvent être multivariés, et leur dynamique peut facilement dépendre des valeurs des fac-

<sup>11</sup>Des travaux récents ont aussi introduit la classe des facteurs d'escompte stochastique fonction exponentielle affine quadratique des facteurs [voir Monfort, Pegoraro (2012)].

<sup>12</sup>La littérature considère plusieurs formes de fonctions  $\gamma_0(X_t)$  et  $\gamma_1(X_t)$  selon que la sensibilité au risque de l'investisseur soit indépendante ou non de la valeur présente des facteurs [ $\gamma_0(X_t) = \gamma_0$ ,  $\gamma_1(X_t) = \gamma_1$ ]. Duffee (2002) souligne ainsi les bonnes propriétés empiriques de la spécification dite "essentiellement affine", où les fonctions  $\gamma_0(X_t)$  et  $\gamma_1(X_t)$  sont linéaires dans les facteurs.

<sup>13</sup>Des hypothèses additionnelles sont nécessaires pour assurer que la transformée de Laplace conditionnelle caractérise la distribution conditionnelle de la variable d'intérêt, et que les moments de la distribution conditionnelle existent [voir Darolles, Gourieroux, Jasiak (2006)].

teurs plusieurs fois retardées [Monfort, Pegoraro (2007)]. Elle est par ailleurs très générale et couvre notamment les processus autorégressif Gaussiens, autorégressif Gamma [voir Gouriéroux, Jasiak (2006)], les processus matriciels autorégressifs Wishart [voir Gouriéroux, Jasiak, Sufana (2012)], et les processus à saut [Duffie, Pan, Singleton (2000)].

Les formules de valorisation des obligations zéro-coupon, et de leurs taux d'intérêt [c'est-à-dire les fonctions  $\alpha(h)$ ,  $\beta(h)$  dans (0.0.35)] sont facilement obtenues sous les hypothèses de dynamique physique CAR des facteurs, et de facteur d'escompte stochastique exponentiel affine. Tout d'abord, (0.0.29), (0.0.36) et (0.0.37) impliquent :

$$\begin{aligned}
 B(t, 1) &= \mathbb{E}^{\mathbb{P}} [M_{t,t+1}] \\
 &= \mathbb{E}^{\mathbb{P}} [\exp(-\gamma_0(X_t) - \gamma_1(X_t)X_t)] \\
 &= \exp(-\gamma_0(X_t) - a[\gamma_1(X_t)]X_t - b[\gamma_1(X_t)]) \\
 &= \exp[-r(t, 1)].
 \end{aligned}$$

D'où :

$$\begin{aligned}
 r(t, 1) &= \alpha'(1)X_t + \beta(1) \\
 &= a[\gamma_1'(X_t)]X_t + b[\gamma_1'(X_t)] + \gamma_0(X_t). \tag{0.0.17}
 \end{aligned}$$

Puis :

$$\begin{aligned}
 B(t, h) &= \mathbb{E}^{\mathbb{P}} [M_{t,t+1}B(t+1, h-1)] \\
 \exp[-h\alpha(h)'X_t - h\beta(h)] &= \mathbb{E}^{\mathbb{P}} [\exp(-\gamma_1'(X_t)X_t - \gamma_0(X_t) - (h-1)\alpha'(h-1)X_{t+1} \\
 &\quad - (h-1)\beta(h-1))] \\
 &= \exp(-r(t, 1) - (h-1)\beta(h-1) + a[\gamma_1'(X_t)]X_t \\
 &\quad + b[\gamma_1'(X_t)] - a[\gamma_1'(X_t)] + (h-1)\alpha'(h-1)X_t \\
 &\quad - b[\gamma_1'(X_t) + (h-1)\alpha'(h-1)]),
 \end{aligned}$$

d'après (0.0.38) et (0.0.37).

D'où par identification :

$$h\alpha'(h) = \alpha'(1) + a[\gamma'_1(X_t) + (h-1)\alpha'(h-1)] - a[\gamma'_1(X_t)],$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + b[\gamma'_1(X_t) + (h-1)\alpha'(h-1)] - b[\gamma'_1(X_t)].$$

Les méthodes de valorisation risque-neutre contrainte et inverse [approches ii) et iii) ci-dessus] valorisent directement le prix des obligations à partir de la dynamique risque-neutre des facteurs. Ces méthodes s'appuient sur un jeu d'hypothèses différent, qui assure la forme affine des taux d'intérêt sans requérir nécessairement la forme exponentielle affine du facteur d'escompte stochastique. La courbe de taux modélisée est en effet affine si et seulement si :

- l'historique des facteurs résume complètement l'information disponible (autrement dit, la connaissance de l'historique des facteurs est suffisante pour valoriser au mieux les obligations);
- la dynamique risque-neutre des facteurs est CAR. Pour les processus CAR(1) nous notons :

$$\mathbb{E}^{\mathbb{Q}}[\exp(-uX_{t+1})|X_t] = \exp(-a^{\mathbb{Q}}[u]X_t - b^{\mathbb{Q}}[u]), u \in \mathbb{R}^N; \quad (0.0.18)$$

- le taux d'intérêt de court-terme est une fonction affine des facteurs :

$$r(t, 1) = \alpha'(1)X_t + \beta(1).$$

Les formules de valorisation s'obtiennent alors de la façon suivante :

$$B(t, h) = \mathbb{E}_t^{\mathbb{Q}}[B(t, 1)B(t+1, h-1)],$$

soit :

$$\begin{aligned}
 \exp[-h\alpha'(h)X_t - h\beta(h)] &= \mathbb{E}_t^{\mathbb{Q}} [\exp(-r(t, 1) - (h-1)\alpha'(h-1)X_{t+1} \\
 &\quad - (h-1)\beta(h-1))] \\
 &= \exp\left(-r(t, 1) - a^{\mathbb{Q}}[(h-1)\alpha'(h-1)]X_t \right. \\
 &\quad \left. - (h-1)\beta(h-1) - b^{\mathbb{Q}}[(h-1)\alpha'(h-1)]\right).
 \end{aligned}$$

D'où par identification :

$$h\alpha'(h) = \alpha'(1) + a^{\mathbb{Q}}[(h-1)\alpha'(h-1)], \quad (0.0.19)$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + b^{\mathbb{Q}}[(h-1)\alpha'(h-1)].$$

Lorsque l'une des dynamiques (risque-neutre ou physique) des facteurs est CAR, la spécification explicite du facteur d'escompte stochastique comme une fonction exponentielle affine des facteurs [approches i) et iii)] permet de déterminer aisément le lien entre les distributions conditionnelles physique et risque-neutre des facteurs par leurs transformées de Laplace conditionnelles. Dans ce cas, l'autre dynamique est nécessairement CAR, et telle que [Bertholon, Monfort, Pegoraro (2008)] :

$$\mathbb{E}_t^{\mathbb{Q}} [\exp(-uX_{t+1})] = \frac{\mathbb{E}_t^{\mathbb{P}} [\exp(-[u + \gamma_1'(X_t)] X_{t+1})]}{\mathbb{E}_t^{\mathbb{P}} [\exp(-\gamma_1(X_t)X_{t+1})]}. \quad (0.0.20)$$

La formules de valorisation (0.0.40) souligne combien les structures par terme de base  $\alpha(h)$ ,  $\beta(h)$  sont nécessairement liées à une dynamique risque-neutre sous-jacente des facteurs. Tous les modèles affines des taux d'intérêt ne sont donc pas nécessairement compatibles avec l'absence d'opportunités d'arbitrage.

En guise d'illustration, intéressons nous au cas des modèles affines à facteurs Gaussiens, i.e. pour lesquels les distributions risque-neutre des facteurs sont Gaussiennes. Par souci de simplification, considérons le cas où  $X_t$  désigne un facteur univarié.

Appliquons la méthode de valorisation inverse, et supposons que le facteur  $X_t$  suit

sous  $\mathbb{Q}$  un processus autorégressif d'ordre 1, affecté par des chocs Gaussiens :

$$X_{t+1} = \mu_{\mathbb{Q}} + \rho_{\mathbb{Q}}X_t + \sigma_{\mathbb{Q}}\varepsilon_{\mathbb{Q},t+1},$$

où le choc  $\varepsilon_{\mathbb{Q},t}$  est un bruit blanc Gaussien de moyenne nulle et de variance 1 sous  $\mathbb{Q}$ . La transformée de Laplace risque-neutre du facteur s'écrit donc :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-uX_{t+1})] = \exp\left(-u\rho_{\mathbb{Q}}X_t - u\mu_{\mathbb{Q}} + \frac{(u\sigma_{\mathbb{Q}})^2}{2}\right),$$

soit dans nos notations :

$$a^{\mathbb{Q}}[u] = u\rho_{\mathbb{Q}} \quad b^{\mathbb{Q}}[u] = u\mu_{\mathbb{Q}} - \frac{(u\sigma_{\mathbb{Q}})^2}{2},$$

ce qui donne les formules récursives de valorisation :

$$h\alpha(h) = \alpha(1) + (h-1)\alpha(h-1)\rho_{\mathbb{Q}},$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + \mu_{\mathbb{Q}}(h-1)\alpha(h-1) - \frac{((h-1)\alpha(h-1)\sigma_{\mathbb{Q}})^2}{2}.$$

La figure 1 ci-dessous présente pour illustration les structures par terme de base  $\alpha(h)$ ,  $\beta(h)$ ,  $h \in \mathbb{N}^+$  pour les paramètres risque-neutre suivants :

$$\mu_{\mathbb{Q}} = 0 \quad \rho_{\mathbb{Q}} = 0.8 \quad \sigma_{\mathbb{Q}} = 0.1.$$

La figure 1 illustre les propriétés de convergence et de décroissance des structures par termes de base  $\alpha(h)$ ,  $\beta(h)$  dans ce modèle affine Gaussien à un facteur. En particulier, l'influence du facteur stochastique  $X_t$  [déterminée par la fonction  $\alpha(h)$ ] sur les taux diminue avec leur maturité : à la limite, les taux de très long-terme modélisés sont constants  $[r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h) = \lim_{h \rightarrow \infty} \alpha(h)X_t + \beta(h) = \lim_{h \rightarrow \infty} \beta(h)]$ .

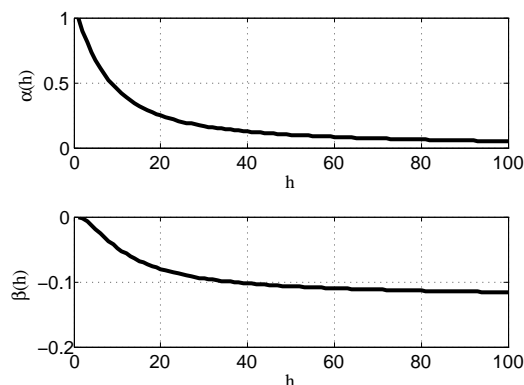


Figure 1: Structures par terme de base, stochastique  $[\alpha(h)]$  et déterministe  $[\beta(h)]$ , pour un modèle affine Gaussien à un facteur, où  $h$  (en abscisse) désigne la maturité résiduelle des taux.

## LES MODÈLES DE TAUX D'INTÉRÊT À TAUX LIMITE STOCHASTIQUES

Tout comme cet exemple, la plupart des modèles de taux d'intérêt en absence d'opportunités d'arbitrage implique un taux limite  $r(t, \infty)$  constant<sup>14</sup>. Certains induisent une explosion des taux des très long-terme, et presque aucun ne suppose un taux limite stochastique<sup>15</sup>. Ainsi, dans la plupart des modèles de taux d'intérêt, pour ne pas dire tous, plus la maturité du taux est grande, moins le taux est volatile. Cette caractéristique n'est pas conforme à la dynamique observée des taux d'intérêt. Par ailleurs, elle contraint fortement les taux extrapolés des courbes de taux observées, extrapolation qui est parfois nécessaire pour la valorisation des flux monétaires de très long-terme [par exemple lors de l'évaluation du

<sup>14</sup>Il semble que ce soit lié aux contraintes très forte induite par l'hypothèse d'absence d'opportunités d'arbitrage sur la modélisation des taux très long : sous cette hypothèse, le taux de très long-terme  $r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h)$  ne peut décroître dans le temps [Dybvig, Ingersoll, Ross (1996), Hubaleck, Klein, Teichmann (2002), Kardaras et Platen (2012)].

<sup>15</sup>Yao (1998), Appendice A, présente une liste partielle de modèles de taux d'intérêt, en soulignant leurs implications pour le taux limite. Ingersoll, Skelton, Weil (1978) proposent un exemple de modèle où le taux limite est stochastique, à variations discrètes. Yao (1998) conjecture un modèle à taux long stochastique, mais sans fournir de formules de valorisation explicites, ni d'évaluation des propriétés empiriques d'un tel modèle.



financement d'un système de retraite, ou d'un projet d'infrastructure]. La figure 2 présente la structure par terme de la volatilité des taux d'intérêt sur les obligations du Trésor des Etats-Unis pour différentes périodes<sup>16</sup>. Elle montre que la volatilité des taux d'intérêt n'est pas nécessairement une fonction décroissante de la maturité des taux (la volatilité des taux est ainsi une fonction constante entre 1995 – 2001, et croissante de la maturité des taux depuis début 2011).

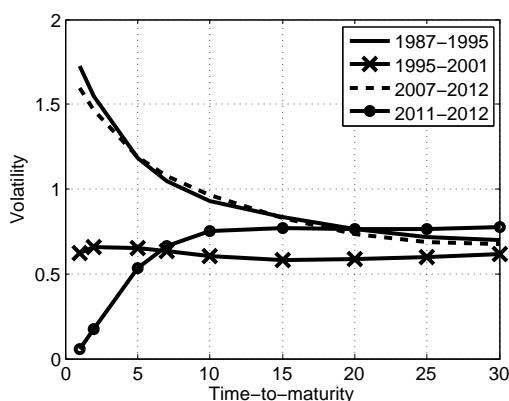


Figure 2: Structure par terme de la volatilité des taux d'intérêt des obligations du Trésor des Etats-Unis. La maturité des taux, en années, se trouve en abscisse, la volatilité des taux (en pourcentage) en ordonnées.

L'objectif de la deuxième partie de cette thèse est de proposer un modèle affine de taux d'intérêt à taux limite stochastique, et d'évaluer ses propriétés empiriques pour la modélisation des taux d'intérêt des obligations du Trésor des Etats-Unis. Pour cela, nous nous concentrons sur les modèles de taux qui incorporent un facteur niveau, dont les variations impactent uniformément l'ensemble des taux d'intérêt, *a fortiori* ceux de long-terme. Le modèle de Nelson et Siegel (1987) appartient à cette classe de modèle. Il spécifie les taux d'intérêt zéro-coupon comme une fonction affine de trois facteurs :

$$r(t, h) = X_{1,t} + \frac{1 - \exp(-\lambda h)}{\lambda h} X_{2,t} + \left( \frac{1 - \exp(-\lambda h)}{\lambda h} - \exp(-\lambda h) \right) X_{3,t}. \quad (0.0.21)$$

<sup>16</sup>Nous utilisons la base de données fournies par Gurkaynack, Sack, Wright (2007). Nous ne présentons pas de résultats pour la période 2002-2006, lorsque le Trésor des Etats-Unis n'émettait pas d'obligations de maturités 30 ans.

Ce modèle est très populaire parmi les praticiens [voir le rapport de la BIS (2005) sur l'utilisation du modèle de Nelson-Siegel dans les banques centrales] du fait de sa capacité à reproduire la plupart des courbes de taux observées. Par ailleurs, les facteurs  $X_{1,t}$ ,  $X_{2,t}$ ,  $X_{3,t}$  y ont une interprétation intuitive de par leurs facteurs de charge  $\alpha_1(h) = 1$ ,  $\alpha_2(h) = \frac{1 - \exp(-\lambda h)}{\lambda h}$ , et  $\alpha_3(h) = \frac{1 - \exp(-\lambda h)}{\lambda h} - \exp(-\lambda h)$ , qui détermine l'influence d'une variation marginale de chaque facteur sur la courbe de taux [nous présentons graphiquement les facteurs de charge dans la figure 3, pour un paramètre  $\lambda$  donné]. Dans ce modèle, le premier facteur ( $X_{1,t}$ ) a un effet de niveau et affecte uniformément les taux, tandis que les deuxième et troisième facteurs ( $X_{2,t}$  et  $X_{3,t}$ ) déterminent les évolutions de la pente et de la courbure de la courbe des taux. Ce modèle est donc particulièrement intéressant pour la modélisation des taux à grande maturité, puisqu'il a la particularité de modéliser un taux de très long terme [ $r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h)$ ] stochastique. Il n'est toutefois pas compatible avec l'hypothèse d'absence d'opportunités d'arbitrage [Bjork et Christensen (1999), Filipovic (1999)]. Une des contributions de cette thèse est de

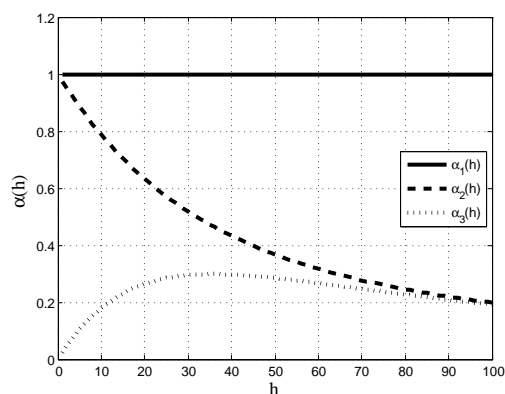


Figure 3: Structures par terme de base  $\alpha(h)$  pour la formule (0.0.42) de Nelson, Siegel (1987), où  $h$  (en abscisse) désigne la maturité résiduelle des taux.

proposer un modèle de taux d'intérêt non-arbitrable, qui conserve les principales propriétés du modèle de Nelson-Siegel, en particulier les taux de très long-terme stochastiques.

## LE CONCEPT DE CHOC DANS LA MODÉLISATION FINANCIÈRE

Cette présentation des principes généraux de valorisation des actifs financiers (en particulier des obligations), nous a permis d'introduire la plupart des concepts utilisés dans cette thèse. Nous avons vu que la valeur qu'un investisseur accorde à un actif financier dépend des flux monétaires futurs qu'il en espère, ainsi que des circonstances (ou états du monde) dans lesquels il percevra ces flux. En pratique, la résolution de ce problème complexe se fait à l'aide de modèles, qui aboutissent à spécifier les prix d'actifs comme fonctions d'un nombre limité de déterminants, ou facteurs (ces déterminants pouvant être choisis de façon structurelle, ou de manière plus ad-hoc).

La première partie de cette thèse porte sur les chocs qui affectent ces déterminants. Le terme de choc apparaît dans la littérature économique et financière dès la fin du 19<sup>ème</sup> siècle [Horton (1886), Giddings (1887)], et désigne une perturbation fréquente et irrégulière se propageant dans le système économique. Les modèles économiques et financiers modernes visent à résumer les fluctuations de l'économie et des marchés financiers par un système d'équations déterminant la dynamique jointe des principales variables d'intérêt. Les chocs se définissent alors comme des perturbations extérieures au système, des "sources d'énergie" qui vont alimenter la dynamique endogène des variables économiques et financières étudiées. L'identification des chocs passés est aujourd'hui devenue fondamentale pour la compréhension des phénomènes économiques et financiers [Duarte et Hoover (2011)].

Nous utilisons le terme de choc pour désigner la part, dans les fluctuations des facteurs, qui ne peut s'expliquer au regard de l'information disponible *ex-ante*. Les chocs permettent donc d'inscrire les facteurs dans une échelle de temps en distinguant à chaque date  $t$  la part de leur dynamique qui relève des événements présents, et celle qui relève des conséquences des événements passés. En guise d'illustration, intéressons nous au cas d'un facteur de valeur  $X_t$  à la date  $t$ , dont la dynamique (physique) peut être modélisée comme un processus autorégressif

d'ordre 1, affecté par des chocs gaussiens :

$$X_{t+1} = \mu + \rho X_t + \sigma \varepsilon_{t+1},$$

où  $|\rho| < 1$ , et où le choc  $\varepsilon_t$  est un bruit blanc gaussien de moyenne nulle et de variance 1. Supposons par exemple  $\rho = 0.8$ ,  $\mu = 0$ ,  $\sigma = 0.1$  et  $X_0 = 1$ . La figure 4 ci-dessous présente une simulation de la trajectoire de 100 périodes (entre  $t = 1$  et  $t = 100$ ) du facteur, ainsi que les réalisations des chocs  $(\varepsilon_t)_{t=1,\dots,100}$ . La survenance

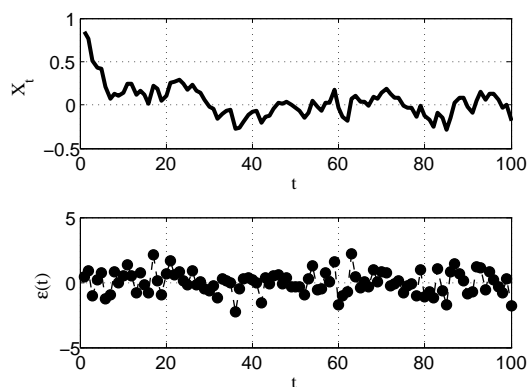


Figure 4: Simulation d'un processus autorégressif gaussien d'ordre 1 sur 100 périodes. Le graphique du haut présente la trajectoire du processus simulé, tandis que celui du bas décrit les chocs affectant le processus à chaque période  $t$  (en abscisse).

des chocs modifie la distribution du facteur, comme illustré par la figure 5 à la suite, qui présente la modification de la distribution de la valeur future du facteur  $X_1, \dots, X_{100}$ , conditionnellement à l'information initiale disponible, en  $t=0$ . L'accumulation des chocs au cours du temps a ici un double effet de position et de dispersion sur la distribution conditionnelle<sup>17</sup>.

Dans le cadre d'un modèle financier, les chocs déterminent la distribution conditionnelle des facteurs déterminant les prix d'actifs, et donc la distribution conditionnelle de la valeur de n'importe quel portefeuille d'actifs financiers. Dans cette thèse, nous faisons explicitement le lien entre la nature des chocs et la distribu-

<sup>17</sup>Dans cet exemple, la nature des chocs ne change pas au cours du temps. Il est vraisemblable que cette hypothèse n'est pas réaliste pour la modélisation des séries financières, qui alternent les périodes de faible volatilité et de forte volatilité.

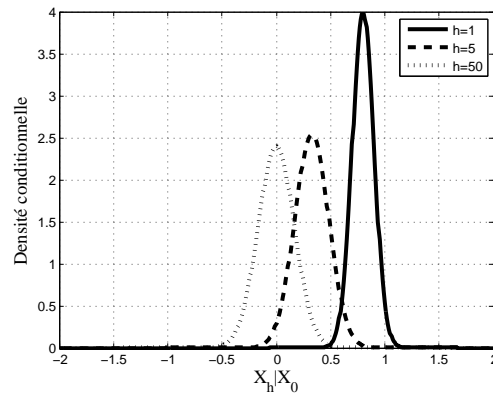


Figure 5: Densité de la distribution de la valeur future du facteur  $X_h$ , conditionnellement à la connaissance de sa valeur initiale  $X_0$ . Les courbes pleine, hachurée, et pointillée désignent respectivement les densités conditionnelles à 1, 5, et 50 périodes.

tion conditionnelle de la valeur future d'un portefeuille financier, et étudions la sensibilité du portefeuille à divers types de chocs.

## PRÉSENTATION DES CHAPITRES DE LA THÈSE

Dans le premier chapitre de cette thèse, nous proposons une définition très générale des chocs affectant les déterminants des prix d'actifs. Nos résultats sont utiles à la gestion du risque financier. Il est en effet courant de mesurer le risque d'un portefeuille en évaluant ses propriétés dans un environnement dégradé, ou de façon équivalente, en mesurant sa sensibilité à certains chocs par des tests de résistance, ou "stress-tests". Les chambres de compensation déterminent par exemple le niveau de garantie demandé à chaque membre (les appels de marge) de cette façon [Pérignon, Lopez, Harris (2011)]. De même, les régulateurs européens (l'Autorité Bancaire Européenne) et américains (la Réserve Fédérale) communiquent depuis quelques années la sensibilité de chaque banque régulée à plusieurs chocs sur les déterminants macroéconomiques des prix d'actifs, et requièrent des banques trop sensibles qu'elles augmentent leurs réserves de fonds propres et/ou modifient leur

stratégie d'investissement.

Dans ce premier chapitre, nous proposons d'enrichir ces exercices par une définition plus générale des chocs utilisés. Alors qu'en pratique, les tests de résistance s'intéressent à la réalisation d'un choc donné, ou déterministe, nous proposons d'étudier les conséquences de chocs stochastiques. Notre approche est ainsi plus robuste, plus difficilement manipulable par le gérant du portefeuille, et permet de prendre en compte de façon cohérente des phénomènes de crises, tels que l'augmentation de la corrélation des prix d'actifs. Dans ce chapitre, nous discutons les deux représentations possibles des chocs, en termes de distribution, ou en termes de variables, et soulignons leurs différences. Nous dérivons analytiquement les conséquences pour un portefeuille de chocs stochastiques sur les facteurs communs aux prix d'actifs composant le portefeuille. Par ailleurs, alors que les exercices de stress-tests se réalisent pour une allocation de portefeuille donnée, nous étudions l'impact des chocs sur les stratégies d'investissement. La littérature financière considère une multitude de critères d'investissement, tels que l'utilité espérée de l'investisseur, ou encore le rendement (moyen ou médian) du portefeuille sous contrainte de risque maximal (où le risque est mesuré par la variance [Markowitz (1952)], ou par un quantile du rendement du portefeuille). Tous ces critères dépendent de la distribution de la valeur future du portefeuille, et donc des chocs altérant cette distribution. Les résultats du premier chapitre présentent, pour plusieurs critères d'investissement, la réponse optimale de l'investisseur en fonction des caractéristiques des chocs. Ils sont donc particulièrement utiles à l'investisseur cherchant à avoir une vision prospective des conséquences de certains chocs, et au régulateur qui souhaite prévoir l'influence d'un choc sur les prises de risque des investisseurs. Finalement, nous illustrons nos résultats par une application, où nous proposons un exemple intuitif de choc stochastique, auquel nous évaluons la sensibilité d'un portefeuille d'obligations souveraines européennes.

Les deuxième et troisième chapitres concernent la valorisation des obligations zéro-coupon, sans risque de défaut de l'émetteur. Les enjeux en termes de gestion du risque financier diffèrent ici du chapitre précédent: alors que le premier chapitre

s'intéressait au risque porté par un investisseur, ou un gestionnaire de fond, les chapitres suivants concernent le risque pris par un teneur de marché qui propose des cotations pour plusieurs actifs (dans notre cas, pour plusieurs obligations de maturités différentes). Ces cotations doivent répondre à deux objectifs : être cohérentes avec les cotations des autres teneurs de marchés, et surtout avoir une cohérence interne, i.e. que l'ensemble des valorisations annoncées par le teneur de marché ne soient pas arbitrables par les autres acteurs de marché. Si la gestion du risque financier est concernée par la distribution physique des prix d'actifs, la valorisation d'un titre est liée à la distribution risque-neutre des déterminants des flux monétaires futurs associés (i.e. à la distribution physique des facteurs sous-jacents à la dynamique des prix, modifiée par le facteur d'escompte stochastique). Les deuxième et troisième chapitres portent donc sur les chocs affectant la distribution risque-neutre des déterminants des prix d'obligations et des taux d'intérêts.

Par ailleurs, nous nous concentrons sur un segment de la courbe de taux encore méconnu de la littérature : celui des obligations à maturité résiduelle ultra longue, supérieure à celle des obligations dites de long-terme de maturité résiduelle 10 ans. La modélisation des taux d'intérêt ultra-long est un défi pour les modèles classiques de courbe de taux, qui supposent dans leur grande majorité une décroissance rapide de la volatilité des taux à mesure que leur maturité augmente. La plupart suppose ainsi un taux limite, i.e. de maturité infinie, constant, à la manière du modèle Gaussien univarié présenté ci-dessus. Cette caractéristique des modèles classiques de taux d'intérêt entre en contradiction avec les données empiriques, qui montrent que la volatilité des taux ultra-long est comparable à celle des taux longs usuels, et parfois même plus grande (voir Figure 2). Parmi les rares modèles de la littérature à taux limite stochastique, certains introduisent un facteur niveau, dont les fluctuations génèrent des mouvements parallèles de la courbe de taux [voir le modèle de Nelson et Siegel (1987) cité ci-dessus]. Le deuxième chapitre s'intéresse à la compatibilité du facteur niveau avec l'hypothèse d'absence d'opportunités d'arbitrage. Nous y montrons qu'un modèle de taux d'intérêt incorporant un facteur niveau requiert des chocs très spécifiques sur les dynamiques physique et

risque-neutre de ce facteur pour être compatible avec l'absence d'opportunités d'arbitrage, la positivité des taux d'inérêt et la non-explosion des taux de très long-terme.

Nous prolongeons ces résultats dans le troisième chapitre, dans lequel nous introduisons une nouvelle classe de modèles à facteurs pour la courbe des taux d'intérêt. La dynamique risque-neutre des facteurs  $y$  est supposée impactée par des chocs bruits blancs, strictement positifs, et à distribution stable<sup>18</sup>. La distribution stable est une distribution statistique qui englobe l'ensemble des distributions restant stable par sommation [dont les distributions Levy, Cauchy, et Gaussienne] : la somme de chocs à distribution stable est elle-même une variable aléatoire à distribution stable [voir Samorodnitsky, Taquq (1994) et Nolan (2007) pour une présentation détaillée des distributions stables]. Les distributions stables sont donc particulièrement utiles pour la valorisation d'obligations, qui nécessite le calcul de la transformée de Laplace d'une somme de chocs aléatoires. Les facteurs à distribution risque-neutre stable ont déjà été utilisés dans la littérature, afin de modéliser la structure par terme des options sur actions américaines [Carr et Wu (2003)], ou de reproduire les variations extrêmes des taux d'intérêts [Raible (2000)]. Nous montrons dans ce troisième chapitre que les distributions stables sont aussi utiles à la modélisation des courbes de taux: elles assurent la positivité des taux modélisés [en ce sens, elles offrent une alternative au modèle de Cox, Ingersoll et Ross (1985)], et permettent d'introduire des modèles à limite stochastique, tout en respectant l'hypothèse d'absence d'opportunités d'arbitrage. Nous introduisons dans ce chapitre un modèle de taux d'intérêt de la famille Nelson-Siegel, non-arbitrable et à taux limite stochastique. Les propriétés empiriques de ce modèle sont illustrées lors d'une application sur les taux des obligations du Trésor des Etats-Unis, dont la maturité peut atteindre trente ans.

---

<sup>18</sup>Ces chocs sont donc aussi positifs sous la mesure de probabilité physique, les supports des distributions conditionnelles risque-neutre et physique des facteurs étant les mêmes (voir ci-dessus).





# General Introduction

---

Financial risk has multiple dimensions: financial regulators distinguish for instance the credit risk (the risk associated with the default of a borrower), from the market risk (the risk of losses on the positions of a financial institution, on and off-balance sheet, due to the fluctuations in market prices) and the operational risk (risk of losses from human failures, inadequate internal processes and models, or external physical events). This thesis is concerned with the so-called "model-risk": the risk of losses due to inadequate internal model for the pricing of financial assets and the measure of investment risk. The recent financial crisis has dramatically demonstrated how critical the control of internal models was for the survival of financial institutions. This thesis is composed of two different parts, though both related to the management of model risk.

Financial models should be robust and coherent. Robust, in the sense that they should not be dramatically affected by a slight change in their underlying assumptions. Coherent, in the sense that they should not identify unlimited investment opportunities with unrealistically high return/low risk (that the economic literature calls arbitrage opportunities).

This introduction is organized in the following way. First, we present the financial assets, to which a large part of the thesis is devoted, in particular the zero-coupon

bonds, and introduce most of the notations, which are useful for the reading of this manuscript. We recall the general principles of asset pricing, focusing particularly on the modeling of the term structure of the prices and interest rates of the zero-coupon bonds, under no-arbitrage assumptions<sup>19</sup>. We build on this presentation to define the concept of shocks in financial modeling, and the related issues for financial risk management. We conclude this introduction by a quick presentation of the three chapters of this thesis.

## THE CHARACTERISTICS OF BONDS CONTRACTS

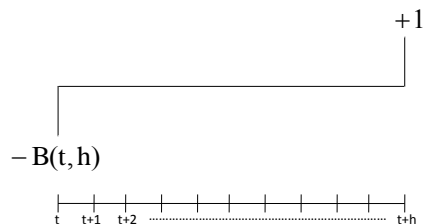
A financial asset is a contract, which gives to its owner the right (or the duty) to receive (pay) monetary flows in the future. A bond for instance, is a contract according to which the bond's issuer promises to pay to the contract's owner monetary flows, according to a rule, which is set at the issuance of the bond. A bond differs from a stock, for which the shareholders vote each year the amount of flows (the dividends) paid to the stocks' owners. A nominal zero-coupon bond is an elementary financial asset, which pays to its owner a unique flow of one monetary unit at the date of maturity  $T$  of the bond<sup>20</sup>. We denote  $B(t, h)$  the price at date  $t$  of a zero-coupon bond with residual maturity<sup>21</sup>  $h = T - t$ . The schemes 1 and 2 present the monetary flows associated to the owning of a zero-coupon bond (assuming the bond's issuer does not default), whether the bond is held up to the maturity of the bond (scheme 1), or whether the bond is resold before maturity (scheme 2):

---

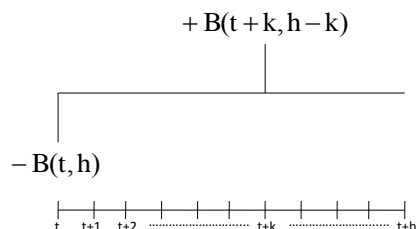
<sup>19</sup>This introduction draws liberally on the books of Cochrane (2001) and Gouriéroux, Jasiak (2001).

<sup>20</sup>This thesis focuses on the pricing of nominal bond. To not increase the notations, the term "bond" refers to "nominal bond" thereafter.

<sup>21</sup>In this thesis, most of the results are presented in discrete time, i.e.  $h \in \mathbb{N}^+$ .



Scheme 1 : Monetary flows associated with the holding, up to maturity, of a zero-coupon bond with residual maturity  $h$  periods, with price  $B(t, h)$  at date  $t$ .



Scheme 2 : Monetary flows associated with the reselling at date  $t + k$  of the zero-coupon bond bought at date  $t$ .

The term structure of interest rates, or yield curve, collects the rates of all zero-coupon bond (denoted  $r(t, h)$  for a bond with residual maturity  $h$  at date  $t$ ) issued by the same entity, defined as<sup>22</sup>:

$$r(t, h) = -\frac{1}{h} \log B(t, h), \forall h \in \mathbb{N}^+. \quad (0.0.22)$$

Finally, we denote  $B^f(t, h_1, h_2)$  the *forward* price of a zero-coupon bond, according to which the contract holder commits himself at date  $t$  to pay  $B^f(t, h_1, h_2)$  at date

---

<sup>22</sup>The interest rate of a zero-coupon bond can be defined in two ways:

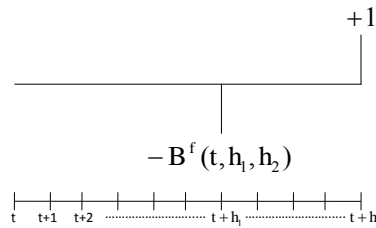
- periodically compounded interest rate:  $r(t, h) = B(t, h)^{-1/h} - 1$ ,
- continuously compounded interest rate:  $r(t, h) = -(1/h) \log B(t, h)$ ,

Without loss of generality, the results of this thesis are written for continuously compounded rates.

$t + h_1$  to the issuer, to receive one monetary unit at date  $t + h_2$  (see scheme 3 below). The *forward* rate  $r^f(t, h_1, h_2)$  is then :

$$r^f(t, h_1, h_2) = -\frac{1}{h_2 - h_1} \log B^f(t, h_1, h_2). \quad (0.0.23)$$

For the sake of simplicity, we denote  $r^f(t, h) = -\log B^f(t, h, h + 1)$  the *forward* short-term rate. The collection of *forward* short-term rates  $r^f(t, h)$  with different commitment dates  $h \in \mathbb{N}^+$  is called a term structure of *forward* interest rates.



Scheme 3 : Monetary flows for a *forward* contract regarding the purchase of a zero-coupon bond at date  $t + h_1$ , with residual maturity  $h_2$ , at a price

$$B^f(t, h_1, h_2) \text{ determined at date } t.$$

The pricing of zero-coupon bonds without risk of issuer default is fundamental in finance: many assets can be decomposed (and priced) as a portfolio of several zero-coupon bonds. For instance, a nominal bond with maturity 30 years, paying fixed coupons semi-annually and the principal *in fine*, can be decomposed in 61 zero-coupon bonds (30x2 associated with the coupon payments + 1 for the principal payment). These zero-coupon bonds may be named STRIPS and traded as such on the financial markets (as is the case with the zero-coupon bonds created from the stripping of the coupon bonds issued by the US Treasury<sup>23</sup>).

<sup>23</sup>See Sack (2000) for a detailed presentation of the STRIPS market in the US.

## GENERAL PRINCIPLES FOR THE PRICING OF FINANCIAL ASSETS

From an economic point of view, a financial asset allows an investor to smooth his consumption over time (for saving purpose), and over the different future states of the world (for insurance purpose). The value that the investor grants to a given asset depends on the flows associated with it, and on the uncertainty surrounding these flows.

The fundamental pricing equation expresses the value of any asset, for a given investor, as a conditional expectation:

$$P_{i,t} = \mathbb{E}_t \left[ M_{i,t+1}^j g_{i,t+1} \right], \quad (0.0.24)$$

where  $P_{i,t}$  indicates the price of the asset  $i$  at date  $t$ ,  $\mathbb{E}_t$  is the expectation operator conditional to the available information a date  $t$ ,  $g_{i,t+1}$  is the payoff associated with the holding of asset  $i$  at date  $t + 1$  (which can be the reselling price of the asset), and  $M_{i,t+1}^j$  refers to the so-called "stochastic discount factor" of investor  $j$ .

The stochastic discount factor is central in equation (0.0.24) and weights the future flows according to the impatience and risk appetite of the investor. In order to discuss intuitively the concept of stochastic discount factor, let us consider an economy *à la* Lucas (1978), where the investor receives at each date  $t$  an external income  $e_t$ , that he used to consume a unique quantity of physical goods  $c_t$ , with price  $q_t$ , and to invest in a quantity  $\omega_t$  of assets, with price  $P_t$  (where  $\omega_t$  and  $P_t$  are vectors)<sup>24</sup>. Besides, the financial assets does not generate intermediate flows: their reselling price are the only flows expected by the investor. The investor chooses his consumption profile ( $c_t$ ), his portfolio allocation ( $\omega_t$ ) at each date  $t \in \mathbb{N}^+$ , in order to maximize the expectation of his intertemporal utility function, subject to his budget constraint<sup>25</sup>:

$$\begin{aligned} \max_{c_t, \omega_t} \quad & \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \delta^k u(c_{t+k}) \right] \\ \text{s.t.} \quad & q_t c_t + \omega_t' P_t = e_t + \omega_{t-1}' P_t, \forall t \in \mathbb{N}^+, \end{aligned} \quad (0.0.25)$$

<sup>24</sup>To keep the notations simple, we drop the  $j$  upper script, which refers to the investor  $j$ . The results hereafter have to be understood for a given investor.

<sup>25</sup>We assume the intertemporal utility of the investor is time-additive.

where  $u(c_{t+k})$  indicates the utility the investor derives from its consumption at date  $t+k$ , and  $\delta$  refers to the his subjective discount rate.

The resolution of the investor's program (0.0.25) give rises to the Euler equation at each date  $t$ , for every asset  $i$  :

$$P_{i,t} = \mathbb{E}_t \left[ \frac{q_t}{q_{t+1}} \delta \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)} P_{i,t+1} \right]. \quad (0.0.26)$$

The investor buys (saves) or sells (borrows) the financial asset  $i$  until the Euler condition (0.0.26) is satisfied.

In this example, the stochastic discount factor is :

$$M_{t,t+1} = \frac{q_t}{q_{t+1}} \delta \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}. \quad (0.0.27)$$

It is a function of the inflation between  $t$  and  $t+1$   $\left(\frac{q_t}{q_{t+1}}\right)$ , of the investor's impatience ( $\delta$ ), and of the ratio of his marginal utility between  $t$  and  $t+1$   $\left(\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}\right)$ .

Everything else equal, the asset's value is higher if:

- i) the investor is patient,
- ii) inflation is weak,
- iii) the future value of the asset is high, when the ratio of marginal utility is high.

It is generally admitted that the investor's utility increases with his consumption, but at decreasing rate (the investor is progressively "satiated"). A high ratio of marginal utilities  $\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}$  thus indicates a decrease of the investor's consumption between  $t$  and  $t+1$ . iii) points out the value the investor grants to the insurance properties of the asset: an asset, whose value is high in the states of the world, where the investor's consumption drops (i.e.  $\frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}$  high), is more desirable because it allows the investor to smooth his consumption over time.

#### Remarks:

- When the investor's utility function is linear in his consumption, the ratio of marginal utility is constant, and the investor is said "risk-neutral". The

asset's value is in this case equal to its expected discounted future value at the subjective discount rate:

$$P_{i,t+1} = (\delta r)^{-1} \mathbb{E}_t (P_{i,t+1}), \text{ avec } r = \frac{\frac{du}{dc}(c_{t+1})}{\frac{du}{dc}(c_t)}, \forall c_t, c_{t+1}. \quad (0.0.28)$$

- The investor's objective function is here written in terms of consumption level. This model may thus not be directly applied to a portfolio manager, whose objective is a function of the future value of his portfolio.
- The intuitions given in i), ii), and iii) are relevant for a given investor. The observed assets' price on the market depends on the supply and demand of several investors, whose aggregation is a complex problem (even in the simple example above). A large part of the literature get around this difficulty by assuming the existence of a representative agent. The fundamental pricing equation (0.0.24) is then used to relate the observed market prices to a single stochastic discount factor. Under this additional assumption, the above model is commonly called the *Consumption based Capital Asset Pricing Model*, or CCAPM. In the following paragraphs, asset prices can be interpreted as the market prices for the representative agent, or the hypothetical "fair" prices for a given investor.

## GENERAL PRINCIPLES FOR THE PRICING OF ZERO-COUPON BONDS

Let us now consider the peculiar case of zero-coupon bonds. The fundamental pricing equation (0.0.24) gives:

$$B(t, 1) = \mathbb{E}_t [M_{t,t+1}], \quad (0.0.29)$$

for the short-term bond, and:

$$\begin{aligned} B(t, h) &= \mathbb{E}_t [M_{t,t+1} B(t+1, h-1)] \\ &= \mathbb{E}_t (M_{t,t+1} \mathbb{E}_{t+1} [M_{t+1,t+2} B(t+2, h-2)]) \\ &= \mathbb{E}_t [M_{t,t+1} \dots M_{t+h-1,t+h}], \end{aligned} \quad (0.0.30)$$



for the longer-term bonds, by the Theorem of iterated expectations.

The price of a zero-coupon bond is thus the conditional expectation of the product of future stochastic discount factors, over the time-to-maturity of the bond. By combining (0.0.30) with (0.0.27) in the above example, the bond's value becomes a function of the subjective discount rate, of the inflation, and of the ratio of marginal utility over the residual maturity of the bond :

$$B(t, h) = \mathbb{E}_t \left[ \delta^h \frac{q_t}{q_{t+h}} \frac{\frac{du}{dc}(c_{t+h})}{\frac{du}{dc}(c_t)} \right]. \quad (0.0.31)$$

The pricing of the term structure of zero-coupon bonds' rate (i.e. of all zero-coupon bonds ordered by their residual maturities) is thus related to the dependence of the stochastic discount factors  $M_{t,t+1}$ ,  $M_{t+1,t+2}, \dots$ ,  $M_{t+h-1,t+h}$ : it will differ dramatically when the stochastic discount factors are independent (in a probabilistic sense) with each others, or not. Conversely, the term structure of bond prices observed on the market gives useful indications about the properties of the stochastic discount factor of the representative agent [see Alvarez, Jermann (2005), Bachus, Chernov and Zin (2011)].

One can also combine (0.0.29) and (0.0.24), such as :

$$P_{i,t} = \mathbb{E}_t \left[ \frac{M_{t,t+1}}{\mathbb{E}_t(M_{t,t+1})} B(t, 1) P_{i,t+1} \right], \quad (0.0.32)$$

where  $\frac{M_{t,t+1}}{\mathbb{E}_t(M_{t,t+1})}$  defines a conditional probability density function. As a consequence, the fundamental pricing equation (0.0.24) can be rewritten as an expectation under a modified probability measure, denoted  $\mathbb{Q}$  (while the "physical" probability measure would be denoted  $\mathbb{P}$ ) :

$$P_{i,t+1} = \mathbb{E}_t^{\mathbb{Q}} [B(t, 1) P_{i,t+1}]. \quad (0.0.33)$$

Under  $\mathbb{Q}$ , any asset price is equal to the expectation of its discounted future value. This is why the probability measure  $\mathbb{Q}$  is usually called the risk-neutral probability measure.

Intuitively, such change of probability measure comes from the fact that a risk-averse investor can equivalently be represented as an investor, who

overweights the unfavorable states of the world (the ones with high stochastic discount factor  $M_{t,t+1}$ , when the investor is very sensitive to a marginal fluctuation in the asset price), and underweights the favorable states of the world.

In the specific case of zero-coupon bonds, asset prices are expectations under  $\mathbb{Q}$  of the discounted future reselling value of the bonds :

$$\begin{aligned} B(t, h) &= \mathbb{E}_t^{\mathbb{Q}} [B(t, 1)B(t + 1, h - 1)] \\ &= \mathbb{E}_t^{\mathbb{Q}} [B(t, 1) \dots B(t + h - 1, 1)], \end{aligned}$$

which differs from the "fair price" of a bond for a risk-neutral investor, denoted  $\tilde{B}(t, h)$  :

$$\tilde{B}(t, h) = \mathbb{E}_t^{\mathbb{P}} [B(t, 1) \dots B(t + h - 1, 1)]. \quad (0.0.34)$$

The bonds' risk premium, or term premium, is sometimes defined as the difference between the bond price  $B(t, h)$ , and the "fair price" for a risk-neutral investor  $\tilde{B}(t, h)$ , or, more frequently, as the difference between the bond's rate  $r(t, h) = -\frac{1}{h} \log B(t, h)$  and the "risk-neutral" rate  $\tilde{r}(t, h) = -\frac{1}{h} \log \tilde{B}(t, h)$ <sup>26</sup>.

---

<sup>26</sup>Sometimes, the risk premium is defined as the spread between the market price and the one, which would be consistent with the Theory of rational expectations, according to which [Cochrane (2001), p. 355] :

- i) the zero-coupon rate of a bond with residual maturity  $h$  is equal to the expected, under  $\mathbb{P}$ , average of future short-term rates,
- ii) the *forward* short-term rate  $r^f(t, h)$  is equal to the expected, under  $\mathbb{P}$ , future short term rate  $\mathbb{E}^{\mathbb{P}} [r(t + h, 1)]$ ,
- iii) the expected, under  $\mathbb{P}$ , return on bonds  $\log \left( \frac{B(t+1, h-1)}{B(t, h)} \right)$  is the same, whatever the bond's maturity.

The Theory of rational expectations and risk-neutrality are close, but not equivalent, since the risk premium derived from the Theory of rational expectations ignores the "convexity" term, which appears, when one transforms bond prices (i.e. exponential functions of interest rates) in rates.

## NO-ARBITRAGE REDUCED-FORM ASSET PRICING MODELS

The CCAPM example above sets precisely the form of the stochastic discount factor (0.0.27). This specification relies heavily on the underlying model's assumptions, and other forms of the stochastic discount factor can be obtained from different assumptions (on the investor's objective function for instance).

It is actually not essential for the pricing of financial assets to specify the stochastic discount factor. Many works indeed demonstrated that the assumption of no-arbitrage is enough for proving its existence. An arbitrage occurs, when it is possible to set a self-financed portfolio, whose return is at least nil, sometimes strictly positive. Assuming the absence of arbitrage opportunities is rather standard for highly liquid markets (liquid in the sense that it is possible to buy or sell an important quantity of securities without affecting the market prices), and is now a standard assumption of the modern asset pricing models<sup>27</sup>.

The absence of arbitrage opportunities is a necessary and sufficient condition for the existence of at least one strictly positive stochastic discount factor  $M_{t,t+1} > 0, \forall t$  [Harrison et Kreps (1979), Hansen et Richard (1987)]. Equivalently, under no-arbitrage, it is always possible to price financial assets as an expectation of their discounted future value, under a modified risk-neutral probability measure  $\mathbb{Q}$ , which is equivalent to the physical probability measure  $\mathbb{P}$  [Harrison et Kreps (1979)]. This modified probability measure is weakly constrained: it mostly has to be such that zero-probability events under  $\mathbb{P}$  have a zero-probability under  $\mathbb{Q}$ , and conversely. In other words, impossible events in the physical world shall have a nil price, and conversely, zero-price events shall not occur.

The academic literature on asset pricing (particularly the literature on bonds

---

<sup>27</sup>In economics, general equilibrium models are also making implicitly this assumption, an arbitrage opportunity is by definition a disequilibrated situation, that rational economic agents would endlessly take advantage on.

pricing) thus followed two complementary routes. The first one sets up structural models, specifying precisely the form of the stochastic discount factor as well as the determinants of the asset price (the representative investor's consumption in the CCAPM example above). The work of Piazzesi and Schneider (2007), among many others, is an illustration of this approach, which is useful for the economic understanding of the absolute and relative asset prices, and of their variations over time.

Unfortunately, these models have still difficulty properly matching the observed dynamics of asset prices, which justifies the use of the second type of asset pricing models. This second approach, initiated by Ross (1976) on stock prices, takes advantage of the rather loose constraints on the form of the stochastic discount factor to model asset prices, or their associated monetary flows, as an ad-hoc function of a small number of pricing factors, while respecting the assumption of no-arbitrage opportunities. In this approach, the asset pricing models are much more flexible and better reproduce the empirical properties of observed asset prices.

We previously presented how "physical" probability measure  $\mathbb{P}$ , stochastic discount factor  $M_{t,t+1}$ , and "risk-neutral" probability measure  $\mathbb{Q}$  were related: setting up two of them automatically specify the third one. Three approaches are thus acceptable for the pricing of financial assets in reduced form [voir Bertholon, Monfort, Pegoraro (2008)] :

i) the **direct** modeling of the joint conditional distribution of the stochastic discount factor and the monetary flows generated by the securities, from which the risk-neutral distribution of the discounted future monetary flows (and thus the asset's price) is derived.

ii) the **risk-neutral constrained direct** modeling, which specifies both physical and risk-neutral conditional distribution of the monetary flows. The stochastic discount factor comes out from the confrontation of both distri-

bution.

iii) the **back** modeling, which sets up the risk-neutral conditional distribution of the monetary flows, and the form of the stochastic discount factor. The physical conditional distribution of the flows is determined as a by-product.

Each approach has advantages and limitations, and should be used according with the modeling objectives: controlling the risk-neutral dynamics of the pricing factors, that is, of the monetary flows associated to the securities [approaches ii) and iii)], can make the pricing formula easier. Besides, it can be desirable to explicitly determine the physical dynamics of the pricing factors/monetary flows [approaches i) and ii)] in accordance with the asset prices dynamics observed in the past. Finally, specifying explicitly the stochastic discount factor [approaches i) iii)] allows to explicitly constrain the (modeled) investor's attitude towards risk.

## NO-ARBITRAGE AFFINE MODELS OF THE TERM STRUCTURE OF ZERO-COUPON INTEREST RATES

With respect to bond price modeling, no-arbitrage affine reduced-form models of the term structure of interest rates have become increasingly popular since the seminal work of Duffie and Kan (1996). These reduced-form models constrain the form of the stochastic discount factor and the physical (or risk-neutral) dynamics of the bonds' pricing factors such that the implied zero-coupon rates are affine functions of the pricing factors [see Gourieroux, Monfort, Polimenis (2006)<sup>28</sup> and Duffie, Filipovic, Schachermayer (2003) for a complete description of affine term structure models in discrete and con-

---

<sup>28</sup>In particular Gourieroux, Monfort, Polimenis (2006), Section 4 for the pricing of bonds without default risk.

tinuous time, respectively]:

$$r(t, h) = \alpha'(h)X_t + \beta(h), h \in \mathbb{N}^+, \quad (0.0.35)$$

where  $X_t$  indicates the pricing factors. The functions  $\alpha(h)$ ,  $\beta(h)$  are baseline term structures, respectively stochastic and deterministic, which for a given factors' value  $X_t$ , combine to give the model-implied yield curve at date  $t$ .

These models usually specify the stochastic discount factor as an exponential affine function of the pricing factors<sup>29</sup> :

$$M_{t,t+1} = \exp[-\gamma_0(X_t) - \gamma_1'(X_t)X_t], \quad (0.0.36)$$

where the functions  $\gamma_0(X_t)$ ,  $\gamma_1(X_t)$  set the sensitivity towards risk of the investor<sup>30</sup>.

This specification ensures the positivity of the stochastic discount factor (essential for the absence of arbitrage opportunities), for all realizations of the pricing factors, and is very general: most of the structural asset pricing models assume such form of the stochastic discount factor [Gourieroux, Monfort (2007)].

The second fundamental assumption of affine term structure models concerns the physical or risk-neutral dynamics of the factors, which has to be CAR: i.e. the conditional Laplace Transform of the factors has to be an exponential affine function of their past and present values<sup>31</sup> [see Darolles, Gourieroux, Jasiak (2006)]. For instance, when the  $K$  factors  $X_t$  are CAR of order 1 under the physical probability measure  $\mathbb{P}$  :

$$\mathbb{E}^{\mathbb{P}} [\exp(-u'X_{t+1})|X_t] = \exp(-a[u]X_t - b[u]), u \in \mathbb{R}^K. \quad (0.0.37)$$

---

<sup>29</sup>Recent works have introduced the class of exponential quadratic affine stochastic discount factors [Monfort, Pegoraro (2012)].

<sup>30</sup>The literature considers several functional forms for the functions  $\gamma_0(X_t)$  et  $\gamma_1(X_t)$ , assuming for instance the independence between the investor's risk sensitivity and the pricing factors [ $\gamma_0(X_t) = \gamma_0$ ,  $\gamma_1(X_t) = \gamma_1$ ]. Among these forms, Duffee (2002) highlights the good properties of the "essentially affine" specification, where the function  $\gamma_0(X_t)$  and  $\gamma_1(X_t)$  are linear in the pricing factors.

<sup>31</sup>Additional assumptions are necessary to ensure the Laplace transform fully characterize the conditional distribution of the factors, and that the moments of their distribution exist [see Darolles, Gourieroux, Jasiak (2006)].

The class of CAR processes is very flexible. For instance, the processes could be multivariate, and the dependence of the conditional factors dynamics to several lags can easily be handled [Monfort, Pegoraro (2007)]. Besides, it encompasses the most popular types of processes commonly used in the literature on term structure modeling such as the autoregressive Gaussian process, the autoregressive Gamma process [Gourieroux, Jasiak (2006)], the matrix autoregressive Wishart process [Gourieroux, Sufana (2012)], and processes with jump [Duffie, Pan, Singleton (2000)].

The zero-coupon bond pricing formula, that is the function  $\alpha(h)$  and  $\beta(h)$  in (0.0.35), are easily obtained under the assumption of CAR "physical" dynamics of the factors, and exponential affine stochastic discount factor. First, (0.0.29), (0.0.36) and (0.0.37) imply :

$$\begin{aligned}
 B(t, 1) &= \mathbb{E}^{\mathbb{P}} [M_{t,t+1}] \\
 &= \mathbb{E}^{\mathbb{P}} [\exp(-\gamma_0(X_t) - \gamma_1(X_t)X_t)] \\
 &= \exp(-\gamma_0(X_t) - a[\gamma_1(X_t)]X_t - b[\gamma_1(X_t)]) \\
 &= \exp[-r(t, 1)].
 \end{aligned}$$

Thus :

$$\begin{aligned}
 r(t, 1) &= \alpha'(1)X_t + \beta(1) \\
 &= a[\gamma_1'(X_t)]X_t + b[\gamma_1'(X_t)] + \gamma_0(X_t). \quad (0.0.38)
 \end{aligned}$$

Moreover :

$$\begin{aligned}
 B(t, h) &= \mathbb{E}^{\mathbb{P}} [M_{t,t+1}B(t+1, h-1)] \\
 \exp[-h\alpha(h)'X_t - h\beta(h)] &= \mathbb{E}^{\mathbb{P}} [\exp(-\gamma_1'(X_t)X_t - \gamma_0(X_t) - (h-1)\alpha'(h-1)X_{t+1} \\
 &\quad - (h-1)\beta(h-1))] \\
 &= \exp(-r(t, 1) - (h-1)\beta(h-1) + a[\gamma_1'(X_t)]X_t \\
 &\quad + b[\gamma_1'(X_t)] - a[\gamma_1'(X_t)] + (h-1)\alpha'(h-1)X_t \\
 &\quad - b[\gamma_1'(X_t)] + (h-1)\alpha'(h-1)),
 \end{aligned}$$

from (0.0.38) and (0.0.37).

Therefore by identification :

$$h\alpha'(h) = \alpha'(1) + a[\gamma'_1(X_t) + (h-1)\alpha'(h-1)] - a[\gamma'_1(X_t)],$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + b[\gamma'_1(X_t) + (h-1)\alpha'(h-1)] - b[\gamma'_1(X_t)].$$

The risk-neutral constrained direct and back modeling [approaches ii) and iii)] derive asset prices from the risk-neutral conditional distribution of the factors. These approaches rely on a different set of assumptions, which ensure the affine specification of the interest rates, without assuming necessarily an exponential affine form of the stochastic discount factor. The modeled interest rates are indeed affine function of the pricing factors if and only if:

- the past and present values of the factors summarize the available information [said differently, the knowledge of the past and present values of the factors are enough to price at best the bonds];
- the risk-neutral dynamics of the factors is CAR. For the CAR(1) process, we denote :

$$\mathbb{E}^{\mathbb{Q}}[\exp(-uX_{t+1})|X_t] = \exp\left(-a^{\mathbb{Q}}[u]X_t - b^{\mathbb{Q}}[u]\right), u \in \mathbb{R}^N; \quad (0.0.39)$$

- the short-term interest rate is affine in the factors :

$$r(t, 1) = \alpha'(1)X_t + \beta(1).$$

The pricing formula are then computed in the following way:

$$B(t, h) = \mathbb{E}_t^{\mathbb{Q}}[B(t, 1)B(t+1, h-1)],$$

that is :

$$\begin{aligned} \exp[-h\alpha'(h)X_t - h\beta(h)] &= \mathbb{E}_t^{\mathbb{Q}}[\exp(-r(t, 1) - (h-1)\alpha'(h-1)X_{t+1} \\ &\quad - (h-1)\beta(h-1))] \\ &= \exp\left(-r(t, 1) - a^{\mathbb{Q}}[(h-1)\alpha'(h-1)]X_t \right. \\ &\quad \left. - (h-1)\beta(h-1) - b^{\mathbb{Q}}[(h-1)\alpha'(h-1)]\right). \end{aligned}$$



Thus by identification :

$$h\alpha'(h) = \alpha'(1) + a^{\mathbb{Q}}[(h-1)\alpha'(h-1)], \quad (0.0.40)$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + b^{\mathbb{Q}}[(h-1)\alpha'(h-1)].$$

When one of the dynamics (risk-neutral or physical) is CAR, the explicit specification of the stochastic discount factor as an exponential affine function [approches i) and iii)] allows to determine explicitly the link between risk-neutral and physical conditional factors distribution through their conditional Laplace transform. In this case, the other dynamics is necessarily CAR, and such as [Bertholon, Monfort, Pegoraro (2008)] :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-uX_{t+1})] = \frac{\mathbb{E}_t^{\mathbb{P}}[\exp(-[u + \gamma_1'(X_t)]X_{t+1})]}{\mathbb{E}_t^{\mathbb{P}}[\exp(-\gamma_1(X_t)X_{t+1})]}. \quad (0.0.41)$$

The pricing formula (0.0.40) stresses how the baseline term structures  $\alpha(h)$ ,  $\beta(h)$  are necessarily related to an underlying risk-neutral factors dynamics. Therefore, all affine term structure models are not necessarily consistent with no-arbitrage opportunities assumption.

As an illustration, let us consider the case of affine term structure models with Gaussian factors, i.e. for which the risk-neutral conditional distribution of the factors is Gaussian. For the sake of simplicity, let us consider one-factor term structure model [we still denote the univariate pricing factor  $X_t$ ]. Let us apply the back modeling approach, and assume the factor  $X_t$  is under  $\mathbb{Q}$  an autoregressive Gaussian process of order 1:

$$X_{t+1} = \mu_{\mathbb{Q}} + \rho_{\mathbb{Q}}X_t + \sigma_{\mathbb{Q}}\varepsilon_{\mathbb{Q},t+1},$$

where the shock  $\varepsilon_{\mathbb{Q},t}$  is a Gaussian white noise with zero mean and variance 1 under  $\mathbb{Q}$ . The conditional Laplace transform of the factor is thus :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-uX_{t+1})] = \exp\left(-u\rho_{\mathbb{Q}}X_t - u\mu_{\mathbb{Q}} + \frac{(u\sigma_{\mathbb{Q}})^2}{2}\right),$$

equivalently in our notations :

$$a^{\mathbb{Q}}[u] = u\rho_{\mathbb{Q}} \quad b^{\mathbb{Q}}[u] = u\mu_{\mathbb{Q}} - \frac{(u\sigma_{\mathbb{Q}})^2}{2},$$

which gives the recursive pricing formula :

$$h\alpha(h) = \alpha(1) + (h-1)\alpha(h-1)\rho_{\mathbb{Q}},$$

$$h\beta(h) = \beta(1) + (h-1)\beta(h-1) + \mu_{\mathbb{Q}}(h-1)\alpha(h-1) - \frac{((h-1)\alpha(h-1)\sigma_{\mathbb{Q}})^2}{2}.$$

Figure 1 below presents the corresponding baseline term structures  $\alpha(h)$ ,  $\beta(h)$ ,  $h \in \mathbb{N}^+$  for the following risk-neutral parameters :

$$\mu_{\mathbb{Q}} = 0 \quad \rho_{\mathbb{Q}} = 0.8 \quad \sigma_{\mathbb{Q}} = 0.1.$$

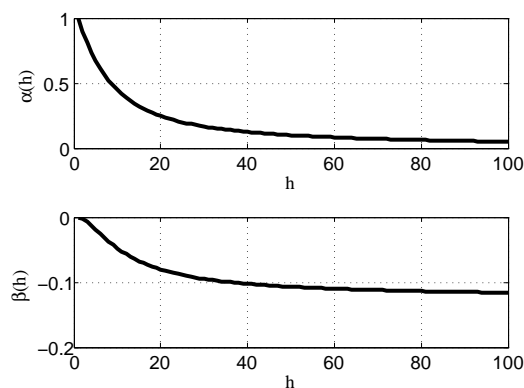


Figure 1: Stochastic  $[\alpha(h)]$  and deterministic  $[\beta(h)]$  baseline term structures for a 1-factor affine Gaussian term structure model, where  $h$  indicates the interest rates' residual maturity.

Figure 1 illustrates the properties [convergent, decreasing] of the baseline term structures  $\alpha(h)$ ,  $\beta(h)$  in this 1-factor affine Gaussian model. In particular, the influence on the rates of the stochastic factor  $X_t$  [set by the  $\alpha(h)$  function] is decreasing in the rates' maturity: ultimately, the very long-term rates are constant in this model  $[r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h) = \lim_{h \rightarrow \infty} \alpha(h)X_t + \beta(h) = \lim_{h \rightarrow \infty} \beta(h)]$ .

## TERM STRUCTURE MODELS WITH STOCHASTIC LONG-TERM RATES

Most of the arbitrage free term structure models in the literature share with this example the feature of constant very long-term rate<sup>32</sup>. Some induce explosive long-term rate, but almost none allows limiting rate  $r(t, \infty)$  to be stochastic<sup>33</sup>. Hence, in most if not all of the term structure models in the literature, the longer the rate, the less volatile it is. This feature is at odd with the observed rates' dynamics. Besides, it has strong implications for the rates extrapolated from observed term structure [this extrapolation is sometimes required for the pricing of very long-term cash-flows, such as the ones generated by some infrastructure projects, or by retirement systems]. Figure 2 presents the term structure of volatility of the US Treasury bonds' yield, for different periods<sup>34</sup>. It shows that the rates with the longest maturity are not necessarily less volatile than the shorter-term ones (the volatility of yields is indeed constant in the rates' maturity between 1995-2001, and increasing since the beginning of 2011).

The aims of the second part of this dissertation are to propose an affine term structure model with stochastic limiting rate, and to test its empirical properties on a dataset of US Treasury bonds' rates, which includes bonds with 30 years maturity. For this purpose, we focus on term structure models

---

<sup>32</sup>This properties of most of the no-arbitrage term structure models [including the above example] seems related with the no-arbitrage assumption, which constraints more the long-term rates than the short-term ones: under no-arbitrage, the limiting long-term rates  $r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h)$  cannot decrease [Dybvig, Ingersoll, Ross (1996), Hubaleck, Klein, Teichmann (2002), Kardaras and Platen (2012)].

<sup>33</sup>Yao (1998), Appendix A, for a (partial) list of term structure models and their implication for the long-term rate. See also Ingersoll, Skelton, Weil (1978) for an example of term structure model with stochastic limiting rate  $r(t, \infty)$ , whose variations are discrete. Yao (1998) conjectured a term structure model with stochastic limiting rate, without providing explicit pricing formula, nor testing the model's empirical properties.

<sup>34</sup>These descriptive statistics are built from the dataset provided by Gurkaynack, Sack, Wright (2007). We do not present statistics for the 2002-2006 period, when the US Treasury stopped the issuance of very long-term bonds, with 30 years maturity.

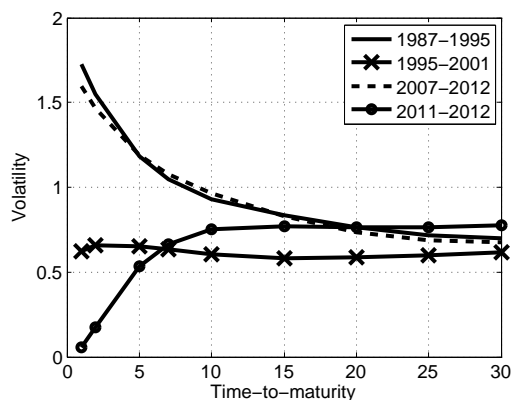


Figure 2: Term structure of volatility of Treasury bonds' yield, for different periods. The x-axis indicates the residual maturity of interest rates in years. The y-axis stands for the yields' volatility, in percent.

incorporating a level factor (whose variations affects uniformly all yields, and *a fortiori* the longest ones).

The Nelson-Siegel (1987) model belongs to this class. It specifies the zero-coupon interest rates as an affine function of 3 factors :

$$r(t, h) = X_{1,t} + \frac{1 - \exp(-\lambda h)}{\lambda h} X_{2,t} + \left( \frac{1 - \exp(-\lambda h)}{\lambda h} - \exp(-\lambda h) \right) X_{3,t}. \quad (0.0.42)$$

This model is very popular among practitioners [see the BIS report (2005) about the use of the Nelson-Siegel model in central banks], because of its ability to reproduce most of the observed term structure. Moreover, the factors loadings  $\alpha_1(h) = 1$ ,  $\alpha_2(h) = \frac{1 - \exp(-\lambda h)}{\lambda h}$ , and  $\alpha_3(h) = \frac{1 - \exp(-\lambda h)}{\lambda h} - \exp(-\lambda h)$ , which determine the impact on the whole term structure of a marginal variation in the factors, give the factors an intuitive interpretation [we present in Figure 3 the factor loadings for a given  $\lambda$  value]. In this model, the first factor ( $X_{1,t}$ ) has a "level" effect and affects uniformly the term structure, while the second and third factors ( $X_{2,t}$  et  $X_{3,t}$ ) set the dynamics of the "slope" and the "curvature" of the yield curve. The Nelson-Siegel model is particularly interesting for the modeling of interest rates with very long maturity, since the limiting rate is stochastic in this model

$[r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h)]$ . This model is however not consistent with no-arbitrage assumption [Bjork et Christensen (1999), Filipovic (1999)]. One

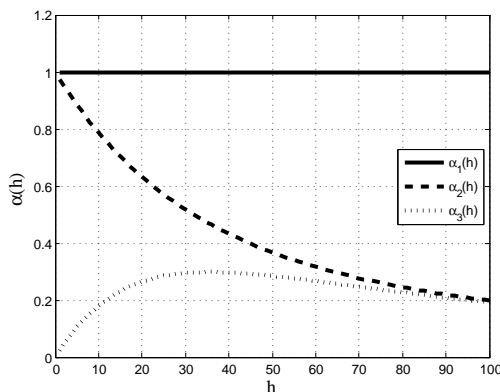


Figure 3: Baseline term structure  $\alpha(h)$  in Nelson-Siegel (1987) term structure model (0.0.42).  $h$  stands for the rates' residual maturities.

of the contribution of this thesis is to propose a no-arbitrage affine term structure model, with stochastic limiting rate, which preserves the main properties of the Nelson-Siegel model.

## SHOCKS IN FINANCIAL MODELS

Through the above presentation of the general principles for the pricing of financial assets (bonds in particular), we introduced most of the concepts used in this thesis. We show that the value an investor grants to a financial asset depends on the future monetary flows he expects, and on the circumstances (or states of the world) at the moment he would receive these flows. In practice, solving this complex problem requires models, which ultimately express asset prices as a function of a limited number of pricing factors.

The first part of this thesis focuses on the shocks, which impact these pricing factors. The concept of shock appears in the economic and financial literature in the late 19th century [Horton (1886), Giddings (1887)], and indicates a frequent and irregular perturbation of the economic system. Economic

and financial models aim to summarize the fluctuations in the economy and financial markets through a system of equations determining the joint dynamics of the main economic and financial variables. "Shocks" are seen as perturbations, which are external to the system, and feed the endogenous dynamics of main variables. The identification of past shocks has now become fundamental for the understanding of the economy [Duarte and Hoover (2011)].

We define as "shocks" the part in the factors fluctuations, which cannot be explained with respect to the *ex-ante* available information. The use of shocks thus introduces a timeline for the understanding of factors dynamics, by decomposing factors' value at each date  $t$  into parts related to present and past events. As an illustration, let us consider an univariate factor with value  $X_t$  at date  $t$ , whose (physical) dynamics is modeled as an autoregressive Gaussian process of order 1:

$$X_{t+1} = \mu + \rho X_t + \sigma \varepsilon_{t+1},$$

with  $|\rho| < 1$ , and where the shock  $\varepsilon_t$  is a Gaussian white noise with zero mean and unitary variance. Let us for instance assume  $\rho = 0.8$ ,  $\mu = 0$ ,  $\sigma = 0.1$  and  $X_0 = 1$ . Figure 4 below presents a simulated trajectory of the factor over 100 periods (between  $t = 1$ , and  $t = 100$ ), as well as the shock' realizations  $(\varepsilon_t)_{t=1, \dots, 100}$ . Each shock realization modifies the conditional distribution of the factor's future values, as illustrated in Figure 5 below. In this example, the accumulation of shocks changes the location and the scale of the conditional distribution<sup>35</sup>.

With respect to financial models, shocks determine the conditional distribution of future pricing factors, and thus the conditional distribution of any portfolio of assets. In this thesis, we explicitly relate the characteristics of shocks with the conditional distribution of any financial portfolio, and inves-

---

<sup>35</sup>In this example, the characteristic of the shocks does not change over time. It is likely this assumption is not realistic as far as financial series, which alternate period of high and low volatility, are concerned.

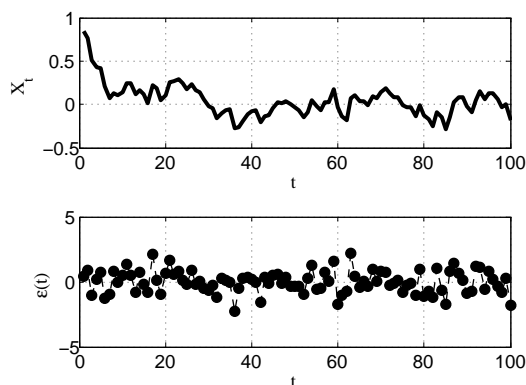


Figure 4: Simulation of an autoregressive Gaussian process of order 1 over 100 periods. The top graph present the factor's trajectory, while the shock's realizations are plotted in the bottom panel, at each date  $t$  (in the x-axis).

tigate the sensitivity of the portfolio to a change in the nature of the shocks.

## PRESENTATION OF THE CHAPTERS OF THE THESIS

In the first chapter of this dissertation, we propose a generalized definition of the shocks impacting asset prices. Our results are useful for the financial risk management. It is indeed common to measure the risk associated with the holding of a portfolio by assessing its properties in a deteriorated environment, or equivalently, its sensitivity to shocks, through stress-tests exercises. For instance, clearinghouses determine the required level of margin calls by stressing the portfolio of clearinghouse members [Pérignon, Lopez, Harris (2011)]. Besides, regulators in Europe [the European Banking Authority] or in the US [the Federal Reserve] assess for several years the sensitivity of regulated banks to several macroeconomic shocks, asking over-sensitive banks to raise their capital ratio and/or change their investment strategy.

In this first chapter, we propose a richer strategy for stress-tests, by considering a larger class of shocks, than the one commonly used. In practice, stress-test exercises consider the realization of a given shock, as if it was

deterministic. In this chapter, we propose to investigate the consequences of stochastic shocks. Our approach is thus more robust, more difficult to manipulate by the portfolio manager, and is able to handle coherently crisis phenomenon, such as a rise in the correlation of asset prices. We present the two possible representations of the shock, in terms of distribution and in terms of variables, and highlight their differences. Moreover, we derive analytically the consequences for a portfolio, of stochastic shocks on the common factors driving the prices of assets in the portfolio. Second, while most of the stress-test exercises are conducted for a given portfolio allocation, we investigate the consequences of shocks for the optimal investment strategy. The literature in finance considers several investment criterion, such as the investor's expected utility, the portfolio's (mean or median) return subject to a upper risk limit (measuring the risk by the variance [Markowitz (1952)] or by a quantile of the portfolio's return). Each of these criterion relies on the conditional distribution of the future value of the portfolio, and thus depends on the shocks affecting this distribution. The first chapter presents, for several investment criterion, the optimal response of the investor with respect to the statistical properties of the shocks. These results are particularly useful for investor looking for the consequences of some hypothetical shocks, and for the regulators, who aim to forecast the impact of shocks on the risk-taking behavior of investors. Finally, we illustrate our results, by stressing a portfolio of European sovereign bonds, and proposing a meaningful example of stochastic shock.

The second and third chapters focus on the pricing of non-defaultable zero-coupon bonds with very long time-to-maturity. This is a different issue with respect to financial risk management: while the first chapter was interested in measuring the risk of a portfolio, the last chapters focus on the risk borne by a market maker, who provides quotes for several assets (in the case we are interested in, for several bonds with different maturities). These quotes have to be consistent with the quotes of the other market makers, and also



internally consistent, i.e. such as the system of quotes announced by the market makers is arbitrage-free. To some extent, while the first chapter is concerned with the physical distribution of asset prices, these last two chapters are rather related to their risk-neutral distribution.

Moreover, we focus on a peculiar segment of the yield curve, and consider specifically bonds with very long-term residual maturity (more than ten years). The modeling of very long-term rates is a challenge for classic term structure models, which usually assumes that the volatility of the rates is decreasing in their maturity. As mentioned above, most of the no-arbitrage term structure models assume a constant limiting rate, in contradiction with empirical data, which show a volatility of very long-term rates similar to the shorter-term rates, sometimes even higher (see Figure 2). Among the rare models with stochastic limiting rate, some incorporate a level factor, whose fluctuations impact uniformly the whole yield curve and *a fortiori* the very long-term rates [see the Nelson-Siegel model presented above]. The second chapter investigates the compatibility of the level factor with the assumption of no-arbitrage. We show in this chapter that a term structure model, which incorporates a level factor, require very specific shocks to the physical and risk-neutral dynamics of the factor to be compatible with the positivity of interest rates, the finiteness of the limiting rates, and no-arbitrage restrictions.

We extend these results in the third chapter, in which we introduce a new class of affine term structure models. The risk-neutral factors process is assumed to be impacted by strictly positive white noises, whose distribution is stable<sup>36</sup>. The stable distribution encompasses all distribution remaining stable by summation [in particular the Levy, Cauchy, and Gaussian distributions]: the distribution of the sum of stable shock is also stable [see Samorodnitsky, Taqqu (1994), Nolan (2007) for a detailed presentation of the stable distribution]. Stables distributions are thus particularly useful for

---

<sup>36</sup>These shocks are thus also positive under the physical probability measure, both measure having same support, see above.

the pricing of bonds, which require the computation of Laplace transform of sum of random shocks. Factor-models with stable distribution have been introduced for the pricing of the term structure of options on US stocks [Carr, Wu (2003)], or for reproducing extreme fluctuations in interest rates [Raible (2000)]. We show in this third chapter that stable distributions are also useful for the modeling of the term structure of interest rates: they ensure the positivity of the modeled interest rates [in this sense they provide an alternative to the popular Cox, Ingersoll, Ross (1985) term structure model], and allow to introduce no-arbitrage term structure models with stochastic limiting rate. The empirical properties of this model are illustrated on a dataset of US Treasury bond prices, whose maturity can reach thirty years.



Part I

---

The Methodology of Stress-Test  
Exercises



# Shock on Variable or Shock on Distribution with Application to Stress-Tests\*

---

## 1.1 INTRODUCTION

The comparison of risks or the analysis of the effects of shocks on a risky portfolio value are problems concerning a comparison of two distributions. However, these questions are often presented in the economic and finance literature in terms of stochastic variables. This is an abuse of language. It is likely introduced to facilitate the understanding of the notion by the standard reader, but it can also imply misleading interpretations and errors in implementing the notion.

To illustrate this practice, let us consider the notion of second-order stochastic dominance. Rothschild, Stiglitz (1970), Theorem 2, have proposed three equivalent characterizations of this notion. Loosely speaking they consider two variables  $Y_0, Y_1$  with respective cumulative distributions  $F_0$  and  $F_1$ . For expository purpose, let us interpret  $Y_0$  and  $Y_1$  as the prices of two assets 0 and 1, or equivalently, as the value of two portfolios completely invested

---

\*This chapter is based on a joint work with Christian Gouriéroux. This chapter have greatly benefited from comments and suggestion by Jean-Paul Laurent, Martin Schweizer, and participants at the 2012 SoFiE conference held in Oxford, the 2012 International Risk Forum held in Paris, and at the Banque de France and CREST internal seminars. We thank J.P. Renne, B. Saes-Escorbiac, and A. Touchais for kindly providing us with the sovereign bond data.

in asset 0 and 1, respectively. The investment  $Y_0$  in asset 0 dominates the investment  $Y_1$  in asset 1 at the second-order if one of the following equivalent conditions i)-iii) is satisfied :

$$\text{i) } \int_0^K [1 - F_0(y)] dy \geq \int_0^K [1 - F_1(y)] dy, \forall K.$$

Let us consider the price, under the historical distribution<sup>37</sup>, of a European call option written on the investment  $Y_0$  with strike  $K$ , that is, the price of a derivative with payoff equal to  $Y_0 - K$ , if  $Y_0 \geq K$ , equal to 0, otherwise. We have :

$$\begin{aligned} \mathbb{E}([Y_0 - K]^+) &= \int_K^{+\infty} (y - K) f_0(y) dy \\ &= - \int_K^{+\infty} (y - K) d[1 - F_0(y)] \\ &= - [(y - K)(1 - F_0(y))]_K^{+\infty} + \int_K^{+\infty} [1 - F_0(y)] dy \\ &= \int_K^{+\infty} [1 - F_0(y)] dy \\ &= E(Y_0) - \int_0^K [1 - F_0(y)] dy, \text{ since } Y_0 > 0. \end{aligned}$$

Thus, condition i) means that, for a given strike  $K$ , a European call option contract written on the price of asset 0 is always cheaper than the similar contract written on asset 1, whenever  $Y_0$  and  $Y_1$  have the same mean. Therefore, the investment  $Y_0$  dominates  $Y_1$  under i), since the price of the insurance against a drop in investment's value is lower for asset 0.

$$\text{ii) } \mathbb{E}[U(Y_0)] \geq \mathbb{E}[U(Y_1)], \text{ for any bounded concave function } U.$$

Condition ii) means that any risk-averse agent, i.e. with concave utility function, would prefer asset 0 to asset 1.

---

<sup>37</sup>Equivalently, this is the price for a risk-neutral investor.

iii) There exists a variable  $Z$  such that :  $Y_1 = Y_0 + Z$ , with  $\mathbb{E}[Z|Y_0] = 0$ .

The first and second characterizations show clearly that the concept of stochastic dominance concerns the distributions. The third characterization seems to be of a different type. It says that we can pass from  $Y_0$  to  $Y_1$  by adding a stochastic shock with zero conditional mean. It concerns variables themselves, but on an extended space, since the comparison of the marginal distributions of  $Y_0$  and  $Y_1$  implicit in i) and ii) is now replaced by a condition, which involves the joint distribution of  $(Y_0, Y_1)$  on the product space. As seen from the proof in Rothschild, Stiglitz (1970), the third characterization can only be obtained after having constructed an artificial product space [see also Strassen (1965), Armbruster (2011)].

In Section 1.2, we consider a parametric family of cumulative distribution functions  $(F_\delta, \delta \in I)$ , and we show that they can always be considered as the marginal distributions of a family of variable  $Y_\delta$  defined by a stochastic equation  $Y_\delta = h(Y_0, \varepsilon; \delta)$ , where  $\varepsilon$  is a variable independent of  $Y_0$ . As in Rothschild, Stiglitz, the result requires the construction of an artificial product space. The “equivalent” representation of the family of distributions allows to understand why shocks can be defined in terms of either parameter, or distribution, or variable.

The effect of a small change of  $\delta$  on the distribution is usually studied by considering a Taylor expansion of distribution  $F_\delta$  with respect to  $\delta$ . In Section 1.3 we show how an equivalent expansion can be performed in terms of variables and relate the two types of expansions.

Stress-tests are subjective statements on the characteristics of asset prices’ distribution, which complement statistical risk management. First, they address model risk, by asking the risk manager to assess the consequences of distressed environment on its portfolio of assets. Stress-test exercises hence incite risk managers to be aware of the limitations of the statistical risk mod-



els they use, and perhaps rely on misspecified risk factors' distribution<sup>38</sup>.

The literature on stress-testing instill in its infancy. In the current practices, the considered distressed environment are deterministic and hand-picked, which makes the exercise highly subjective and undermines its credibility. Recent works have proposed a more systematic search of the distressed environment. Breuer and Csiszar (2012) suggest to look for the worst risk factors' distribution according to an arbitrary loss function, under a constraint on the distance between the stressed factors' distribution and the one estimated on historical data. Meucci (2012) proposes to search for the closest asset prices' distribution according to its distance with the historical one, under subjective constraints on the distressed prices' distribution.

In this chapter, we interpret stress-test exercises as the application of a shock on the risk factors. In this view, stress-tests are consistent with the existing quantitative framework, which is compulsory for their implementation and the interpretation of their results. Moreover, we consider the risk/portfolio manager's investment criterion function, which allows us to take into account her/his optimal reaction to the shock. For this purpose we extend the results of Sections 1.2 and 1.3 to some functionals of distribution  $F_{\delta}$  in Section 1.4, where we consider the effects of shocks on both the objective function of the investor and on her/his optimal portfolio allocation.

We enrich the above mentioned references by deriving carefully the transmission of the shocks on the risk factors' distribution to the asset prices' distribution. Some examples of specification of shocks are given in Section 1.5. We emphasize in this section the advantages of a joint specification of the shocks, both in terms of distribution and variable<sup>39</sup>. As an illustration, an example of stress-tests on portfolios of sovereign bonds is provided in Section 1.6. We consider sovereign bonds for different countries on period 2001-2011, and extract the underlying factors by a principal component

---

<sup>38</sup>See the "Principles for sound stress-testing practices and supervision" edited by the Basel Committee (2009).

<sup>39</sup>Breuer and Csiszar (2012), Meucci (2012) specify the shock in terms of distribution only.

analysis. The distribution of these factors on periods 2001-2007 and 2007-2011 shows a significant change due to the recent financial crisis. Then we consider period 2001-2007 as a benchmark and shock the first factor by contaminating the benchmark distribution with crisis specific distribution. We analyze the effects of this contamination on a crystallized portfolio and an optimally updated portfolio, both for shocks in distribution and in variable. Section 1.7 concludes. The technical proofs are gathered in Appendices.

## 1.2 FAMILY OF DISTRIBUTIONS OR FAMILY OF VARIABLES

The aim of this section is to relate a parametric modeling written in terms of distributions and a parametric modeling written in terms of stochastic variables. For expository purpose, we consider continuous distributions on  $\mathbb{R}$  and a scalar parameter  $\delta$  (the extension to multivariate distributions is given in Appendix 1.1 i)).

The first type of modeling defines a family of distributions  $\{F_\delta, \delta \in I\}$ , where  $I$  denotes an interval of  $\mathbb{R}$  and  $F_\delta$  the cumulative distribution function. The second modeling is based on some relationship between variables :  $Y_\delta = h(Y_0, \varepsilon; \delta)$ ,  $\delta \in I$ , say, where  $Y_0$  and  $\varepsilon$  are independent variables with  $Y_0$  following  $F_0$  and  $\varepsilon$  following the uniform distribution on  $[0, 1]$ . The questions solved in this section are the following ones :

i) Given the parametric family  $\{F_\delta, \delta \in I\}$ , is it always possible to find a variable  $\varepsilon$  and a function  $h$  such that the marginal distribution of  $Y_\delta$  is  $F_\delta$ , for any  $\delta \in I$ ?

ii) Is such a function  $h$  unique, if it exists?

### 1.2.1 COPULA

Whereas the modeling in distribution involves the one-dimensional space  $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ , the modeling in variable involves a two dimensional space  $(\mathbb{R}^2, \mathbb{B}(\mathbb{R}^2))$ . Thus, we have first to introduce such a bidimensional artificial space. Let us denote by  $(U, V)$  a pair of variables on this space with marginal distributions which are uniform on  $[0, 1]$ , and a joint c.d.f.  $C(u, v) = \mathbb{P}[U < u, V < v]$ ,  $u, v \in [0, 1]$ . The variables are usually called rank variables and function  $C$  is the copula cumulative function.

Let us first consider two values, 0 and 1, say, of parameter  $\delta$ , and the associ-

ated distributions  $F_0$  and  $F_1$ . By Sklar theorem [Sklar (1959)], the variables :

$$Y_0 = F_0^{-1}(U), Y_1 = F_1^{-1}(V), \quad (1.2.1)$$

have the joint c.d.f. :

$$\mathbb{P}[Y_0 < y_0, Y_1 < y_1] = C[F_0(y_0), F_1(y_1)]. \quad (1.2.2)$$

In particular the marginal distribution of  $Y_0$  (resp.  $Y_1$ ) is  $F_0$  (resp.  $F_1$ ).

Let us now consider the conditional c.d.f. of  $Y_1$  given  $Y_0$ . We have [see Joe (1997), p.245] :

$$F_{1|0}(y_1|Y_0) \equiv \mathbb{P}[Y_1 < y_1|Y_0] = \frac{\partial C}{\partial u}[F_0(Y_0), F_1(y_1)]. \quad (1.2.3)$$

Thus, by using the inverse transform method, the variable :

$$\varepsilon = F_{1|0}(Y_1|Y_0) = \frac{\partial C}{\partial u}[F_0(Y_0), F_1(Y_1)], \quad (1.2.4)$$

follows a uniform distribution on  $[0, 1]$  and is independent of  $Y_0$ . We deduce the expected expression in terms of variables :

$$\begin{aligned} Y_1 &= \left(F_{1|0}(\bullet|Y_0)\right)^{-1}(\varepsilon) \\ &= \left(\frac{\partial C}{\partial u}[F_0(Y_0), F_1(\bullet)]\right)^{-1}(\varepsilon) \\ &= h(Y_0, \varepsilon; 1), \text{ say.} \end{aligned} \quad (1.2.5)$$

Thus we have shown the existence of function  $h$ . Moreover, by increasing the dimension of the space, we allow for a variety of choices of copula  $C$  and then of function  $h$ . Finally, equation (1.2.4) can also be written as :

$$\varepsilon = \frac{\partial C}{\partial u}(U, V),$$

which shows how the uniform variable  $\varepsilon$  depends on the basic uniform variables  $U$  and  $V$  in a complicated nonlinear way.

### 1.2.2 EXTENSION TO FAMILIES

The result of Section 1.2.1 can be applied to any pair of parameter values  $(0, \delta)$ , associated distributions  $(F_0, F_\delta)$ , and variables  $(Y_0, Y_\delta)$ . We get :

$$\varepsilon_\delta = \frac{\partial C_\delta}{\partial u} [F_0(Y_0), F_\delta(Y_\delta)], \delta \in (0, 1),$$

and

$$Y_\delta = \left( \frac{\partial C_\delta}{\partial u} [F_0(Y_0), F_\delta(\bullet)] \right)^{-1} (\varepsilon_\delta), \delta \in (0, 1), \quad (1.2.6)$$

where the copula is now indexed by parameter  $\delta$ . All variables  $\varepsilon_\delta$  are uniformly distributed on  $[0, 1]$  and independent of  $Y_0$ . In practice the shock defined in terms of variable involves an innovation  $\varepsilon$ , which is independent of  $\delta$ . This is easily derived if we consider equality in distribution. Indeed, equation (1.2.6) implies the equality in distribution

$$Y_\delta \stackrel{d}{=} h(Y_0, \varepsilon; \delta), \delta \in [0, 1], \quad (1.2.7)$$

where

$$h(Y_0, \varepsilon; \delta) = \left( \frac{\partial C_\delta}{\partial u} [F_0(Y_0), F_\delta(\bullet)] \right)^{-1} (\varepsilon), \quad (1.2.8)$$

$\varepsilon$  is independent of  $\delta$ , of  $Y_0$ , and uniformly distributed on  $[0, 1]$ .

### 1.2.3 SHOCKS

Let us now discuss the introduction of shocks.

i) The shock can be defined by means of parameter  $\delta$ . For instance, the parameter can pass from value 0 to value  $\delta$ , say.

ii) The corresponding effect on distribution will be the change from  $F_0$  to  $F_\delta$ .

iii) Let us finally consider the variable interpretation. Equality (1.2.7) in distribution has to be replaced by an equality in terms of variables. In fact,

we can define

$$Y_\delta = h(Y_0, \varepsilon; \delta),$$

with  $h$  satisfying (1.2.8), whenever the following coherency condition is satisfied :

$$Y_0 = h(Y_0, \varepsilon; 0). \quad (1.2.9)$$

Then the shock on variables is  $Y_\delta - Y_0 = h(Y_0, \varepsilon; \delta) - Y_0$ . For a given  $Y_0$ , this is in general a stochastic shock due to the effect of the uniform stochastic variable  $\varepsilon$ .

The coherency condition (1.2.9) implies restrictions on the choice of copula in (1.2.8). Let us for instance consider the Gaussian copula with correlation parameter  $\rho(\delta)$ , and variables  $Y_0, Y_\delta$ , whose marginal distributions are  $F(y; 0), F(y; \delta)$ , respectively. We have (see Appendix 1.4 i) :

$$\varepsilon = \Phi \left( \frac{\Phi^{-1} [F(Y_\delta; \delta)] - \rho(\delta)\Phi^{-1} [F(Y_0; 0)]}{\sqrt{1 - \rho^2(\delta)}} \right),$$

where  $\Phi$  is the standard Gaussian c.d.f. Thus :

$$\Phi^{-1} [F(Y_\delta; \delta)] = \rho(\delta)\Phi^{-1} [F(Y_0; 0)] + \sqrt{1 - \rho^2(\delta)}\Phi^{-1} (\varepsilon). \quad (1.2.10)$$

Equation (1.2.10) shows that the coherency condition restricts the Gaussian copula to be such that  $\rho(0) = 1$ .

Finally note that the definition of the direction of the shock is more accurate with the specification in variable. Indeed, it specifies  $Y_\delta - Y_0$  with respect to  $Y_0$ . Thus it requires the specification of the joint distribution of  $Y_0$  and  $Y_\delta$ , whereas the specification in terms of distribution demands the unconditional distributions only. This explains why different specifications of shocks in terms of variable can lead to a same specification in terms of distribution.

### 1.3 LOCAL ANALYSIS

We have seen in Section 1.2 how to link the approaches in distributions and in variables in a global analysis of shocks. However, extreme effects can result from small shocks when the system is nonlinear. In applications to finance, these nonlinearities are due to derivatives (call options, credit derivatives) included in the portfolio as well as the nonlinear portfolio management strategies. This explains why a global analysis has to be completed by a local analysis.

The effect of a small change<sup>40</sup> in  $\delta$  is usually treated by considering appropriate Taylor expansions. These expansions can be done from the distributions themselves, or from the interpretation in terms of variables. We consider below these expansions in a neighborhood of  $\delta = 0$ .

#### 1.3.1 EXPANSION OF THE DISTRIBUTION

Let us denote  $f(y; \delta)$  the density function corresponding to  $F_\delta$ . Under standard regularity conditions, we get the Taylor expansion at order  $p$  :

$$f(y; \delta) = f(y; 0) + \sum_{j=1}^p \frac{\delta^j}{j!} \frac{\partial^j f(y; 0)}{\partial \delta^j} + o(\delta^p), \quad (1.3.1)$$

where  $o(\cdot)$  denotes a deterministic negligible term.

In particular, we get at second-order :

$$\begin{aligned} f(y; \delta) &= f(y; 0) + \delta \frac{\partial f(y; 0)}{\partial \delta} + \frac{\delta^2}{2} \frac{\partial^2 f(y; 0)}{\partial \delta^2} + o(\delta^2) \\ &= f(y; 0) \left\{ 1 + \delta \frac{\partial \log f(y; 0)}{\partial \delta} + \frac{\delta^2}{2} \left[ \frac{\partial^2 \log f(y; 0)}{\partial \delta^2} + \left( \frac{\partial \log f(y; 0)}{\partial \delta} \right)^2 \right] \right\} + o(\delta^2), \end{aligned}$$

[see Chesher (1983), (1984) for an application of this expansion for testing neglected heterogeneity].

---

<sup>40</sup>A small change in  $\delta$  implies a large change on variable  $Y_0$ , if the stochastic direction of the shock concerns extreme risks (see Section 1.5).

### 1.3.2 EXPANSION IN TERMS OF VARIABLE

Let us now consider the model :

$$Y_\delta = h(Y_0, \varepsilon; \delta).$$

We could apply a Taylor expansion at order  $p$  to get :

$$Y_\delta = Y_0 + \sum_{j=1}^p \left( \frac{\delta^j}{j!} \frac{\partial^j h(Y_0, \varepsilon; 0)}{\partial \delta^j} \right) + o_{\mathbb{P}}(\delta^p), \quad (1.3.2)$$

where  $o_{\mathbb{P}}(\cdot)$  denotes a negligible term in probability.

However, such an expansion would be difficult to interpret in terms of distributions. Instead, we consider below the approximate computation of an expectation  $\mathbb{E}[g(Y_\delta)]$ , where  $g$  is an infinitely differentiable function with compact support. We get :

$$\begin{aligned} \mathbb{E}[g(Y_\delta)] &= \mathbb{E}[g(h(Y_0, \varepsilon; \delta))] \\ &= \mathbb{E}[g(Y_0)] + \sum_{j=1}^p \left\{ \frac{\delta^j}{j!} \mathbb{E} \left[ \frac{\partial^j}{\partial \delta^j} g(h(Y_0, \varepsilon; \delta)) \right]_{\delta=0} \right\} + o(\delta^p). \end{aligned}$$

Then, we can apply Faa di Bruno's formula [see Faa di Bruno (1855), Johnson (2002), Spindler (2005)], which provides the  $j$ -th derivative of a composite function.

**Lemma 1 :** Faa di Bruno's formula

$$\frac{d^n}{dt^n} g[h(t)] = \sum_{\text{Dio}_n} \frac{n!}{k_1! k_2! \dots k_n!} g^{(k)}[h(t)] \left( \frac{h^{(1)}(t)}{1!} \right)^{k_1} \left( \frac{h^{(2)}(t)}{2!} \right)^{k_2} \dots \left( \frac{h^{(n)}(t)}{n!} \right)^{k_n},$$

where  $g^{(k)}$  denotes the  $k^{\text{th}}$  derivative of function  $g$ , and where the sum is over all nonnegative integer solutions of the Diophantine equation :  $k_1 + 2k_2 + \dots + nk_n = n$ , and where  $k = k_1 + \dots + k_n$ .



We deduce that :

$$\mathbb{E} [g(Y_\delta)] = \mathbb{E} [g(Y_0)] + \sum_{j=1}^p \left\{ \frac{\delta^j}{j!} \sum_{k=1}^j \mathbb{E} \left[ g^{(k)}(Y_0) A_{j,k}(Y_0, \varepsilon) \right] \right\} + o(\delta^p),$$

where

$$A_{j,k}(Y_0, \varepsilon) = \sum_{\text{Dio}_{j,k}} \left\{ \frac{j!}{k_1! k_2! \dots k_j!} \left( \frac{1}{1!} \frac{\partial h(Y_0, \varepsilon, 0)}{\partial \delta} \right)^{k_1} \dots \left( \frac{1}{j!} \frac{\partial^j h(Y_0, \varepsilon, 0)}{\partial \delta^j} \right)^{k_j} \right\}, \quad (1.3.3)$$

and the sum is over  $k_1, \dots, k_j$  such that  $k_1 + 2k_2 + \dots + jk_j = j$  and  $k_1 + \dots + k_j = k$ .

We can also write by the Iterated Expectation Theorem :

$$\mathbb{E} [g(Y_\delta)] = \mathbb{E} [g(Y_0)] + \sum_{j=1}^p \left\{ \frac{\delta^j}{j!} \sum_{k=1}^j \mathbb{E} \left[ g^{(k)}(Y_0) a_{j,k}(Y_0) \right] \right\} + o(\delta^p), \quad (1.3.4)$$

where

$$a_{j,k}(Y_0) = \mathbb{E} [A_{j,k}(Y_0, \varepsilon) | Y_0]. \quad (1.3.5)$$

Finally the following Lemma is proved in Appendix 1.2.

**Lemma 2 :** We have

$$\mathbb{E} \left[ g^{(k)}(Y_0) a(Y_0) \right] = (-1)^k \mathbb{E} \left[ \frac{g(Y_0)}{f(Y_0; 0)} \frac{d^k}{dy^k} [a(Y_0) f(Y_0; 0)] \right],$$

for any  $k$  and function  $a$ .

Thus, we can rewrite equation (1.3.4) as :

$$\mathbb{E} [g(Y_\delta)] = \mathbb{E} [g(Y_0)] + \sum_{j=1}^p \left\{ \frac{\delta^j}{j!} \sum_{k=1}^j (-1)^k \mathbb{E} \left[ \frac{g(Y_0)}{f(Y_0; 0)} \frac{d^k}{dy^k} [a_{j,k}(Y_0) f(Y_0; 0)] \right] \right\} + o(\delta^p). \quad (1.3.6)$$

The knowledge of the expectation  $\mathbb{E} [g(Y_\delta)]$  for all infinitely differentiable functions with compact support characterizes the distribution of  $Y_\delta$ . Thus,

by comparing expansions (1.3.1) and (1.3.6), we see how the Taylor expansion in terms of distribution can be interpreted in terms of variable.

**Proposition 1 :** We have

$$\frac{\partial^j}{\partial \delta^j} [f(y; 0)] = \sum_{k=1}^j (-1)^k \frac{\partial^k}{\partial y^k} [a_{j,k}(y) f(y; 0)],$$

where  $a_{j,k}$  is given in (1.3.5) and (1.3.3).

For instance, the first and second-order derivatives of a composite function are :

$$\frac{d}{dt} g[h(t)] = g^{(1)}[h(t)] h^{(1)}(t),$$

$$\frac{d^2}{dt^2} g[h(t)] = g^{(2)}[h(t)] (h^{(1)}(t))^2 + g^{(1)}[h(t)] h^{(2)}(t).$$

Thus we get :

$$\begin{aligned} \mathbb{E}[g(Y_\delta)] &= \mathbb{E}[g(Y_0)] - \delta \mathbb{E} \left[ g^{(1)}(Y_0) \frac{\partial h(Y_0, \varepsilon, 0)}{\partial \delta} \right] \\ &+ \frac{\delta^2}{2} \mathbb{E} \left[ g^{(2)}(Y_0) \left( \frac{\partial h(Y_0, \varepsilon, 0)}{\partial \delta} \right)^2 + g^{(1)}(Y_0) \frac{\partial^2 h(Y_0, \varepsilon, 0)}{\partial \delta^2} \right] + o(\delta^2). \end{aligned}$$

We deduce the following second-order expansion of the p.d.f :

$$\begin{aligned} f(y; \delta) &= f(y; 0) - \delta \frac{d}{dy} \left[ f(y; 0) \mathbb{E} \left[ \frac{\partial h(Y_0, \varepsilon, 0)}{\partial \delta} \middle| Y_0 = y \right] \right] \quad (1.3.7) \\ &+ \frac{\delta^2}{2} \left\{ \frac{d^2}{dy^2} \left[ f(y; 0) \mathbb{E} \left[ \left( \frac{\partial h(Y_0, \varepsilon, 0)}{\partial \delta} \right)^2 \middle| Y_0 = y \right] \right] \right. \\ &\quad \left. - \frac{d}{dy} \left[ f(y; 0) \mathbb{E} \left[ \frac{\partial^2 h(Y_0, \varepsilon, 0)}{\partial \delta^2} \middle| Y_0 = y \right] \right] \right\} \\ &+ o(\delta^2). \end{aligned}$$

The expansion in terms of variable is greatly simplified when the shock in variable is linear in  $\delta$ .

**Corollary 1 :** Let us assume that  $Y_\delta = Y_0 + \delta Z(Y_0, \varepsilon)$ , say, and denote  $\mu_p(Y_0) = \mathbb{E}[Z^p(Y_0, \varepsilon) | Y_0]$  the conditional power moments of the stochastic

shock  $Z$ . We get :

$$f(y; \delta) = f(y; 0) + \sum_{j=1}^p \left[ \frac{\delta^j}{j!} (-1)^j \frac{d^j}{dy^j} [f(y; 0)\mu_j(y)] \right] + o(\delta^p).$$

This specific expansion has been first derived in the literature by Martin, Wilde (2002), Theorem C, based on the analysis of the moment generating function of variable  $Y_\delta$ .

For instance, let us assume that  $Z(Y_0, \varepsilon) = a(Y_0)U$ , with  $U = \Phi^{-1}(\varepsilon)$  and  $\Phi$  is the c.d.f. of the standard normal. Then, we get :

$$\mu_p(Y_0) = a^p(Y_0)\mathbb{E}(U^p) = a^p(Y_0)2^{-p/2} \frac{p!}{(p/2)!}, \text{ for } p = 2n, n \in \mathbb{N} - \{0\}, 0, \text{ otherwise.}$$

We deduce that :

$$f(y; \delta) = f(y; 0) + \sum_{j=1}^p \left[ \frac{\delta^{2j}}{2^j(j!)} \frac{d^{2j}}{dy^{2j}} [f(y; 0)a^{2j}(y)] \right] + o(\delta^{2p}).$$

The expansion of the p.d.f for a shock in terms of variable has been derived above in an indirect way. In special cases, it is possible to get it in a direct way, but the computation is rather cumbersome (see Appendix 1.3).

## 1.4 SHOCK ON RISK MANAGEMENT

In this section, we apply the previous characterizations of shocks to risk management problems, and derive the closed-form expressions of the local effects of a shock, both in terms of distribution and variable.

### 1.4.1 RISK MANAGEMENT

Let us consider a given investor, who allocates his/her wealth  $W_t$  at time  $t$  among  $N$  risky assets and 1 risk-free asset. The vector  $P_t$  collects the prices at  $t$  of the risky assets, while the risk-free asset worths 1 at  $t$  and pays  $1 + r_{f,t}$  to the investor in the following period, where  $r_{f,t}$  denotes the risk-free rate. The investor's budget constraint at date  $t$  is :

$$W_t = \alpha_f + \alpha' P_t, \quad (1.4.1)$$

where  $\alpha_f$  is the amount invested in the risk-free asset and  $\alpha$  is the allocation in the risky assets. The portfolio value at the next period is :

$$\begin{aligned} W_{t+1} &= \alpha_f (1 + r_{f,t}) + \alpha' P_{t+1} \\ &= W_t (1 + r_{f,t}) + \alpha' [P_{t+1} - P_t (1 + r_{f,t})] \\ &= \tilde{W}_t + \alpha' Y_{t+1}, \end{aligned} \quad (1.4.2)$$

where  $Y_{t+1}$  is the vector of excess gains on the period  $[t, t + 1]$ , and  $\tilde{W}_t$  is the future value of a portfolio entirely invested in the risk-free asset.

Let us consider a family of distributions  $F_\delta$  for the excess gains on asset prices  $Y_{t+1}$ , and denote  $f_\delta = f(y_{t+1}; \delta)$  the corresponding density function. The results can be written in terms of either distribution, or variable. For this latter approach, we have  $Y_{t+1} = h(Y_0, \varepsilon; \delta)$ .

We assume that the investor chooses his/her optimal asset allocation with respect to a general criterion  $G$  on the portfolio allocation  $\alpha$  and on the distribution  $f_\delta$  of excess gains  $Y_{t+1}$ . We denote  $\alpha^*(f_\delta)$  the optimal allocation,

corresponding to distribution  $f_\delta$  :

$$\alpha^*(f_\delta) = \arg \max_{\alpha} G(\alpha, f_\delta). \quad (1.4.3)$$

Several criteria have been considered in the literature. Some are written below in terms of both distribution and variable.

**i) Expected utility approach :**

*Distribution based approach :*

$$G(\alpha, f_\delta) = \mathbb{E}_\delta \left( U \left[ \tilde{W}_t + \alpha' Y_{t+1} \right] \right),$$

where the utility function  $U[\bullet]$  is increasing and concave in investor's wealth.

*Variable based approach :*

$$G(\alpha, f_\delta) = \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta) \right] \right).$$

**ii) Mean-Variance approach**

*Distribution based approach :*

$$G(\alpha, f_\delta) = \mathbb{E}_\delta(\alpha' Y_{t+1}) - \frac{\gamma}{2} \mathbb{V}_\delta(\alpha' Y_{t+1}),$$

where  $\mathbb{E}_\delta(\bullet)$  and  $\mathbb{V}_\delta(\bullet)$  are the expectation and variance operators with respect to distribution  $f_\delta$ , and  $\gamma$  has the interpretation of an absolute risk-aversion..

*Variable based approach :*

$$G(\alpha, f_\delta) = \mathbb{E}[\alpha' h(Y_0, \varepsilon; \delta)] - \frac{\gamma}{2} \mathbb{V}[\alpha' h(Y_0, \varepsilon; \delta)],$$

where the expectation and variance operators are taken with respect to both random variables  $Y_0$  and  $\varepsilon$ .

**iii) Mean-Expected Shortfall approach [mean-ES thereafter] :**

*Distribution based approach :*

$$G(\alpha, f_\delta) = \mathbb{E}_\delta(\alpha' Y_{t+1}) - \zeta \mathbb{E}_\delta^2(\alpha' Y_{t+1} | \alpha' Y_{t+1} < Q_{q,\delta}(\alpha' Y_{t+1})),$$

where  $Q_{q,\delta}$  is the quantile function of  $\alpha'Y_{t+1}$  with respect to distribution  $f_\delta$ , at risk level  $q$ .

*Variable based approach :*

$$G(\alpha, f_\delta) = \mathbb{E}[\alpha'h(Y_0, \varepsilon; \delta)] - \zeta \mathbb{E}^2[\alpha'h(Y_0, \varepsilon; \delta) | \alpha'h(Y_0, \varepsilon; \delta) < Q_q(\alpha'h(Y_0, \varepsilon; \delta))],$$

where  $Q_q$  is the quantile function of  $\alpha'h(Y_0, \varepsilon; \delta)$ .

The mean-ES approach is usually described by means of a constrained optimization problem :

$$\begin{aligned} & \max_{\alpha} \mathbb{E}_\delta(\alpha'Y_{t+1}) \\ \text{s.t. } & \mathbb{E}_\delta[\alpha'Y_{t+1} | \alpha'Y_{t+1} < Q_{q,\delta}(\alpha'Y_{t+1})] < ES_0, \end{aligned}$$

where  $ES_0$  is a risk threshold. This problem is equivalent to :

$$\begin{aligned} & \max_{\alpha} a[\mathbb{E}_\delta(\alpha'Y_{t+1})] \\ \text{s.t. } & b[\mathbb{E}_\delta(\alpha'Y_{t+1} | \alpha'Y_{t+1} < Q_{q,\delta}(\alpha'Y_{t+1}))] < b[ES_0], \end{aligned}$$

where  $a$  and  $b$  are two increasing functions. Therefore, there exist a multiplicity of Lagrangean associated with this constrained optimization, that are :

$$a[\mathbb{E}_\delta(\alpha'Y_{t+1})] - \zeta b[\mathbb{E}_\delta(\alpha'Y_{t+1} | \alpha'Y_{t+1} < Q_{q,\delta}(\alpha'Y_{t+1}))],$$

where  $\zeta$  denotes the Lagrange multiplier. As seen later on, we have chosen  $a = \text{identity}$ ,  $b(y) = y^2$  to get a Lagrangean, which is a strictly concave function of the portfolio allocation.

#### 1.4.2 EXPANSION IN TERMS OF DISTRIBUTION

We consider below the expansion of the optimal allocation and of the criterion function in terms of distribution.

#### 1.4.2.1 Expansion of the criterion function

As a functional of  $f$ ,  $G(\alpha, f)$  admits under suitable conditions a first-order Taylor expansion :

$$G(\alpha, f_\delta) = G(\alpha, f_0) + \int \frac{\partial G(\alpha, f(y; 0))}{\partial f} df_\delta(y) + o(df_\delta),$$

where  $\frac{\partial G(\alpha, f(\bullet; 0))}{\partial f}$  is the functional (Hadamard) derivative<sup>41</sup> of  $G(\alpha, \bullet)$  at  $f_0$ .

#### 1.4.2.2 Expansion of the optimal allocation

The optimal allocation satisfies the first-order condition :  $\frac{\partial G(\alpha^*(f_\delta), f_\delta)}{\partial \alpha} = 0$ . This is an implicit condition, which can be expanded to derive the local behavior of the optimal allocation. We get :

$$\frac{\partial^2 G(\alpha^*(f_0), f_0)}{\partial \alpha \partial \alpha'} d\alpha^*(f_\delta) + \int \frac{\partial^2 G(\alpha^*(f_0), f(y; 0))}{\partial \alpha \partial f} df_\delta(y) + o(df_\delta) = 0.$$

Therefore,

$$d\alpha^*(f_\delta) = \left[ -\frac{\partial^2 G(\alpha^*(f_0), f_0)}{\partial \alpha \partial \alpha'} \right]^{-1} \int \frac{\partial^2 G(\alpha^*(f_0), f(y; 0))}{\partial \alpha \partial f} df_\delta(y) + o(df_\delta). \quad (1.4.4)$$

---

<sup>41</sup>Let us consider a functional  $h \rightarrow \Phi(h)$  defined on a Banach space  $\mathcal{H}$  with norm  $\|\cdot\|_{\mathcal{H}}$ , and taking its values in  $(\mathbb{R}, \|\cdot\|)$ . This functional is Hadamard differentiable if and only if for any vectors  $h_0, h^*$  in  $\mathcal{H}$ , there exists a real  $\Phi^{(1)}(h_0, h^*)$  such that :

$$\lim_{\delta \rightarrow 0} \left\| \frac{\Phi(h_0 + \delta h_\delta^*) - \Phi(h_0)}{\delta} - \Phi^{(1)}(h_0, h^*) \right\| = 0,$$

for any sequence  $h_\delta^*$  such that  $\lim_{\delta \rightarrow 0} \|h_\delta^* - h^*\|_{\mathcal{H}} = 0$ .

$\Phi^{(1)}(h_0, h^*)$  is called the Hadamard derivative of functional  $\Phi$  at  $h_0$  in the direction  $h^*$  [see e.g. Van der Vaart (1998)].

When a functional  $\Phi$  is Hadamard differentiable, we get a first-order Taylor expansion of the type :

$$\Phi(h_0 + dh) = \Phi(h_0) + \left\langle \frac{\partial \Phi(h_0)}{\partial h}, dh \right\rangle + o(\|dh\|_{\mathcal{H}}),$$

where  $\left\langle \frac{\partial \Phi(h_0)}{\partial h}, \cdot \right\rangle$  is a linear form on  $\mathcal{H}$ . This linear form is characterized by the Hadamard derivative  $h^* \rightarrow \Phi^{(1)}(h_0, \cdot)$ .

### 1.4.2.3 Expansion of the optimal value of the criterion

Finally, the optimal value of the criterion is :

$$G^*(f_\delta) = G(\alpha^*(f_\delta), f_\delta).$$

Its expansion is :

$$\begin{aligned} G^*(f_\delta) &= G(\alpha^*(f_0), f_0) + \frac{\partial G(\alpha^*(f_0), f_0)}{\partial \alpha} d\alpha^*(f_\delta) + \int \frac{\partial G(\alpha^*(f_0), f(y; 0))}{\partial f} df_\delta(y) \\ &\quad + o(df_\delta) \\ &= G(\alpha^*(f_0), f_0) + \int \frac{\partial G(\alpha^*(f_0), f(y; 0))}{\partial f} df_\delta(y) + o(df_\delta), \end{aligned} \quad (1.4.5)$$

since  $\frac{\partial G(\alpha^*(f_0), f_0)}{\partial \alpha} = 0$  at the optimum, which is the envelop theorem [see e.g. Sydsaeter, Hammond (2008)].

### 1.4.2.4 Examples

As an illustration, let us consider the objective functions presented in Section 1.4.1. For each of them, we provide below their relevant derivatives :

#### i) Expected utility approach :

$$\frac{\partial G(\alpha, f(y; 0))}{\partial f} = U[\tilde{W}_t + \alpha'y]$$

$$\frac{\partial^2 G(\alpha, f(y; 0))}{\partial f \partial \alpha} = U^{(1)}[\tilde{W}_t + \alpha'y]y, \quad \frac{\partial^2 G(\alpha^*(f_0), f_0)}{\partial \alpha \partial \alpha'} = \mathbb{E}_0 \left( U^{(2)}[\tilde{W}_t + \alpha^*(f_0)Y_{t+1}]Y_{t+1}Y'_{t+1} \right).$$

where  $U^{(k)}[\bullet]$  is the  $k$ -th order derivative of  $U[\bullet]$ .

#### ii) Mean-Variance approach :

$$\frac{\partial G(\alpha, f(y; 0))}{\partial f} = \alpha'y \left\{ 1 - \frac{\gamma}{2} [\alpha'y - 2\mathbb{E}_0(\alpha'Y_{t+1})] \right\},$$

$$\frac{\partial^2 G(\alpha, f(y; 0))}{\partial f \partial \alpha} = y - \gamma \{ \alpha'y [y + \mathbb{E}_0(Y_{t+1})] - y\mathbb{E}_0(\alpha'Y_{t+1}) \}, \quad \frac{\partial^2 G(\alpha^*(f_0), f_0)}{\partial \alpha \partial \alpha'} = -\gamma \mathbb{V}_0[Y_{t+1}].$$

#### iii) Mean-ES approach :

The computation of the relevant derivatives for the mean-ES approach requires the first and second-order derivatives of the expected shortfall :

$$ES_{q, \delta, \alpha} = \mathbb{E}_\delta(\alpha'Y_{t+1} | \alpha'Y_{t+1} < Q_{q, \delta}(\alpha'Y_{t+1})).$$



They have been derived in Scaillet (2004), and Bertsimas, Lauprete, Samarov (2004) following the approach used for the derivatives of the VaR [Gourieroux, Laurent, Scaillet (2000)]. We have :

$$\begin{aligned}\frac{\partial ES_{q,\delta,\alpha}}{\partial \alpha} &= \mathbb{E}_\delta (Y_{t+1} | \alpha' Y_{t+1} < Q_{q,\delta}(\alpha' Y_{t+1})), \\ \frac{\partial^2 ES_{q,\delta,\alpha}}{\partial \alpha \partial \alpha'} &= \frac{1}{q} f_{\alpha' Y_{t+1}, \delta} (Q_{q,\delta}(\alpha' Y_{t+1})) \mathbb{V}_\delta (Y_{t+1} | \alpha' Y_{t+1} = Q_{q,\delta}(\alpha' Y_{t+1})),\end{aligned}$$

where  $f_{\alpha' Y_{t+1}, \delta}$  denotes the probability density function of  $\alpha' Y_{t+1}$ . In particular, the Hessian of the expected shortfall is a symmetric positive semi-definite matrix, with a kernel generated by the vector  $\alpha$  of portfolio allocation.

We deduce :

$$\begin{aligned}\frac{\partial G(\alpha, f(y; 0))}{\partial \alpha} &= \mathbb{E}_0 (Y_{t+1}) - 2\zeta ES_{q,0,\alpha} \frac{\partial ES_{q,0,\alpha}}{\partial \alpha}, \\ \frac{\partial^2 G(\alpha, f(y; 0))}{\partial \alpha \partial \alpha'} &= -2\zeta \left\{ \frac{\partial ES_{q,0,\alpha}}{\partial \alpha} \frac{\partial ES_{q,0,\alpha}}{\partial \alpha'} + ES_{q,0,\alpha} \frac{\partial^2 ES_{q,0,\alpha}}{\partial \alpha \partial \alpha'} \right\},\end{aligned}$$

where the Hessian of the objective function is now invertible. Furthermore :

$$\begin{aligned}\frac{\partial G(\alpha, f(y; 0))}{\partial f} &= \alpha' y - 2\frac{\zeta}{q} ES_{q,0,\alpha} \mathbb{1}_{\alpha' y \leq Q_{q,0}(\alpha' Y_{t+1})} (\alpha' y - Q_{q,0}(\alpha' Y_{t+1})), \\ \frac{\partial^2 G(\alpha, f(y; 0))}{\partial f \partial \alpha} &= y - 2\frac{\zeta}{q} \frac{\partial ES_{q,0,\alpha}}{\partial \alpha} \mathbb{1}_{\alpha' y \leq Q_{q,0}(\alpha' Y_{t+1})} (\alpha' y - Q_{q,0}(\alpha' Y_{t+1})) \\ &\quad - 2\frac{\zeta}{q} ES_{q,0,\alpha} \mathbb{1}_{\alpha' y \leq Q_{q,0}(\alpha' Y_{t+1})} (\alpha \alpha')^{-1} \alpha (\alpha' y - Q_{q,0}(\alpha' Y_{t+1})).\end{aligned}$$

### 1.4.3 EXPANSION IN TERMS OF VARIABLE

The main difficulty in the variable based approach is to find a specification of the criterion function that includes the main examples encountered in the literature. We consider below a criterion equal to an optimum of expected utilities :

$$\begin{aligned}G(\alpha, f_\delta) &= \text{opt}_\theta \mathbb{E}_\delta (U [\tilde{W}_t + \alpha' Y_{t+1}; \theta]) \\ &= \text{opt}_\theta \mathbb{E} (U [\tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta); \theta]),\end{aligned}\tag{1.4.6}$$

with  $h(Y_0, \varepsilon, 0) = Y_0$ .

A more general criterion would be :

$$G(\alpha, f_\delta) = a \left\{ \mathbb{E} \left( U_0 \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta) \right] \right) \right\} - b \left\{ \underset{\theta}{\text{opt}} \mathbb{E} \left( U_1 \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta); \theta \right] \right) \right\}, \quad (1.4.7)$$

where  $a$  and  $b$  are increasing functions. This latter criterion provides a trade off between two types of expected utility criteria. Then :

i) The expected utility criterion is obtained, when no parameter  $\theta$  is introduced.

ii) The mean-variance criterion can be written as [see Schweizer, p.1 (2010)] :

$$\begin{aligned} G(\alpha, f_\delta) &= \mathbb{E}_\delta(\alpha' Y_{t+1}) - \frac{\gamma}{2} \mathbb{V}_\delta(\alpha' Y_{t+1}) \\ &= \mathbb{E}_\delta(\alpha' Y_{t+1}) - \frac{\gamma}{2} \mathbb{E}_\delta(\alpha' Y_{t+1} - \mathbb{E}_\delta(\alpha' Y_{t+1}))^2 \\ &= \max_{\theta} \mathbb{E}_\delta \left( \alpha' Y_{t+1} - \frac{\gamma}{2} [\alpha' Y_{t+1} - \theta]^2 \right), \end{aligned} \quad (1.4.8)$$

that is, an optimum of expected quadratic utility functions.

iii) The general criterion in (1.4.6) allows for robust management strategies, in which the preferences are represented by a fixed utility function  $u$ , while different scenarios are considered by the investor. These scenarios can be introduced by means of a parametric family  $g(y_0, \varepsilon; \theta)$  of possible distributions of the pair  $(Y_0, \varepsilon)$ . Then the objective function becomes :

$$\begin{aligned} G(\alpha, f_\delta) &= \min_{\theta} \mathbb{E}_\theta \left( u \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta) \right] \right) \quad (1.4.9) \\ &= \min_{\theta} \int \int u \left[ \tilde{W}_t + \alpha' h(y_0, \varepsilon; \delta) \right] g(y_0, \varepsilon; \theta) dy_0 d\varepsilon \\ &= \min_{\theta} \int \int u \left[ \tilde{W}_t + \alpha' h(y_0, \varepsilon; \delta) \right] \frac{g(y_0, \varepsilon; \theta)}{g(y_0, \varepsilon; 0)} g(y_0, \varepsilon; 0) dy_0 d\varepsilon \\ &= \min_{\theta} \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta); \theta \right] \right), \end{aligned}$$

where  $U \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta); \theta \right] = u \left[ \tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta) \right] \frac{g(Y_0, \varepsilon; \theta)}{g(Y_0, \varepsilon; 0)}$ .

Thus the manager follows a maximin procedure by maximizing with respect to the portfolio allocation the expected utility associated with the least favorable scenario. This interpretation in terms of scenario is in the spirit of

the representation of coherent risk measures derived in Artzner, Delbaen, Eber and Heath (1999), Proposition 4.1.

In this framework,  $\theta$  stands for a given scenario considered by the investor, whereas  $\delta$  indexes the shock. A shock on the distribution (or on the variable) modifies the distribution of excess gains, whatever the scenario chosen by the investor. In particular, the robustness of an optimal portfolio to the least favorable scenario says nothing about the sensitivity of the optimal allocation, or of the optimized criterion to a perturbation on the true distribution of excess gains.

iv) Let us finally consider the mean-ES criterion. We have [see Koenker, p.289 (2005)] :

$$Q_{q,\delta}(\alpha'Y_{t+1}) = \arg \min_{\theta} \mathbb{E}_{\delta} \left( (\alpha'Y_{t+1} - \theta) (q - \mathbf{1}_{\alpha'Y_{t+1} < \theta}) \right),$$

where  $\mathbf{1}_{\bullet}$  is the indicator function, and :

$$\begin{aligned} & \min_{\theta} \mathbb{E}_{\delta} \left( (\alpha'Y_{t+1} - \theta) (q - \mathbf{1}_{\alpha'Y_{t+1} < \theta}) \right) \\ &= \mathbb{E}_{\delta} \left( (\alpha'Y_{t+1} - Q_{q,\delta}(\alpha'Y_{t+1})) (q - \mathbf{1}_{\alpha'Y_{t+1} < Q_{q,\delta}(\alpha'Y_{t+1})}) \right) \\ &= q\mathbb{E}_{\delta}(\alpha'Y_{t+1}) - qES_{q,\delta,\alpha} \end{aligned}$$

We deduce that :

$$\begin{aligned} G(\alpha, f_{\delta}) &= \mathbb{E}_{\delta}(\alpha'Y_{t+1}) - \zeta \mathbb{E}_{\delta}^2(\alpha'Y_{t+1} | \alpha'Y_{t+1} < Q_{q,\delta}(\alpha'Y_{t+1})) & (1.4.10) \\ &= \mathbb{E}_{\delta}(\alpha'Y_{t+1}) - \zeta \left\{ \min_{\theta} \left( -\frac{1}{q} \mathbb{E}_{\delta} \left[ (\alpha'Y_{t+1} - \theta)(q - \mathbf{1}_{\alpha'Y_{t+1} < \theta}) \right] + \mathbb{E}_{\delta}(\alpha'Y_{t+1}) \right) \right\}^2. \end{aligned}$$

#### 1.4.3.1 Expansion of the criterion function

Let us focus on the criterion function given in (1.4.6) and expand it with respect to  $\delta$  in a neighborhood of  $\delta = 0$ . The case of the mean-ES is treated in Appendix 1.5. We get :

$$\begin{aligned} G(\alpha, f_{\delta}) &= \operatorname{opt}_{\theta} \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha' (Y_0 + dh(Y_0, \varepsilon; \delta)); \theta \right] \right) \\ &\cong \operatorname{opt}_{\theta} \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha' Y_0 \right] + \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha' Y_0; \theta \right] \alpha' dh(Y_0, \varepsilon; \delta) \right) \end{aligned}$$

where  $dh(Y_0, \varepsilon; \delta) = \frac{\partial h}{\partial \delta}(Y_0, \varepsilon, 0)\delta$ .

The first-order condition providing the optimal value of  $\theta$ , denoted  $\theta^*(\alpha, f_\delta)$ , is :

$$\mathbb{E} \left( \frac{\partial U}{\partial \theta} [\tilde{W}_t + \alpha' h(Y_0, \varepsilon; \delta); \theta^*(\alpha, f_\delta)] \right) = 0 \quad (1.4.12)$$

We deduce the expansion of the first-order condition :

$$\begin{aligned} & \mathbb{E} \left( \frac{\partial^2 U}{\partial \theta \partial W} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \alpha' dh(Y_0, \varepsilon; \delta) \right) \\ & + \mathbb{E} \left( \frac{\partial^2 U}{\partial \theta \partial \theta'} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) d\theta^*(\alpha, f_\delta) \cong 0, \end{aligned}$$

which is equivalent to :

$$\begin{aligned} d\theta^*(\alpha, f_\delta) &= - \left\{ \mathbb{E} \left( \frac{\partial^2 U}{\partial \theta \partial \theta'} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) \right\}^{-1} \\ & \mathbb{E} \left( \frac{\partial^2 U}{\partial \theta \partial W} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \alpha' dh(Y_0, \varepsilon; \delta) \right) \\ & + o(\delta) \end{aligned} \quad (1.4.13)$$

Finally, we deduce :

$$\begin{aligned} G(\alpha, f_\delta) &= \mathbb{E} \left( U [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_\delta)] \right) \\ & + \mathbb{E} \left( \frac{\partial U}{\partial W} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_\delta)] \alpha' dh(Y_0, \varepsilon; \delta) \right) + o(\delta) \\ &= \mathbb{E} \left( U [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) + \mathbb{E} \left( \frac{\partial U}{\partial \theta'} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) d\theta^*(\alpha, f_\delta) \\ & + \mathbb{E} \left( \frac{\partial U}{\partial W} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \alpha' dh(Y_0, \varepsilon; \delta) \right) + o(\delta), \end{aligned}$$

that is,

$$\begin{aligned} G(\alpha, f_\delta) &= \mathbb{E} \left( U [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) \\ & + \mathbb{E} \left( \frac{\partial U}{\partial W} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \alpha' dh(Y_0, \varepsilon; \delta) \right) \\ & + o(\delta), \end{aligned} \quad (1.4.14)$$

since  $\mathbb{E} \left( \frac{\partial U}{\partial \theta} [\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)] \right) = 0$  by the optimality of  $\theta^*(\alpha, f_0)$ . This latter simplification is just the envelop theorem.

### 1.4.3.2 Expansion of the optimal allocation

In the variable based approach, the optimal allocation  $\alpha^*(f_\delta)$  is defined as :

$$\alpha^*(f_\delta) = \arg \max_{\alpha} \operatorname{opt}_{\theta} \mathbb{E}_{\delta} \left( U \left[ \tilde{W}_t + \alpha' Y_{t+1}; \theta \right] \right),$$

or :

$$\alpha^*(f_\delta) = \arg \max_{\alpha} \mathbb{E}_{\delta} \left( U \left[ \tilde{W}_t + \alpha' Y_{t+1}; \theta^*(\alpha, f_\delta) \right] \right).$$

The first-order condition is :

$$\mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha^{*'}(f_\delta); \theta^*(\alpha^*(f_\delta), f_\delta) \right] h(Y_0, \varepsilon; \delta) \right) = 0,$$

by the envelop theorem applied to  $\theta^*(\alpha^*(f_\delta), f_\delta)$ .

We deduce the expansion of the first-order condition :

$$\begin{aligned} 0 &\cong \mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 Y_0' \right) d\alpha^*(f_\delta) \\ &+ \mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 \alpha^{*'}(f_0) dh(Y_0, \varepsilon; \delta) \right) \\ &+ \mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha^{*'}(f_0); \theta^*(\alpha^*(f_\delta), f_\delta) \right] dh(Y_0, \varepsilon; \delta) \right) \\ &+ \mathbb{E} \left( \frac{\partial^2 U}{\partial W \partial \theta'} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 \left[ \frac{\partial \theta^*(\alpha^*(f_\delta), f_\delta)}{\partial \alpha'} d\alpha^*(f_\delta) + d\theta_f^*(\alpha^*(f_\delta), f_\delta) \right] \right) \end{aligned}$$

where  $d\theta_f^*(\alpha^*(f_\delta), f_\delta)$  is the partial Hadamard differential of  $\theta^*(\alpha, f)$  with respect to  $f$ .

As a consequence :

$$\begin{aligned} d\alpha^*(f_\delta) &= - \left\{ \mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 Y_0' \right) \right. \\ &\quad \left. + \mathbb{E} \left( \frac{\partial^2 U}{\partial W \partial \theta'} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 \frac{\partial \theta^*(\alpha^*(f_0), f_0)}{\partial \alpha'} \right) \right\}^{-1} \\ &\quad \left\{ \mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 \alpha^{*'}(f_0) dh(Y_0, \varepsilon; \delta) \right) \right. \\ &\quad \left. + \mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] dh(Y_0, \varepsilon; \delta) \right) \right. \\ &\quad \left. + \mathbb{E} \left( \frac{\partial^2 U}{\partial W \partial \theta'} \left[ \tilde{W}_t + \alpha^{*'}(f_0) Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 d\theta_f^*(\alpha^*(f_\delta), f_\delta) \right) \right\} + o(\delta) \end{aligned} \tag{1.4.15}$$

### 1.4.3.3 Expansion of the optimal value of the criterion

Finally, the optimal value of the criterion is :

$$\begin{aligned} G(\alpha^*(f_\delta), f_\delta) &= \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha^{*'}(f_0)Y_0; \theta^*(\alpha^*(f_0), f_0) \right] \right) \\ &+ \mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha^{*'}(f_0)Y_0; \theta^*(\alpha^*(f_0), f_0) \right] \alpha^{*'}(f_0) dh(Y_0, \varepsilon; \delta) \right) \\ &+ \mathbb{E} \left( \frac{\partial U}{\partial \theta'} \left[ \tilde{W}_t + \alpha^{*'}(f_0)Y_0; \theta^*(\alpha^*(f_0), f_0) \right] d\theta_f^*(\alpha^*(f_\delta), f_\delta) \right) + o(\delta), \end{aligned} \quad (1.4.16)$$

since  $\mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha^{*'}(f_0)Y_0; \theta^*(\alpha^*(f_0), f_0) \right] Y_0 \right)$  and  $\mathbb{E} \left( \frac{\partial U}{\partial \theta'} \left[ \tilde{W}_t + \alpha^{*'}(f_0)Y_0; \theta^*(\alpha^*(f_0), f_0) \right] \right)$  are both equal to zero, due to the optimality of  $\alpha^*(f_0)$  and  $\theta^*(\alpha^*(f_0), f_0)$ .

### 1.4.3.4 Examples

Let us now compute the relevant derivatives for the criteria of Section 1.4.1.

For the sake of exposition, we will denote in the following examples :

$$U_\delta^* = U \left[ \tilde{W}_t + \alpha^{*'}(f_\delta)h(Y_0, \varepsilon; \delta); \theta^*(\alpha^*(f_\delta), f_\delta) \right], \quad \alpha_\delta^* = \alpha^*(f_\delta), \quad \theta_\delta^* = \theta^*(\alpha^*(f_\delta), f_\delta).$$

#### i) Expected utility approach :

The expected utility criterion does not require any parameter  $\theta$ . Thus :

$$G(\alpha^*(f_\delta), f_\delta) = \mathbb{E} \left( U \left[ \tilde{W}_t + \alpha_0^{*'} Y_0 \right] \right) + \mathbb{E} \left( \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha_0^{*'} Y_0 \right] \alpha_0^{*'} dh(Y_0, \varepsilon; \delta) \right) + o(\delta)$$

$$\begin{aligned} d\alpha^*(f_\delta) &= - \left\{ \mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha_0^{*'} Y_0 \right] Y_0 Y_0' \right) \right\}^{-1} \\ &\mathbb{E} \left( \frac{\partial^2 U}{\partial W^2} \left[ \tilde{W}_t + \alpha_0^{*'} Y_0 \right] Y_0 \alpha_0^{*'} dh(Y_0, \varepsilon; \delta) + \frac{\partial U}{\partial W} \left[ \tilde{W}_t + \alpha_0^{*'} Y_0 \right] dh(Y_0, \varepsilon; \delta) \right) \\ &+ o(\delta). \end{aligned}$$

#### ii) Mean-Variance approach :

The relevant derivatives in the mean-variance approach are :

$$\mathbb{E} \left( \frac{\partial U_0^*}{\partial W} \alpha^*(f_0) dh(Y_0, \varepsilon; \delta) \right) = \mathbb{E} \left( \left[ 1 - \gamma \alpha_0^{*'} Y_0 \right] \alpha_0^{*'} dh(Y_0, \varepsilon; \delta) \right) + \gamma \mathbb{E} \left( \alpha_0^{*'} Y_0 \right) \mathbb{E} \left( \alpha_0^{*'} dh(Y_0, \varepsilon; \delta) \right)$$

$$\mathbb{E} \left( \frac{\partial U_0^*}{\partial \theta'} d\theta_f^*(\alpha^*(f_\delta), f_\delta) \right) = \gamma \mathbb{E} \left( \left[ \alpha_0^{*'} Y_0 - \mathbb{E} \left( \alpha_0^{*'} Y_0 \right) \right] \mathbb{E} \left( \alpha_0^{*'} dh(Y_0, \varepsilon; \delta) \right) \right) = 0.$$

Regarding the expansion of the optimal allocation, we get :

$$\mathbb{E} \left( \frac{\partial^2 U_0^*}{\partial W^2} Y_0 Y_0' \right) + \mathbb{E} \left( \frac{\partial^2 U_0^*}{\partial W \partial \theta'} Y_0 \frac{\partial \theta_0^*}{\partial \alpha'} \right) = -\gamma \mathbb{V}(Y_0),$$

$$\begin{aligned}\mathbb{E}\left(\frac{\partial^2 U_0^*}{\partial W^2} Y_0 \alpha_0^{*'} dh(Y_0, \varepsilon; \delta)\right) &= -\gamma \mathbb{E}\left(Y_0 \alpha_0^{*'} dh(Y_0, \varepsilon; \delta)\right), \\ \mathbb{E}\left(\frac{\partial^2 U_0^*}{\partial W \partial \theta'} Y_0 d\theta_f^*(\alpha^*(f_\delta), f_\delta)\right) &= -\gamma \mathbb{E}(Y_0) \mathbb{E}\left(\alpha_0^{*'} dh(Y_0, \varepsilon; \delta)\right), \\ \mathbb{E}\left(\frac{\partial U_0^*}{\partial W} dh(Y_0, \varepsilon; \delta)\right) &= \mathbb{E}\left(dh(Y_0, \varepsilon; \delta) - \gamma \left[\alpha_0^{*'} Y_0 - \mathbb{E}\left(\alpha_0^{*'} Y_0\right)\right] dh(Y_0, \varepsilon; \delta)\right).\end{aligned}$$

**iii) Robust portfolio approach :**

Regarding robust portfolio, the criterion  $G$  in (1.4.9) can be expanded in the following way :

$$\begin{aligned}G(\alpha, f_\delta) &= \mathbb{E}\left(U\left[\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)\right]\right) \\ &+ \mathbb{E}\left(\frac{\partial u}{\partial W}\left[\tilde{W}_t + \alpha' Y_0; \theta^*(\alpha, f_0)\right] \frac{g(Y_0, \varepsilon, \theta^*(\alpha, f_0))}{g(Y_0, \varepsilon)} \alpha' dh(Y_0, \varepsilon; \delta)\right) + o(\delta),\end{aligned}$$

where  $\theta^*(\alpha, f_0)$  is the optimal scenario minimizing the initial criterion  $\mathbb{E}\left(u\left[\tilde{W}_t + \alpha' Y_0\right]\right)$ .

## 1.5 SYSTEMATIC SHOCK

Stress-tests are regularly performed to check the resistance of the financial system. They consist in applying shocks to the balance sheets of the financial institutions (assimilated to risky portfolios). The aim of this section is to discuss the notion of shock by means of a systematic factor.

### 1.5.1 SHOCKS ON TAILS

The calibration of the shocks is an important stage in the implementation of stress-test exercises. In practice, the shocks are calibrated to be extreme, that is, they lead to realizations that can deviate significantly from the usual observations of the variables of interest<sup>42</sup>. The simplest form of shock is obtained through a translation of variable of interest, or equivalently thanks to a shift on its distribution (see Figure 1.5.1 in the univariate case). For instance, in its 2011 stress-testing exercise on European banks, the European Banking Authority studied an increase of +1% on the average default rate of loans, equal to about +1.5%, in banks' portfolios (see the EBA 2011 Aggregate Report)<sup>43</sup>.

With this definition of shock the approaches in terms of variable and distribution look similar. However this definition shows some deficiencies: the shock is deterministic, and concerns the mean of the distribution, not the higher moments generally used to capture risk.

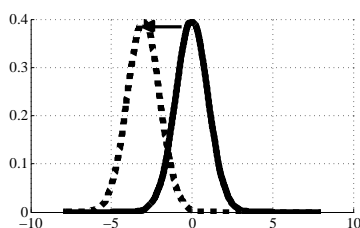


Figure 1.1: Shift of a distribution. The solid line stands for the baseline univariate distribution.

<sup>42</sup>However, as noted above, small shocks may imply large effects in a nonlinear framework. Thus, it is important to consider the effects as a function of  $\delta$ , that is, the impulse response function [see the discussion in Section 1.4.2.3].

<sup>43</sup>Such an increase is rather extreme, and should rather be fixed taking into account the quality of the loans in the balance-sheet of each bank. Typically an increase of +1% corresponds to a downgrade from AAA to BB, but only from BB to B, and can lead to a stable C rating.



In this Section, we investigate more sophisticated forms of shocks that have been considered in the literature, in terms of either variable, or distribution. We present the link between both approaches in the final part of this Section.

### *1.5.1.1 Shocks on tails in terms of variable*

Richer formulations of the shocks are conceivable. Let us first consider a variable based interpretation, where the shock on  $Y_0$  is denoted  $Y_\delta - Y_0 = h(Y_0, \varepsilon; \delta) - Y_0$ . Different forms of shocks on tails are obtained from various dependence structure between the shock and the baseline variable  $Y_0$ . Let us consider a shock such that :

$$Y_\delta = h(Y_0, \varepsilon; \delta) = Y_0 + \delta a(Y_0)b(\varepsilon), \quad (1.5.1)$$

where  $a$  ( $a$  larger than 0) and  $b$  are given functions.

We focus on risks in tails by selecting a function  $a$  taking large values in the left and/or right tails of  $Y_0$ . The symmetry or asymmetry of extreme shocks can be managed by an appropriate choice of function  $b$ . As an illustration, we consider below the following specification :

$$h(Y_0, \varepsilon; \delta) = Y_0 + \delta Y_0^2 \exp[-\zeta_y Y_0] \Phi^{-1}(\varepsilon) \exp[-\zeta_\varepsilon \Phi^{-1}(\varepsilon)], \quad (1.5.2)$$

where  $\varepsilon$  is an independent variable with Uniform distribution (see Section 1.2),  $\Phi$  is the c.d.f. of the standard Gaussian distribution, and  $\zeta_y, \zeta_\varepsilon$  are two positive scalars.

Four variants are presented in Figure 1.5.1.1 : symmetric extreme shock on both tails of  $Y_0$  ( $\zeta_y = 0$  and  $\zeta_\varepsilon = 0$ , top left panel in Figure 2), symmetric extreme shock concentrated on the left tail of  $Y_0$  ( $\zeta_y = 0.5$  and  $\zeta_\varepsilon = 0$ , bottom left panel), and asymmetric shocks on both tails ( $\zeta_y = 0$  and  $\zeta_\varepsilon = 2.5$ , top right panel), or on the left tail of  $Y_0$  ( $\zeta_y = 0.5$  and  $\zeta_\varepsilon = 2.5$ , bottom right panel). The corresponding unconditional distributions of  $Y_\delta = h(Y_0, \varepsilon; \delta)$  are given in Figure 1.3, which emphasizes the impact of the shock on the thickness of the tails of the distributions, its asymmetry, or its number of modes.

The expansion of the unconditional p.d.f. of  $Y_\delta$  in (1.5.2) is derived by

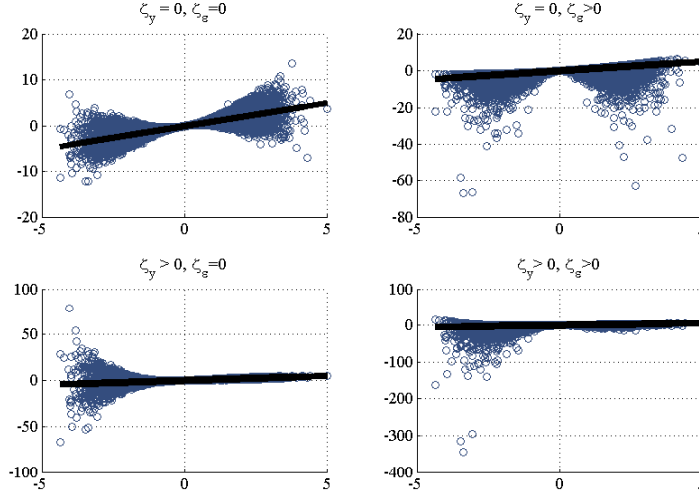


Figure 1.2: Conditional distributions of the shocked variable  $h(Y_0, \varepsilon; \delta)$  with respect to  $Y_0$  for different  $a$  and  $b$  functions in (1.5.2). The solid line stands for  $h(Y_0, \varepsilon; 0) = Y_0$ , and  $\Phi^{-1}(\varepsilon)$  has a standard Gaussian distribution.

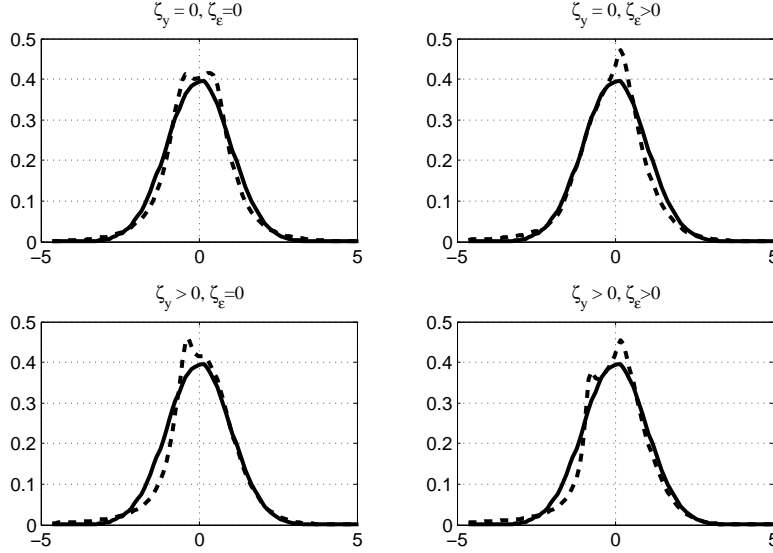


Figure 1.3: Unconditional distributions of the shocked variable  $h(Y_0, \varepsilon; \delta)$  for different  $a$  and  $b$  functions in (1.5.2), where  $Y_0$  and  $\Phi^{-1}(\varepsilon)$  have standard independent Gaussian distributions.

Corollary 1. At second-order in  $\delta$ , we get :

$$\begin{aligned}
 f(y; \delta) &= \phi(y) + \delta \zeta_\varepsilon \exp[\zeta_\varepsilon^2] \left( 2y - y^3 - \zeta_y y \right) \exp[-\zeta_y y] \phi(y) \\
 &+ \frac{\delta^2}{2} \left( 1 + 4\zeta_\varepsilon^2 \right) \exp[2\zeta_\varepsilon^2] \left( 2y^3 - \frac{y^5}{2} - \zeta_y y^4 \right) \exp[-2\zeta_y y] \phi(y),
 \end{aligned}$$

under the assumption of a standard normal variable  $Y_0$ , where  $\phi(y)$  is the p.d.f. of the standard normal distribution. The first and second-order moments of  $U \exp[-\zeta_\varepsilon U]$ , where  $U = \Phi^{-1}(\varepsilon)$ , are derived by considering the

first and second-order derivatives of the Laplace transform of the standard Gaussian variable.

*1.5.1.2 Shocks on tails in terms of distribution*

Such shocks have been introduced in the literature on robust statistics, which studies the contamination of a baseline distribution by “outliers”.

**i) Contamination**

A standard specification, introduced in Huber (1964), presents the contaminated distribution as a mixture of a baseline c.d.f.  $F(y; 0)$  and a contaminating c.d.f.  $\Xi(y)$  :

$$F(y; \delta) = (1 - \delta)F(y; 0) + \delta\Xi(y), \text{ with } 0 \leq \delta \leq 1. \quad (1.5.3)$$

This specification is a special case of shocks in terms of distribution, for which the first-order expansion in  $\delta$  is exact :  $f(y; \delta) = f(y; 0) + \delta [\xi(y) - f(y; 0)]$ , where  $\xi(y)$  is the p.d.f. of the contaminating distribution. A left tail contamination is illustrated in Figure 1.4, with Gaussian distributions, with different means and the same variance for  $F(\bullet; 0)$  and  $\Xi(\bullet)$ .

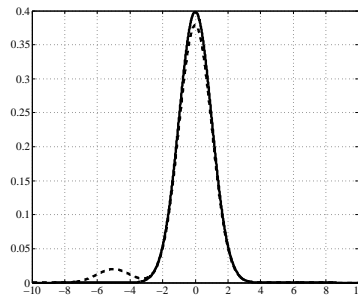


Figure 1.4: Left tail contaminated distribution. The solid line stands for the baseline distribution, the dashed line for the contaminated distribution.

**ii) Contamination in terms of variable**

The conversion of contaminated distribution in terms of variable relies on the general link between families of distributions and families of variables presented in Section 1.2. Let us for instance apply the Gaussian copula of

Section 1.2.3 to the specification (1.5.3), where the baseline and contaminated variables are denoted  $Y_0$  and  $Y_\delta$ , respectively. We have (see Appendix 1.4 ii)) :

$$\Phi\left(\sqrt{1-\rho^2(\delta)}U + F_0(Y_0)\right) = (1-\delta)F_0(Y_\delta) + \delta\Xi(Y_\delta), \quad (1.5.4)$$

where  $U$  is standard normal and  $\rho(0) = 1$  to ensure the coherency condition, that is,  $Y_\delta = Y_0$ , when  $\delta = 0$ .

### 1.5.1.3 How to reconcile shocks on tails in terms of variable and distribution

Let us focus on the previous example of contaminated distribution. Several forms of correlation parameter  $\rho(\delta)$ , and thus several Gaussian copulas, are consistent with contaminated distribution (1.5.3), once they satisfy the coherency condition  $\rho(0) = 1$ . This example stresses the difficulty in identifying a unique variable-based specification of a shock on tails expressed in terms of distribution.

Besides, the linearity in  $\delta$  of shocks in terms of variable as considered in Section 1.5.1.1 do not necessarily imply the linearity in  $\delta$  of the shocks in terms of distributions as in (1.5.3). Indeed, let us consider the following approximation of the correlation parameter :

$$\rho(\delta) = 1 - \delta^2 r, \text{ where } r = - \left. \frac{\partial \rho(\delta)}{\partial \delta} \right|_{\delta=0},$$

where the first term in the expansion of  $\rho$  is of order  $\delta^2$  to ensure the constraint  $\rho(\delta) \leq 1, \forall \delta$ . We prove in Appendix 1.4 iii) that :

$$Y_\delta = Y_0 + \delta Z + o(\delta), \text{ say,} \quad (1.5.5)$$

where the variable  $Z$  is given by :

$$Z \approx \frac{\sqrt{2r}\phi(\Phi^{-1}(F(Y_0;0)))U + F(Y_0;0) - \Xi(Y_0)}{f(Y_0;0)}. \quad (1.5.6)$$

The expansion (1.5.5)-(1.5.6) highlights the double effect of the contamination of a distribution in terms of variable: i) a drift effect  $\frac{F(Y_0;0) - \Xi(Y_0)}{f(Y_0;0)}$ , and ii) a conditionally heteroscedastic effect  $\frac{\sqrt{2r}\phi(\Phi^{-1}(F(Y_0;0)))U}{f(Y_0;0)}$  similar to the effects in (1.5.1), that will impact the tails of the distribution of the contaminated variable. This heteroscedastic effect depends on the curvature of function

$\rho(\bullet)$  in the neighborhood of  $\delta = 0$ , while the drift effect is negative whenever the contaminating distribution  $\Xi$  stochastically dominates at order 1 the baseline distribution  $F(Y_0; 0)$ . We still have the multiplicity of interpretations of the shock in terms of variable, since the curvature effect  $r$  is unconstrained. In terms of variable, larger is this curvature, larger is the weight on the volatility with respect to the drift.

The comparison of (1.5.6) with the linear interpretation of the shock in the variable-based approach in (1.5.1), for which drift effects are missing, emphasizes the differences in both specifications, in spite of their common linear in  $\delta$  formulation.

### 1.5.2 SHOCKS ON A SYSTEMATIC FACTOR

For a vast majority of asset classes, statistical analyses of the price dynamics reveal a limited number of linear, or nonlinear, latent factors, which explain most of the variation in asset prices, or asset returns. In a linear dynamic framework, it is known for instance that one factor drives most of the returns on US Treasury bonds [see e.g. Cochrane, Piazzesi (2005)], while the literature identifies few common linear factors for equity returns [see e.g. Fama, French (1992)].

These systematic factors represent the common dynamic patterns among multiple asset prices, and characterize the dependence structure between assets. For regulators, the identification of linear or nonlinear systematic factors is of crucial importance, since their variations have the biggest impact on institutions' portfolios, as opposed to idiosyncratic factors, whose risk can be diversified away by the holding of a large number of assets. As an illustration, let us consider the Basel 2 credit risk model. Inspired from the Value-of-the-Firm model [Merton (1974)], it decomposes in a linear way the log asset/liability  $Y_j^*$  of a firm  $j$  in common factor  $X$  and specific (or idiosyncratic) components  $\eta_j$  :

$$Y_j^* = \sqrt{\rho}X + \sqrt{1-\rho}\eta_j,$$

with positive asset correlation  $\rho$ , and deduces the default indicator as :

$$Y_j = \mathbb{1}_{Y_j^* < 0} = \mathbb{1}_{\sqrt{\rho}X < -\sqrt{1-\rho}\eta_j}.$$

In the basic model, this default indicator is directly related to the payoff of a Credit Default Swap (CDS) written on firm  $j$ . Thus, the probability

of very large cumulated losses on the institution's portfolio of CDS highly depends in a nonlinear way on the share of common/systematic factors in the distribution of firm's log asset/liability.

We present below the local expansion of the distribution of variables of interest with respect to "systematic shocks", that are shocks on the distribution of systematic factors, both in terms of distribution and variable.

### 1.5.2.1 Local analysis of systematic shock in terms of distribution

Let us decompose the joint distribution of variables of interest  $f(y; \delta)$  into two parts: one corresponding to the marginal distribution of the common factors, and the other one referring to the distribution of the variables of interest, conditional on the common factors. A systematic shock would hit the marginal distribution of the common factor  $X$  without modifying the conditional distribution of  $Y$  given  $X$ . Thus we get :

$$f(y; \delta) = \int f_1(y|x)f_2(x; \delta)dx. \quad (1.5.7)$$

A direct application of the results in Section 1.3.1 gives the following local expansion of the distribution of the variable of interest :

$$f(y; \delta) = f(y; 0) + \sum_{j=1}^p \frac{\delta^j}{j!} \int f_1(y|x) \frac{\partial^j f_2(x; 0)}{\partial \delta^j} dx + o(\delta^p). \quad (1.5.8)$$

In particular, at second-order, we get :

$$f(y; \delta) = f(y; 0) + \delta \int f_1(y|x) \frac{\partial f_2(x; 0)}{\partial \delta} dx + \frac{\delta^2}{2} \int f_1(y|x) \frac{\partial^2 f_2(x; 0)}{\partial \delta^2} dx + o(\delta^2).$$

### 1.5.2.2 Local analysis of systematic shock in terms of variable

Let us now consider the (vector of) variables of interest  $Y_\delta$  as a function of common factors  $X_\delta$ , and independent variables  $\eta$  :

$$Y_\delta = b(X_\delta, \eta), \quad (1.5.9)$$

where the common factors are such that :

$$X_\delta = h(X_0, \varepsilon; \delta). \quad (1.5.10)$$

The model involves two types of basic impulses, that are the variable  $\eta$  representing the idiosyncratic component and the variable  $\varepsilon$  used to define

shock on the systematic factor. Equation (1.5.9) is compatible with nonlinear effects of both types of components and possibly cross-effects of systematic and idiosyncratic components.

From Proposition 1, we get at second-order :

$$\begin{aligned} f(y; \delta) &= f(y; 0) - \delta \int f_1(y|x) \frac{\partial}{\partial x} \left[ f_2(x; 0) \mathbb{E} \left[ \frac{\partial h(X_0, \varepsilon; \delta)}{\partial \delta} | X_0 = x \right] \right] dx \\ &+ \frac{\delta^2}{2} \int f_1(y|x) \left\{ \frac{\partial^2}{\partial x^2} \left[ f_2(x; 0) \mathbb{E} \left[ \left( \frac{\partial h(X_0, \varepsilon, 0)}{\partial \delta} \right)^2 | X_0 = x \right] \right] \right. \\ &\quad \left. - \frac{\partial}{\partial x} \left[ f_2(x; 0) \mathbb{E} \left[ \frac{\partial^2 h(X_0, \varepsilon, 0)}{\partial \delta^2} | X_0 = x \right] \right] \right\} dx + o(\delta^2). \end{aligned}$$

For instance, let us consider a Value-of-the-Firm model for an homogeneous population of  $N$  different firms, where  $Y_\delta^*$  is the vector of firms' log asset/liability :

$$Y_\delta^* = \sqrt{\rho} \mathbf{1}' X_\delta + \sqrt{1 - \rho} \eta,$$

where  $\mathbf{1}$  is a  $N \times 1$  vector of ones, and  $\eta$  has a standard multivariate normal distribution.

Let us assume that the single common factor is such that :

$$X_\delta = X_0 + \delta a(X_0) U,$$

where  $U$  is a random variable independent from  $X_0$ , whose second-order moments exists.

We get at second-order :

$$\begin{aligned} f_2(x; \delta) &= f_2(x; 0) - \delta \mathbb{E}(U) \frac{d}{dx} [a(x) f_2(x; 0)] + \frac{\delta^2}{2} \mathbb{E}(U^2) \frac{d}{dx} [a^2(x) f_2(x; 0)] + o(\delta^2), \\ f(y^*; \delta) &= f(y^*; 0) - \delta \mathbb{E}(U) \int f_1(y^*|x) \left( a^{(1)}(x) f_2(x; 0) + a(x) f_2^{(1)}(x; 0) \right) dx \\ &+ \frac{\delta^2}{2} \mathbb{E}(U^2) \int f_1(y^*|x) \left( 2a(x) a^{(1)}(x) f_2(x; 0) + a^2(x) f_2^{(1)}(x; 0) \right) dx + o(\delta^2). \end{aligned}$$

Then this expansion can be used to deduce the expansion of the distribution of the default indicators.

### 1.5.2.3 Systematic shock with extreme effects

This section helps to precise the notion of extreme shock, which remains rather vague in the literature, and calls for richer specifications of shock

than the “shock-in-mean” usually considered in stress-test exercises. When performing stress-tests, we are interested in shocks on a systematic factor with extreme impacts on asset portfolio values.

### **i) Linear dynamic factor model**

In a linear dynamic framework, large effects on the tail of the distribution of variable of interest  $Y$  are obtained by introducing shocks on the tails of the factor distribution. This is why the literature usually focus on factor tails (see e.g. Section 1.5.1).

### **ii) Nonlinear dynamic factor model**

The situation is very different in the nonlinear framework encountered for portfolios of derivative assets (see e.g. the example of CDS portfolio), or for portfolios managed in some optimal way, for which the relevance of the chosen shocks depends on the nonlinear link between the distribution of portfolio’s value and the distribution of the systematic factor.

In such nonlinear factor model, a rather small shock on a factor value can sometimes have a large impact on the portfolio. As an illustration, let us consider an investor, who allocates her/his wealth among two assets, when the asset prices are almost perfectly correlated. Such a large correlation would incite her/him to lever up her/his wealth by selling the less profitable asset to buy the other one. Let us now consider a shock on the joint asset distribution changing the asset correlation to a negative value close to  $-1$ , whereas keeping the expected returns and volatilities at the same level. Under a Gaussian assumption on returns, the joint Gaussian distribution is modified much more in its central part than in its tails. However, the portfolio allocation crystallized at its previous level will be very sensitive to this change, since now the risks lie in the same direction, and the leverage effect is exactly at the opposite of what has to be done. This is exactly the situation encountered by N. Leeson that implied the failure of the Barings, or in the LTCM default.

This example shows that a shock on even a central part of the distribution can have extreme consequences. Therefore, it is important to detect the type of shock on the factor with such huge consequences. For this purpose, it is



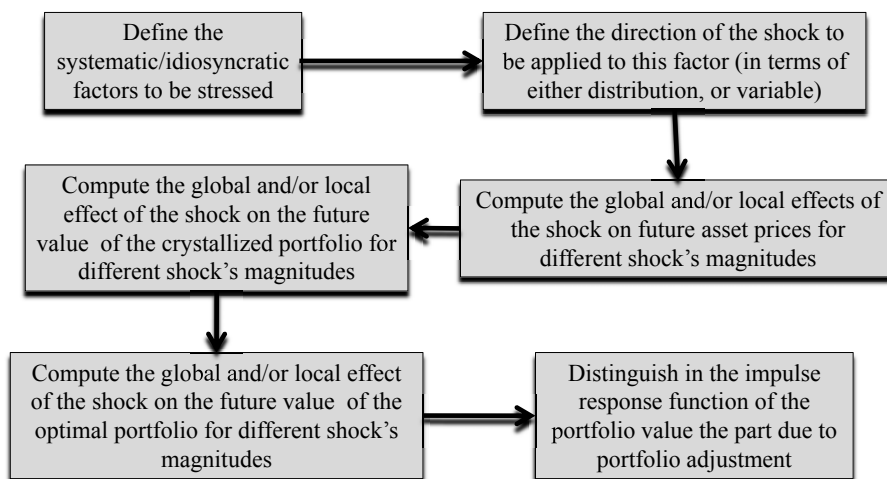
recommended to define a direction of the shock in terms of either variable, or distribution, and to measure the consequences for different levels of  $\delta$ , that is, to consider an impulse response function [see e.g. Koop, Pesaran, Potter (1996), and Gouriéroux, Jasiak (2005)].

## 1.6 STRESS-TESTING THE EUROPEAN SOVEREIGN BOND MARKET

The aim of this section is to avoid the limitations of the current implementation of stress-tests, in which the shocks are assumed deterministic and the portfolio is crystallized. In the application presented below, we stress the portfolio of a financial institution invested in European sovereign bonds. We propose an intuitive direction of the shock in terms of distribution and variable and show how the riskiness of investor's portfolio, either crystallized or optimally updated, evolves with the size of the shock on the distribution of excess gains.

### 1.6.1 STRESS-TEST

A stress-test requires the application of the different notions introduced in Sections 1.2, 1.3, and 1.5 along the lines of scheme 4.



**Scheme 4: A stress-test**

In the following parts of the section, we apply these different steps to a portfolio of European sovereign bonds.

## 1.6.2 EXCESS GAINS ON THE EUROPEAN SOVEREIGN BOND MARKET

As an illustration, let us consider an investor with mean-variance objective function, who invests her/his wealth in the European sovereign bond market. For simplicity, we consider zero-coupon bonds, with face value 1, and maturities 10 years. We restrict our sample to six countries representing the variety of the euro area sovereign bond market, that are Germany, France, Italy, Spain, Ireland and Greece. The sample covers the period from July 2001 to June 2011. We assume that the investor has a monthly horizon, and has access to a risk-free asset, which pays the 1M Eonia swap rate<sup>44</sup> after a one-month holding period. The corresponding monthly excess gains in euros<sup>45</sup> are plotted in Figure 1.5.

We identify from principal components analysis one systematic factor, which

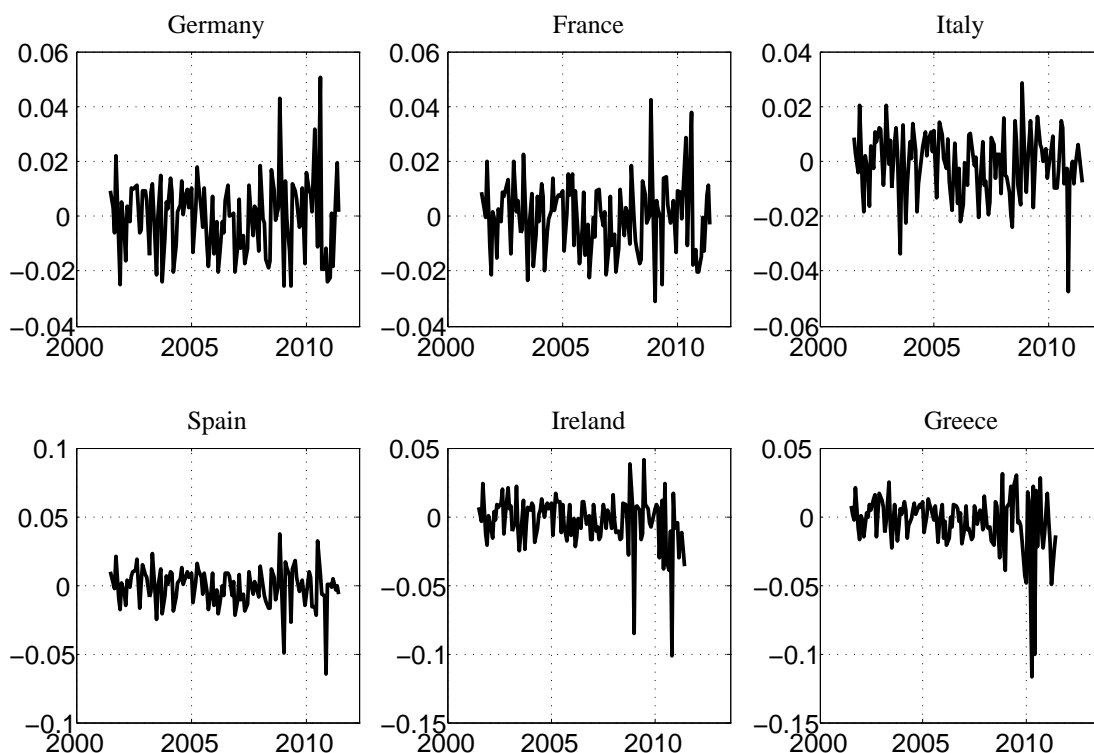


Figure 1.5: Monthly excess gains on 10Y European sovereign bonds.

explains about 95% of the variance of the historical excess gains. The first

<sup>44</sup>The Eonia swap rate is the fixed rate in a swap contract, whose variable leg is pegged to the Euro OverNight Index Average.

<sup>45</sup>We did not take into account the 1-month maturity effect on bond prices in the computation of excess gains.

factor weights uniformly the 10Y rates of all countries (see Appendix 1.6). Loosely speaking, this factor is a kind of Eurozone systematic risk. We will shock this “Euro factor”<sup>46</sup>.

As put forward in Figure 1.6, the first factor has different distributions

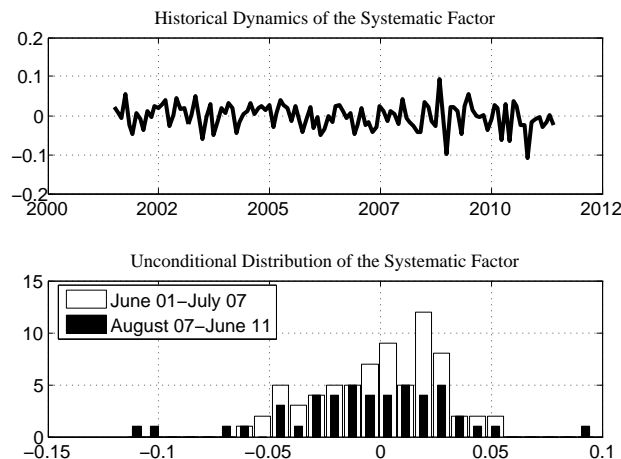


Figure 1.6: Historical dynamics and histograms of the first systematic factor on periods 2001-2007, and 2007-2011.

before and after July 2007: it features fatter tails, becomes bimodal and seems less asymmetric. We rely on this change to define a systematic shock by contamination.

### 1.6.3 CONTAMINATION ON THE SYSTEMATIC FACTOR

More precisely, we identify the distribution on the 2001-2007 period of the systematic factor  $X$  as our baseline distribution, denoted  $F(x; 0)$ , while the distribution on 2007-2011 plays the role of the contaminating distribution  $\Xi(x)$ .

Then we consider a specification of the shock on the factor’s distribution as in Section 1.5.1.2 :

$$F(x; \delta) = (1 - \delta)F(x; 0) + \delta\Xi(x), \text{ with } 0 \leq \delta \leq 1. \quad (1.6.1)$$

Let us now consider the contamination in terms of variable. We know that there exists an infinite number of specifications in terms of variable providing

---

<sup>46</sup>It would also be possible to consider specific shocks, for instance on the value of the Greek sovereign debt.

the specification (1.6.1) of the unconditional distribution of  $Y_\delta$ . Instead of selecting one of these specifications, we consider a linear specification providing locally the equivalence. As shown in Section 1.5.1.3, equation (1.5.6), we can express at first-order the empirical contaminated distribution in terms of variable as :

$$X_\delta = X_0 + \delta Z + o(\delta), \quad (1.6.2)$$

where  $X_0$  and  $X_\delta$  are the baseline and contaminated factors, with

$$Z \approx \frac{\sqrt{2r}\phi(\Phi^{-1}(F(X_0; 0)))U + F(X_0; 0) - \Xi(X_0)}{f(X_0; 0)}, \quad (1.6.3)$$

$U$  is standard Gaussian, and the curvature parameter  $r$  is set to  $r = 2$ . We take advantage of representation (1.6.3) to simulate a set of 1000 contaminated variables, by drawing independently 1000 times  $X_0$  in the set of realized  $X_0$  and  $U$  in the standard normal distribution, and by taking  $f(\bullet; 0)$ ,  $F(\bullet; 0)$ , and  $\Xi(\bullet)$  at their empirical counterparts  $\hat{f}(\bullet; 0)$ ,  $\hat{F}(\bullet; 0)$ , and  $\hat{\Xi}(\bullet)$ . The corresponding c.d.f. of the simulated contaminated variables are plotted in Figure 3.16 for different  $\delta$  values. Even if the contamination models in distribution and variable are equivalent in a neighborhood of  $\delta = 0$ , the comparison of Figures 3.16 and 3.16 shows that the two specifications imply different types of stochastic shocks for larger values of  $\delta$ . In particular, the effect on tails is more important with the contamination approach (1.6.2) written in terms of variable.

#### 1.6.4 IMPACT OF THE SYSTEMATIC SHOCK ON PORTFOLIO CHARACTERISTICS

##### **i) Simulation of contaminated excess-gains**

Let us consider a simple factor model for the vector of excess gains :

$$P_\delta = \mu + \beta' X_\delta + \Sigma^{1/2}\eta, \quad (1.6.4)$$

where  $\Sigma$  is diagonal,  $\eta$  is a standard zero-mean vector, and  $\beta$  collects the systematic factor loadings for each bond's excess gain.

The parameters  $\mu$ ,  $\beta$ ,  $\Sigma$  are estimated from the Seemingly Unrelated Regression (SUR) of excess gains  $P_t$  on the first factor deduced from the principal

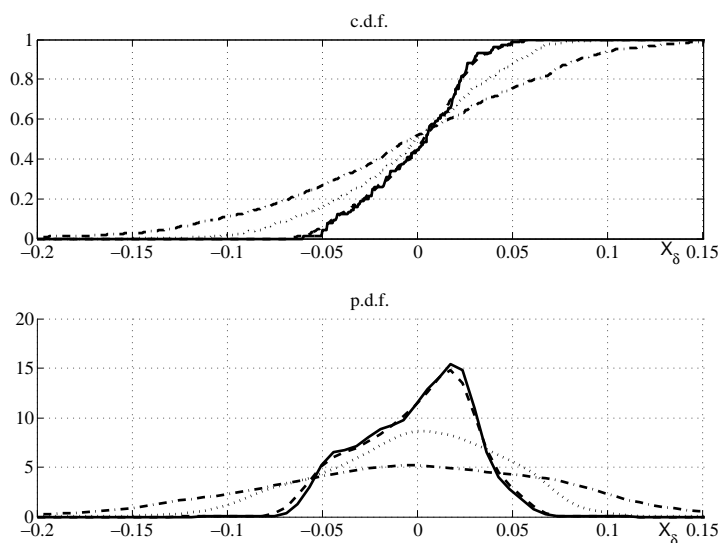


Figure 1.7: Empirical c.d.f. and p.d.f. of 1000 simulated contaminated factor values  $X_\delta$  in terms of variable for various  $\delta$ . The solid line stands for the c.d.f. of the baseline factor  $X_0$ , while dashed, dotted, and dash-dotted lines represent the contaminated empirical distribution for  $\delta = 0.1, 0.5, 1$ .

component analysis. Then the distribution of  $\eta$  is approximated by the historical distribution of the residuals :

$$\hat{\eta}_t = \hat{\Sigma}^{-1/2} (P_t - \hat{\mu} - \hat{\beta}' \hat{X}_t).$$

The simulation of contaminated excess gains is based on model (1.6.4) after replacement of the parameters by their estimates. Let us consider a contamination in terms of variable (the contamination in terms of distribution is presented with the corresponding results in Appendix 1.6). We draw independently a value  $\eta^s$  in the empirical distribution of the residuals  $\hat{\eta}_t$ , and a value  $X_\delta^s$  in terms of variable as described in Section 1.6.3. Then the simulated contaminated excess gain is :

$$P_\delta^s = \hat{\mu} + \hat{\beta}' X_\delta^s + \hat{\Sigma}^{1/2} \eta^s.$$

This procedure is repeated  $S = 1000$  times. The empirical distribution of the simulated contaminated variables  $P_\delta^s$ ,  $s = 1, \dots, S$ , provides an estimate of the theoretical distribution of the contaminated excess-gains.

## ii) The contaminated mean-variance allocation

We can now derive the mean-variance allocation for an investor, who adjusts herself/himself to the contaminated excess gain distribution. From the estimated distribution of the contaminated excess-gains derived in 5.4 i), we deduce for each magnitude  $\delta$  of the shock the mean and variance/covariance matrix of contaminated excess gains, from which we deduce the optimal mean-variance allocation, denoted  $\alpha^*(\delta)$ <sup>47</sup>. The optimal portfolio allocation as a function of the magnitude of the shock  $\delta$  is given in Figure 1.8. This figure highlights the nonlinear effects of  $\delta$  on the optimal allocation.

This nonlinearity is a direct consequence of the specification of the stochastic shock and of the mean-variance portfolio management. Indeed, we deduce from (1.6.2)-(1.6.4) that:

$$\begin{aligned}\mathbb{E}(P_\delta) &= \mu + \beta' \mathbb{E}(X_\delta) = \mu + \beta' \mathbb{E}(X_0) + \beta' \mathbb{E}(Z), \\ \mathbb{V}(P_\delta) &= \beta' \mathbb{V}(X_\delta) \beta + \Sigma = \beta' \left[ \mathbb{V}(X_0) + \delta \text{Cov}(X_0, Z) + \delta \text{Cov}(Z, X_0) + \delta^2 \mathbb{V}(Z) \right] \beta + \Sigma.\end{aligned}$$

Therefore the mean-variance allocation:  $\alpha^*(\delta) = \gamma^{-1} \mathbb{V}(P_\delta)^{-1} \mathbb{E}(P_\delta)$ , has components, which are ratios of polynomials in  $\delta$  of degree 12.

Despite this nonlinear pattern, we observe that these optimal portfolios are short in German, French, Italian and Greek bonds [resp. long in Irish, Spanish bonds] for any value of  $\delta$ .

### **iii) Impulse responses for crystallized and optimally adjusted portfolios**

Let us finally compare the properties of the portfolio value under the systematic shock. We consider two portfolio managements, that are i) the mean-variance portfolio crystallized at its optimal level before contamination, and ii) the mean-variance portfolio adjusted for contamination<sup>48</sup>. The properties of the risky part of these portfolios are represented in Figures 1.9-3.21 by their mean, variance (volatility), Sharpe performance, VaR and expected shortfall (at the 1%, 5% and 10% levels for the two last summaries).

Figures 1.9-3.21 emphasize the significant impact of the portfolio management on the characteristic of the portfolio. The performance of the optimal

---

<sup>47</sup>We set investor's wealth at 100, and the level of her/his risk aversion,  $\gamma = 2$ .

<sup>48</sup>As usual in such stress-tests, we assume that the portfolio updating, that is the demand updating by the banks, has no effect on the asset price dynamics.

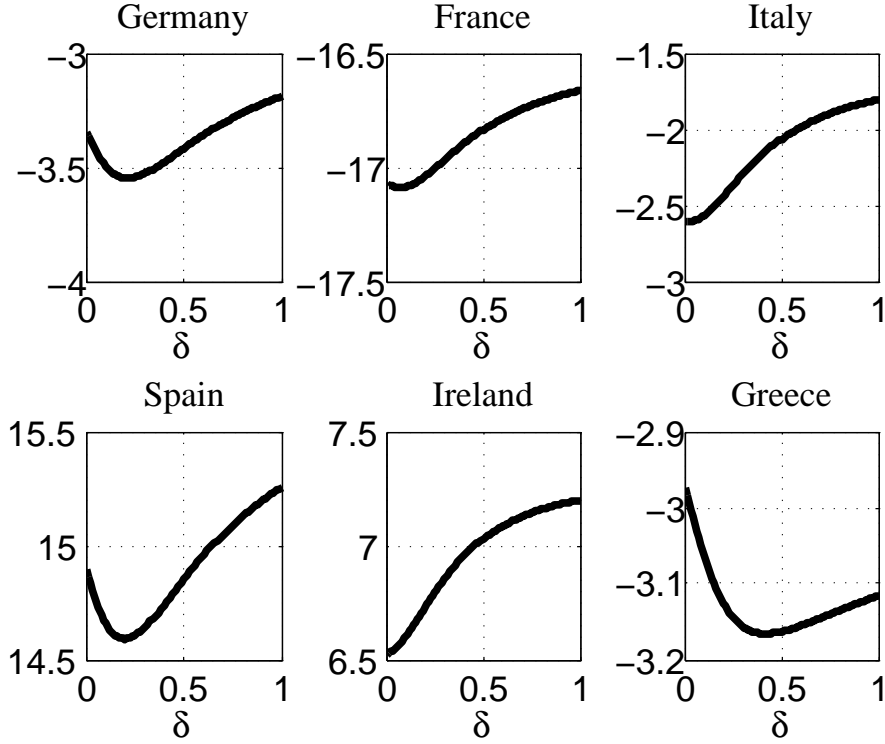


Figure 1.8: Contaminated mean-variance allocation as a function of  $\delta$ .

portfolio significantly dominates the crystallized portfolio (see the Sharpe ratio of both portfolios in Figure 1.9). In this example, the risky part of the crystallized portfolio becomes more volatile and features higher VaR and expected shortfall than the optimal portfolio.

The local behavior of the VaR for the crystallized and mean-variance portfolios is easily analyzed. The shock in variable is :

$$P_\delta = P_0 + \delta\beta'Z + o(\delta).$$

The mean-variance allocation can be expanded as :

$$\begin{aligned} \alpha^*(\delta) &= [\mathbb{V}(P_\delta)]^{-1} \mathbb{E}(P_\delta) \\ &= [\mathbb{V}(P_0 + \delta\beta'Z)]^{-1} \mathbb{E}(P_0 + \delta\beta'Z) + o(\delta) \\ &= \{\mathbb{V}(P_0) + \delta[Cov(P_0, \beta'Z) + Cov(\beta'Z, P_0)]\}^{-1} [\mathbb{E}(P_0) + \delta\mathbb{E}(\beta'Z)] + o(\delta) \\ &= \{[\mathbb{V}(P_0)]^{-1} - \delta[\mathbb{V}(P_0)]^{-1}[Cov(P_0, \beta'Z) + Cov(\beta'Z, P_0)][\mathbb{V}(P_0)]^{-1}\} \\ &\quad [\mathbb{E}(P_0) + \delta\mathbb{E}(\beta'Z)] + o(\delta) \\ &= [\mathbb{V}(P_0)]^{-1} \mathbb{E}(P_0) + \delta\{[\mathbb{V}(P_0)]^{-1} \mathbb{E}(\beta'Z) \\ &\quad - [\mathbb{V}(P_0)]^{-1}[Cov(P_0, \beta'Z) + Cov(\beta'Z, P_0)][\mathbb{V}(P_0)]^{-1} \mathbb{E}(P_0)\} + o(\delta) \\ &= \alpha_0 + \delta\alpha_1, \text{ say.} \end{aligned}$$



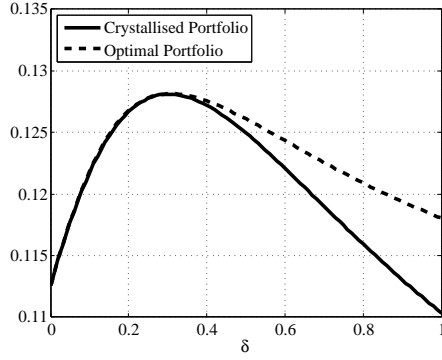


Figure 1.9: Impulse response of the Sharpe ratio of crystallized and Mean-Variance portfolios (contamination in variable).

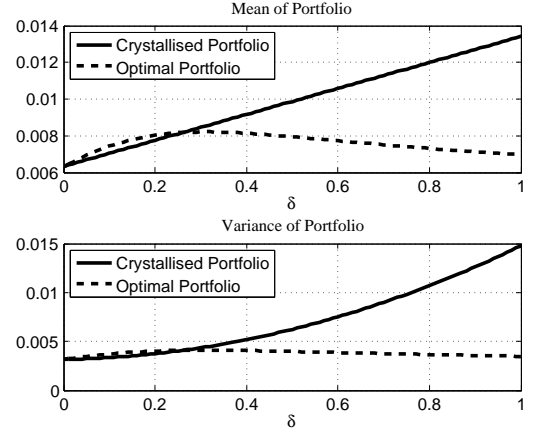


Figure 1.10: Impulse response of the Mean and Variance of crystallized and Mean-Variance portfolios (contamination in variable).

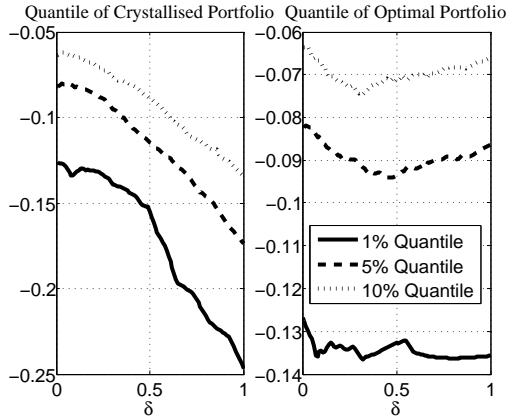


Figure 1.11: Impulse response of the VaR of crystallized and Mean-Variance portfolios (contamination in variable).

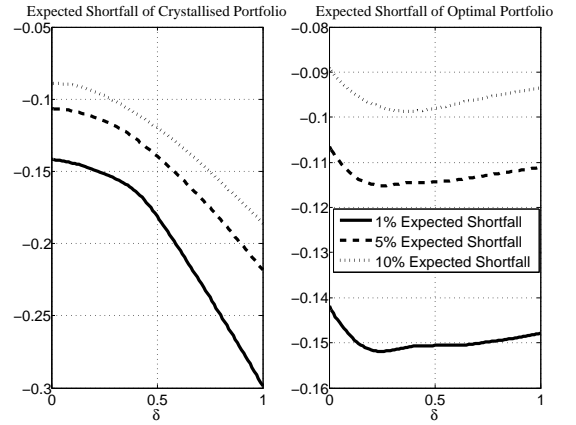


Figure 1.12: Impulse response of the Expected-shortfall of crystallized and Mean-Variance portfolios (contamination in variable).

The value of the mean-variance portfolio is :

$$Y_\delta^* = \alpha^*(\delta)' P_\delta = \alpha_0' P_0 + \delta (\alpha_0' P_1 + \alpha_1' P_0) + o(\delta),$$

whereas the value of the crystallized portfolio is equal to :

$$\tilde{Y}_\delta = \alpha_0' (P_0 + \delta P_1) + o(\delta).$$

We deduce the expansion of the VaR of the mean-variance portfolio :

$$\begin{aligned} VaR_q(Y_\delta^*) &= VaR_q(Y_0^*) + \delta \mathbb{E}[\alpha'_1 P_1 + \alpha'_1 P_0 | Y_0^* = VaR_q(Y_0^*)] + o(\delta) \\ &= VaR_q(\tilde{Y}_\delta) + \delta \alpha'_1 \mathbb{E}[P_0 | Y_0^* = VaR_q(Y_0^*)] + o(\delta), \end{aligned}$$

by using the expression of the derivative of the VaR [Gourieroux, Laurent, Scaillet (2000)].

Therefore, the difference between the two VaR's is equivalent to :

$$\begin{aligned} \frac{1}{\delta} [VaR_q(Y_\delta^*) - VaR_q(\tilde{Y}_\delta)] &\approx \mathbb{E}[P_0 | Y_0^* = VaR_q(Y_0^*)]' [\mathbb{V}(P_0)]^{-1} \beta' \mathbb{E}(Z) \\ &\quad - \mathbb{E}[P_0 | Y_0^* = VaR_q(Y_0^*)]' [\mathbb{V}(P_0)]^{-1} Cov(P_0, Z) \beta [\mathbb{V}(P_0)]^{-1} \mathbb{E}(P_0) \\ &\quad - \mathbb{E}[P_0 | Y_0^* = VaR_q(Y_0^*)]' [\mathbb{V}(P_0)]^{-1} \beta' Cov(Z, P_0) [\mathbb{V}(P_0)]^{-1} \mathbb{E}(P_0). \end{aligned}$$

This difference can be of any sign, especially in our framework in which  $Z$  and  $Y_0$  are dependent [see equation (1.6.3)].

Finally, we present in Figures 1.13-1.14 the conditional distribution of the contaminated portfolio value  $Y_\delta$  with respect to its initial/non-contaminated value  $Y_0$ , which can only be derived from a specification of shock in terms of variable. Figures 1.13-1.14 emphasize the heteroscedasticity of the shock considered in this exercise: the main effect is concentrated on the central part of the initial distribution. Moreover, the comparison of Figures 1.13-1.14 highlights how the optimization of the portfolio allocation circumscribes the shock's impact on the portfolio value, even for large shock's magnitude. This feature is consistent with the low sensitivity to the shock of several risk measures for the optimally adjusted portfolio in this example.

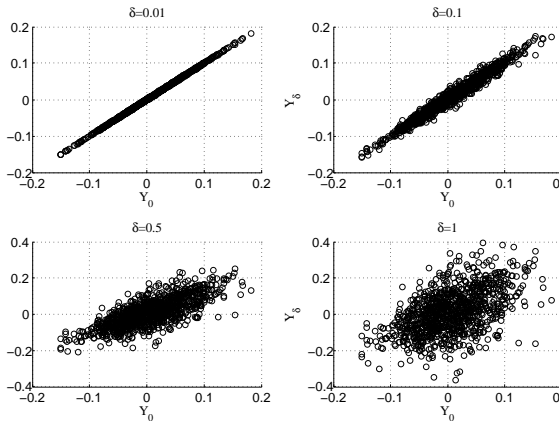


Figure 1.13: Impulse response of the conditional distribution of the crystallized portfolio  $Y_\delta|Y_0$  (contamination in variable).

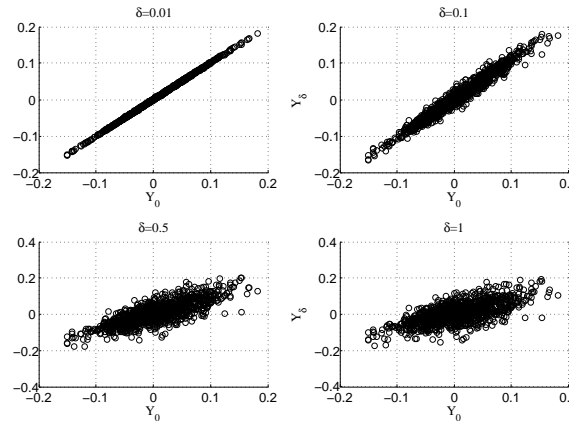


Figure 1.14: Impulse response of the conditional distribution of the optimally adjusted portfolio  $Y_\delta|Y_0$  (contamination in variable).

## 1.7 CONCLUDING REMARKS

We have discussed the links between a modeling of shocks in terms of distribution and variable, both for global and local shocks. We have seen that a multiplicity of specifications of the shock in terms of variable can lead to a same specification of the shock in terms of distribution. Moreover, the link between these specifications is not obvious: for instance, a linear shock in terms of distribution does not imply a linear shock in terms of variable. Therefore, a prudential approach may consist in considering carefully joint interpretations of a shock, both in terms of distribution and variable.

Such shocks can be introduced on systematic factors to perform stress-tests. This methodology of systematic shock has been illustrated for portfolios of European sovereign bonds. This highlights the different sensitivity to systematic shock of crystallized and optimally updated portfolios.

---

Part II

---

Pricing Very Long-Term Bonds under  
No-Arbitrage Assumptions



This second part focuses on the modeling of the very long-end of the yield curve of interest rates. Governments, non-financial private companies, banks and insurance companies issue fixed-income securities with very long-term maturities, such as coupon bonds with 50 to 100 years maturity, consols (i.e. coupon bonds with infinite time-to-maturity), or longevity bonds (i.e. bonds whose interest payments are indexed on some mortality index). The size of the market is expanding since 2000, fed by the demand of institutional investors with very long-term liabilities (pension funds for instance). However, the modeling of the very long-term rates have received little attention in the literature so far, and most of the term structure models considered, in particular the popular class of Affine Term Structure Models (ATSM henceforth), introduced by Duffie, Kan (1996), assume a simple behavior of the very long term rate, which usually converges to a constant level as its time-to-maturity increases [see the simple example of this phenomenon in the General Introduction of this dissertation]. These models may not be appropriate to take into account the observed volatility of the long-term rates [see Figure 2 in the General Introduction], nor to discount very long-term cash-flows, which requires an extrapolation of the observed term structures. In our research on term structure models with stochastic very long-term rates, we focus on term structure models, which incorporate a level factor, whose variations affect uniform all yields, and *a fortiori* the longest ones.

This part of the thesis aims to discuss the relevant alternatives for the pricing of fixed income securities with very long-term maturities (from 10 years to infinity), under the assumption of no-arbitrage opportunities in the government bond market. It is indeed crucial for financial institutions to rely on such arbitrage-free models: the contrary would incite the institution's traders to make unlimited investments in positions they believe are risk-free, but are not.

The following chapters stress how challenging it is to design term structure models, which are able to allow limiting rate to be stochastic, while being arbitrage-free [Chapter 2], and introduce a new class of term structure model for this purpose [Chapter 3].



# Term Structure Models with Level Factor, a Long-Term Rates Perspective\*

---

## 2.1 INTRODUCTION

The dynamic analysis of the term structure of interest rates reveals the existence of a limited number of underlying factors. It is usual to interpret sequentially these factors as a level factor, a slope (or steepness) factor, a curvature (or butterfly) factor, and so on, even if these notions have not been precisely defined in the literature [see e.g. Litterman, Scheinkman (1991), Jones (1991)].<sup>49</sup> This factor interpretation has also been extended to the field of option pricing [see Rogers, Tehranchi (2008) for a study of parallel shifts in the term structure of implied volatilities].

The aim of this chapter is to consider an arbitrage-free factor model of the term structure of interest rates, where one of the factors has a level interpretation (loosely speaking, any shock on the factor  $X$  will imply a parallel shift in the whole term structure). This specification is interesting for our purpose since it implies a stochastic limiting rate, and may be a valid alternative to reproduce the observed volatility of very long-term rate<sup>50</sup>, or to extrapolate

---

\*This chapter is based on a joint work with Christian Gourieroux. We thank Bruno Feunou, Jean-Sebastien Fontaine, Zorana Grbac, Jonathan Ingersoll, Alain Monfort, Steve Ross, and the participants at the 2011 Mathematical Finance Days, HEC Montreal, the Banque de France and Bank of Canada internal seminars for their helpful comments.

<sup>49</sup>"Level, slope and curvature factor loadings at the core of (term structure) models have their origin in the somewhat arbitrary and atheoretical field of yield curve fitting" [Krippner (2009)].

<sup>50</sup>See Figure 2 in the general introduction of this thesis, which stresses the high volatility of very long-term rates.



the observed term structures. The aim of this chapter is to investigate the consistency with no-arbitrage assumption of such term structure model. The compliance with no-arbitrage is of great importance for asset pricing modeling. In particular, it is crucial for market makers to rely on arbitrage-free asset pricing models for their quotes, to avoid the market participants to benefit from unlimited free lunches opportunities.

Let us first consider a single factor model :

$$r(t, h) = g(X_t, h), \quad \forall t \geq 0, h \geq 0,$$

where  $r(t, h)$  is the continuously compounded rate at date  $t$  for time-to-maturity  $h$ , and  $X$  is the single factor. A level factor is such that any shock  $\delta x$ , say, affecting the factor impacts the term structure by a drift independent of time-to-maturity. Thus we have :

$$g(X_t + \delta x, h) = g(X_t, h) + d(X_t, \delta x), \quad \forall X_t, h, \delta x,$$

where  $d(X_t, \delta x)$  denotes the drift, which can depend on the state  $X_t$  and of the magnitude of the shock. Without loss of generality, we can assume that  $X_t$  can take the value zero. Thus, we deduce that :

$$g(X_t + \delta x, h) = g(0, h) + d(0, \delta x), \quad \forall h, \delta x,$$

or equivalently that  $g$  can be decomposed as :

$$g(X_t + \delta x, h) = g(0, h) + d(0, X_t).$$

It admits an additive decomposition into a function of the factor and a function of the time-to-maturity. Since the level factor is defined up to a nonlinear transformation, any 1-factor model with a level factor can always be written as<sup>51</sup> :

$$r(t, h) = X_t + c(h), \tag{2.1.1}$$

In Section 2.2, we consider a discrete time term structure model with a single factor interpretable as a level factor. Then we introduce buy and hold strategies based on two zero-coupon bonds and derive the necessary

---

<sup>51</sup>Model (2.1.1) has been written for the continuously compounded rate. If we denote by  $r^*(t, h)$  the rate which is not continuously compounded, we have :  $\exp[-hr(t, h)] = [1 + r^*(t, h)]^{-h}$ , or equivalently  $r^*(t, h) = \exp[r(t, h)] - 1 = \exp[X_t + c(h)] - 1$ . Thus the notion of level factor depends on the definition of the rate. We keep the continuously compounded definition in this dissertation, which is compatible with the existing literature.

and sufficient conditions for no-arbitrage: the sequence  $[c(h)]$  has to be a sequence of Cesaro means of a nonnegative increasing function. In Section 2.3, we discuss the implications of this result on the behavior of the long-term interest rate. Section 2.4 exhibits all risk-neutral dynamics compatible with parallel shifts of the yield curve. We prove that they correspond to strong random walks and we explain how the behavior of the long-term interest rate depends on the distribution of the innovation of this random walk <sup>52</sup>. Section 2.5 extends our results to 2-factor models, by adding a slope factor to the term structure model with level factor. Section 2.6 concludes. The history of parallel and affine shifts of the term structure in the financial literature is presented in Appendix 2.1, and technical derivations are gathered in the other appendices.

---

<sup>52</sup>The main result of this Section contradicts Theorem 4 in Ingersoll, Skelton, Weil (1978). We will see later on why their result is incomplete.

## 2.2 NO-ARBITRAGE CONDITION FOR BUY-AND-HOLD STRATEGIES BASED ON TWO ZERO-COUPON BONDS.

In Sections 2.2-2.4, we consider a discrete time term structure model, i.e.  $t \in \mathbb{N}, h \in \mathbb{N} - \{0\}$ , with a single level factor [see (2.1.1)]. In decomposition (2.1.1), the factor is defined up to an additive constant. Therefore, without loss of generality, we can always assume :

**Assumption A.1 :**  $c(1) = 0$ .

Under Assumption A.1, the factor coincides with the short term interest rate :  $X_t = r(t, 1)$ .

We also assume that model (2.1.1) really is a single factor model :

**Assumption A.2 :** The support of the conditional distribution of  $(X_t)$  given  $X_0$  is not reduced to a single point.

Finally, we assume nonnegative rates, with a short term interest rate, which can reach value zero.

**Assumption A.3 :**

- i) The lower bound of the support of the distribution of  $X_t$  given  $X_0$  is zero;
- ii)  $c(h) \geq 0, \forall h \in \mathbb{N} - \{0\}$ .

Let us consider at date  $t$  a portfolio of two zero-coupon bonds with time-to-maturity  $h_1$  and  $h_2, h_2 > h_1$ , respectively. Its price at date  $t$  is :

$$\Pi_t(h_1, h_2, \alpha) = \alpha_1 B(t, h_1) + \alpha_2 B(t, h_2),$$

where  $B(t, h) = \exp[-hr(t, h)]$  denotes the price of the zero-coupon bond, and  $(\alpha_1, \alpha_2)$  the allocations.

The value of this portfolio at date  $t + k, k \leq h_1$ , is :

$$\Pi_{t+k}(h_1 - k, h_2 - k, \alpha) = \alpha_1 B(t + k, h_1 - k) + \alpha_2 B(t + k, h_2 - k).$$

The no-arbitrage condition is the impossibility to ensure a positive future value with zero or negative initial endowment . This is equivalent to :

$$\begin{aligned} & \{\min_{t+k}[\Pi_{t+k}(h_1 - k, h_2 - k, \alpha) \geq 0]\} \\ \Rightarrow & \{\min_t \Pi_t(h_1, h_2, \alpha) \geq 0\}, \quad \forall \alpha, \forall k \leq h_1, h_2, \end{aligned} \quad (2.2.1)$$

where  $\min_t$  is the minimum taken over the admissible values of the state variable of date  $t$ .

Condition (2.1) provides a restriction on zero-coupon prices, only if  $\alpha_1$  and  $\alpha_2$  have opposite sign. Thus, without loss of generality, we can choose  $\alpha_1 = 1, \alpha_2 = -\alpha, \alpha > 0$ , say.

**Proposition 2.1 :** Under model (2.1.1) and Assumptions A.1-A.3, the buy and hold strategies based on two zero-coupon bonds do not feature arbitrage opportunity if and only if the function  $c^*(h) = hc(h)$  is such that :  $c^*(h + 1) - c^*(h)$  is a nonnegative increasing function of  $h$ .

**Proof :** We have :

$$\begin{aligned} \Pi_{t+k}(h_1 - k, h_2 - k, \alpha) &= \exp[-(h_1 - k)X_{t+k} - c^*(h_1 - k)] \\ &\quad - \alpha \exp[-(h_2 - k)X_{t+k} - c^*(h_2 - k)] \\ &= B(t + k, h_1 - k) \\ &\quad \{[1 - \alpha \exp[-(h_2 - h_1)X_{t+k}]] \exp[-c^*(h_2 - k) + c^*(h_1 - k)]\}. \end{aligned}$$

Since  $\alpha \geq 0, h_2 \geq h_1$ , we deduce that :

$$\min_{t+k} \Pi_{t+k}(h_1 - k, h_2 - k, \alpha) \geq 0 \text{ if and only if } 1 - \alpha \exp[-c^*(h_2 - k) + c^*(h_1 - k)] \geq 0.$$

Therefore,  $\min_{t+k} \Pi_{t+k}(h_1 - k, h_2 - k, \alpha)$  is nonnegative if and only if  $\alpha \leq \exp[c^*(h_2 - k) - c^*(h_1 - k)]$ . Similarly,  $\min_t \Pi_t(h_1, h_2, \alpha)$  is nonnegative positive if and only if  $\alpha \leq \exp[c^*(h_2) - c^*(h_1)]$ .

Thus the no-arbitrage condition is satisfied if and only if,

$$\{\alpha \leq \exp[c^*(h_2 - k) - c^*(h_1 - k)]\} \Rightarrow \{\alpha \leq \exp[c^*(h_2) - c^*(h_1)]\},$$

which is equivalent to :

$$c^*(h_2 - k) - c^*(h_1 - k) \leq c^*(h_2) - c^*(h_1), \forall k \leq h_1 \leq h_2. \quad (2.2.2)$$

i) It is easily checked that condition (2.2) above is equivalent to the fact that the function  $c^*(h_2 + k) - c^*(h_1 + k)$  is increasing in  $k$  for any  $h_2 \geq h_1$ .

ii) Finally, by noting that :

$$\begin{aligned} c^*(h_2 + k) - c^*(h_1 + k) &= [c^*(h_2 + k) - c^*(h_2 - 1 + k)] + [c^*(h_2 - 1 + k) - c^*(h_2 - 2 + k)] \\ &\quad + \dots + [c^*(h_1 + 1 + k) - c^*(h_1 + k)], \end{aligned}$$

we get the increasingness condition.

To prove the nonnegativity, we have to check that  $c^*(2) - c^*(1) = 2c(2)$  is nonnegative (since  $c^*(1) - c^*(0) = 0$ ). This is a direct consequence of Assumption A.3.

QED

**Corollary 2.1 :** The no-arbitrage condition of Proposition 2.1 is satisfied if and only if the sequence  $[c(h)]$  is a sequence of Cesaro means of a nonnegative increasing function.

**Proof :** We have :

$$c(h) = c^*(h)/h = \frac{1}{h} \sum_{l=1}^h \Delta c^*(l),$$

with  $\Delta c^*(l) = c^*(l) - c^*(l-1)$ .

The result follows from Proposition 2.1.

QED

**Corollary 2.2 :** Under model (2.1.1), Assumption A1-A3 and no-arbitrage, function  $c^*$  is superadditive, that is,

$$c^*(h_1) + c^*(h_2) \leq c^*(h_1 + h_2), \forall h_1, h_2 \in \mathbb{N} - \{0\}.$$

**Proof :** Indeed, let us consider the special case of inequality (2.2) for  $k = h_1$ . We get :

$$c^*(h_2 - h_1) \leq c^*(h_2) - c^*(h_1), \quad \forall h_1 \leq h_2.$$

QED

This condition was expected. Indeed, under Assumption A.3 the lower bound of the support of  $r(t, h)$  is equal to  $c(h)$ . It has been proved in Gourieroux-Monfort (2011) that  $h$  times this lower bound, that is,  $c^*(h) = hc(h)$  is necessarily superadditive under no-arbitrage condition.

No-arbitrage requires the function  $\Delta^*(h) = c^*(h) - c^*(h - 1)$  to be non negative increasing for all  $h \in \mathbb{N} - \{0\}$ . Let us for instance assume that  $c^*(\bar{h}) - c^*(\bar{h} - 1) > c^*(\bar{h} + 1) - c^*(\bar{h}) > 0$  for a single time-to-maturity  $\bar{h}$ . Then, the portfolio composed at time  $t$  by  $\alpha_1 = 1$  zero-coupon bond with maturity  $\bar{h}$  and  $\alpha_2 = -\frac{B(t, \bar{h})}{B(t, \bar{h} + 1)}$  bond with maturity  $\bar{h} + 1$  is worthless at time  $t$ , but have a positive value  $\Pi_{t+1}$  at time  $t + 1$  with probability 1 :

$$\begin{aligned} \Pi_{t+1} &= B(t + 1, \bar{h} - 1) \left[ 1 - \frac{\alpha_2}{\alpha_1} \exp \left( -X_t - c^*(\bar{h}) + c^*(\bar{h} - 1) \right) \right] \\ &= B(t + 1, \bar{h} - 1) \left[ 1 - \exp \left( -X_{t+1} + \left[ c^*(\bar{h} + 1) + c^*(\bar{h}) \right] - \left[ c^*(\bar{h}) + c^*(\bar{h} - 1) \right] \right) \right] \\ &> 0, \text{ for all } X_{t+1} \geq 0. \end{aligned}$$

## 2.3 BEHAVIOR OF THE LONG-TERM INTEREST RATE

In this section, we investigate the implication of level factor for the modeling of long-term rates under no-arbitrage.

**Proposition 2.2 :** Under model (2.1.1) and Assumptions A.1-A.3, we get one of the two following cases :

i)  $r(t, \infty) = +\infty$  :

ii)  $r(t, \infty) = X_t + c_\infty$ , where  $c_\infty$  is a given positive constant.

**Proof :** Since  $\Delta c^*(h)$  is nonnegative increasing, we have either  $\liminf_{h \rightarrow \infty} \Delta c^*(h) = \infty$ , or  $\liminf_{h \rightarrow \infty} \Delta c^*(h) = \limsup_{h \rightarrow \infty} \Delta c^*(h) = c_\infty < \infty$  say. Since  $\Delta c^*(h)$  is a non-negative increasing function, we deduce that the Cesaro mean  $[c(h)]$  is such that :

$$c(h) = \frac{1}{h} \sum_{l=1}^h \Delta c^*(l) \leq \Delta c^*(h), \forall h,$$

$$\text{and } c(h) \geq \frac{1}{h} \sum_{l=k+1}^h \Delta c^*(l) \geq \frac{h-k}{h} \Delta c^*(k), \forall k \leq h.$$

These two inequalities explain why the sequences  $[c(h)]$  and  $[\Delta c^*(h)]$  have the same asymptotic behavior. For instance, let us assume that  $\lim_{h \rightarrow \infty} \Delta c^*(h) = +\infty$ . Then, from the second inequality, we get :

$$\liminf_{h \rightarrow \infty} c(h) \geq \Delta c^*(k), \forall k,$$

which implies  $\liminf_{h \rightarrow \infty} c(h) \geq +\infty$ . We deduce that  $\lim_{h \rightarrow \infty} c(h) = +\infty$ . When  $\lim_{h \rightarrow \infty} \Delta c^*(h) = c_\infty$ , the joint use of the two inequalities shows that  $\liminf_{h \rightarrow \infty} c(h)$  and  $\limsup_{h \rightarrow \infty} c(h)$  exist and are equal to  $c_\infty$ .

QED

Proposition 2.2 shows that the case, where the long-term interest rate does not exist due, for instance, to a periodic asymptotic behavior of function  $c$ , has been eliminated.

Proposition 2.2 concerns the limiting behavior of the long run spot interest rate when the whole term structure moves by parallel shifts. The instantaneous forward interest rate is given by :

$$f(t, h) = hr(t, h) - (h - 1)r(t, h - 1).$$

Under model (2.1.1), the instantaneous forward rate is equal to :

$$f(t, h) = X_t + c^*(h) - c^*(h - 1), \forall t, h.$$

It is not a constant function of time. In particular, if  $\lim_{h \rightarrow \infty} \Delta c^*(h)$  exists, the long run instantaneous forward interest rate also exists and is stochastic.



## 2.4 RISK-NEUTRAL FACTOR DYNAMIC

Let us now analyze the factor dynamics.

**Proposition 2.3 :** Under model (2.1.1), no-arbitrage opportunities and Assumptions A.1-A.3, the factor process is a Markov process under the risk-neutral probability  $\mathbb{Q}$  and we have :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-hX_{t+1})] = \exp[-hX_t + c^*(h) - c^*(h+1)].$$

**Proof :** Under no-arbitrage, we have :

$$B(t, h+1) = \mathbb{E}_t^{\mathbb{Q}}\{\exp[-r(t, 1)]B(t+1, h)\}, \forall h,$$

or, equivalently :

$$\exp[-(h+1)r(t, h+1)] = \exp[-r(t, 1)]\mathbb{E}_t^{\mathbb{Q}}\{\exp[-hr(t+1, h)]\}, \forall h.$$

By decomposition (2.1.1), we deduce :

$$\exp[-(h+1)X_t - c^*(h+1)] = \exp(-X_t)\mathbb{E}_t^{\mathbb{Q}}\{\exp[-hX_{t+1} - c^*(h)]\},$$

or :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-hX_{t+1})] = \exp[-hX_t + c^*(h) - c^*(h+1)], \forall h.$$

For a nonnegative variable, the knowledge of the Laplace transform for negative integer characterizes the distribution <sup>53</sup>. We deduce that the conditional distribution of  $X_{t+1}$  given its past depends on the past by means of the most recent observation. This is the Markov property and Proposition 2.3 follows.

QED

---

<sup>53</sup>Indeed, let us denote  $Z = \exp(-X)$ . Variable  $Z$  takes values in  $(0, 1)$ . Thus, for any argument  $u$ , the series  $\sum_{h=0}^{\infty} \frac{E(Z^h)(iu)^h}{h!}$  is uniformly absolutely convergent. We deduce that the characteristic function  $\psi(u) = E[\exp(iuZ)]$  exists [see Feller (1971), Vol2, p430].

The conditional log-Laplace transform is an affine function of the current value of the process. This is exactly the definition of a Compound Autoregressive (CaR) process [see Darolles, Gouriéroux, Jasiak (2006)], also called Affine process in continuous time [Duffie, Kan (1996), Duffie, Filipovic, Schachermayer (2003)].

**Proposition 2.4**<sup>54</sup> : Under model (2.1.1), no-arbitrage opportunities, and Assumptions A.1-A.3, the level factor process is a strong random walk under  $\mathbb{Q}$  :

$$X_{t+1} = X_t + \varepsilon_{t+1},$$

where  $(\varepsilon_t)$  is under  $\mathbb{Q}$  a sequence of nonnegative i.i.d. variables with Laplace transform :

$$\psi_\varepsilon(h) = \mathbb{E}^{\mathbb{Q}}[\exp(-h\varepsilon_t)] = \exp[c^*(h) - c^*(h+1)].$$

**Proof :** Let us denote  $\varepsilon_{t+1} = X_{t+1} - X_t$ . From Proposition 2.3, we deduce that :  $\mathbb{E}_t^{\mathbb{Q}}[\exp(-h\varepsilon_{t+1})] = \exp[c^*(h) - c^*(h+1)]$ . This shows that the conditional distribution of  $\varepsilon_{t+1}$  is independent of the past and provides the form of its Laplace transform. Moreover,  $\varepsilon$  is nonnegative, since by Assumption A.3,  $X_t$  can be arbitrary close to zero. In this case  $\varepsilon_{t+1} = X_{t+1}$ , which is nonnegative.

QED

Since  $\varepsilon$  is nonnegative,  $\psi_\varepsilon(h)$  is smaller than 1 and a decreasing function of  $h$ . We deduce that  $c^*(h+1) - c^*(h)$  is a nonnegative increasing function of  $h$  (which is Proposition 2.1). We also get the following Corollaries :

**Corollary 2.3 :** Model (2.1.1) is compatible with the no-arbitrage condition if and only if function  $c^*$  is such that :  $\exp[c^*(h) - c^*(h+1)]$  is the Laplace transform of a positive variable.

---

<sup>54</sup>Proposition 2.4 contradicts Theorem 4 in Ingersoll, Skelton, Weil (1978), where it is said that any parallel shift in a term structure is not arbitrage free. The random walk models in Proposition 2.4 are both compatible with parallel shift and no-arbitrage. This contradiction is due to a misleading proof in ISW (1978), p635, 13, where the effect of diminishing time-to-maturity is omitted when computing the future value of the portfolio of zero-coupon bonds. In some sense, they have implicitly assumed a flat term structure  $c(h) = 0$  [see the discussion in Appendix 2.1].

**Corollary 2.4 :** Under model (2.1.1), no-arbitrage opportunities, and Assumptions A.1-A.3, the factor process is a non decreasing function of time: the term structure cannot make uniform downward move.

Let us now come back to the behavior of the long-term interest rate. We have the following proposition :

**Proposition 2.5 :** For a strong random walk under  $\mathbb{Q}$ , the long-term interest rate exists, if and only if :

$$\lim_{h \rightarrow \infty} \{-\log \mathbb{E}[\exp(-h\varepsilon)]\} = -\log P[\varepsilon = 0] = c_\infty < \infty;$$

then the long run interest rate is equal to :

$$r(t, \infty) = X_t + c_\infty.$$

**Proof :** The first condition concerning  $\lim_{h \rightarrow \infty} \{-\log \mathbb{E}[\exp(-h\varepsilon)]\} = c_\infty < \infty$  is a direct consequence of Proposition 2.4 and the proof of Proposition 2.2.

Moreover, we have

$$\mathbb{E}[\exp(-h\varepsilon)] = P[\varepsilon = 0] + \int \mathbf{1}_{x>0} \exp(-xh) dF(x).$$

But  $\lim_{h \rightarrow \infty} \exp(-hx) = 0, \forall x > 0$ , and since  $\exp(-hx) \in (0, 1)$ , we deduce by Beppo-Levi theorem that  $\lim_{h \rightarrow \infty} \int \mathbf{1}_{x>0} \exp(-hx) dF(x) = 0$ . The result follows.

QED

In this framework, the long-term rate exists, is stochastic and provides the same information as the underlying factor. This contradicts Lemma 3 in El Karoui, Frachot, Geman (1998), which asserts that the long-term yield (if it exists) cannot be stochastic in a one-factor model.

The need for an innovation with point mass at zero explains the strange behavior of the long run interest rate, in affine models with a level factor following a Gaussian random walk, even if this factor is not positive [see e.g. Christensen, Diebold, Rudebusch (2010)]. In this framework, the long run interest rate is equal to  $-\infty$ .

To illustrate Corollary 2.3, let us consider a random walk with a Poisson distributed innovation  $\varepsilon_t \sim \mathcal{P}(\lambda)$ . We have :

$$\begin{aligned}\psi_\varepsilon(h) &= \exp\{-\lambda[1 - \exp(-h)]\}, \\ -\log P[\varepsilon = 0] &= \lambda,\end{aligned}$$

and the interest rate with time-to-maturity  $h$  is :

$$r(t, h) = X_t + \lambda\left\{\frac{h-1}{h} + \frac{1}{h}\left[1 - \frac{1 - \exp(-h)}{1 - \exp(-1)}\right]\right\}.$$

We check that :  $\lim_{h \rightarrow \infty} \psi_\varepsilon(h) = r(t, \infty) - X_t = -\log P[\varepsilon = 0] = \lambda$ .

The results above concern the risk-neutral dynamics. It is known that the historical and risk-neutral dynamics are weakly linked [see Rogers (1977)]. For instance, the historical dynamic of  $(X_t)$  is not necessarily affine, and does not necessarily feature a unit root. Nevertheless, the historical and risk-neutral distributions have a same support : in particular the process  $(X_t)$  can never fall under the historical probability<sup>55</sup> and the probability that  $X_{t+1} = X_t$  is nonzero if the long run interest rate exists. Similarly, when it exists, the long run interest rate is also an nondecreasing function of time. Therefore under model (2.1.1), either the long-term spot interest rate does not exist, or if it exists it can never fall.<sup>56</sup>

---

<sup>55</sup>The empirical literature on term structure models with level factor identify on the contrary a decreasing trend in the level factor dynamics [see for instance Diebold, Rudebush Aruoba (2006), p 312 fig. 2].

<sup>56</sup>Several authors argue that this property of the long-term spot rate is a consequence of no-arbitrage [Dybvig, Ingersoll, Ross (1996), El Karoui, Frachot, Geman (1998)], but prove this property under additional assumptions. These assumptions can be a predetermined long interest rate [DIR (1996)], or a long rate satisfying a diffusion equation [EFG (1998)].

## 2.5 TERM STRUCTURE MODEL WITH LEVEL AND SLOPE FACTORS

At this stage, we may think that the deficiencies of the level factor modeling is specific of the single factor model and might disappear if a second factor is introduced, such as a slope factor. We will see below that this is not the case. The steps of the proof are as follows :

- i) We first show that we have necessarily an affine term structure (Section 2.5.1).
- ii) Then, we show that, except for some special baseline slope functions, the risk-neutral factor dynamics is necessarily affine (Section 2.5.2).
- iii) In Sections 2.5.3-2.5.4, we check that an affine dynamics for the level and slope factor is not arbitrage free.
- iv) Finally, we consider the cases of special baseline functions and non-affine factor dynamics in Section 2.5.5.

### 2.5.1 THE AFFINE TERM STRUCTURE

To highlight the arguments, let us consider now a continuous time model with two factors. Then we can write :

$$r(t, h) = g(X_t, Z_t, h), \quad \forall t, h \geq 0. \quad (2.5.1)$$

In order to allow for independent shocks on the level and slope factor, we need conditions on the joint support of variables  $X_t, Z_t$ . In particular, to assimilate the magnitude of the shock  $\delta x$  (resp.  $\delta z$ ) with an increase in  $X$  (resp.  $Z$ ), we need a property of invariance of the support. This condition is summarized in Assumption A\*1 below.

**Assumption A\*1 :**

i) The support of the historical (risk-neutral) conditional distribution of  $X_t, Z_t$  given its past  $\underline{X}_{t-h}, \underline{Z}_{t-h}$  is  $\mathcal{X} \times \mathcal{Z}$ , for any  $t, h \geq 0$ .

ii) The supports  $\mathcal{X}$  and  $\mathcal{Z}$  are additive groups.

Since the support of the historical and risk-neutral conditional distributions are the same, the condition is valid for both of them.

**Proposition 2.6 :** Under Assumption A\*1, a two factor model with a level and a slope factor can always be written as :

$$r(t, h) = X_t + Z_t \beta(h) + \gamma(h), \quad \forall t, h \geq 0.$$

where  $\beta(\cdot)$  is an increasing function,  $\beta(0) = 1, \gamma(0) = 1$ .

**Proof :**  $X_t$  and  $Z_t$  are level and slope factors if and only if they can be shocked separately (under the historical distribution), with a drift and a slope effects, respectively, on the term structure.

i) Let us first consider a shock  $\delta x$  on  $X_t$ . By definition of the level factor, we get :

$$g(X_t + \delta x, Z_t, h) = g(X_t, Z_t, h) + d(X_t, Z_t, \delta x), \quad \forall X_t, Z_t, \delta x, h,$$

where the drift effect can depend on the environment. Without loss of generality, we can assume that the level factor can take value zero. Thus we get :

$$g(\delta x, Z_t, h) = g(0, Z_t, h) + d(0, Z_t, \delta x), \quad \forall Z_t, \delta x, h,$$

or equivalently, we can write :

$$g(X_t, Z_t, h) = g_1(X_t, Z_t) + g_2(Z_t, h), \quad \text{say.} \quad (2.5.2)$$

ii) Let us now apply a shock  $\delta z$  on  $Z_t$ . We get :

$$\begin{aligned} g(X_t, Z_t + \delta z, h) &= g_1(X_t, Z_t + \delta z) + g_2(Z_t + \delta z, h), \\ &= g_1(X_t, Z_t) + g_2(Z_t, h) + s(X_t, Z_t, \delta z) \beta(h), \quad \forall X_t, Z_t, \delta z, h, \end{aligned}$$

by denoting  $\beta(h)$  the baseline slope effect on the term structure.

$\beta(\cdot)$  has to be monotonous, for instance increasing, for the slope interpretation, and the magnitude of the slope effect can depend on the environment. We can always assume that  $Z_t$  can take the value zero. Then we get :

$$g(X_t, \delta z, h) = g_1(X_t, 0) + g_2(0, h) + s(X_t, 0, \delta z)\beta(h), \quad \forall X_t, \delta z, h,$$

or equivalently :

$$g(X_t, Z_t, h) = g_1(X_t) + g_2(X_t, Z_t)\beta(h) + \gamma(h), \quad \text{say.} \quad (2.5.3)$$

iii) Let us finally consider the expression (2.5.3) and apply a shock on the level factor. The effect of this shock equals to :

$$g_1(X_t + \delta x) - g_1(X_t) + [g_2(X_t + \delta x, Z_t) - g_2(X_t, Z_t)]\beta(h),$$

has to be independent of  $h$ . This implies that  $g_2(X_t, Z_t)$  is independent of  $X_t$ .

To summarize we can write :

$$g(X_t, Z_t, h) = g_1(X_t) + g_2(Z_t)\beta(h) + \gamma(h),$$

or equivalently :

$$g(X_t, Z_t, h) = X_t + Z_t\beta(h) + \gamma(h),$$

since  $X_t$  (resp.  $Z_t$ ) is defined up to a transformation.

Finally, if  $\gamma(0) \neq 0$ ,  $\beta(0) \neq 1$ , we can always perform a drift in the definition of factor  $X$ :  $X_t \rightarrow X_t + \gamma(0)$ , and introduce a multiplicative scale on factor  $Z$ :  $Z_t \rightarrow Z_t\beta(0)$ , to satisfy the conditions  $\beta(0) = 1$ ,  $\gamma(0) = 0$ .

QED

Then, the level and slope interpretations of the factors imply an affine term structure model, with a constant baseline term structure for the level factor, and an increasing baseline term structure for the slope factor.

The instantaneous interest rate  $r_t = r(t, 0) = X_t + Z_t$  is the sum of the level and slope factors.

### 2.5.2 THE RISK-NEUTRAL FACTOR DYNAMICS

We will now use the affine term structure in Proposition 2.6 to restrict the specification of the risk-neutral factor dynamics. To simplify the discussion, we assume that the factor process satisfies a stochastic differential system.

**Assumption A\*2** : Under the risk-neutral distribution, the bivariate process  $Y_t = (X_t, Z_t)'$  satisfies the stochastic differential equation :

$$dY_t = \mu(Y_t)dt + \Sigma^{1/2}(Y_t)dW_t,$$

where  $(W_t)$  is a standard bivariate Brownian motion,  $\mu(\cdot)$ ,  $\Sigma(\cdot)$  are the infinitesimal drift and volatility, respectively.

By assuming a continuous time model without jumps, we avoid the limit case of the special random walk encountered in Section 2.4, Proposition 2.5.

By applying the pricing formula :

$$B(t, h) = \exp[-hr(t, h)] = \mathbb{E}^{\mathbb{Q}} \left[ \exp \left( - \int_t^{t+h} r_u du \right) | Y_t \right],$$

where  $\mathbb{Q}$  denotes the risk-neutral distribution, we deduce the relationship between  $\mu, \Sigma$  and the expression of the interest rate [see e.g. Duffie (2001), Chapter 7]. Let us denote:

$$r(t, h) = g(Y_t, h),$$

where function  $g$  satisfies the partial differential equation (see Appendix 2.2) :

$$\begin{aligned} g(y, h) - g(y, 0) + h \frac{\partial g(y, h)}{\partial h} &= h \frac{\partial g(y, h)}{\partial y'} \mu(y) + \frac{1}{2} h T r \left[ \Sigma(y) \frac{\partial^2 g(y, h)}{\partial y \partial y'} \right] \\ &\quad - \frac{h^2}{2} \frac{\partial g(y, h)}{\partial y'} \Sigma(y) \frac{\partial g(y, h)}{\partial y}, \quad \forall y, h. \end{aligned} \quad (2.5.4)$$

In our framework, we have:  $g(y, h) = x + \beta(h)z + \gamma(h)$ , with  $\beta(0) = 1$ ,  $\gamma(0) = 0$ .

Therefore, differential system (2.5.4) reduces to :



$$\begin{aligned}
 [\beta(h) - 1]z + \gamma(h) + h \left[ \frac{d\beta(h)}{dh} z + \frac{d\gamma(h)}{dh} \right] &= h \begin{bmatrix} 1, & \beta(h) \end{bmatrix} \mu(y) \\
 &- \frac{h^2}{2} \begin{bmatrix} 1, & \beta(h) \end{bmatrix} \Sigma(y) \begin{bmatrix} 1 \\ \beta(h) \end{bmatrix}, \forall y, h.
 \end{aligned} \tag{2.5.5}$$

For given  $y$ , we get an infinite dimensional linear system of equations in  $\mu_1(y), \mu_2(y), \sigma_{11}(y), \sigma_{12}(y), \sigma_{22}(y)$ , that are the elements of  $\mu(y), \Sigma(y)$ , respectively. Then, following Duffie, Kan (1996), we deduce the necessary form of the drift and volatility functions.

**Proposition 2.7 :** If the baseline slope  $\beta(h)$  is not an affine function of  $h$ , and is not proportional to  $\beta(h) = 1 + \frac{\sqrt{h^2 + (a_2 + a_3 h)^2} - (a_2 + a_3 h)}{h}$ , with  $a_2 > 0$ , the drift and volatility functions are necessarily affine functions of  $z$  only under Assumptions A\*1-A\*2, and no-arbitrage opportunity.

**Proof :** see Appendix 2.3.

Thus, we get an affine risk-neutral dynamics for the factor process :

$$\begin{pmatrix} dX_t \\ dZ_t \end{pmatrix} = (\mu_0 + \mu_1 Z_t) + (\Sigma_0 + \Sigma_1 Z_t)^{1/2} dW_t, \text{ say,} \tag{2.5.6}$$

where  $\mu_0, \mu_1$  are bivariate vectors and  $\Sigma_0, \Sigma_1$  are (2,2) symmetric matrices.

As noted in Duffie, Kan (1996), p. 386, an affine term structure of interest rate implies an affine risk-neutral factor dynamics under some rank conditions. In our framework the rank condition is equivalent to the conditions on the baseline slope function in Proposition 2.7.

### 2.5.3 CONSTRAINTS ON THE AFFINE FACTOR DYNAMICS

Let us now discuss the constraints implied by the positivity of the factor volatility matrix and by the nonnegativity of the instantaneous rate.

**Assumption A\*3 :** The instantaneous interest rate  $r(t, 0) = X_t + Z_t$  is nonnegative, and can reach value zero.

By introducing this new assumption, we are limiting the set of possible factor values to  $\{\mathcal{X} \times \mathcal{Z}\} \cap \{(x, z) : x + z \geq 0\}$ . In other words, we only allow for shocks on either level, or slope factor, keeping nonnegative the instantaneous rate.

**Proposition 2.8 :** Under Assumptions A\*1-A\*3 and no-arbitrage opportunities, process  $(Z_t)$  is a drifted Cox-Ingersoll-Ross process.

**Proof :**

i) First note that  $\Sigma_1 \neq 0$ . Otherwise,  $Y_t = (X_t, Z_t)'$  would be a multivariate Ornstein-Uhlenbeck process, and  $r(t, 0)$  would be conditionally Gaussian, which contradicts Assumption A\*3.

ii) Let us now denote  $\sigma_1(2, 2)$  the (2,2) element of  $\Sigma_1$ . If  $\sigma_1(2, 2) = 0$ , the process  $(Z_t)$  would be an Ornstein-Uhlenbeck process, and would take any value in  $(-\infty, +\infty)$ .

Since  $\sigma_0(1, 1) + \sigma_1(1, 1)Z_t$  has to be nonnegative for any value of  $Z_t$ , we deduce that  $\sigma_1(1, 1) = 0$ . Moreover the positivity of  $\Sigma_0 + \Sigma_1 z_t$  for large  $z_t$  implies  $\Sigma_1 \gg 0$  and thus  $\sigma_1(1, 2) = 0$  by Cauchy-Schwarz.

We deduce that the condition  $\sigma_1(2, 2) = 0$  implies  $\Sigma_1 = 0$ , which contradicts Assumption A\*3 by i).

iii) To summarize the slope process  $(Z_t)$  satisfies the stochastic differential equation :

$$dZ_t = (\mu_0(2) + \mu_1(2)Z_t) dt + (\sigma_0(2, 2) + \sigma_1(2, 2)Z_t)^{1/2} dW_t^*,$$

where  $(W_t^*)$  is a one-dimensional Brownian motion and  $\sigma_1(2, 2) > 0$ . Therefore,  $(Z_t)$  is necessarily a drifted CIR process defined on the interval  $\mathcal{Z}^* = \left(-\frac{\sigma_0(2, 2)}{\sigma_1(2, 2)}, \infty\right)$ .

QED

#### 2.5.4 THE ABSENCE OF SOLUTION WITH AFFINE LEVEL AND SLOPE FACTORS

**Proposition 2.9 :** Under Assumptions A\*1-A\*3, a model with level and slope factors with an affine risk-neutral dynamics is not arbitrage free.

**Proof :**

Under Assumptions A\*1-A\*3, and the absence of arbitrage opportunity, the factor process satisfies the affine dynamics (2.5.6). Thus, we get an affine term structure model in which the sensitivity coefficients of the factors, that are 1 and  $\beta(h)$  satisfy a Riccati equation. In our framework this equation is [see Duffie (2001), Chapter 7, eq. (31)] :

$$\left( \begin{array}{c} 0 \\ \frac{d\beta(h)}{dh} \end{array} \right) = \left( \begin{array}{c} 1 \\ 1 \end{array} \right) - \mu_1\beta(h) - \frac{1}{2} \left( \begin{array}{c} 0 \\ [1, \beta(h)] \Sigma_1 \left[ \begin{array}{c} 1 \\ \beta(h) \end{array} \right] \end{array} \right),$$

The second equation :

$$\frac{d\beta(h)}{dh} = 1 - \mu_1(2)\beta(h) - \frac{1}{2} [1, \beta(h)] \Sigma_1 \left[ \begin{array}{c} 1 \\ \beta(h) \end{array} \right], \text{ with } \sigma_1(2, 2) > 0,$$

is the equation corresponding to a drifted Cox-Ingersoll-Ross process, whose solution involves a rational function of exponential functions of  $h$  [see e.g. Duffie (2001), Chapter 7, eq(11)]<sup>57</sup>. The first equation is  $1 = \mu_1(1)\beta(h)$ . This leads to a contradiction since function  $\beta(h)$  is not constant.

QED

---

<sup>57</sup>The expression of  $\beta(h)$  does not correspond to the expression of the sensitivity coefficient of  $(Z_t)$  in a term structure model with single factor  $(Z_t)$  satisfying the dynamics of Proposition 2.8, except if  $\sigma_1(1, 1) = \sigma_1(1, 2) = 0$ . This shows that factor  $(X_t)$  generally matters even if  $(Z_t)$  admits an exogenous dynamics.

### 2.5.5 NON-AFFINE LEVEL AND SLOPE FACTORS

Finally, let us consider the special patterns of the slope function  $\beta$  appearing in Proposition 2.7. From differential equation (2.5.5), we get :

$$\begin{aligned} \frac{d[h\gamma(h)]}{dh} &= -[\beta(h) - 1]z - h\frac{d\beta(h)}{dh}z + h \left[ 1, \beta(h) \right] \mu(y) \\ &- \frac{h^2}{2} \left[ 1, \beta(h) \right] \Sigma(y) \begin{bmatrix} 1 \\ \beta(h) \end{bmatrix}, \forall y, h. \end{aligned} \quad (2.5.7)$$

i) Let us first consider the case of a slope affine function of  $h$ , i.e.  $\beta(h) = b_0 + b_1h$ , with nonnegative real  $b_1$  to ensure that the slope baseline term structure is increasing in  $h$ . In this case, the linear baseline term structure  $\beta(h)$  forces very long-term rates to be unboundedly large (and positive). In particular, the limiting interest rate  $r(t, \infty) = \lim_{h \rightarrow \infty} r(t, h)$  is infinite, at any time  $t$ . It is interesting to discuss this special case, since the affine slope baseline was historically the first proposed specification [see Appendix 2.1 iv)].

ii) Let us now consider the function  $\beta(h) = 1 + \frac{\sqrt{h^2 + (a_2 + a_3h)^2} - (a_2 + a_3h)}{h}$ , with  $a_2 > 0$ . This function tends to a finite limit when  $h$  tends to infinity. We deduce from equation (2.5.7), that for large  $h$ , the derivative  $\frac{d[h\gamma(h)]}{dh}$  is of order  $h^2$ , whenever there exists a state  $y$  such that  $\Sigma(y) \neq 0$ . Under this condition, which ensures a nondegenerate 2-factor model,  $\gamma(h)$  is the dominant term in the expression of  $r(t, h)$ . In particular, the long-term rate can be unbounded.

To summarize, let us introduce the following assumption of finiteness of the long-term rate.

**Assumption A\*4 :**  $\mathbb{P} \left[ \limsup_{h \rightarrow \infty} |r(t, h)| < \infty \right] = 0$ .

We have the Proposition below :

**Proposition 2.10 :** Under Assumptions A\*1-A\*4, a model with level and slope factors is not arbitrage free.

## 2.6 CONCLUDING REMARKS

A large part of the term structure literature interprets the first factors as a level factor, a slope factor, a curvature factor, respectively. Initially this interpretation relies on the pattern of the weights that each factor assigns to the different maturities: the level factor has almost equal weights, the slope factor has monotonic weights, but this interpretation has also been used to discuss the immunization of bond portfolios with respect to specific shocks on the term structure (see the discussion in Appendix 2.1). Level factor models have also the interesting property to imply stochastic limiting rate. However, the literature does not check if these interpretations of the first factors is coherent, that is compatible with no-arbitrage opportunity.

The aim of this chapter was to investigate the consistency of such term structure model with no-arbitrage assumptions. To discuss this point, we consider successively a single factor model with a level factor, and a 2-factor model with level and slope factors. None of these models are compatible with the positivity of interest rates, the finiteness of the long-term rate and no-arbitrage restrictions.

In particular, we show that under no-arbitrage assumptions, term structure models with level factor require very specific factor dynamics to avoid explosive long-term rates. The following chapter builds on this result to propose, by restricting the level factor's dynamics, an arbitrage-free term structure model, where the limiting rate is stochastic.

# Stable Term Structure Models with Stochastic Long-Term Rates.\*

---

## 3.1 INTRODUCTION

This chapter focuses on the modeling of the very long-end of the yield curve of interest rates. Following the previous chapter, it aims to introduce a new class of term structure model, in which the implied yield curve's limit is stochastic, while most of the existing term structure models assume a constant very long-term rate.

As a consequence, these models imply a relatively fast decreasing volatility in the maturity of the rates, while it seems that the volatility of the ultra long-term yields (with 30 years time-to-maturity or more) is fairly comparable with the rates of commonly labeled “long” bonds with 10 years to maturity, even higher on some periods [see Figure 2 in the General Introduction of this dissertation]. Besides, these models may not be appropriate for the pricing of very long-term cash-flows<sup>58</sup>, which requires to extrapolate observed term structures.

It is however challenging to escape from the constant yield curve's limit due to the assumption of no-arbitrage. This is for instance emphasized by the internal inconsistencies of the Brennan and Schwartz (1982) model of long-term rate [see Hogan (1993)]<sup>59</sup>, or by the fact that this limit cannot

---

\*This chapter is derived from an ongoing working paper with Christian Gourieroux.

<sup>58</sup>Such as the ones generated by long-term infrastructure projects, or retirement systems.

<sup>59</sup>Notice that the long-term rate in the Brennan and Schwartz model (1982) is the yield on a consol bond, which pays perpetually fixed coupons. In this dissertation, we focus on zero coupon

decrease over time under no-arbitrage<sup>60</sup> [see Dybvig, Ingersoll, Ross (1996), Hubalek, Klein, Teichman (2002), and more recently Kardaras and Platen (2012)]. No-arbitrage term structure model with stochastic long-term rate are rare in the literature [see Yao (1998) for a (partial) list of term structure models and their implication for long-term rates]. Ingersoll, Skelton, Weil (1978) provided an example of term structure model with level factor and stochastic limiting rate, whose variations are discrete. Moreover, Yao (1998) conjectured that a decreasing-in-time pricing factor could be a way to model stochastic limiting rate under no-arbitrage assumptions, though without providing explicit pricing formula, nor testing the empirical properties of such model. In this chapter, we build on these two contributions and introduce a new class of Affine Term Structure Models (ATSM henceforth), based on a conditional positive stable risk-neutral distribution for the short-term rate, in which the long-term end-point is stochastic, definite, with continuous variations.

Stable positive processes are presented and discussed in Section 3.2. We first define the positive stable variable, as a limit of a Poisson compound variable with an infinitely large number of small jumps. This highlights the importance of the Laplace transform of the stable variable, with a simple exponential form. Then, we consider the moving average process based on a positive stable white noise. We discuss the stationarity conditions and derive the transitions of the process at any horizon as drifted positive stable distributions. The Stable AutoRegressive process of order 1 [SAR(1)] is presented as an illustration.

The term structure models based on positive stable distribution are introduced in Section 3.3. We first introduce the Stable Autoregressive Term Structure Model (SARTSM), when the short-term interest rate follows a SAR(1) model under the risk-neutral probability. Then, we consider a random walk model with the scale parameter of the stable innovation being time-decreasing to construct an almost flat term structure.

Affine Term Structure Models, which focus on the whole set of existing maturities in the government bond market (e.g. from 1 year to 30 years in the US) usually incorporate 4 to 5 stochastic factors, including two “curvature” factors, whose influence on the yield curve peaked on definite time invariant

---

bonds.

<sup>60</sup>It differs from most of the securities for which the constraints implied by no-arbitrage are rather loose [see Cochrane (2001)].

segments of the yield curve. These two curvature factors give to the model the flexibility to generate yield curves with 2 humps [as we sometimes observe in the US, see Dubecq, Gourieroux (2010)], and to shock specific segments of the yield curve. In Section 3.4, we review two ATSM with 5 factors (including 2 curvature factors), which feature different limiting behavior of the yields: the yield curve's end-point is diverging in the first model, stochastic and definite in the second one. Both models are very similar to the existing literature on ATSM, in the sense that the factors are assumed to be a VAR(1) process under the risk-neutral probability measure. Moreover, they both can be interpreted as arbitrage-free Nelson Siegel term structure models, the first one being the discrete time equivalent of the Arbitrage Free Generalized Nelson Siegel Term Structure Model introduced by Christensen, Diebold, Rudebush (2009). The empirical properties of both models are presented in Section 3.5 on a dataset of US T-bonds. In this application, we use the same Fama-Bliss dataset as Christensen, Diebold, Rudebush (2009), in order to make our results comparable to their estimates. We show that the model with stable factor, though more constrained, performs similarly to its unconstrained Gaussian counterpart, while providing stochastic and definite long-term rates. Section 3.6 concludes. The proofs are gathered in Appendices.



## 3.2 STABLE POSITIVE PROCESS

In this section, we first recall how to construct a positive stable variable by considering the limit of an appropriate sequence of compound variables. This interpretation is useful to derive the Laplace transform of the stable distribution, but also for simulation purpose. Then, we discuss moving average processes with stable positive innovations.

### 3.2.1 POSITIVE STABLE DISTRIBUTION

The construction of the positive stable distribution is based on the following lemma, whose proof is given in Appendix 3.1 [see also Feller (1971)].

**Lemma 1 :** Let us consider the compound variable :  $Y_\delta = \sum_{k=1}^{Z_\delta} X_{k,\delta}$ , where, for any given value  $\delta$ , the variables  $Z_\delta$ ,  $X_{\delta,k}$ ,  $k = 1, \dots$ , are independent, with  $Z_\delta$  following a Poisson distribution with parameter  $c\delta^{-\alpha}/\Gamma(1 - \alpha)$ , and  $X_{\delta,k}$  a drifted Pareto distribution with survival function :

$$\mathbb{P}[X_\delta > x] = \begin{cases} \delta^\alpha/x^\alpha, & \text{if } x > \delta, \\ 1, & \text{otherwise,} \end{cases}$$

with  $0 < \alpha < 1$ . Then,  $Y_\delta \xrightarrow{d} Y$ , when  $\delta$  tends to zero, where  $\xrightarrow{d}$  denotes the convergence in distribution, and the Laplace transform of the limiting distribution is :

$$\mathbb{E}[\exp(-uY)] = \exp(-cu^\alpha), u \geq 0.$$

The variables  $Y_\delta$  and  $Y$  take positive values and their distribution is characterized by their Laplace transform computed for nonnegative real argument  $u$  [Feller (1971)]. In the rest of the chapter, we denote  $\mathcal{S}(\alpha, c)$  the distribution with Laplace transform  $\exp(-cu^\alpha)$ . This distribution is a special case of asymmetric stable distribution with scale parameter  $c$  and characteristic exponent parameter  $\alpha$  (also called index of stability). Parameter  $c$  (resp.  $\alpha$ ) has to be strictly positive (resp. between 0 and 1) <sup>61</sup>. These distributions feature heavy tail, which magnitude is measured by exponent  $\alpha$ ; the larger

---

<sup>61</sup>As discussed in Nolan (2009), Section 3.1.3, different parameterizations of the family of stable distributions are considered in the literature. We have retained a parametrization convenient for the use of Laplace transform.

the exponent, the smaller the tail <sup>62</sup>. In the limiting case  $\alpha = 1$ , we find the point mass at  $c$ . Lemma 1 shows that the heavy tail of variable  $Y$  is obtained by summing a very large number ( $Z_\delta$ ) of small values ( $X_{\delta,k}$ ).

The density function of distribution  $\mathcal{S}(\alpha, 1)$  is given by [Feller (1971), p583] :

$$f_\alpha(y) = -\frac{1}{\pi y} \sum_{k=1}^{\infty} \left\{ \frac{\Gamma(k\alpha + 1)}{k!} (-y^{-\alpha})^k \sin(\alpha k\pi) \right\}, \quad (3.2.1)$$

and is bell-shaped [Gawronski (1984)]. In the special case of the Levy distribution corresponding to  $\alpha = 1/2$ , the density is given by :

$$f_{1/2}(y) = \frac{1}{2\sqrt{\pi}} \frac{1}{y^{3/2}} \exp\left(-\frac{1}{4y}\right), y > 0. \quad (3.2.2)$$

A Levy variable is equal to the inverse of the square of a zero-mean Gaussian variable, up to a scale factor.

### 3.2.2 MOVING AVERAGE PROCESS WITH POSITIVE STABLE INNOVATION

The analysis of moving average processes of stable innovations is based on the following Lemma, which is a direct consequence of the expression of the Laplace transform. <sup>63</sup>

**Lemma 2 :** Let us consider independent variables  $\varepsilon_h, h = 1, \dots, H$  with the same distribution  $\mathcal{S}(\alpha, c)$  and a sequence of positive scalars  $a_h, h = 1, \dots, H$ .

Then the distribution of the linear combination  $\sum_{h=1}^H a_h \varepsilon_h$  follows the stable

distribution  $\mathcal{S}(\alpha, c \sum_{h=1}^H (a_h)^\alpha)$ .

In the rest of this section, we consider a stable positive white noise  $(\varepsilon_t, t \in \mathbb{Z})$ , where  $\varepsilon_t \sim \mathcal{S}(\alpha, c)$ .

---

<sup>62</sup>There exist stable distributions with  $1 \leq \alpha \leq 2$ , whose Laplace transform exists, such as the Gaussian distributions ( $\alpha = 2$ ) [see Gupta and Waymire (1993)]. These distributions are however supported on the the whole real line.

<sup>63</sup>The property given in Lemma 2 is often used to define stable random variables [see e.g. Nolan (2009), Section 3.1.2].

**i) Infinite moving average stable process**

Let us now consider the process  $(y_t)$  defined by :

$$y_t = \sum_{h=0}^{\infty} a_h \varepsilon_{t-h},$$

where  $(a_h)$  is a sequence of positive scalars. By using Lemma 2, this process is well-defined if and only if the series  $a_h^\alpha$  is convergent :

$$\sum_{h=0}^{\infty} a_h^\alpha < \infty. \tag{3.2.3}$$

This condition implies restrictions on the behavior of  $a_h$  for large  $h$ . For instance, it is satisfied if coefficients  $a_h$  are geometrically decreasing. For a power decrease  $a_h \sim \gamma/h^d$ , say, it implies  $d > 1/\alpha$ . This shows the trade-off between the tail magnitude and the long memory effect. The existence of process  $(y_t)$  requires a shorter memory, when the tail is heavier.

Under stationarity condition <sup>64</sup> (2.3), all conditional distributions are drifted stable distributions with the same exponent  $\alpha$ . More precisely, let us denote by  $I_t$  the information including the current and past values of the white noise, the conditional distribution of  $y_{t+h}$  given  $I_t$  is such that :

$$y_{t+h} - \sum_{k=h}^{\infty} a_k \varepsilon_{t-k} | I_t \sim \mathcal{S}(\alpha, c \sum_{k=0}^{h-1} (a_k)^\alpha). \tag{3.2.4}$$

This conditional distribution is supported on  $(\sum_{k=h}^{\infty} a_k \varepsilon_{t-k}, \infty)$ .

When horizon  $h$  increases, the lower bound of the support has a reduced number of terms, whereas the scale effect increases.

**ii) Stable Autoregressive SAR(1) process**

For an AR(1) process with stable innovations,

$$y_t = \rho y_{t-1} + \varepsilon_t, \varepsilon_t \sim IIS(\alpha, c), \tag{3.2.5}$$

with  $0 \leq \rho < 1$ , we have  $a_h = \rho^h$ , and

$$\sum_{k=0}^{h-1} (a_k)^\alpha = \sum_{k=0}^{h-1} (\rho^{k\alpha}) = \frac{1 - \rho^{\alpha h}}{1 - \rho^\alpha}.$$

---

<sup>64</sup>This is stationarity in distribution, since first-and second-order moments do not exist.

We deduce that the conditional distribution of  $y_{t+h}$  given  $I_t$  is such that :

$$y_{t+h} - \rho^h y_t | I_t \sim \mathcal{S} \left( \alpha, c \frac{1 - \rho^{\alpha h}}{1 - \rho^\alpha} \right), \quad (3.2.6)$$

whereas the stationary distribution is such that :

$$y_t \sim \mathcal{S} \left( \alpha, \frac{c}{1 - \rho^\alpha} \right). \quad (3.2.7)$$

The ratio of the unconditional scale of  $y$  by the innovation scale, i.e.  $1/(1 - \rho^\alpha)$ , is an increasing function of the tail magnitude.

From (2.6), the conditional Laplace transform of process  $(y_t)$  at horizon  $h$  is given by :

$$E_t[\exp(-uy_{t+h})] = \exp[-u\rho^h y_t - c \frac{1 - \rho^{\alpha h}}{1 - \rho^\alpha} u^\alpha]. \quad (3.2.8)$$

Thus, the conditional log-Laplace transform of  $y_{t+h}$  is an affine function of the current factor value. The SAR(1) process is an example of Compound Autoregressive (CAR) process [see Darolles, Gouriéroux, Jasiak (2006), Gouriéroux (2009)]. This explains the associated Affine term structure model (see Section 3.3.1 and the discussion of its continuous time analogue in Appendix 3.3).

This conditional distribution has a support with lower bound  $\rho^h y_t$ . Therefore, when  $y_t$  is observed, we know that :

$$y_{t+h} > \rho^h y_t \iff \rho < (y_{t+h}/y_t)^{1/h}.$$

Since this inequality is valid for any  $t$  and  $h$ , we deduce that :

$$\rho < \min_{t,h} [y_{t+h}/y_t]^{1/h}. \quad (3.2.9)$$

This inequality can be used in practice to get an estimated upper bound for the  $\rho$  value and a consistent estimator of  $\rho$  if the number of observation dates tends to infinity.

To illustrate the behavior of the SAR(1) process, we simulate trajectories with length  $T = 200$ , for a same history of stable shocks and for the following values of the parameters :

$$c = 1, \alpha = 0.5, 0.8, 0.9, \rho = 0, 0.5, 0.9.$$

We clearly observe the two specific features of stable processes, that are extreme values, which can arise in cluster when  $\rho$  is positive large, and the memory effect, especially visible for large  $\rho$  and  $\alpha$ .

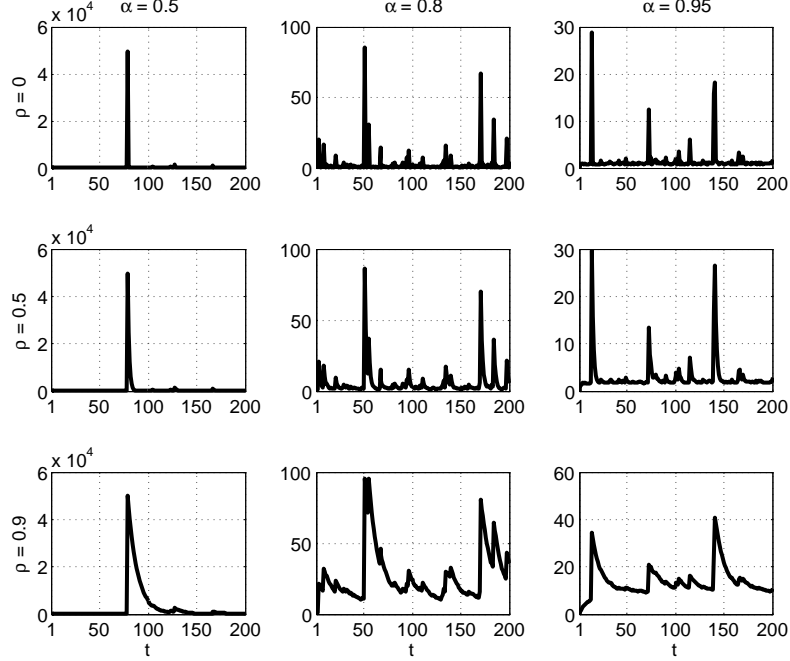


Figure 3.1: Trajectories of the SAR(1) process. Simulated trajectories of the SAR(1) process ( $y_t$ ) with length  $t = 1, \dots, 200$  for unitary scale parameter  $c = 1$ , different values of parameters  $\alpha$  and  $\rho$  and a same history of shocks.

We compute for each trajectory above, the trajectory of the historical mean  $\bar{y}_t = \frac{1}{t}(y_1 + \dots + y_t)$ ,  $t = 1, \dots, T$ . We have :

$$\bar{y}_t = \frac{1}{t} \frac{1 - \rho^t}{1 - \rho} y_1 + \sum_{k=1}^{t-1} \left\{ \frac{1}{t} \frac{1 - \rho^k}{1 - \rho} \varepsilon_k \right\},$$

and conditional on  $y_1$  :

$$\bar{y}_t - \frac{1}{t} \frac{1 - \rho^t}{1 - \rho} y_1 | y_1 \sim \mathcal{S}\left(\alpha, \frac{c}{t^\alpha} \sum_{k=1}^{t-1} \left(\frac{1 - \rho^k}{1 - \rho}\right)^\alpha\right).$$

We see from the above expression that the scale parameter of the mean distribution diverges asymptotically  $\left(\lim_{t \rightarrow \infty} \frac{c}{t^\alpha} \sum_{k=1}^{t-1} \left(\frac{1 - \rho^k}{1 - \rho}\right)^\alpha = +\infty, \forall \alpha, 0 < \alpha < 1\right)$ . Thus, the trajectories of  $\bar{y}_t$  do not asymptotically converge as a consequence of heavy tail, despite the strong stationarity of the process (see Figure 3.2).

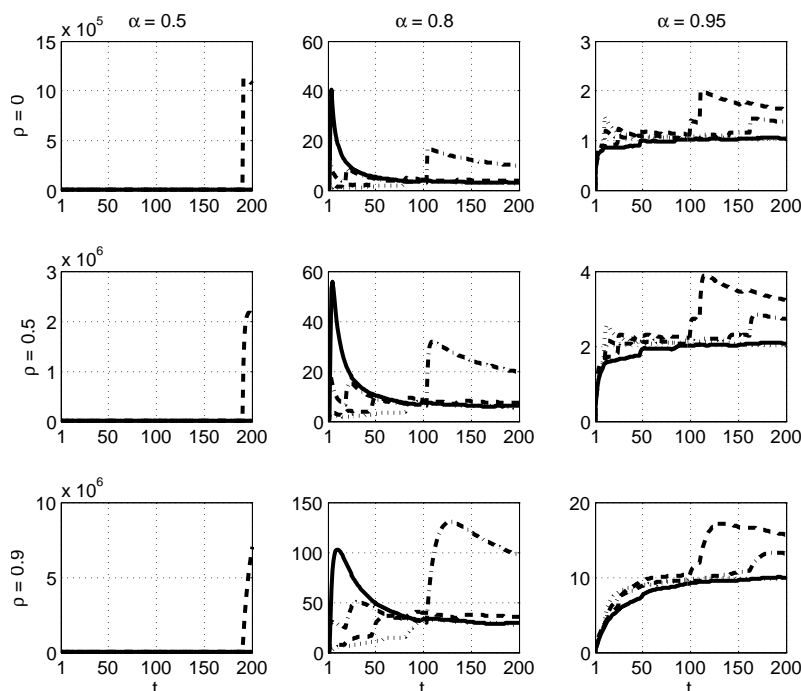


Figure 3.2: Trajectories of historical means. Simulated trajectories of the historical mean  $\bar{y}_t = \frac{1}{t}(y_1 + \dots + y_t), t = 1, \dots, 200$  of the SAR(1) process  $y_t$  for unitary scale parameter  $c = 1$ , different values of the parameters  $\alpha$  and  $\rho$ . Four different trajectories are plotted for each graph.

This is an important difference with other positive processes such as the (time-discretized) CIR process for instance. To highlight the asymptotic behavior of the historical mean, we provide in Figure 3.3 the histogram<sup>65</sup> of  $\bar{y}_{200} - y_1 \frac{1-\rho^{200}}{200(1-\rho)}$  based on 200 trajectories of a SAR(1) process. We observe that the limiting distribution of  $\bar{y}_{200}$  is right skewed and can weight extreme values. For comparison, we provide in Figure 3.4 the similar histogram for an autoregressive gamma process of order 1 (ARG(1)), which is the exact time discretization of the CIR process [see Gouriéroux, Jasiak (2006)] for different values of the parameters<sup>66</sup> :

$$c = 1, d = 0.1, 1, 5, b = 0, 0.5, 0.9.$$

The selected parameter values have been chosen to get a stationary ARG(1)

<sup>65</sup>We subtract the deterministic term in the mean formula [see above] in order to remove the location effect.

<sup>66</sup>As defined by Gouriéroux and Jasiak (2006), the process  $(y_t^*)$  is an autoregressive gamma process of order 1 (ARG(1)) if and only if, conditioned on  $y_{t-1}^*$ , there exists a Poisson variable  $z_t \sim \mathcal{P}(by_{t-1}^*)$  such that the conditional distribution of  $y_t^*$  given  $z_t$  is Gamma  $\gamma(d + z_t, c)$ .

process, for which the historical mean is consistent asymptotically normal. The finite sample distributions are provided in Figure 3.4. We remove any location effect subtracting from each empirical mean its theoretical counterpart <sup>67</sup>.

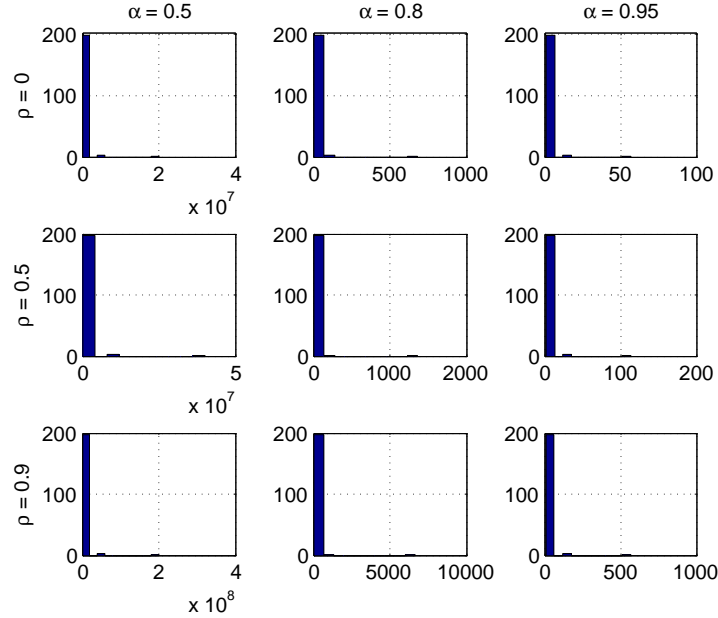


Figure 3.3: Histogram of the historical mean [SAR(1)]. Histogram of the historical mean  $\bar{y}_{200} - y_1 \frac{1-\rho^{200}}{200(1-\rho)}$  based on 200 different trajectories of the SAR(1) process  $y_t$ , for  $c = 1$ , and different values of parameters  $\alpha$  and  $\rho$ .

Finally, we report in Table 1, in Appendix 3.5, the estimated values  $\hat{\rho}_T = \min_{1 \leq t \leq t+h \leq T} (y_{t+h}/y_t)^{1/h}$ ,  $h \leq T$  for different numbers of observations  $T = 10, 50, 100, 200$ . We see that the higher the tail magnitude  $\alpha$  and the persistence parameter  $\rho$ , the weaker is the accuracy of the statistic  $\hat{\rho}_T$ .

<sup>67</sup>The marginal distribution of an ARG(1) process is a Gamma distribution  $\gamma(d, \frac{c}{1-bc})$  [see Gouriéroux, Jasiak (2006)].

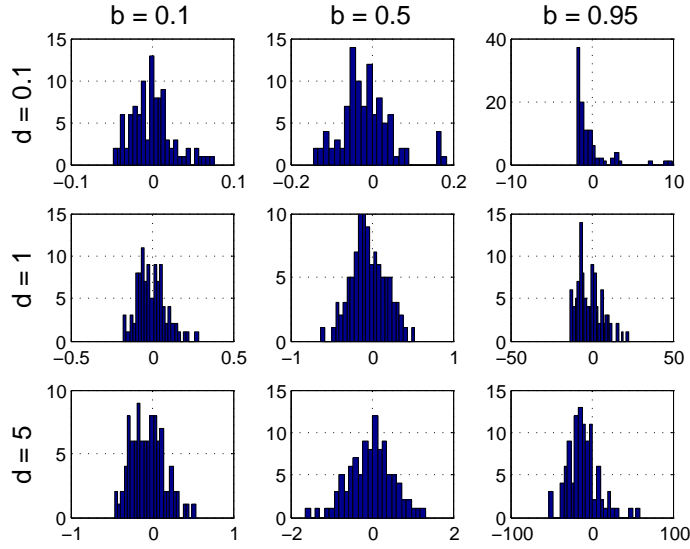


Figure 3.4: Histogram of the historical mean [ARG(1)]. Histogram of the centered historical means  $\bar{y}_{200}^* - E(y^*)$  based on 200 different trajectories of the ARG(1) process  $y_t^*$ , for different values of parameters  $d$  and  $b$ , with  $c = 1$ .

### 3.3 STABLE TERM STRUCTURE MODELS

The explicit form of the conditional Laplace transform for the moving average in positive stable innovations, at any horizon, can be used to derive term structure models. In this section,  $y_t$  denotes the value at date  $t$  of the (nominal) geometric short-term interest rate, that is, the rate for period  $(t, t + 1)$ . Then, the price at  $t$  of the zero-coupon bond of term  $h$  is given by :

$$B(t, h) = \mathbb{E}_t^{\mathbb{Q}}[\exp\{-(y_t + \dots + y_{t+h-1})\}], \quad (3.3.1)$$

where  $\mathbb{E}_t^{\mathbb{Q}}$  denotes the risk-neutral conditional distribution given  $I_t$ . The associated geometric rate at time-to-maturity  $h$  is :

$$r(t, h) = -\frac{1}{h} \log B(t, h). \quad (3.3.2)$$

In the first subsection, we introduce the term structure corresponding to the stable AR(1) processes, and discuss it as an alternative to the standard Cox, Ingersoll, Ross (CIR) process [Cox, Ingersoll, Ross (1985)]. In the second subsection, we show how the stable processes can be used to construct term structure models with a stochastic long-term rate.



### 3.3.1 STABLE AUTOREGRESSIVE TERM STRUCTURE MODEL (SARTSM)

Let us assume that the short-term interest rate  $y_t$  follows a SAR(1) model [see eqn. (2.5)] under the risk-neutral probability. We have :

$$y_{t+k} = \rho^k y_t + \rho^{k-1} \varepsilon_{t+1} + \dots + \varepsilon_{t+k}, k = 0, 1, \dots,$$

and

$$\begin{aligned} y_t + \dots + y_{t+h-1} &= \left( \sum_{k=0}^{h-1} \rho^k \right) y_t + \left( \sum_{k=0}^{h-2} \rho^k \right) \varepsilon_{t+1} + \dots + \varepsilon_{t+h-1}, \\ &= \frac{1 - \rho^h}{1 - \rho} y_t + \frac{1 - \rho^{h-1}}{1 - \rho} \varepsilon_{t+1} + \dots + \frac{1 - \rho}{1 - \rho} \varepsilon_{t+h-1}, h = 1, 2, \dots \end{aligned}$$

We deduce from Lemma 2, that the conditional distribution of the cumulated interest rate is such that :

$$(y_t + \dots + y_{t+h-1}) - \frac{1 - \rho^h}{1 - \rho} y_t | I_t \sim \mathcal{S} \left( \alpha, c \sum_{k=1}^{h-1} \left( \frac{1 - \rho^k}{1 - \rho} \right)^\alpha \right). \quad (3.3.3)$$

Therefore, we have:

$$\begin{aligned} &\mathbb{E}_t^{\mathbb{Q}} \{ \exp[-(y_t + \dots + y_{t+h-1})] \} \\ &= \exp \left\{ -\frac{1 - \rho^h}{1 - \rho} y_t - c \sum_{k=1}^{h-1} \left( \frac{1 - \rho^k}{1 - \rho} \right)^\alpha \right\}, \end{aligned}$$

by using the expression of the conditional Laplace transform with  $u = 1$ . By applying the definition of the rate at term  $h$  [see (3.1), (3.2)], we deduce the term structure associated with the SAR(1).

**Proposition 3.1 :** The SARTSM is such that :

$$r(t, h) = \frac{1}{h} \frac{1 - \rho^h}{1 - \rho} y_t + \frac{c}{h} \sum_{k=1}^{h-1} \left( \frac{1 - \rho^k}{1 - \rho} \right)^\alpha.$$

The general term of the series in the right hand side  $\left( \frac{1 - \rho^k}{1 - \rho} \right)^\alpha$  tends to  $\frac{1}{(1 - \rho)^\alpha}$ , when  $k$  tends to infinity, since  $0 < \rho < 1$ . The Cesaro sum computed from this series tends to the same limit, which provides the result below.

**Corollary 3.1 :** In the SARTSM, the long run interest rate exists and is constant equal to :

$$r(t, \infty) = \frac{c}{(1 - \rho)^\alpha}$$

Therefore the SARTSM can also be written as :

$$r(t, h) = \frac{1}{h} \frac{1 - \rho^h}{1 - \rho} r(t, 1) + r(t, \infty) \frac{1}{h} \sum_{k=1}^{h-1} (1 - \rho^k)^\alpha.$$

The SARTSM shares common feature with the (time discretized) CIR model. Both of them are affine term structure models, account for the positivity of nominal rates and involve three parameters. However, these parameters have different interpretations. They are a mean, a volatility and a mean-reversion in the usual CIR. These interpretations are not possible for a SAR(1), since the mean and volatility no longer exist. We have a scale parameter ( $c$ ), a long run parameter ( $c/(1 - \rho)^\alpha$ ), but  $\rho$  has a completely different interpretation. In the CIR model, this parameter implies some tendency to get closer to the mean. In the SAR(1) model parameter  $\rho$  is introduced, as a zero-reverting parameter, to avoid increasing short-term rate as discussed at the end of Section 3.2.2.

The pattern of the SARTSM is provided in Figure 3.5. The short-term rate  $r(t, 1)$  is set at 3%, and the long-term rate at either  $r(t, \infty) = 3\%$ , or  $r(t, \infty) = 8\%$ . The values of the other parameters are  $\rho = 0, 0.5, 0.9, \alpha = 0.5, 0.8, 0.9$  as in Figure 1. The term structure can feature monotonic, or hump-shaped patterns, depending on the respective locations of the short and long run interest rates. When  $\rho = 0$ , the SARTSM reduces to :

$$r(t, h) = \frac{1}{h} r(t, 1) + r(t, \infty),$$

and no longer depends on exponent  $\alpha$ .

If parameter  $\rho$  is small, we have  $(1 - \rho^k)^\alpha \sim 1 - \alpha\rho^k$ , and the SARTSM is equivalent to :

$$\begin{aligned} r(t, h) &= \frac{1}{h} \frac{1 - \rho^h}{1 - \rho} r(t, 1) + r(t, \infty) \left( 1 - \frac{\alpha}{h} \frac{1 - \rho^h}{1 - \rho} \right) + o_h(\rho) \\ &= r(t, \infty) + \frac{1}{h} \frac{1 - \rho^h}{1 - \rho} [r(t, 1) - \alpha r(t, \infty)] + o_h(\rho), \end{aligned}$$

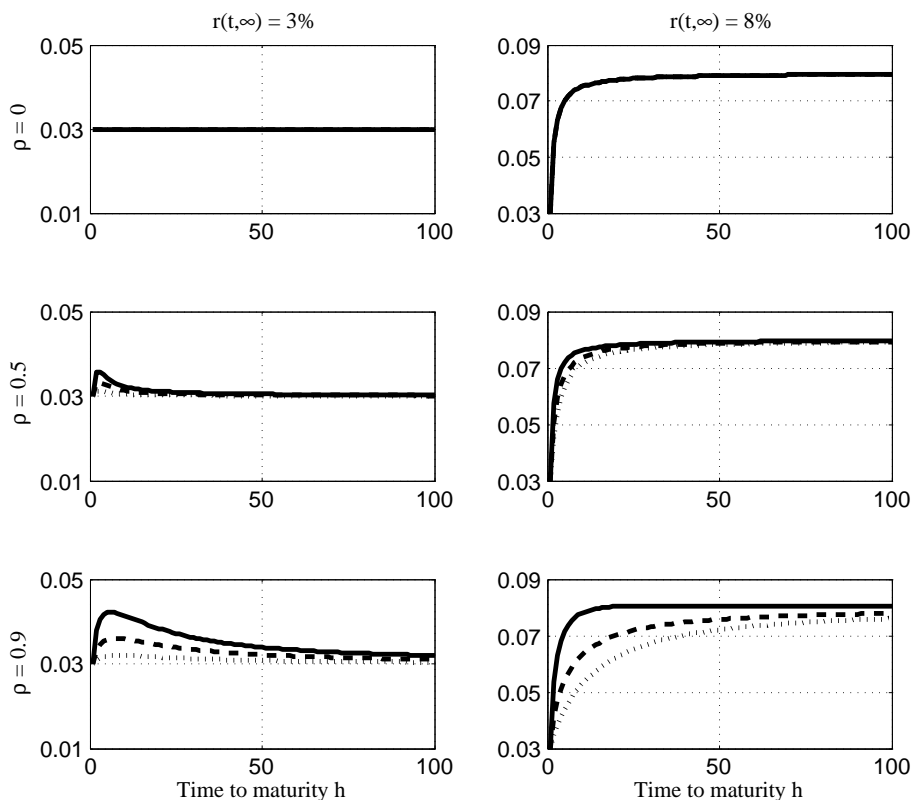


Figure 3.5: Pattern of the SARTSM. Pattern of SARTSM  $r(t, h)$  with  $h = 1, \dots, 100$ , for different values of  $r(t, \infty)$ ,  $\rho$  and  $\alpha$ , when the short-term rate  $r(t, 1)$  is set at 3%. The solid, dashed and dotted lines stand for  $\alpha = 0.5, 0.8, 0.9$ , respectively.

where the negligible term depends on time-to-maturity. This shows the importance of the location of the percentage spread  $r(t, \infty)/r(t, 1)$  with respect to  $1/\alpha$ , which measures the (price of) extreme risk.

### 3.3.2 A STOCHASTIC AND ALMOST FLAT TERM STRUCTURE

It is known that a flat term structure :  $r(t, h) = r_t, \forall h$ , say, is compatible with no arbitrage, only if level  $r_t$  is constant. As a consequence, any term structure model including a stochastic "level component" will allow for perfect arbitrage opportunity [see Chapter 2]. Since such models have been commonly introduced in the applied literature [see e.g. Litterman, Scheinkman (1991), Christensen, Diebold, Rudebusch (2011)], it is important to reconcile their "stochastic level" component with the no-arbitrage condition. We show below how to construct a stochastic arbitrage free almost flat term struc-

ture, in the sense that the short and long-term interest rates are equal and stochastic, by considering a stable distribution framework.

**Proposition 3.2 :** Let us consider the (risk-neutral) stable random walk :

$$y_t = y_{t-1} + \exp(-\gamma t)\varepsilon_t, \gamma > 1,$$

with  $\varepsilon_t \sim IIS(\alpha, c)$ . Then we have :

$$r(t, h) = y_t + d(t, h),$$

with  $d(t, h) = \frac{1}{h} \sum_{k=1}^{h-1} [c^\alpha (h-k)^\alpha \exp[-\alpha\gamma(t+k)]]$ . In particular :

$$\lim_{h \rightarrow \infty} d(t, \infty) = 0, \forall t,$$

or equivalently :  $r(t, \infty) = r(t, 1), \forall t$ .

**Proof :** See Appendix 3.2.

Under the above risk-neutral dynamic, the instantaneous forward interest rate given by :

$$f(t, h) = hr(t, h) - (h-1)r(t, h-1),$$

tends also to  $r(t, 1)$ , when  $h$  tends to infinity.

The random walk used to derive this very special term structure is mixing the standard explosion phenomenon due to the unit root with a tendency to become asymptotically constant due to a scale parameter tending to zero in time. Such a single factor TSM implies short-term interest rate increasing in time and thus cannot be used as such in practice. However, it can be introduced in a multiple factor term structure model to replace the usual "stochastic level component" (see Section 3.4.2).

The Term Structure Model above is also interesting from a theoretical point of view. Indeed in an arbitrage free term structure model, the (stochastic) long-term forward rate can never fall [see Dybvig, Ingersoll, Ross (1996), Hubalek, Klein, Teichman (2002), and more recently Kardaras and Platen (2012)]. This increasingness condition is satisfied by  $f(t, \infty) = r(t, 1)$ . In the example of Proposition 3.2, this limiting rate is stochastic and finite<sup>68</sup>.

---

<sup>68</sup>This provides an answer to the dilemma faced by El Karoui, Frachot, Geman (1998), since this long-term yield is neither constant, nor infinite.

### 3.4 DISCRETE TIME NELSON-SIEGEL TERM STRUCTURE MODEL

In this section, we compare different 5-factor affine term structure models (ATSM) of the type :

$$\begin{aligned} r(t, h) &= a_1(h, \theta)y_{1,t} + a_2(h, \theta)y_{2,t} + a_3(h, \theta)y_{3,t} + a_4(h, \theta)y_{4,t} + a_5(h, \theta)y_{5,t} + d(h, \theta) \\ &= a'(h, \theta)y_t + d(h, \theta), \end{aligned} \tag{3.4.1}$$

where we denote  $y_t = (y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t}, y_{5,t})'$  the factors,  $r(t, h)$  the zero-coupon rate with maturity  $h$  at time  $t$ , and where the sensitivity coefficients depend on unknown parameters  $\theta$ . The factors can be Gaussian, or stable under the risk-neutral distribution.

Gaussian factor models have been used to be compatible with the Nelson-Siegel-Svensson (NSS) model, which is very popular among practitioners (see Appendix 3.4 for a short description of the Nelson Siegel family of term structure models). We first consider 2-factor and 3-factor Gaussian affine term structures, with a triangular autoregressive matrix whose eigenvalues are equal. Because of the shape of the corresponding factor loadings' functions, these models are arbitrage-free 2-factor or 3-factor NSS type model in a discrete-time setting, where the factors can be identified as slope and curvature factors. Then, we consider 5-factor models by combining two 2-factor NSS model with a Gaussian level factor *à la* Christensen, Diebold, Rudebusch (2009, 2011), with the inconvenience of negative infinite long-term interest rate, or with an almost flat stable term structure as presented in Section 3.3.

#### 3.4.1 THE TRIANGULAR GAUSSIAN ATSM

The Gaussian ATSM is defined by the short-term rate equation :

$$r(t, 1) = \alpha'y_t + \beta, \tag{3.4.2}$$

where the  $n$ -dimensional factor satisfies the Gaussian Vector Autoregressive model :

$$y_t = \Phi y_{t-1} + \Sigma^{1/2} \varepsilon_t, \varepsilon_t \sim IIN(0, Id), \tag{3.4.3}$$

under the risk-neutral probability.

Therefore, we get :

$$y_t + \dots + y_{t+h-1} = (Id - \Phi)^{-1}(Id - \Phi^h)y_t + \sum_{k=1}^{h-1} (Id - \Phi)^{-1}(Id - \Phi^k)\varepsilon_{t+k},$$

and by (3.3.1), (3.3.2), and the expression of the Laplace Transform of the Gaussian distribution :

$$\begin{aligned} r(t, h) &= \beta + \frac{1}{h}\alpha'(Id - \Phi)^{-1}(Id - \Phi^h)y_t \\ &\quad - \frac{1}{2h} \sum_{k=1}^{h-1} \alpha'(Id - \Phi)^{-1}(Id - \Phi^k)\Sigma(Id - \Phi^k)(Id - \Phi)^{-1} \end{aligned} \quad (3.4.4)$$

whenever the autoregressive matrix has no eigenvalue equal to one.

The underlying factors are defined up to a one-to-one linear transformation. Therefore, without loss of generality, we can assume that the autoregressive matrix is written under its Jordan representation, that is, with its eigenvalues on the main diagonal, 0 or 1 on the diagonal just above the main diagonal, and zero elements anywhere else.

Let us now consider an autoregressive matrix with triangular Jordan form :

$$\Phi = \begin{pmatrix} \lambda & 1 & 0 & \\ 0 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, \quad (3.4.5)$$

where  $|\lambda| < 1$ .

The power of this triangular matrix is easily derived. For instance, we have :

$$\Phi^h = \begin{pmatrix} \lambda^h & h\lambda^{h-1} \\ 0 & \lambda^h \end{pmatrix}, \text{ for } n = 2,$$

and

$$\Phi^h = \begin{pmatrix} \lambda^h & h\lambda^{h-1} & h(h-1)\lambda^{h-1}/2 \\ 0 & \lambda^h & h\lambda^{h-1} \\ 0 & 0 & \lambda^h \end{pmatrix}, \text{ for } n = 3.$$

By replacing in formula (3.4.4), we get the following term structures (see Appendix 3.4).

**Proposition 3.3 :**

i) The 2-factor triangular term structure with unique eigenvalue  $\lambda$  is such that :

$$r(t, h) = a'(h, \theta)y_t + d(h, \theta), \text{ with}$$

$$\begin{aligned} a_1(h, \theta) &= \frac{1-\lambda^h}{h} \frac{\alpha_1}{1-\lambda}, \\ a_2(h, \theta) &= \frac{1-\lambda^h}{h} \left( \frac{\alpha_1}{(1-\lambda)^2} + \frac{\alpha_2}{1-\lambda} \right) - \lambda^h \frac{\alpha_1}{\lambda(1-\lambda)}, \end{aligned}$$

$$hd(h, \theta) = (h-1)d(h-1, \theta) + \beta - \frac{1}{2}(h-1)^2 a'(h-1, \theta) \Sigma a(h-1, \theta),$$

with  $d(0, \theta) = 0$ .

ii) The 3-factor triangular term structure with unique eigenvalue  $\lambda$  is such that :

$$r(t, h) = a'(h, \theta)y_t + d(h, \theta), \text{ with}$$

$$\begin{aligned} a_1(h, \theta) &= \frac{1-\lambda^h}{h} \frac{\alpha_1}{1-\lambda}, \\ a_2(h, \theta) &= \frac{1-\lambda^h}{h} \left( \frac{\alpha_1}{(1-\lambda)^2} + \frac{\alpha_2}{1-\lambda} \right) - \lambda^h \frac{\alpha_1}{\lambda(1-\lambda)}, \\ a_3(h, \theta) &= \frac{1-\lambda^h}{h} \left( \frac{\alpha_1}{(1-\lambda)^3} + \frac{\alpha_2}{(1-\lambda)^2} + \frac{\alpha_3}{1-\lambda} \right) - \lambda^h \left( \alpha_1 \frac{3\lambda-1}{2\lambda^2(1-\lambda)^2} + \frac{\alpha_2}{\lambda(1-\lambda)} \right) \\ &\quad - h\lambda^h \frac{\alpha_1}{2\lambda^2(1-\lambda)}, \end{aligned}$$

$$hd(h, \theta) = (h-1)d(h-1, \theta) + \beta - \frac{1}{2}(h-1)^2 a'(h-1, \theta) \Sigma a(h-1, \theta),$$

with  $d(0, \theta) = 0$ .

The first two factor loadings of the 2-factor and 3-factor triangular TSM are the same. The first factor loading is a monotonous function in the time-to-maturity  $h$ , usually interpreted as a "slope" factor. Moreover, the second factor loading (and also the third one in the 3-factor model) is humped because of the equality in the eigenvalues of the autoregressive matrix<sup>69</sup>. In this sense, the corresponding factors could be identified as "curvature" factors. In the 3-factor model, the last two factors have a "curvature" interpretation with

---

<sup>69</sup>This equality generates "resonance" among the factors: their cross correlation function reaches a maximum at intermediate lags [see Gouriéroux, Jasiak (2001)]

different hump locations (the third factor loading attains its maximum at a higher time-to-maturity than the second one).

Besides, up to a linear transformation, it is possible to build the 2-factor (and 3-factor) triangular TSM on a basis of  $\frac{1-\lambda^h}{h}, \lambda^h$  (and  $h\lambda^h$  in the 3-factor case) which changes the factors' interpretation. In this case, the transformed factors are such that :

$$\begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1}{1-\lambda} & \frac{\alpha_1}{(1-\lambda)^2} + \frac{\alpha_2}{1-\lambda} \\ 0 & \frac{\alpha_1}{\lambda(1-\lambda)} \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix},$$

or in the 3-factor triangular TSM :

$$\begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \\ \tilde{y}_{3t} \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1}{1-\lambda} & \frac{\alpha_1}{(1-\lambda)^2} + \frac{\alpha_2}{1-\lambda} & \frac{\alpha_1}{(1-\lambda)^3} + \frac{\alpha_2}{(1-\lambda)^2} + \frac{\alpha_3}{1-\lambda} \\ 0 & \frac{\alpha_1}{\lambda(1-\lambda)} & -\left(\alpha_1 \frac{3\lambda-1}{2\lambda^2(1-\lambda)^2} + \frac{\alpha_2}{\lambda(1-\lambda)}\right) \\ 0 & 0 & -\frac{\alpha_1}{2\lambda^2(1-\lambda)} \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix}.$$

Another possibility is to rotate the factors in order to express the yields  $r(t, h)$  as function of the short-term rate  $r(t, 1)$ . This leads to the following representation :

$$\begin{aligned} r(t, h) &= \frac{1-\lambda^h}{h} r(t, 1) + \left( \frac{1-\lambda^h}{h} \frac{\alpha_1}{(1-\lambda)^2} - \lambda^h \frac{\alpha_1}{\lambda(1-\lambda)} \right) y_{2t} \\ &\quad - \frac{1-\lambda^h}{h} \beta + d(h, \theta), \end{aligned}$$

or :

$$\begin{aligned} r(t, h) &= \frac{1-\lambda^h}{h} r(t, 1) + \left( \frac{1-\lambda^h}{h} \frac{\alpha_1}{(1-\lambda)^2} - \lambda^h \frac{\alpha_1}{\lambda(1-\lambda)} \right) y_{2t} \\ &\quad + \left( \frac{1-\lambda^h}{h} \left( \frac{\alpha_1}{(1-\lambda)^3} + \frac{\alpha_2}{(1-\lambda)^2} \right) - \lambda^h \left( \alpha_1 \frac{3\lambda-1}{2\lambda^2(1-\lambda)^2} + \frac{\alpha_2}{\lambda(1-\lambda)} \right) - h\lambda^h \frac{\alpha_1}{2\lambda^2(1-\lambda)} \right) y_{3t} \\ &\quad - \frac{1-\lambda^h}{h} \beta + d(h, \theta). \end{aligned}$$

### 3.4.2 THE EXTENSIONS OF THE NELSON-SIEGEL-SVENSSON TERM STRUCTURE MODEL

The 3-factor NSS model with a unique eigenvalue is pretty constrained, since all factor loadings' functions  $a(h)$  (and in particular the locations of both humps in the curvature factors loadings' functions) are set by only one parameter  $\lambda$ . In order to gain flexibility we consider a combination of two 2-factor NSS model with eigenvalues  $\lambda_1$  and  $\lambda_2$ , to which we add a level factor. By doing so, we follow CDR (2009), which will make our empirical results (presented in the next section) comparable with their application.



It has been shown [see Dubecq Gourieroux (2010)] that a “level” factor, which makes the whole yield curve move by parallel shifts, is necessarily a random walk process under the risk-neutral probability measure. In this chapter, we consider two alternatives for the conditional distribution of the innovations, which hit the level factor process :

i) We assume the innovations to be i.i.d Gaussian, which makes our model a discrete time equivalent of the so-called Arbitrage-Free Generalized Nelson-Siegel model put forward by Christensen, Diebold, Rudebush (2009). The corresponding term structure is such that :

$$\begin{aligned}
 a_1(h, \theta) &= \alpha_1, \\
 a_2(h, \theta) &= \frac{1-\lambda_1^h}{h} \frac{\alpha_2}{1-\lambda_1}, \\
 a_3(h, \theta) &= \frac{1-\lambda_1^h}{h} \left( \frac{\alpha_2}{(1-\lambda_1)^2} + \frac{\alpha_3}{1-\lambda_1} \right) - \lambda_1^h \frac{\alpha_2}{\lambda_1(1-\lambda_1)}, \\
 a_4(h, \theta) &= \frac{1-\lambda_2^h}{h} \frac{\alpha_4}{1-\lambda_2}, \\
 a_5(h, \theta) &= \frac{1-\lambda_2^h}{h} \left( \frac{\alpha_4}{(1-\lambda_2)^2} + \frac{\alpha_5}{1-\lambda_2} \right) - \lambda_2^h \frac{\alpha_4}{\lambda_2(1-\lambda_2)}, \\
 hd(h, \theta) &= (h-1)d(h-1, \theta) + \beta - \frac{1}{2}(h-1)^2 a'(h-1, \theta) \Sigma a(h-1, \theta)
 \end{aligned}$$

with  $d(0, \theta) = 0$ .

This specification (or a 3-factor version of it without the second slope and curvature factors  $(y_{4,t}, y_{5,t})$ ) has been used recently in several papers [see for instance Bauer (2009), Christensen, Lopez, Rudebush (2009), Fontaine, Garcia (2009)]. However, such model implies a long-term interest rate equal to  $-\infty$ .

ii) A model obtained by combining two 2-factor triangular Gaussian ATSM with distinct eigenvalues  $\lambda_1, \lambda_2$ , with the almost flat stable ATSM introduced in Section 3.3.2. Formally :

$$y_t = \Phi y_{t-1} + \Omega_t \varepsilon_t,$$

with

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \quad \Omega_t = \begin{pmatrix} \exp(-\gamma t) & 0 \\ 0 & \Sigma^{1/2} \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{25,t} \end{pmatrix},$$

with  $\varepsilon_{1,t} \sim IIS(\alpha, c)$  and  $\varepsilon_{25,t} = \begin{pmatrix} \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \end{pmatrix} \sim IIN(0, Id)$  under the risk-neutral probability.

The corresponding term structure is such that :

$$\begin{aligned}
 a_1(h, \theta) &= \alpha_1, \\
 a_2(h, \theta) &= \frac{1-\lambda_1^h}{h} \frac{\alpha_2}{1-\lambda_1}, \\
 a_3(h, \theta) &= \frac{1-\lambda_1^h}{h} \left( \frac{\alpha_2}{(1-\lambda_1)^2} + \frac{\alpha_3}{1-\lambda_1} \right) - \lambda_1^h \frac{\alpha_2}{\lambda_1(1-\lambda_1)}, \\
 a_4(h, \theta) &= \frac{1-\lambda_2^h}{h} \frac{\alpha_4}{1-\lambda_2}, \\
 a_5(h, \theta) &= \frac{1-\lambda_2^h}{h} \left( \frac{\alpha_4}{(1-\lambda_2)^2} + \frac{\alpha_5}{1-\lambda_2} \right) - \lambda_2^h \frac{\alpha_4}{\lambda_2(1-\lambda_2)}, \\
 hd(t, h, \theta) &= hd_1(t, h, \theta) + hd_2(h, \theta), \\
 hd_1(t, h, \theta) &= \exp(-\alpha\gamma)(h-1)d_1(t, h-1, \theta) + c^\alpha(h-1)^\alpha \exp[-\alpha\gamma(t+1)] \\
 &= \exp(-\alpha\gamma)hd_1(t-1, h, \theta), \\
 hd_2(h, \theta) &= (h-1)d_2(h-1, \theta) + \beta \\
 &\quad - \frac{1}{2}(h-1)^2 \begin{pmatrix} a_2(h-1, \theta) \\ a_3(h-1, \theta) \\ a_4(h-1, \theta) \\ a_5(h-1, \theta) \end{pmatrix}' \Sigma \begin{pmatrix} a_2(h-1, \theta) \\ a_3(h-1, \theta) \\ a_4(h-1, \theta) \\ a_5(h-1, \theta) \end{pmatrix},
 \end{aligned}$$

with  $d_1(t, 0, \theta) = d_2(0, \theta) = 0, \forall t$ .

As the previous model, this specification keeps the popular "level", "slope" and "curvature" factor interpretations, but does not provide divergent long-term rates.

### 3.5 APPLICATION

The data we use in our application are monthly unsmoothed Fama Bliss zero-coupon yields on US T-bonds from January 1987 to December 2002. For the sake of comparison, the maturities of zero-coupon bonds in our sample are the same as in CDR (2009) and range from 3 months to 30 years. Figure 3.6 illustrates the time series of the 1 year, 5 years, 10 years and 30 years rates.

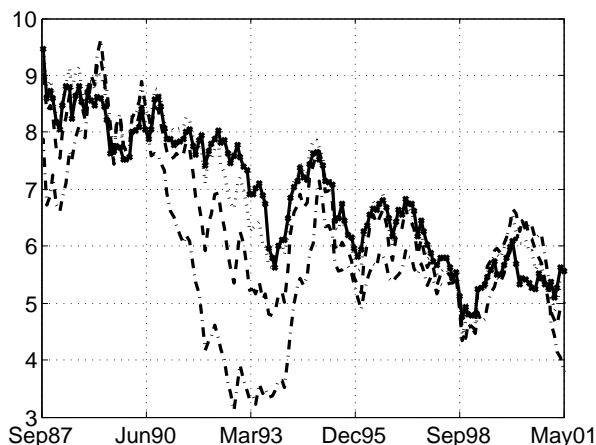


Figure 3.6: Times Series of the unsmoothed Fama-Bliss zero-coupon US T-bond yields from January 1987 to December 2002. The yields shown have 1 year (dash-dot line), 5 years (dashed line), 10 years (dotted line), 30 years (solid line).

#### 3.5.1 METHODOLOGY

We estimate both models with Kalman filter. In both models, we assume that the Gaussian factors (under the risk-neutral measure) are also Gaussian Vector AutoRegressive (VAR) process of order 1 under the physical measure, denoted  $\mathbb{P}$ . The conditional distribution of Gaussian factors under  $\mathbb{P}$  features a different mean and autoregressive matrix (denoted  $\mu_{\mathbb{P}}$ ,  $\Phi_{\mathbb{P}}$ , respectively) compared to their risk-neutral counterparts ( $\mu$ ,  $\Phi$ ). However, the conditional variance of the shocks to the Gaussian factors is constrained to be the same under both measure. We follow in this respect the common practice in the literature [see Ang, Piazzesi, Wei (2006), Dai, Singleton (2000), CDR (2009)], which considers reduced form ATSM with Gaussian factors, exponential affine stochastic discount factor, and essentially affine specification of the market price of risk [see Duffee (2002)].

The estimation of the model with stable level factor is more complex than the fully Gaussian model, due to the increasingness of the stable level factor under the pricing measure. Indeed, because of no-arbitrage assumption, this constraint holds also under the physical measure [see Harrison, Kreps (1979), Hansen, Richards (1987)]. We solve this problem by approximating the stable level factor as an increasing random walk with Gaussian innovation, while constraining the (drift and volatility of the) conditional distribution of the Gaussian shocks to be such that the probability of a decrease in the stable level factor is highly unlikely<sup>70</sup>. Furthermore, we allow the stable level factor to have a causal effect on the dynamics of the other (Gaussian) factors, but not the reverse. Finally, we also constrain the volatility of the innovations, which hit the stable factor under the  $\mathbb{P}$ -measure to be deterministically decreasing in time (similarly to the risk-neutral shocks, whose scale decreases in time). Formally, we denote  $\sigma_{11,t} = c_{\mathbb{P}} \exp[-\gamma_{\mathbb{P}}t]$ , where  $\sigma_{11,t}$  indicates the time-varying conditional  $\mathbb{P}$ -volatility of the stable factor, and  $c_{\mathbb{P}}, \gamma_{\mathbb{P}}$  are strictly positive parameters.

### 3.5.2 RESULTS

Figures 3.7 to 13 summarize the results (we collect the parameter estimates in Table 2 to 5 in Appendix 3.6).

Figure 3.7 to 3.9 display the time series of residuals and Root Mean Squared Errors (RMSE) for the 1 year, 5 years, 10 years and 30 years time-to-maturity, and the average of residuals and RMSE across all maturities (we recall the RMSE obtained by CDR (2009) in their application). These figures show that, in spite of the strong constraint put on the dynamics of the stable level factor, the term structure model with stable innovations on the random walk competes neck and neck with the unconstrained fully Gaussian model, and the continuous-time equivalent of the fully Gaussian model put forward by CDR (2009). Remarkably, our both models performs better than CDR (2009) on the very long-term rates, presumably because of the strong constraint CDR put on the  $\mathbb{P}$ -autoregressive matrix of the factors  $\Phi_{\mathbb{P}}$  (CDR constraints  $\Phi_{\mathbb{P}}$  to be diagonal, while we do not).

---

<sup>70</sup>In practice, we simply penalize the log-likelihood function when the filtered estimates of the stable factor decreases.

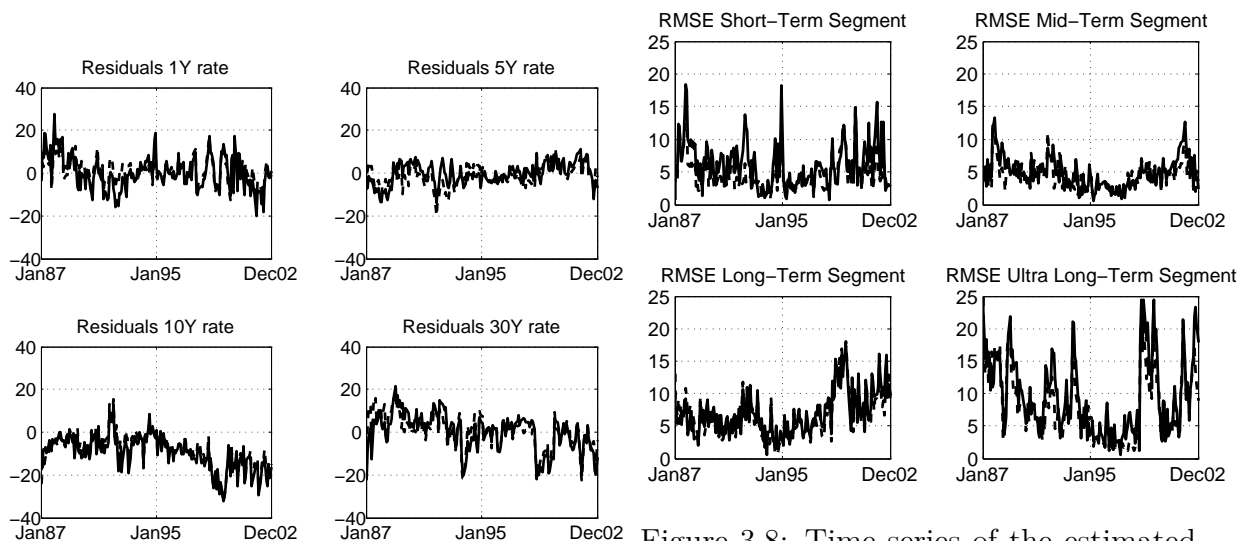


Figure 3.7: Time series of the estimated pricing errors for the model with stable level factor (solid line), and the Fully Gaussian model (dashed line), from January 1987 to December 2002. The models are estimated on the level of yields. The reported values are in basis points.

Figure 3.8: Time series of the estimated Root Mean Squared Errors for the model with stable level factor (solid line), and the Fully Gaussian model (dashed line), from January 1987 to December 2002. The models are estimated on the level of yields. The reported values are in basis points.

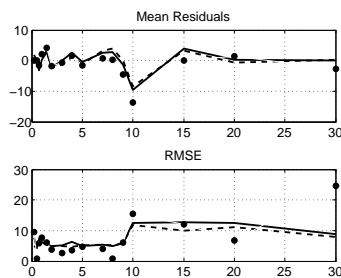


Figure 3.9: Time average of the pricing errors and Root Mean Squared Errors on the level of yields, for all maturities (from 3 months to 30 years) for the model with stable level factor (solid line), and fully Gaussian model (dashed line). The mean of pricing errors and Root Mean Squared Errors obtained by Christensen, Diebold, Rudebush (2009) are reported as black circles. The x-axis stands for the yield maturity in years. The reported values are in basis points.

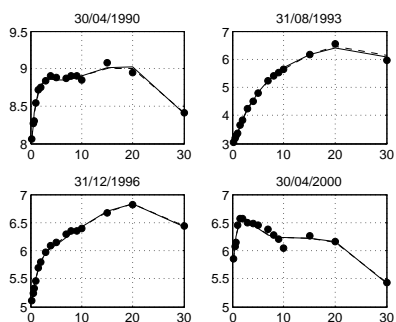


Figure 3.10: Fitted yield curve for 4 specific dates. The observed yields are indicated as black circles, and the fitted yield curve are from the model with stable factor (solid line) and the fully Gaussian model (dashed line). The reported values are in percentage points.

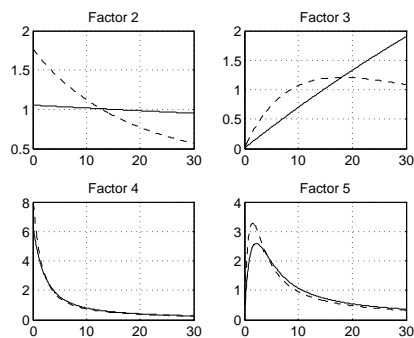


Figure 3.11: Normalized Factor loadings  $a(h)$  for the second to fifth factor. For each factor, the factor loading function  $a(h)$  is normalized by its average value over the whole set of maturities. The x-axis stands for the yield's maturity in years. The reported values are in percentage points. The solid, and dashed lines, indicate the factor loading functions for the model with stable level factor, and for the Fully Gaussian model, respectively.

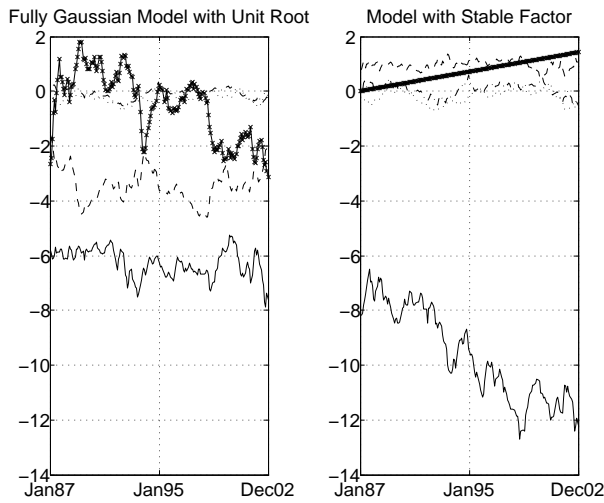


Figure 3.12: Time series of the estimated factors from January 1987 to December 2002 for the fully Gaussian model (left plot) and for the model with stable level factor (right plot). All factors are scaled by the average value over all maturities of the corresponding factor loading function. The reported values are in percentage points. For each model, the solid line with black cross, the solid line, the dashed line, the dotted line, and the dash-dot line, indicate the time series of the first to fifth factors, respectively.

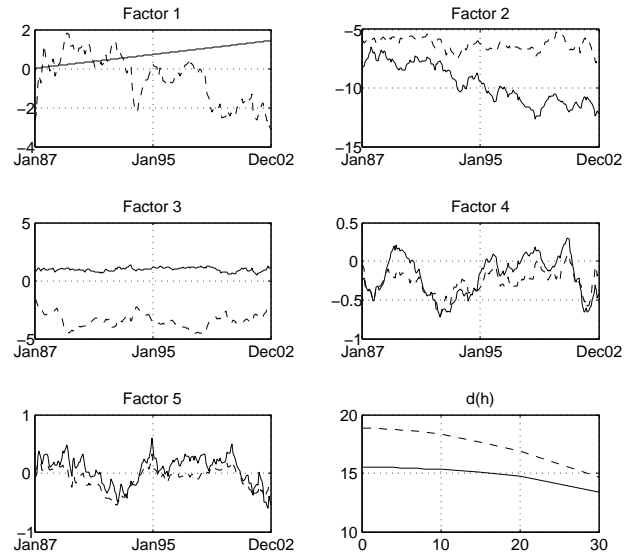


Figure 3.13: Time series of the estimated factors from January 1987 to December 2002. All factors are scaled by the average value over all maturities of the corresponding factor loading function. The reported values are in percentage points. The solid, and dashed lines, indicate the factor loading functions for the model with stable level factor, and for the Fully Gaussian model, respectively. The graph on the south-east corner presents the deterministic baseline term structure  $d_2(h, \theta)$  and  $d(h, \theta)$  for the model with stable level factor (solid line) and the fully Gaussian model (dashed line), respectively.

Figures 3.10 to 3.14 illustrate the difference in the characteristics of both models. Figure 3.10 shows the ability of both models to replicate observed yield curves, even the ones featuring 2 humps as in April 1990. The second and third factors, which, with the stable level factor, have the biggest impact on the very long term rates [see figure 3.11], compensate for the stochastic increase of the stable factor displayed in Figure 3.14. Figure 3.12 to 3.13 put forward how the stable level factor change the estimated dynamics of the other slope and curvature factors.

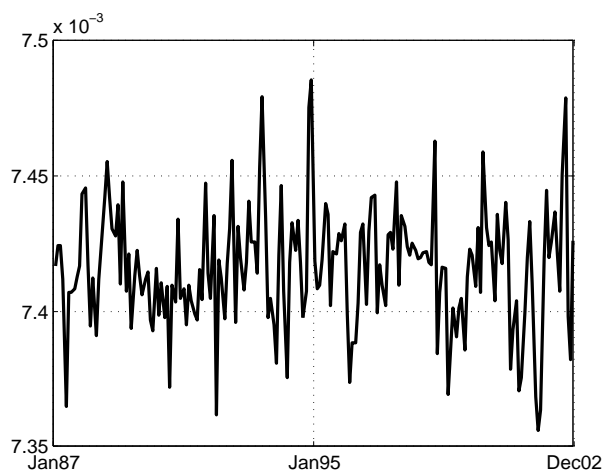


Figure 3.14: Time series of the estimated stable factor's variation from February 1987 to December 2002. The variations are scaled by the average value over all maturities of the level factor loading function. The reported values are in percentage points.

### 3.6 CONCLUDING REMARKS

In this chapter, we consider discrete time multifactor term structure models, with one Stable AutoRegressive factor among the pricing factors. The Stable Term Structure models feature interesting properties. They provide an alternative for the modeling of interest rates subject to positivity constraints. Second, they can be used to specify arbitrage-free term structure models, with stochastic long-term rates. We estimate such model on a dataset of US T-bonds, and show that they perform similarly, though much constrained, to their popular Gaussian counterparts.

In the literature, the popular level, slope and curvature interpretation of the factors have been related to macroeconomic variables such as economic growth, or inflation rate [see Diebold, Rudebusch, and Aruoba (2006)]. Our specification of the level factor modifies the whole interpretation of the factors, and thus, would have interesting implications for the macro-finance literature, that lie out of the scope of this chapter.





## General Conclusion

---

In the first part of this dissertation, we analyze the current methodology of stress-test exercises. Stressing a portfolio of assets, that is, investigating the consequences of a given (deteriorated) environment on a portfolio of assets amounts to compare two distributions for the portfolio's value. We investigate carefully the shock, which links these two distributions. First, we distinguish two representations of the shock on the common factors driving asset prices. Such shock can indeed be defined either in terms of distribution, or in terms of variables. We highlight both representations, and analyze the differences between them. Second, we propose to investigate the consequences of stochastic shocks when performing stress-test exercises, contrary to the current practice, which measure the risk of a portfolio through its reaction to a given realization of the pricing factors, that is, to a deterministic shock. We claim our approach is more robust, more difficult to manipulate, and is able to handle coherently second-order crisis phenomenon, such as a rise in the correlation of asset prices. Besides, our method allows to consider the reaction of the portfolio manager to the shock. We derive carefully the sensitivity of the portfolio distribution to a shock on the distribution of the pricing factors, for both representations (in terms of distribution, and in terms of variables). We thus study analytically the optimal reaction of

the portfolio manager to the shock, for several investment criteria. As an illustration, we perform a stress-test exercise for a portfolio of bond, in which we propose an intuitive definition for the stochastic shock. We study the reaction of the portfolio for different sizes of the shock, as well as the optimal reaction of a portfolio manager with mean-variance objective function, and show how taking into account the manager's reaction change dramatically the diagnosis on the riskiness of the portfolio.

The methodology introduced in this chapter opens the way for questions, which should be part of a future research agenda. First, the measure of the risk of a portfolio of assets is the main concern of this chapter. Therefore, we do not consider explicitly for the considered portfolio the possibility that its variations could trigger feedback effects on asset prices (the portfolio we are considering are not "systemic" in this sense). The assessment of the consequences of the variations in the value of systemic portfolio (or systemic financial institutions) have become an important question in the on-going research projects in finance [see e.g. Brownlees, Engle (2012), Acharya, Pedersen, Philippon, and Richardson (2010)]. This question has however not been addressed in a stress-testing perspective, and constitutes a relevant extension of our results. Second, though this chapter is merely concerned with the statistical distribution of the value of a portfolio under shocks, one could derive implications for asset pricing models, i.e. to take into account the robustness of financial assets for their pricing [see e.g. Boyle, Feng, Tian, and Wang (2007) for an example of robust pricing model]. Finally, while this chapter highlights the implications of the portfolio's manager adjustment to the shocks for the measure of the portfolio's risk, we have unfortunately little knowledge of the strategy of financial institutions. We circumvent this limitation by considering classic investment objective functions considered in the finance literature. Nevertheless, stress-testing exercises and the methodology put forward in this chapter would greatly benefit from a better understanding of financial institutions' behavior, which constitutes a promising research

agenda in itself [in this respect, the recent work of Begenau, Piazzesi, Schneider (2012) on banks' interest risk exposure represents an interesting first step].

In the second part of the thesis, we sought to provide term structure models, which are well suited for the pricing of very long-term bonds. To do so, we investigate the ability of term structure models to allow the modeled limiting rate (i.e. with infinite time-to-maturity) to be stochastic, while satisfying the assumption of absence of arbitrage opportunities. We first consider factor models of the yield curve, where one of the factor has a level interpretation, i.e. whose variations affect uniformly all interest rates. These models are particularly interesting for our purpose, since they are one of the few existing term structure model, which imply stochastic limiting rate. Historically, term structure models incorporating a level factor have been among the first introduced by the literature, and some versions remain very popular (the Nelson-Siegel model in particular). The second chapter is devoted to the compatibility of such models with the assumption of no-arbitrage. This chapter highlights the strong constraint put on the dynamics of the level factor by the assumptions of no-arbitrage, positivity of interest rates, and finiteness of long-term rates. In the third chapter, we build on this result and introduce a new class of no-arbitrage term structure models, which assumes the conditional distribution of the pricing factors to be stable, and allows the modeled limiting rate to be stochastic. This model is estimated on a time-series of US Treasury bond rates, and its performance is compared with a standard Nelson-Siegel type of arbitrage-free term structure model, whose long-term rates are explosive. This chapter shows that no-arbitrage term structure model with stochastic limiting rate have similar performance to other standard models, though constraining much more the factors' dynamics.

Nevertheless, we are confronted with several limits in this part of the dissertation that could be addressed through further investigations that fit into a research agenda. First, the economic interpretation of our results lies out of the scope of this thesis. The level of very long-term rate is however an important economic issue, in particular regarding the discounting of very long-term (nominal) financial flows. This question is crucial for the assessment of the economic value of projects as diverse as the construction of nuclear plants and railroads, or sustainable development policy. The market price of long-term bond should have something to say on this issue. Hence, it would be interesting to establish a parallel between our models with stochastic limiting rate, and general equilibrium models, which derive the long-term rate from the preferences of economic agents [see the thesis of Guéant (2007) on this branch of the literature]. Second, these chapters take the assumption of no-arbitrage opportunities as granted. As we explained, this is because we aim to provide term structure models for market makers, whose quotes have to be internally consistent. Nevertheless, the economic interpretation of the factors' dynamics that we estimate in the third chapter should take into account the potential limits to arbitrage, which may occur in the specific market of long-term government bonds, as highlighted by Greenwood, Vayanos (2012) on the UK Gilts' market.

In spite of the several limits our works are concerned with, we believe that we provide interesting answers and also that the questions we raise are as many alleys for future researches in these exciting topics.

Acharya, V., V., Pedersen, L., H., Philippon, T., and M., P., Richardson (2010): "Measuring Systemic Risk". AFA 2011 Denver Meetings Paper.

Alvarez, F., and U., J., Jermann (2005) : "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth". *Econometrica*, 73, 6, 1977-2016.

Andersen, T., and J., Lund (1997) : "Estimating Continuous-Time Stochastic Volatility Models of the Short Term Interest Rate", *Journal of Econometrics*, 77, 343-377.

Ang, A., Piazzesi, M., and, M. Wei (2006) : "What does the yield curve tell us about GDP growth?", *Journal of Econometrics* 2006, 131, 359-403.

Armbruster, B. (2011) : "'A Short Proof of Strassen's Theorem'". Northwestern University Working Paper.

Artzner, P., Delbaen, F., Eber, J., M., and D. Heath (1999) : "Coherent Measures of Risk". *Mathematical Finance*, 9, 3, 203-228.

Backus, D., Chernov, M., and S., E., Zin (2011) : "Sources of Entropy in Representative Agent Models". NBER Working Paper No. 17219.

Bank of International Settlements (2005) : "Zero-Coupon Yield Curves: Technical Documentation". BIS Working Paper, 25.

Bank of International Settlements (2008) : "Principles for Sound Stress Test-

ing Practices and Supervision". BIS Publication.

Bauer, M. (2009) : "Term Premia and the News", University of California Working Paper.

Begenau, J., Piazzesi, M., and M., Schneider (2012): "Banks' Risk Exposures". Working paper.

Bertholon, H., Monfort, A., and F., Pegoraro (2009) : "Econometric Asset Pricing Modelling". *Journal of Financial Econometrics*, 6, 4, 407-458;

Bertsimas, D., Lauprete, J., G., and A., Samarov (2004) : "Shortfall as a Risk Measure : Properties, Optimization, and Application." *Journal of Economic Dynamics and Control*, 28, 1353-1381.

Bjork, T., and P., Christensen (1999) : "Interest Rate Dynamics and Consistent Forward Rate Curves". *Mathematical Finance*, 9, 323-348.

Boyle, P., Feng, S., Tian, W., and T., Wang (2007): "Robust Stochastic Discount Factor". *The Review of Financial Studies*, 21, 3, 1077-1122.

Brennan, M., J., and E., S., Schwartz (1982): "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency". *Journal of Financial and Quantitative Analysis*, 17, 3, 301-329.

Breuer, T., and I., Csiszar (2012): "Systematic Stress Tests with Entropic Plausibility Constraints". *Journal of Banking and Finance*, Forthcoming.

Brownlees, C., and R., Engle (2012): "Volatility, Correlation and Tails for Systemic Risk Measurement". Working paper.

Carr, P., and L., Wu (2003) : "The Finite Moment Log-Stable Process and Option Pricing ". Journal of Finance, 58,2, 753-778.

Chesher, A. (1983) : "The Information Matrix Test : Simplified Calculation via a Score Test". Economic Letters, 13, 45-48.

Chesher, A. (1984) : "Testing for Neglected Heterogeneity". Econometrica, 52, 865-872.

Christensen, J., Diebold, F., and G., Rudebusch (2011) : "The Affine Arbitrage Free Class of Nelson-Siegel Term Structure Model", Journal of Econometrics, 164, 4-20.

Christensen, J., Diebold, F., and G., Rudebusch (2009) : "An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model," The Econometrics Journal, 12, 33-64.

Christensen, J., Lopez, J., and G., Rudebusch (2008) : "Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields", Fed. Reserve Bank of San Francisco, DP 2008-34.

Cochrane, J., and M., Piazzesi (2005) : "Bond Risk Premia". American Economic Review, 94, 1, 138-160.



Cox, J., Ingersoll, J., and S., Ross (1985) : "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-408.

Dai, Q., and K. Singleton (2000) : "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-1978.

Darolles, S., Gouriéroux, C., and J., Jasiak (2006) : "Structural Laplace Transform and Compound Autoregressive Models", *Journal of Time Series Analysis*, 27, 477-503.

Deelstra, G (2000) : "Long Term Returns in Stochastic Interest Rate Models : Application", *Astin Bulletin*, 30.

Delbaen, F., and G., Deelstra (1995) : "Long Term Returns in Stochastic Interest Rate Models", *Insurance : Mathematics and Economics*, 17, 163-169.

Diebold, F.X., Rudebusch, G.D., and B., Aruoba (2006) : "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach", *Journal of Econometrics*, 131, 309-338.

Duarte, P., G., and K., D., Hoover (2011) : "Observing Shocks". CHOPE Working Paper No. 2011-09 .

Dubecq, S., and C., Gouriéroux (2010) : "An Analysis of Ultra Long-Term Yields", CREST Working Paper.

*Concluding Remarks*

Duffee, G. (2002) : "Term Premia and Interest Rate Forecasts in Affine Models", *Journal of Finance*, 51, 527-552.

Duffie, D. (2001) : "Dynamic Asset Pricing Theory". Princeton University Press.

Duffie, D., Filipovic, D., and W., Schachermayer (2003) : "Affine Processes and Applications in Finance", *Annals of Applied Probability*, 13, 984-1053.

Duffie, D., Pan, J., and K., Singleton (2000) : "Transform Analysis and Asset Pricing for Affine Jump-Diffusions", *Econometrica*, 68, 6, 1343-1376.

Duffie, D., and R., Kan (1996) : "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, 6, 4, 379-406.

Durand, D. (1942) : "Basic Yields of Corporate Bonds, 1900-1942", Technical Paper 3, National Bureau of Economic Research.

Dybvig, P., Ingersoll, J., and S., Ross (1996) : "Long Forward and Zero Coupon Rates Can Never Fall", *Journal of Business*, 69, 1-25.

El Karoui, N., Frachot, A., and H., Geman (1998) : "On the Behaviour of Long Zero Coupon Rates in a No-Arbitrage Framework", *Review of Derivatives Research*, 1, 351-389.

European Banking Authority (2011) : "EU-Wide Stress Test Aggregate Report".

Faa di Bruno, F. (1855): "Sullo Sviluppo delle Funzioni". *Annali di Scienze Matematiche e Fisiche*, 6, 479-480.

Fabozzi, F. (1991) : "Duration, Convexity and Other Bond Risk Measures", New-Hope, Pennsylvania.

Fama, E., F., and K., R., French (1992) : "The Cross-Section of Expected Stock Returns". *Journal of Finance*, 47, 2, 427-465.

Feller, W. (1971) : "An Introduction to Probability Theory and its Application", 2nd edition, vol 2, Wiley New-York.

Filipovic, D. (1999) : "A Note on the Nelson-Siegel Family", *Mathematical Finance*, 9, 349-359.

Fisher, L. (1966) : "An Algorithm for Finding Exact Rates of Return", *Journal of Business*, 39, 111-118.

Fisher, L., and R., Weil (1971) : "Coping with the Risk of Interest Rate Fluctuations : Returns to Bondholders from Naive and Optimal Strategies", *Journal of Business*, 44, 408-431.

Fontaine, J. S. and R., Garcia (2009) : "Bond Liquidity Premia", Bank of Canada, DP 2009-28.

Gourieroux, C., and J., Jasiak (2001) : "Financial Econometrics", Princeton Series in Finance.

Gourieroux, C., and J., Jasiak (2005) : "Nonlinear Impulse Response Function". *Annales d'Économie et de Statistique*, 78, 1-33.

Gourieroux, C., and J., Jasiak (2006) : "Autoregressive Gamma Processes", *Journal of Forecasting*, 25, 129-152.

Gourieroux, C., Jasiak, J., and R., Sufana (2012) : "The Wishart Autoregressive Process of Multivariate Stochastic Volatility". *Journal of Econometrics*, 150, 2, 167-181.

Gourieroux, C., Laurent, J-P., and O., Scaillet (2000) : "Sensitivity Analysis of Value-at-Risk". *Journal of Empirical Finance*, 7, 225-245.

Gourieroux, C., and A., Monfort (2007) : "Econometric Specification of Stochastic Discount Factor Models". *Journal of Econometrics*, 136, 2, 509-530..

Gourieroux, C., and A., Monfort (2011) : "Domain Restrictions on Interest Rate Implied by No Arbitrage", *Mathematical Finance*, 21, 2, 281-291.

Gourieroux, C., and A., Monfort (1995): "Statistic and Econometric Models". Vol.2, Cambridge University Press.

Gourieroux, C., Monfort, A., and V., Polimenis (2006) : "Affine Models for Credit Risk Analysis", *Journal of Financial Econometrics*, 4, 494-530.

Giddings, F., H. (1887) : "The Theory of Profit Sharing". *Quarterly Journal*

of Economics, 1, 376-376.

Greenwood, R., and D., Vayanos (2012): "Bond Supply and Excess Bond Returns". Working paper.

Guéant, O. (2007): "Mean Field Games and Applications to Economics. Secondary topic: Discount rates and sustainable development". Thesis dissertation. Université Paris-Dauphine.

Gupta, V., and E., Waymire (1993) : "A Statistical Analysis of Mesoscale Rainfall as a Random Cascade". Journal of Applied Meteorology, 32, 251-267.

Gurkaynak, R. S., Sack, B., and J. H., Wright (2007): "The U.S. Treasury yield curve: 1961 to the present". Journal of Monetary Economics, 54, 8, 2291-2304.

Hansen, L., and S., Richard (1987) : "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing". Econometrica, 55, 587-614.

Harrison, M., and D., Kreps (1979) : "Martingales and Arbitrage in Multi-period Securities Markets", Journal of the Economic Theory, 20, 381-408.

Hicks, S. (1939) : "Value and Capital, Oxford, UK, Clarendon Press.

Hogan, M. (1993): "Problems in certain Two-Factor Term Structure Models".

Annals of Applied Probability, 3, 222-239.

Hopewell, M., and G., Kaufman (1973) : "Bond Price Volatility and Term to Maturity : A Generalized Respecification", American Economic Review, 749-753.

Horton, S., D. (1886) : "Silver Before Congress in 1886". Quarterly Journal of Economics, 1, 45-75.

Huang, H., Marcantognini, A., and N., Young (2005) : "Chain Rules for Higher Order Derivatives". Working Paper.

Hubalek, F., Klein, I., and J., Teichmann (2002) : "A General Proof of the Dybvig-Ingersoll-Ross Theorem : Long Forward Rates Can Never Fall", Mathematical Finance, 12, 447-451.

Huber, J., P. (1964) : "Robust Estimation of a Location Parameter". The Annals of Mathematical Statistics, 35, 1, 73-101.

Hull, J., C. (2005) : "Options, Futures and Other Derivatives", Prentice Hall, 6th edition.

Ingersoll, J., Skelton, J., and R., Weil (1978) : "Duration Forty Years Later", Journal of Financial and Quantitative Analysis, 13, 627-650.

Joe, H. (1997) : "Multivariate Models and Dependence Concepts". Chapman & Hall, London.

Johnson, W. (2002) : "The Curious History of Faa di Bruno's Formula". American Mathematical Monthly, 109, 217-234.

Jones, F. (1991) : "Yield Curve Strategies", Journal of Fixed Income, 43-51.

Kardaras, C., and E., Platen (2012): "On the Dybvig-Ingersoll-Ross Theorem". Mathematical Finance, 22, 4, 729-740.

Koop, G., Pesaran, H., and S., Potter (1996): "Impulse Response Analysis in Nonlinear Multivariate Models". Journal of Econometrics, 74, 119-147.

Krippner, L. (2009) : "A Theoretical Foundation for the Nelson and Siegel Class of Yield Curve Models", Bank of New-Zealand DP.

Litterman, R., and J., Scheinkman (1991) : "Common Factors Affecting Bond Returns", Journal of Fixed Income, 1, 54-61.

Lucas, R. (1978) : "Asset Prices in an Exchange Economy". Econometrica, 4, 1429-1445.

Macaulay, F. (1931) : "The Smoothing of Time Series", New-York, NBER.

Macaulay, F. (1938) : "The Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856", New-York, NBER.

Markowitz, H. (1952) : "Portfolio Selection". Journal of Finance, 7, 77-99.

*Concluding Remarks*

Martin, R., and T., Wilde (2002) : "Unsystematic Credit Risk", Risk Magazine, 15, 123-128.

Merton, R., C. (1974) : "On the Pricing of Corporate Debt : the Risk Structure of Interest Rates". Journal of Finance, 29, 449-470.

Meucci, A. (2012): "Stress Testing with Fully Flexible Causal Inputs". Risk Magazine, Technical Paper.

Monfort, A., and F., Pegoraro (2007) : "Switching VARMA Term Structure Models". Journal of Financial Econometrics, 5,1, 105-153.

Monfort, A., and F., Pegoraro (2012) : "Asset Pricing with Second-Order Esscher Transforms" . The Journal of Banking and Finance, 36, 1678-1687.

Nelson, C., R., and A., F., Siegel (1987) : "Parsimonious Modeling of Yield Curves", The Journal of Business, 60, 4, 473-489.

Nolan, J. (2009) : "Stable Distributions : Models for Heavy Tailed Data", Springer Verlag.

Pérignon, C., Lopez, J., C., and J., Harris (2011) : "Clearing House, Margin Requirements, and Systemic Risk". Review of Futures Markets, 19, 39-54.

Piazzesi, M., and M., Schneider (2007) : "Equilibrium Yield Curves". NBER Macroeconomics Annual 2006, published in 2007, Cambridge MA: MIT Press, 389-442.



Raible, S. (2000) : "Lévy Processes in Finance: Theory, Numerics, and Empirical Facts". University of Fribourg. PhD dissertation.

Redington, F. (1952) : "Review of the Principles of Life-Office Valuations", Journal of the Institute of Actuaries, 18, 286-315.

Rogers, L. (1977) : "The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates", Mathematical Finance, 7, 157-176.

Rogers, L., and M., Tehranchi (2010) : "Can the Implied Volatility Move by Parallel Shifts?" , Finance and Stochastics, 14, 235-248.

Ross, S. (1976) : "The Arbitrage Theory of Capital Asset Pricing". Journal of Economic Theory, 13, 341-360.

Rothschild, M., and J., E. Stiglitz (1970) : "Increasing Risk: 1. A Definition", Journal of Economic Theory, 2, 225-283.

Sack, B. (2000) : "Using Treasury STRIPS to Measure the Yield Curve". Federal Reserve Working Paper.

Samorodnitsky, G., and M., S., Taqqu (1994) : "Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance". Chapman & Hall

Sklar, A. (1959) : "Fonctions de répartition à N dimensions et leurs marges". Publications de l'Institut de Statistiques de l'Université de Paris, 8, 229-231.

*Concluding Remarks*

Spindler, K. (2005) : "A Short Proof of the Formula of Faa di Bruno".  
Elemente der Mathematik, 60, 33-35.

Strassen, V. (1965) : "The Existence of Probability Measure with Given  
Marginals". The Annals of mathematical Statistics, 36, 423-439.

Svensson, L. (1995) : "Estimating Forward Interest Rates with the Extended  
Nelson and Siegel Model", Quarterly Review, Sveriges Riskbank, 3, 13-26.

Sydsaeter, K., and P., Hammond (2008) : "Essential Mathematics for Eco-  
nomic Analysis". 3<sup>rd</sup> ed., Prentice Hall.

Yao, Y. (1998): "Term Structure Models: A perspective from the Long Rate".  
North American Actuarial Journal, 3, 3, 122-138.



---

Part III

Appendix



# Appendix 1.1

## Shocks to Multivariate Distribution

### i) Multivariate Gaussian copula for bidimensional variables

**Definition a.1 :** The Gaussian copula for bidimensional variables is the function defined on  $[0, 1]^4$  by :

$$C(u_1, u_2, v_1, v_2) = \Psi \left[ \Phi^{-1}(u_1), \Phi^{-1}(u_2), \Phi^{-1}(v_1), \Phi^{-1}(v_2), R \right], \quad (3.6.1)$$

where  $\Psi(x_1, x_2, y_1, y_2, R)$  is the joint c.d.f. of the four-dimensional Gaussian distribution  $\mathcal{N} \left( \mathbf{0}, \begin{pmatrix} Id_2 & R \\ R & Id_2 \end{pmatrix} \right)$ . The block matrix  $R$  is such that  $R = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}$ , and  $\mathbf{0}$  is a 4x1 vector of zeros.

The block correlation matrix is constrained by the positive definiteness of the variance-covariance matrix, which imposes that :

$$\det \begin{pmatrix} 1 & 0 & \rho_{11} \\ 0 & 1 & \rho_{21} \\ \rho_{11} & \rho_{21} & 1 \end{pmatrix} = 1 - \rho_{11}^2 - \rho_{21}^2 > 0,$$

and

$$\begin{aligned} \det \begin{pmatrix} 1 & 0 & \rho_{11} & \rho_{12} \\ 0 & 1 & \rho_{21} & \rho_{22} \\ \rho_{11} & \rho_{21} & 1 & 0 \\ \rho_{12} & \rho_{22} & 0 & 1 \end{pmatrix} &= \det(Id) \det(Id - R'R) \\ &= \det(Id - R'R) \\ &= (1 - \rho_{11}^2 - \rho_{21}^2)(1 - \rho_{12}^2 - \rho_{22}^2) - (\rho_{11}^2 \rho_{12}^2 + \rho_{21}^2 \rho_{22}^2) > 0, \end{aligned}$$

by the block description of the determinant [see e.g. Gouriéroux, Monfort (1995), property A.5].

This copula defines a joint c.d.f. for the bivariate vectors:  $U = (U_1, U_2)'$  and  $V = (V_1, V_2)'$ , which have the same marginal uniform distribution on  $[0, 1]^2$ .

## ii) Derivation of the recursive form

We have now to derive the conditional c.d.f. of  $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$  given  $U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$  corresponding to this copula, in order to get interpretations in terms of variables. For this purpose, let us consider two standard Gaussian bivariate variables  $Y_0$  and  $Y_1$  :

$$Y_0 = \begin{pmatrix} Y_{01} \\ Y_{02} \end{pmatrix}, \quad Y_1 = \begin{pmatrix} Y_{11} \\ Y_{12} \end{pmatrix},$$

such that  $Cov(Y_0, Y_1) = R$ .

Then the variables  $\varepsilon_1 = F(Y_{11}|Y_0)$  and  $\varepsilon_2 = F(Y_{12}|Y_0, Y_{11})$  follow uniform distributions on  $[0, 1]$ , are independent of each other, and independent of  $Y_0$ .

By computing directly the conditional c.d.f., we get:

$$\varepsilon_1 = \Phi\left(\frac{Y_{11} - \mathbb{E}(Y_{11}|Y_0)}{\mathbb{V}^{1/2}(Y_{11}|Y_0)}\right) \quad (3.6.2)$$

$$\varepsilon_2 = \Phi\left(\frac{Y_{12} - \mathbb{E}(Y_{12}|Y_0, Y_{11})}{\mathbb{V}^{1/2}(Y_{12}|Y_0, Y_{11})}\right). \quad (3.6.3)$$

The conditional means and variances can be computed explicitly. We get:

$$\mathbb{E}(Y_{11}|Y_0) = \rho_{11}Y_{01} + \rho_{12}Y_{02} \equiv a_1Y_{01} + a_2Y_{02}, \quad (3.6.4)$$

$$\mathbb{V}(Y_{11}|Y_0) = 1 - \rho_{11}^2 - \rho_{12}^2 \equiv \sigma_1^2, \quad (3.6.5)$$

$$\begin{aligned} \mathbb{E}(Y_{12}|Y_0, Y_{11}) &= Cov \left[ Y_{12}, \begin{pmatrix} Y_0 \\ Y_{11} \end{pmatrix} \right] \left[ \mathbb{V} \begin{pmatrix} Y_0 \\ Y_{11} \end{pmatrix} \right]^{-1} \begin{pmatrix} Y_0 \\ Y_{11} \end{pmatrix} \\ &= \begin{pmatrix} \rho_{12} & \rho_{22} \end{pmatrix} \left[ Id - \begin{pmatrix} \rho_{11} \\ \rho_{21} \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{21} \end{pmatrix} \right]^{-1} \begin{pmatrix} Y_0 \\ Y_{11} \end{pmatrix} \\ &\equiv b_1Y_{01} + b_2Y_{02} + b_3Y_{11}, \end{aligned} \quad (3.6.6)$$

by applying the block inversion formula [see e.g. Gouriéroux, Monfort (1995) Property A.4]. Moreover :

$$\mathbb{V}(Y_{12}|Y_0, Y_{11}) = 1 - \begin{pmatrix} \rho_{12} & \rho_{22} \end{pmatrix} \left[ Id - \begin{pmatrix} \rho_{11} \\ \rho_{21} \end{pmatrix} \begin{pmatrix} \rho_{11} & \rho_{21} \end{pmatrix} \right]^{-1} \begin{pmatrix} \rho_{12} \\ \rho_{22} \end{pmatrix} \equiv \sigma_2^2, \quad (3.6.7)$$

by applying the block inversion formula.

We deduce that:

$$\begin{cases} Y_{11} = a_1Y_{01} + a_2Y_{02} + \sigma_1\Phi^{-1}(\varepsilon_1), \\ Y_{12} = b_1Y_{01} + b_2Y_{02} + b_3Y_{11} + \sigma_2\Phi^{-1}(\varepsilon_2), \end{cases}$$

or equivalently the conditional distribution of  $V$  given  $U$  is such that we get the recursive system of equations:

$$\begin{cases} \Phi^{-1}(V_1) = a_1\Phi^{-1}(U_1) + a_2\Phi^{-1}(U_2) + \sigma_1\Phi^{-1}(\varepsilon_1), \\ \Phi^{-1}(V_2) = b_1\Phi^{-1}(U_1) + b_2\Phi^{-1}(U_2) + b_3\Phi^{-1}(V_1) + \sigma_2\Phi^{-1}(\varepsilon_2), \end{cases}$$

where  $\varepsilon_1, \varepsilon_2$  are independent uniform variables independent of  $U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$ .

Let us finally consider a bivariate vectors  $Y_0 = \begin{pmatrix} Y_{01} \\ Y_{02} \end{pmatrix}$ ,  $Y_1 = \begin{pmatrix} Y_{11} \\ Y_{12} \end{pmatrix}$  with marginal distributions  $F_0(y_0)$ ,  $F_1(y_1)$ , say. The marginal distribution



of the bivariate vector  $Y_0$  can be decomposed into the marginal distribution of  $Y_{01}$  with c.d.f.  $F_{01}(y_{01})$  and the conditional distribution of  $Y_{02}$  given  $Y_{01}$  with conditional c.d.f.  $F_{02|01}(y_{02}|y_{01})$ . Similar notations are introduced for the second bivariate vector, that are  $F_{11}(y_{11})$  and  $F_{12|11}(y_{12}|y_{11})$ .

The basic uniform variable  $U_1, U_2, V_1, V_2$  can be chosen such that:

$$\begin{aligned} U_1 &= F_{01}(Y_{01}), & U_2 &= F_{02|01}(Y_{02}|Y_{01}), \\ V_1 &= F_{11}(Y_{11}), & V_2 &= F_{12|11}(Y_{12}|Y_{11}). \end{aligned}$$

Then the two bidimensional stochastic variables  $Y_0$  and  $Y_1$  can be linked by a Gaussian copula (3.6.1).

**Proposition a.2 :** Let us consider a pair of bidimensional stochastic variables  $Y_0 = \begin{pmatrix} Y_{01} \\ Y_{02} \end{pmatrix}$ ,  $Y_1 = \begin{pmatrix} Y_{11} \\ Y_{12} \end{pmatrix}$ , such that the marginal c.d.f of  $Y_{01}$  is  $F_{01}(y_{01})$  and the conditional distribution of  $Y_{02}$  given  $Y_{01}$  is  $F_{02|01}(y_{02}|y_{01})$  [resp.  $F_{11}(y_{11})$  and  $F_{12|11}(y_{12}|y_{11})$ ].

Let us denote:

$$\begin{aligned} U_1 &= F_{01}(Y_{01}), & U_2 &= F_{02|01}(Y_{02}|Y_{01}), \\ V_1 &= F_{11}(Y_{11}), & V_2 &= F_{12|11}(Y_{12}|Y_{11}). \end{aligned}$$

Vectors  $Y_0$  and  $Y_1$  admit a Gaussian copula if and only if the conditional distribution of  $Y_1$  given  $Y_0$  can be represented by the system of equations:

$$\begin{aligned} \Phi^{-1}[F_{11}(Y_{11})] &= a_1\Phi^{-1}[F_{01}(Y_{01})] + a_2\Phi^{-1}[F_{02|01}(Y_{02}|Y_{01})] + \sigma_1\Phi^{-1}(\varepsilon_1) \\ \Phi^{-1}[F_{12|11}(Y_{12}|Y_{11})] &= b_1\Phi^{-1}[F_{01}(Y_{01})] + b_2\Phi^{-1}[F_{02|01}(Y_{02}|Y_{01})] + b_3\Phi^{-1}[F_{11}(Y_{11})] + \sigma_2\Phi^{-1}(\varepsilon_2) \end{aligned}$$

where  $\varepsilon_1, \varepsilon_2$  are independent uniform variables, independent of  $Y_0$ , and the coefficients  $a_1, a_2, b_1, b_2, b_3, \sigma_1, \sigma_2$  are function of  $R$  given in (3.6.4)-(3.6.7).

The extension to parametric families of bivariate variables  $Y_\delta$  is obtained by making the matrix  $R(\delta)$  function of  $\delta$ . The coherency condition (1.2.9) then implies  $R(0) = Id_2$ .

## ii) Local analysis for multivariate distribution

Let us consider a multivariate variable  $Y_\delta = h(Y_0, \varepsilon; \delta)$ , where  $Y_\delta, Y_0, \varepsilon$  are vectors of dimension  $N$ ,  $h(Y_0, \varepsilon; \delta) = (h_1(\bullet), \dots, h_N(\bullet))'$ , and the expectation  $\mathbb{E}(g[Y_\delta])$ , where  $g$  is a function of dimension 1. Thus,

$$\begin{aligned}
\mathbb{E}(g[Y_\delta]) &= \mathbb{E}(g[h(Y_0, \varepsilon; \delta)]) \\
&= \mathbb{E}(g(Y_0)) + \delta \mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \frac{\partial h(Y_0, \varepsilon; 0)}{\partial \delta}\right) \\
&\quad + \frac{\delta^2}{2} \mathbb{E}\left(\frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta} \frac{\partial^2 g(Y_0)}{\partial y \partial y'} \frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta}\right) + \frac{\delta^2}{2} \mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \frac{\partial^2 h(Y_0, \varepsilon; 0)}{\partial \delta^2}\right) + o(\delta^2) \\
&= \mathbb{E}(g(Y_0)) + \delta \mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \mathbb{E}\left(\frac{\partial h(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0\right)\right) \\
&\quad + \frac{\delta^2}{2} \mathbb{E}\left(\text{Tr} \left[ \frac{\partial^2 g(Y_0)}{\partial y \partial y'} \mathbb{E}\left(\frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta} \frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0\right)\right]\right) \\
&\quad + \frac{\delta^2}{2} \mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \mathbb{E}\left(\frac{\partial^2 h(Y_0, \varepsilon; 0)}{\partial \delta^2} \middle| Y_0\right)\right) + o(\delta^2)
\end{aligned}$$

Then, we have :

$$\begin{aligned}
\mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \mathbb{E}\left(\frac{\partial h(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0\right)\right) &= \sum_{j=1}^N \mathbb{E}\left[\frac{\partial g(Y_0)}{\partial y_j} \mathbb{E}\left(\frac{\partial h_j(Y_0, \varepsilon; 0)}{\partial \delta}\right)\right] \\
&= - \sum_{j=1}^N \mathbb{E}\left[\frac{g(Y_0)}{f(Y_0; 0)} \frac{d}{dy_j} \left[\mathbb{E}\left(\frac{\partial h_j(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0 = y\right) f(y; 0)\right]\right]
\end{aligned}$$

Similarly, we have :

$$\mathbb{E}\left(\frac{\partial g(Y_0)}{\partial y'} \mathbb{E}\left(\frac{\partial^2 h(Y_0, \varepsilon; 0)}{\partial \delta^2} \middle| Y_0\right)\right) = - \sum_{j=1}^N \mathbb{E}\left(\frac{g(Y_0)}{f(Y_0; 0)} \frac{d}{dy_j} \left[\mathbb{E}\left(\frac{\partial^2 h_j(Y_0, \varepsilon; 0)}{\partial \delta^2} \middle| Y_0 = y\right) f(y; 0)\right]\right),$$

and

$$\begin{aligned}
&\mathbb{E}\left(\text{Tr} \left[ \frac{\partial^2 g[Y_0]}{\partial y \partial y'} \mathbb{E}\left(\frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta} \frac{\partial h'(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0 = y\right)\right]\right) \\
&= \sum_{j=1}^N \sum_{k=1}^N \frac{d}{dy_j dy_k} \mathbb{E}\left(\frac{g(Y_0)}{f(Y_0; 0)} \left[\mathbb{E}\left(\frac{\partial h_j(Y_0, \varepsilon; 0)}{\partial \delta} \frac{\partial h_k(Y_0, \varepsilon; 0)}{\partial \delta} \middle| Y_0 = y\right) f(y; 0)\right]\right)
\end{aligned}$$

Thus the multivariate equivalent of Proposition 1 is :

$$\begin{aligned}\frac{\partial}{\partial \delta}[f(y; 0)] &= -\sum_{j=1}^N \frac{d}{dy_j} \left[ \mathbb{E} \left( \frac{\partial h_j(Y_0, \varepsilon; 0)}{\partial \delta} | Y_0 = y \right) f(y; 0) \right] \\ \frac{\partial^2}{\partial \delta^2}[f(y; 0)] &= \sum_{j=1}^N \sum_{k=1}^N \frac{d}{dy_j dy_k} \left[ \mathbb{E} \left( \frac{\partial h_j(Y_0, \varepsilon; 0)}{\partial \delta} \frac{\partial h_k(Y_0, \varepsilon; 0)}{\partial \delta} | Y_0 = y \right) f(y; 0) \right] \\ &\quad - \sum_{j=1}^N \frac{d}{dy_j} \left[ \mathbb{E} \left( \frac{\partial^2 h_j(Y_0, \varepsilon; 0)}{\partial \delta^2} | Y_0 = y \right) f(y; 0) \right].\end{aligned}$$

## Appendix 1.2

### Proof of Lemma 2

The result is obtained by a sequence of integration by part. Let us denote  $(\underline{y}, \bar{y})$  the support of the function  $g$ . We get :

$$\begin{aligned}\mathbb{E} \left[ g^{(k)}(Y_0) a(Y_0) \right] &= \int_{\underline{y}}^{\bar{y}} g^{(k)}(y) a(y) f(y; 0) dy \\ &= \left[ g^{(k-1)}(y) a(y) f(y; 0) \right]_{\underline{y}}^{\bar{y}} - \int_{\underline{y}}^{\bar{y}} g^{(k-1)}(y) \frac{d}{dy} [a(y) f(y; 0)] dy \\ &= - \int_{\underline{y}}^{\bar{y}} g^{(k-1)}(y) \frac{d}{dy} [a(y) f(y; 0)] dy \quad (\text{since } g(\underline{y}) = g(\bar{y}) = 0) \\ &= (-1)^k \int_{\underline{y}}^{\bar{y}} g(y) \frac{d^k}{dy^k} [a(y) f(y; 0)] dy \quad (\text{by applying recursively the same}) \\ &= (-1)^k \mathbb{E} \left[ \frac{g(Y_0)}{f(Y_0; 0)} \frac{d^k}{dy^k} [a(Y_0) f(Y_0; 0)] \right].\end{aligned}$$

## Appendix 1.3

### An alternative derivation of the expansion in Corollary 1

Let us consider the specification  $Y_\delta = h(Y_0, U; \delta) = Y_0 + \delta a(Y_0)U$ , where  $U$  is a variable independent of  $Y_0$  with p.d.f.  $g(u)$ ,  $\delta \geq 0, a(\bullet) > 0$ . We

assume that :

**Assumption A1 :** Given  $Y_0$  and  $\delta$ , there is an increasing bijective relationship between  $U$  and  $Y_\delta$ ;

**Assumption A2 :** Given  $Y_\delta$  and  $\delta$ , there is an increasing bijective relationship between  $U$  and  $Y_0$ .

Under Assumption  $A_1$ , we can write :

$$u = \frac{y - y_0}{\delta a(y_0)},$$

Then the p.d.f. of  $Y_\delta$  conditional to  $Y_0$  is :

$$f(y|y_0; \delta) = \frac{1}{\delta a(y_0)} g\left(\frac{y - y_0}{\delta a(y_0)}\right),$$

and the unconditional p.d.f of  $Y_\delta$  is :

$$f(y; \delta) = \int [\delta a(y_0)]^{-1} g\left(\frac{y - y_0}{\delta a(y_0)}\right) f(y_0; 0) dy_0. \quad (3.6.8)$$

Under Assumption  $A_2$ , we have :

$$y = y_0 + \delta a(Y_0)u \Leftrightarrow y_0 = c(y, u; \delta), \text{ say.}$$

Let us now consider the change of variable  $y_0 \rightarrow u$  in integral (3.6.8). We get :

$$f(y; \delta) = \int [\delta a(c(y, u; \delta))]^{-1} g(u) f(c(y, u; \delta); 0) \left| \frac{\partial c(y, u; \delta)}{\partial u} \right| du. \quad (3.6.9)$$

At first-order in  $\delta$  we get :  $c(y, u; \delta) \approx y - \delta a(y)u$ . Thus, (3.6.9) becomes :

$$\begin{aligned} f(y; \delta) &\approx \int \frac{a(y)}{a(y - \delta a(y)u)} g(u) f(y - \delta a(y)u; 0) du \\ &\approx \int [1 - \delta a^{(1)}(y)u]^{-1} g(u) [f(y; 0) - \delta f^{(1)}(y; 0)a(y)u] du \\ &\approx \int [1 + \delta a^{(1)}(y)u] g(u) [f(y; 0) - \delta f^{(1)}(y; 0)a(y)u] du \\ &\approx \int g(u) f(y; 0) du - \delta \int [f^{(1)}(y; 0)a(y)u - a^{(1)}(y)u f(y; 0)] du \end{aligned} \quad (3.6.10)$$

,which is the first term in expansion of Corollary 1.

## Appendix 1.4

### From a distribution-based to a variable-based approach with Gaussian copula

The Gaussian copula is given by :

$$C(u, v) = \Psi \left[ \Phi^{-1}(u), \Phi^{-1}(v), \rho \right],$$

where  $\Psi(x, y, \rho)$  is the joint c.d.f. of the bivariate Gaussian distribution  $\mathcal{N} \left( \left( \begin{matrix} 0 \\ 0 \end{matrix} \right), \left( \begin{matrix} 1 & \rho \\ \rho & 1 \end{matrix} \right) \right)$ .

#### i) Derivatives of equation (1.2.4)

Let us consider a Gaussian copula with correlation parameter  $\rho$  and two variables  $Y_0, Y_1$ , whose marginal distributions are standard normal. We have from (1.2.4) :

$$\begin{aligned} \varepsilon &= F_{1|0}(Y_1|Y_0) \\ &= \frac{\partial C}{\partial u} [F_0(Y_0), F_1(Y_1)] \\ &= \Phi \left( \frac{\Phi^{-1}(V) - \rho \Phi^{-1}(U)}{\sqrt{1 - \rho^2}} \right), \end{aligned}$$

where :

$$V = \Phi(Y_1), \quad U = \Phi(Y_0).$$

Therefore :

$$\Phi^{-1}(V) = \rho \Phi^{-1}(U) + \sqrt{1 - \rho^2} \Phi^{-1}(\varepsilon), \quad (3.6.11)$$

where  $\varepsilon$  has a uniform distribution on  $[0;1]$ . The result is easily extended to families of distribution by taking Gaussian copulas with correlation parameter  $\rho(\delta)$  indexed by  $\delta$ .

## ii) Contamination

Let us now consider the contamination of Section 4.1.2 in the Gaussian copula case :

$$F(y; \delta) = (1 - \delta)F(y; 0) + \delta\Xi(y).$$

Thus :

$$\begin{aligned} \varepsilon &= \Phi \left( \frac{\Phi^{-1}(F(Y_\delta; \delta)) - \rho(\delta)\Phi^{-1}(F(Y_0; 0))}{\sqrt{1 - \rho^2(\delta)}} \right) \\ &= \Phi \left( \frac{\Phi^{-1}((1 - \delta)F(Y_\delta; 0) + \delta\Xi(Y_\delta)) - \rho(\delta)\Phi^{-1}(F(Y_0; 0))}{\sqrt{1 - \rho^2(\delta)}} \right), \end{aligned}$$

which gives :

$$\Phi \left( \sqrt{1 - \rho^2(\delta)}U + \rho(\delta)\Phi^{-1}(F(Y_0; 0)) \right) = (1 - \delta)F(Y_\delta; 0) + \delta\Xi(Y_\delta), \quad (3.6.12)$$

where  $U = \Phi^{-1}(\varepsilon)$  is a standard Gaussian variable.

## iii) First-order expansion

Let us now assume that  $\rho(\delta) = 1 - \delta^2 r + o(\delta^2)$ , where  $r = -\left. \frac{\partial \rho(\delta)}{\partial \delta} \right|_{\delta=0}$ , and approximate  $Y_\delta$  at first-order :

$$Y_\delta = Y_0 + \delta Z + o(\delta).$$

More precisely, let us consider the two sides of equation (3.6.12). We get for the left hand side :

$$\begin{aligned} \Phi \left( \sqrt{1 - \rho^2(\delta)}U + \rho(\delta)\Phi^{-1}(F(Y_0; 0)) \right) &\approx \Phi \left( \sqrt{2r\delta^2}U + \Phi^{-1}(F(Y_0; 0)) \right) \\ &\approx F(Y_0; 0) + \delta\sqrt{2r}\phi \left( \Phi^{-1}(F(Y_0; 0)) \right) \end{aligned}$$

and

$$\begin{aligned}(1 - \delta)F(Y_\delta; 0) + \delta\Xi(Y_\delta) &= F(Y_\delta; 0) + \delta(\Xi(Y_\delta) - F(Y_\delta; 0)) \\ &\approx F(Y_0; 0) + \delta Zf(Y_0; 0) + \delta(\Xi(Y_0) - F(Y_0; 0)), \text{ for the right h}\end{aligned}$$

This provides the expression of variable  $Z$ , that is (1.5.6), by identification.

## Appendix 1.5

### Application: contamination in terms of variable

GE	FR	IT	SP	IR	GR
-0.9668	-0.8860	-0.9037	-0.7851	-0.9293	-0.7971

Table 3.1: Mean of excess gains on 2001-2007. All numbers must be divided by 1000.

GE	FR	IT	SP	IR	GR
0.9759	0.3775	-1.2251	-2.1459	-7.3262	-8.9894

Table 3.2: Mean of excess gains on 2007-2011. All numbers must be divided by 1000.

	GE	FR	IT	SP	IR	GR
GE	0.1253	0.1164	0.1206	0.1195	0.1247	0.1131
FR	0.1164	0.1252	0.1196	0.1235	0.1292	0.1173
IT	0.1206	0.1196	0.1278	0.1215	0.1263	0.1153
SP	0.1195	0.1235	0.1215	0.1274	0.1306	0.1201
IR	0.1247	0.1292	0.1263	0.1306	0.1386	0.1238
GR	0.1131	0.1173	0.1153	0.1201	0.1238	0.1166

Table 3.3: Covariance matrix of excess gains on 2001-2007. All numbers must be divided by 1000.

	GE	FR	IT	SP	IR	GR
GE	0.2838	0.2346	0.1140	0.1370	0.1491	0.0496
FR	0.2346	0.2224	0.1253	0.1673	0.2023	0.1023
IT	0.1140	0.1253	0.1567	0.1838	0.2239	0.1436
SP	0.1370	0.1673	0.1838	0.2923	0.3545	0.2524
IR	0.1491	0.2023	0.2239	0.3545	0.6431	0.3645
GR	0.0496	0.1023	0.1436	0.2524	0.3645	0.8449

Table 3.4: Covariance matrix of excess gains on 2007-2011. All numbers must be divided by 1000.



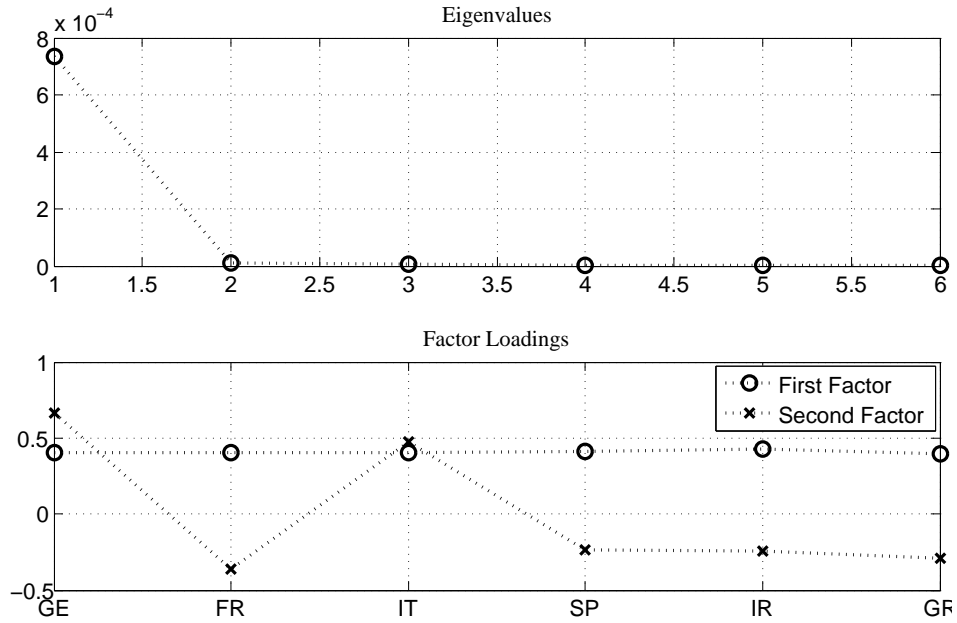


Figure 3.15: Eigenvalues and factor loadings from the principal components analysis of the 2001-2011 excess gains covariance matrix

## Appendix 1.6

### Application: contamination in terms of distribution

Let us perform the same application as in Section 6, when the contamination of the systematic factor is expressed in terms of distribution. In this case, we rely on the specification (1.5.3) to simulate the contaminated factors from a mixture of the baseline distribution of the factor with c.d.f.  $F(x; 0)$ , which is estimated on the 2001-2007 sample, and a contaminating distribution with c.d.f.  $\Xi(x)$  inferred from the 2007-2011 period :

$$F(x; \delta) = (1 - \delta)F(x; 0) + \delta\Xi(x), \text{ with } 0 \leq \delta \leq 1.$$

More precisely, we proceed in three steps :

1. We draw a set of  $S = 1000$  uniform independent variables  $(\omega^s)_{s=1\dots S}$  on  $[0,1]$ , and two sets of  $S = 1000$  non-contaminated and contaminated factors  $[(X_0^s)_{s=1\dots S}$  and  $(X_1^s)_{s=1\dots S}$ , respectively] from the realized factors on 2001-

2007, and 2007-2011.

2. For a given  $\delta$  we compute the variable :

$$\begin{cases} Z_\delta^s = 1, & \text{when } \omega^s \leq \delta, \\ Z_\delta^s = 0, & \text{otherwise.} \end{cases}$$

3. Finally, we compute the contaminated factor :

$$X_\delta^s = (1 - Z_\delta^s)X_0^s + Z_\delta^s X_1^s.$$

The empirical distribution of the simulated factor is presented in Figure 3.16 for  $\delta = 0, 0.1, 0.5,$  and  $1$ , while the properties of the crystallized and optimally adjusted mean-variance portfolios are plotted in Figures 3.16-3.21. As in the contamination in terms of variable, the characteristics of the optimal mean-variance portfolio deviates significantly from the crystallized portfolio's ones. As expected the analysis in terms of variable and distribution look similar for small  $\delta$ , but they can deviate significantly when  $\delta$  becomes larger.

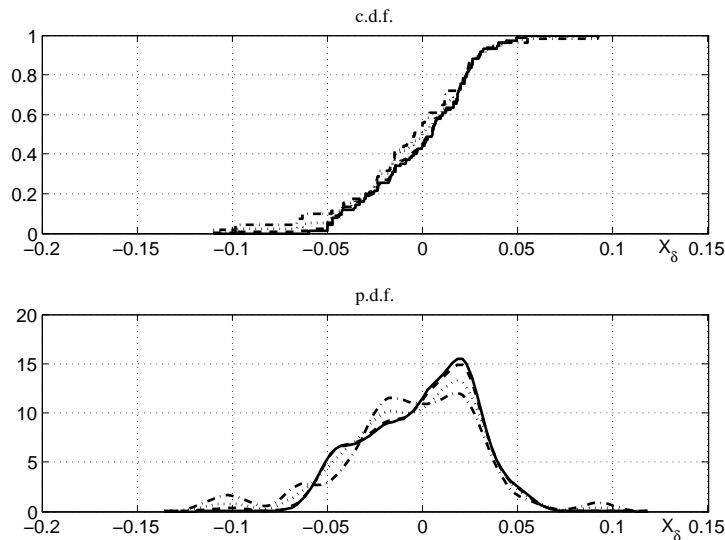


Figure 3.16: Empirical c.d.f and p.d.f. of 1000 simulated factor  $X_\delta$ , contaminated in terms of distribution for various  $\delta$ . The solid line stands for the c.d.f. of the baseline factor  $X_0$ , while dashed, dotted, and dash-dotted lines represent the contaminated empirical distribution for  $\delta = 0.1, 0.5, 1$ .

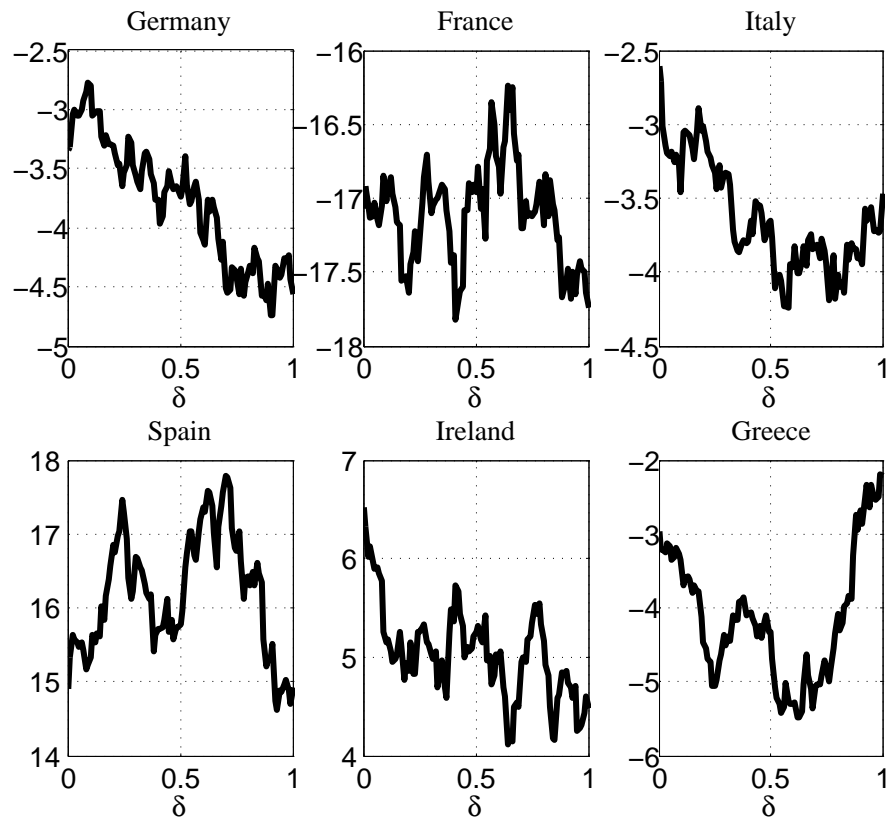


Figure 3.17: Contaminated mean-variance allocation as a function of  $\delta$ .

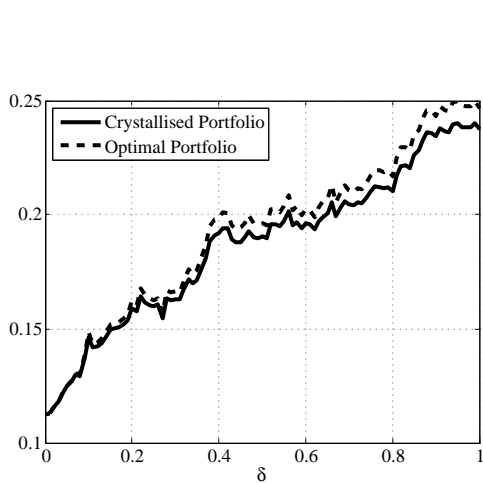


Figure 3.18: Impulse response of the Sharpe ratio of crystallized and Mean-Variance portfolios (contamination in distribution).

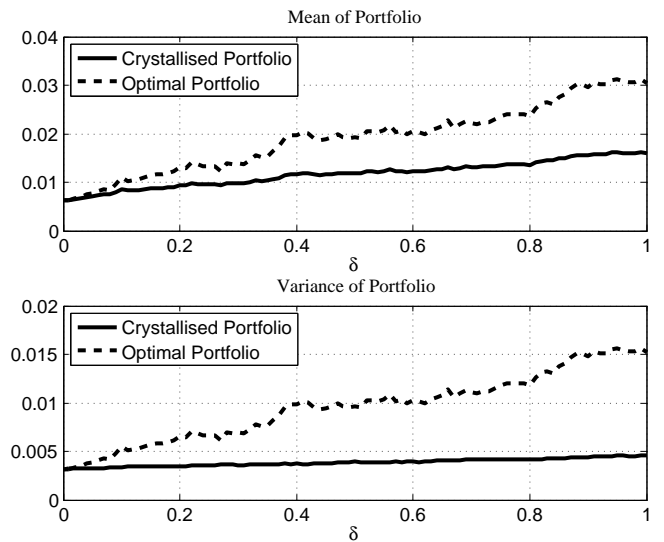


Figure 3.19: Impulse response of the Mean and Variance of crystallized and Mean-Variance portfolios (contamination in distribution).

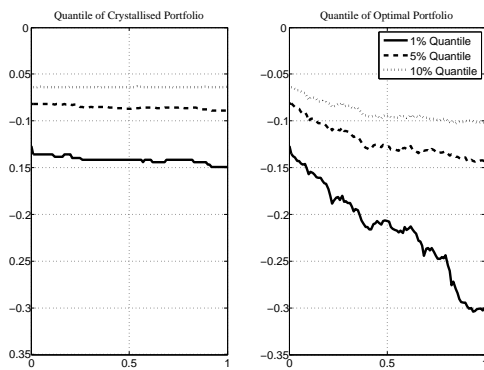


Figure 3.20: Impulse response of the VaR of crystallized and Mean-Variance portfolios (contamination in distribution).

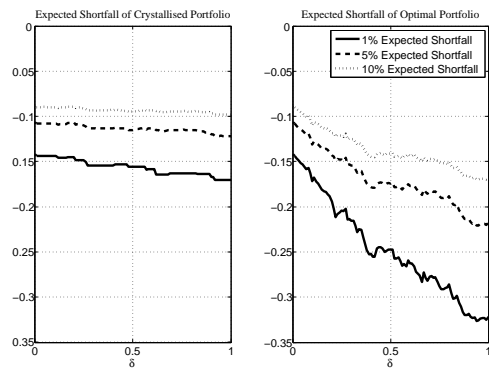


Figure 3.21: Impulse response of the Expected-shortfall of crystallized and Mean-Variance portfolios (contamination in distribution).

**Appendix 2.1**  
**Parallel Shift and Change of Slope in the Term Structure**  
**Literature**

**1. The level in the literature**

**i) From a coupon bond to a zero-coupon bond**

Very early in the literature [Macaulay (1931), (1938), p.48] appears the idea to replace a coupon bond by an "equivalent" zero-coupon bond in order to facilitate the comparison of bonds with different maturities and seasoning. More precisely, let us consider at time  $t$  a coupon bond with nonnegative coupons  $A_h, h = 0, 1 \dots$  at the different times-to-maturity, and a current price  $\Pi_t(A)$ . To create the "equivalent" zero-coupon bond, we have to define the corresponding rate and time-to-maturity. They are usually defined as follows : the equivalent rate, or yield, is the solution  $r_t^I(A)$  of the equation :

$$\Pi_t(A) = \sum_{h=0}^{\infty} A_h \exp[-hr_t^I(A)].$$

The equivalent time-to-maturity is the Macaulay's duration <sup>71</sup> defined by :

$$D_t^I(A) = \sum_{h=0}^{\infty} hA_h \exp[-hr_t^I(A)] / \sum_{h=0}^{\infty} A_h \exp[-hr_t^I(A)].$$

It is equal to the average time-to-maturity of the coupons weighted by the discounted coupons, which corresponds to a modified probability measure with elementary probabilities  $\left( q_t(h) = A_h \exp[-hr_t^I(A)] / \sum_{h=0}^{\infty} A_h \exp[-hr_t^I(A)] \right)$ .

---

<sup>71</sup>The eponym "Macaulay's duration" has been introduced in Fisher, Weil (1971), p416.

In a modern terminology, these two notions are an implied rate and an implied time-to-maturity, since they are computed from a misspecified term structure model, which assumes a flat term structure, possibly varying in time :

$$r(t, h) = X_t, \forall h, \text{ say.} \quad (\text{A.1})$$

## ii) Consistency with no-arbitrage

The flat term structure model (A.1) underlying the derivation and interpretation of the yield and duration hardly coincides with the true term structure. Nevertheless, this misspecified model should be consistent with no-arbitrage restrictions.

From (A.1), we note that the underlying model is a special case of model (1.1) with  $c(h) = 0, \forall h$ . By arguments similar to the arguments in Section 4, we deduce that, under no-arbitrage, the dynamic of  $(X_t)$  is such that :

$$X_{t+1} = X_t, \forall t. \quad (\text{A.2})$$

Therefore, under no-arbitrage, the term structure is flat at all dates if and only if it is also time independent :

$$\begin{aligned} X_t &= X_0, \forall t, \\ \Leftrightarrow r(t, h) &= X_0, \forall t, h \text{ say.} \end{aligned} \quad (\text{A.3})$$

Thus the no-arbitrage restriction induces strong links between the pattern of the term structure (which is flat) and its evolution (which is constant in time).

The underlying model can be stochastic if the initial value is stochastic, but the associated notion of shock is very special. Indeed, a shock on  $X_0$  can be introduced : this shock will have a drift effect not only on the term structure at date  $t$ , but on the term structures of all dates jointly. Under no-arbitrage on the underlying misspecified model, a transitory shock on a term structure, that is a shock specific to date  $t$ , cannot be defined. This shock is systematically permanent.<sup>72</sup>

The aim of Assumption A.2 was to eliminate this very special limiting case.

### iii) The duration as a sensitivity parameter.

It is also well-known that the duration is a measure of the sensitivity of the bond price with respect to shock on the level of interest rate, which does not depend on time-to-maturity due to the assumption of a flat term structure [see e.g. Hicks (1939), Redington (1952), Fisher (1966), Hopewell, Kaufman (1973)].

More precisely, let us shock the flat term structure  $r(t, h) = r_t^I(A)$ ,  $\forall h$ , into  $r^l(t, h) = r_t^I(A) + \delta$ ,  $\forall h$ , and consider the associated value  $\Pi_t^l(A, \delta) = \sum_{h=0}^{\infty} A_h \exp[-h(r_t^I(A) + \delta)]$ , where the upper index  $l$  mentions a shock on level. We have :

$$\left. \frac{\partial \log \Pi_t^l(A, \delta)}{\partial \delta} \right|_{\delta=0} = D_t^I(A). \quad (\text{A.4})$$

## 2. The slope in the literature

---

<sup>72</sup>Theorem 4 in Ingersoll, Skelton, Weil (1978) provides an alternative proof of the result. They show that a transitory shift in a flat term structure is not arbitrage free.

### i) The convexity

The sensitivity analysis can be extended at second-order. The convexity is the second-order derivative of the log-price<sup>73</sup> function with respect to shock  $\delta$  :

$$\begin{aligned}
 C_t^I(A) &= \left. \frac{\partial^2 \log \Pi_t^I(A, \delta)}{\partial \delta^2} \right|_{\delta=0} \\
 &= \frac{\sum_{h=0}^{\infty} h^2 A_h \exp[-hr_t^I(A)]}{\sum_{h=0}^{\infty} A_h \exp[-hr_t^I(A)]} - \left( \frac{\sum_{h=0}^{\infty} h A_h \exp[-hr_t^I(A)]}{\sum_{h=0}^{\infty} A_h \exp[-hr_t^I(A)]} \right)^2 \\
 &= \sum_{h=0}^{\infty} h^2 q_t(h) - \left[ \sum_{h=0}^{\infty} h q_t(h) \right]^2 .
 \end{aligned}$$

Thus the convexity can be interpreted as the variance of the time-to-maturity of the bond under modified probability measure  $[q_t(h)]$ .

### ii) Effect of an affine linear shock

Let us now consider an affine linear shock, whose effect on the yield curve differs with the maturity of the rates, that is,

$$r^s(t, h, \delta) = r_t^I(A) + \delta [h - D_t^I(A)] , \quad \forall h,$$

where the shock on the yield curve is calibrated around the bond's duration  $D_t^I(A)$ , and impacts the slope without affecting the level of the yield curve.

---

<sup>73</sup>Our definition differs from the usual definition of convexity [see for instance Hull (2005), p.92], according to which convexity is equal to the second-order derivative of the price function  $\frac{\partial^2 \Pi_t^I(A, \delta)}{\partial \delta^2}$ . The second-order Taylor expansion is often used to derive approximated prices, that is, to consider  $\Pi_t^I(A, 0) + \delta \frac{\partial \Pi_t^I(A, 0)}{\partial \delta} + \frac{\delta^2}{2} \frac{\partial^2 \Pi_t^I(A, 0)}{\partial \delta^2}$  instead of  $\Pi_t^I(A, \delta)$ . Such a Taylor expansion does not ensure the positivity of the approximated prices. At the opposite, this property is satisfied when the expansion is performed on the log-price as proposed in our definition.



In this case, the bond price becomes  $\Pi_t^s(A, \delta) = \sum_{h=0}^{\infty} A_h \exp \left[ -h \left( r_t^I(A) + \delta \left[ h - D_t^I(A) \right] \right) \right]$ , and the first-order sensitivity of the bond price to the affine linear shock becomes :

$$\begin{aligned} \left. \frac{\partial \log \Pi_t^s(A, \delta)}{\partial \delta} \right|_{\delta=0} &= - \frac{\sum_{h=0}^{\infty} \left\{ h \left[ h - D_t^I(A) \right] A_h \exp \left[ -hr_t^I(A) \right] \right\}}{\sum_{h=0}^{\infty} A_h \exp \left[ -hr_t^I(A) \right]} \\ &= - \sum_{h=0}^{\infty} \left\{ h \left[ h - D_t^I(A) \right] q_t(h) \right\} \\ &= -C_t^I(A), \end{aligned}$$

which is the (opposite of) bond's convexity.

## Appendix 2.2

### Proof of Equation 5.4

By definition, we have :

$$B(t, h) = \exp[-hr(t, h)] = \exp[-hg(Y_t, h)] = F(Y_t, h), \text{ say.}$$

Since  $(Y_t)$  is an Ito process satisfying  $dY_t = \mu(Y_t)dt + \Sigma^{1/2}(Y_t)dW_t$ , the function  $F$  satisfies the following PDE [see Duffie (2001), Chapter 7, eq. (22)] :

$$g(y, 0)F(y, h) = -\frac{\partial F(y, h)}{\partial h} + \frac{\partial F(y, h)}{\partial y'}\mu(y) + \frac{1}{2}Tr \left[ \Sigma(y) \frac{\partial^2 F(y, h)}{\partial y \partial y'} \right], \quad (\text{A.5})$$

with boundary conditions.

Since  $F(y, h) = \exp[-hg(y, h)]$  we get:

$$\frac{\partial F(y, h)}{\partial h} = -F(y, h) \left( g(y, h) + h \frac{\partial g(y, h)}{\partial h} \right),$$

$$\frac{\partial F(y, h)}{\partial y} = F(y, h) \left( -h \frac{\partial g(y, h)}{\partial y} \right),$$

$$\frac{\partial^2 F(y, h)}{\partial y \partial y'} = F(y, h) \left( h^2 \frac{\partial g(y, h)}{\partial y} \frac{\partial g(y, h)}{\partial y'} - h \frac{\partial^2 g(y, h)}{\partial y \partial y'} \right).$$

Therefore equation (A.5) becomes :

$$\begin{aligned} g(y, 0) &= \left( g(y, h) + h \frac{\partial g(y, h)}{\partial h} \right) - \left( h \frac{\partial g(y, h)}{\partial y'} \right) \mu(y) + \frac{1}{2}Tr \left[ \Sigma(y) \left( h^2 \frac{\partial g(y, h)}{\partial y} \frac{\partial g(y, h)}{\partial y'} - h \frac{\partial^2 g(y, h)}{\partial y \partial y'} \right) \right] \\ &= g(y, h) + h \frac{\partial g(y, h)}{\partial h} - h \frac{\partial g(y, h)}{\partial y'} \mu(y) + \frac{1}{2}h^2 \frac{\partial g(y, h)}{\partial y'} \Sigma(y) \frac{\partial g(y, h)}{\partial y} - \frac{1}{2}hTr \left[ \Sigma(y) \frac{\partial^2 g(y, h)}{\partial y \partial y'} \right], \end{aligned}$$

which gives equation (2.5.4).

**Appendix 2.3**  
**Proof of Proposition 7**

i) The coefficients of  $\mu_1(y)$ ,  $\mu_2(y)$ ,  $\sigma_{11}(y)$ ,  $\sigma_{12}(y)$ ,  $\sigma_{22}(y)$  in system (2.5.5) are the functions :

$$h, h\beta(h), -\frac{h^2}{2}, -h^2\beta(h), -h^2\beta^2(h), \quad h \in (0, \infty).$$

When these five functions are linearly independent, system (2.5.5) admits a unique solution  $\mu(y)$ ,  $\Sigma(y)$ , (if a solution exists), which is necessarily affine in  $z$  due to the expression of the left hand side of system (2.5.5).

ii) Thus we have to check if these functions are linearly independent. Let us consider a linear combination :

$$a_0h + a_1h^2 + a_2h\beta(h) + a_3h^2\beta(h) + a_4h^2\beta^2(h) = 0, \quad \forall h \in (0, \infty).$$

This condition implies :

$$a_0 + a_1h + a_2\beta(h) + a_3h\beta(h) + a_4h\beta^2(h) = 0, \quad \forall h \in (0, \infty).$$

By setting  $h = 0$ , we get  $a_0 + a_2 = 0$  and the condition becomes :

$$a_1h + a_2(\beta(h) - 1) + a_3h\beta(h) + a_4h\beta^2(h) = 0, \quad \forall h \in (0, \infty),$$

corresponding to the linear dependence between  $h$ ,  $\beta(h) - 1$ ,  $h\beta(h)$ ,  $h\beta^2(h)$  values.

Let us denote :  $\tilde{\beta}(h) = \beta(h) - 1$ . It is equivalent to consider the linear dependence of functions  $h$ ,  $\tilde{\beta}(h)$ ,  $h\tilde{\beta}(h)$ ,  $h\tilde{\beta}^2(h)$ . This dependence arises if :

–  $\tilde{\beta}(h) = a_1 h$ , where  $a_1$  is nonnegative to ensure that the slope baseline is increasing.

– They can also exist  $a_1, a_2$  such that :

$$h\tilde{\beta}(h) + a_2\tilde{\beta}(h) + a_1h = 0 \Leftrightarrow \tilde{\beta}(h) = -\frac{a_1h}{a_2+h} = -a_1 + \frac{a_1a_2}{a_2+h}$$

For  $\tilde{\beta}$  to be continuous on  $(0, \infty)$ , we need  $a_2 < 0$ . For  $\tilde{\beta}$  to be positive for large value of  $h$  we need  $a_1 > 0$ . Finally we have  $\frac{d\tilde{\beta}(h)}{dh} = -\frac{a_1a_2}{(a_2+h)^2} < 0$ . We deduce that this situation cannot arise for continuous increasing function  $\tilde{\beta}$  with  $\tilde{\beta}(0) = 0$ .

– The last possibility of linear dependence arises when there exist  $a_1, a_2, a_3$  such that :

$$\begin{aligned} h\tilde{\beta}^2(h) + a_3h\tilde{\beta}(h) + a_2\tilde{\beta}(h) + a_1h &= 0 \\ \Leftrightarrow h\tilde{\beta}^2(h) + (a_2 + a_3h)\tilde{\beta}(h) + a_1h &= 0 \end{aligned} \quad (\text{A.6})$$

**Lemma 1 :** A real solution to equation (A.6) exists for any  $h \in (0, \infty)$  if and only if :

i)  $a_1 \leq 0$ ,

or if

ii)  $a_1 > 0, a_2 > 0$  and  $a_3 \geq 2\sqrt{a_1}$ ,

or if

iii)  $a_1 > 0, a_2 = 0$ ,

or if

iii)  $a_1 > 0, a_2 < 0$  and  $a_3 = -2\sqrt{a_1}$ .

**Proof :** A solution exists if and only if

$$\Delta(h) = (a_2 + a_3h)^2 - 4a_1h^2 = (a_3^2 - 4a_1)h^2 + 2a_2a_3h + a_2^2 \geq 0, \quad \forall h \in (0, \infty).$$

In particular by considering the limiting value  $h = 0$  and  $h \rightarrow \infty$ , we see that  $a_3^2 - 4a_1 \geq 0$  and  $a_2^2 \geq 0$ .

i) If  $a_1 = 0$ , we get :  $\Delta(h) = (a_2 + a_3h)^2 \geq 0, \forall a_2, a_3$ .

ii) If  $a_3^2 - 4a_1 = 0$ , we get :  $\Delta(h) = a_2(2a_3h + a_2) \geq 0$ , and  $a_1 \geq 0$ .

$a_1 = 0$  implies  $a_3 = 0$  and  $\Delta(h) = a_2^2 \geq 0, \forall h$ . If  $a_1 > 0$  and  $a_3 = 2\sqrt{a_1} > 0$ ,  $\Delta(h)$  is positive for all  $h$  if and only if  $a_2 > 0$ . Conversely, if  $a_1 > 0$  and  $a_3 = -2\sqrt{a_1} < 0$ ,  $\Delta(h)$  is positive if and only if  $a_2 < 0$ . Finally, if  $a_1 > 0$  and  $a_2 = 0$ , we get  $\Delta(h) = 0, \forall h$ .

iii) If  $a_3^2 - 4a_1 > 0$ ,  $\Delta(h) = 0$  is a quadratic equation, whose discriminant is equal to :

$$\Delta' = (a_2a_3)^2 - a_2^2(a_3^2 - 4a_1) = 4a_1a_2^2.$$

If  $a_1 \leq 0$  the discriminant is negative or null and  $\Delta(h) \geq 0, \forall h \in (-\infty, \infty)$ .

If  $a_1 > 0$ , the equation  $\Delta(h) = 0$  admits the real roots :

$$\frac{-a_2a_3 \pm 2|a_2|\sqrt{a_1}}{a_3^2 - 4a_1}.$$

$\Delta(h)$  is nonnegative for any nonnegative  $h$ , if and only if the maximal real root is nonpositive. Then the condition is :  $-a_2a_3 + 2|a_2|\sqrt{a_1} \leq 0$ .

This inequality can be rewritten according to the sign of  $a_2$ . If  $a_2 > 0$ , we get  $2\sqrt{a_1} < a_3$ , and if  $a_2 < 0$ , we get  $a_3 < -2\sqrt{a_1}$ , which is not compatible with the inequality  $a_3^2 - 4a_1 > 0$ . Finally, if  $a_2 = 0$  the maximal real root is null for all  $a_3$ , and thus  $\Delta(h) \geq 0 \forall h \geq 0$ .

QED

Then, the solution is necessarily the root :

$$\tilde{\beta}(h) = \frac{-(a_2 + a_3h) + \sqrt{(a_2 + a_3h)^2 - 4a_1h^2}}{2h}, \quad (\text{A.7})$$

due to the restriction  $\tilde{\beta}(0) = 0$ .

Let us now compute the derivative of this function. We get :

$$\frac{d\tilde{\beta}(h)}{dh} = \frac{a_2\Delta^{-1/2}(h)}{2h^2} \left[ \Delta^{1/2}(h) - (a_2 + a_3h) \right].$$

This derivative is positive if and only if, either  $a_2 > 0$ ,  $a_1 < 0$ , or  $a_2 < 0$ ,  $a_1 > 0$ . By combining these restrictions with the restrictions of Lemma 1, we get the next Lemma.

**Lemma 2 :** A solution to equation (A.6) exists and is increasing for  $h \in (0, \infty)$ , if and only if  $a_1 < 0$  and  $a_2 > 0$ . This solution is given by :

$$\tilde{\beta}(h) = \frac{\sqrt{(a_2 + a_3h)^2 - 4a_1h^2} - (a_2 + a_3h)}{2h}.$$

### Appendix 3.1

#### Proof of Lemma 1

Let us consider the compound variable :  $Y = \sum_{k=1}^Z X_k$ , where the variables  $Z, X_k, k = 1, \dots, K$  are independent,  $Z$  following a Poisson distribution  $\mathcal{P}(\lambda)$  and  $X_k$  a drifted Pareto distribution with  $\mathbb{P}[X > x] = \frac{\delta^\alpha}{x^\alpha}$ , if  $x > \delta, = 1$ , otherwise.

The Laplace transform of  $Y$  is given by :

$$\begin{aligned}
 & \mathbb{E}[\exp(-uY)] \\
 &= \mathbb{E}\mathbb{E}[\exp(-uY)|Z] \\
 &= \mathbb{E}[\exp(-uX)]^Z \\
 &= \exp\{-\lambda[1 - \mathbb{E}(\exp -uX)]\} \\
 &= \exp\{\lambda \int_{\delta}^{\infty} \frac{\alpha \delta^\alpha}{x^{\alpha+1}} [\exp(-ux) - 1] dx\}.
 \end{aligned}$$

When  $\lambda = c\delta^{-\alpha}$ , the expression above becomes :

$$\mathbb{E}[\exp(-uY)] = \exp\{c \int_{\delta}^{\infty} \frac{\alpha}{x^{\alpha+1}} [\exp(-ux) - 1] dx\},$$

which tends to :

$$\begin{aligned}
 I &= \exp\{c \int_0^{\infty} \frac{\alpha}{x^{\alpha+1}} (\exp(-ux) - 1) dx\}, \\
 &= \exp\{cu^\alpha \int_0^{\infty} \frac{\alpha}{x^{\alpha+1}} [\exp(-x) - 1] dx\},
 \end{aligned}$$

when  $\delta \rightarrow 0$  (Note that the integral exists for  $0 < \alpha < 1$ ).

By integrating by part, we get :

$$\begin{aligned} & \int_0^{\infty} \frac{\alpha}{x^{\alpha+1}} [\exp(-x) - 1] dx \\ &= (-x^{-\alpha} [\exp(-x) - 1])_0^{\infty} - \int_0^{\infty} x^{-\alpha} \exp(-x) dx. \\ &= -\Gamma(1 - \alpha). \end{aligned}$$

We deduce that :

$$I = \exp[-cu^{\alpha}\Gamma(1 - \alpha)],$$

and the result follows.



## Appendix 3.2

### Proof of Proposition 2.

i) **Expression of  $d(t, h)$**

We have :

$$d(t, h) = -\frac{1}{h} \log \mathbb{E}_t^{\mathbb{Q}} \{ \exp[-(h-1)\sigma_{t+1}\varepsilon_{t+1} + \dots + \sigma_{t+h-1}\varepsilon_{t+h-1}] \},$$

with  $\sigma_{t+1} = \exp[-\gamma(t+1)]$ , and  $\varepsilon_t \sim IIS(\alpha, c)$ . Therefore :

$$\begin{aligned} d(t, h) &= \frac{1}{h} \sum_{k=1}^{h-1} [(h-k)^\alpha \sigma_{t+k}^\alpha] \\ &= \frac{1}{h} \sum_{k=1}^{h-1} [c^\alpha (h-k)^\alpha \exp[-\alpha\gamma(t+k)]] \end{aligned}$$

ii) **Asymptotic behavior of  $d(t, h)$ .**

Let us decompose the sum above into the  $[\log h]$  first terms and  $h-1-[\log h]$  last ones, where  $[\log h]$  denotes the integer part of  $\log h$ . We have :

$$\begin{aligned} d(t, h) &= \frac{1}{h} \sum_{k=1}^{[\log h]} \{c^\alpha (h-k)^\alpha \exp[-\alpha\gamma(t+k)]\} + \frac{1}{h} \sum_{k=[\log h]+1}^{h-1} \{c^\alpha (h-k)^\alpha \exp[-\alpha\gamma(t+k)]\} \\ &\leq c^\alpha (h-1)^\alpha \exp[-\alpha\gamma(t+1)] \frac{[\log h]}{h} + \frac{(h-[\log h])^{\alpha+1}}{h} c^\alpha \exp[-\alpha\gamma(t+[\log h])]. \end{aligned}$$

The first term of the upper bound is of order  $h^{\alpha-1} \log h$  and tends to zero

when  $h$  tends to infinity, since  $\alpha < 1$ . The second term is of order  $h^{\alpha(1-\gamma)}$  and tends to zero when  $h$  tends to infinity, whenever  $\gamma > 1$ .

Therefore  $\lim_{n \rightarrow \infty} d(t, h) = d(t, \infty) = 0$ .

## Appendix 3.3

### The Continuous Time Analogue

#### i) Analogue of the SAR(1)

Let us consider a given time-to-maturity  $\tau$ , say, and a time unit  $\delta$  such that  $\tau = \delta h$ . For this time unit eqn. (2.8) becomes :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-uy_{t+\tau})] = \exp[-u\rho_{\delta}^h y_t - c_{\delta} \frac{1 - \rho_{\delta}^{\alpha h}}{1 - \rho_{\delta}^{\alpha}} u^{\alpha}],$$

where parameters  $\rho$  and  $c$  are indexed by  $\delta$ . Let us now choose these parameters such that :

$$\rho_{\delta} = (\rho^*)^{\delta}, \frac{c_{\delta}}{1 - \rho_{\delta}^{\alpha}} = \gamma^*, \text{ say}$$

Then, we get :

$$\mathbb{E}_t^{\mathbb{Q}}[\exp(-uy_{t+\tau})] = \exp[-u(\rho^*)^{\tau} y_t - \gamma^*(1 - (\rho^*)^{\alpha\tau})u^{\alpha}],$$

which provides the continuous time counterpart of the SAR(1) process.

#### ii) Analogue of the SARTSM

Similarly, we get :

$$\begin{aligned}
r(t, \tau) &= \frac{1}{h} \frac{1 - \rho_\delta^h}{1 - \rho_\delta} y_t - \frac{c\delta}{h} \sum_{k=1}^{h-1} \frac{(1 - \rho_\delta^k)^\alpha}{(1 - \rho_\delta)^\alpha} \\
&= \frac{1}{\delta h} \frac{1 - \rho^{*\delta h}}{\frac{1}{\delta}[1 - \rho^{*\delta}]} y_t - \gamma^* \frac{1}{\delta h} \delta \sum_{k=1}^{h-1} (1 - \rho^{*k\delta})^\alpha.
\end{aligned}$$

This quantity tends to :

$$r(t, \tau) = \frac{1 - \rho^{*\tau}}{-\tau \log \rho^*} - \frac{\gamma^*}{\tau} \int_0^\tau (1 - \rho^{*x})^\alpha dx,$$

when  $\delta \rightarrow 0, h \rightarrow \infty$  such that  $\delta h = \tau$ .

We observe that the SARTSM is not equal to the time discretized version of its continuous time analogue. Thus, the SARTSM gives an example of discrete time affine term structure model, which is not the exact time discretization of its continuous time analogue.

## Appendix 3.4

### Nelson Siegel Term Structure Model

The original model introduced in Nelson, Siegel (1987) provides a simple functional form in order to fit the yield curve :

$$r(h) = \beta_0 + \beta_1 \left( \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} \right) + \beta_2 \left( \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} - \exp[-\lambda_1 h] \right), \quad (\text{A.8})$$

which has been extended by Svensson (1995) with the addition of a fourth term, which improves the ability of the model to fit the rates of the long-term bonds :

$$\begin{aligned} r(h) = & \beta_0 + \beta_1 \left( \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} \right) + \beta_2 \left( \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} - \exp[-\lambda_1 h] \right) \\ & + \beta_3 \left( \frac{1 - \exp[-\lambda_2 h]}{\lambda_2 h} - \exp[-\lambda_2 h] \right). \end{aligned} \quad (\text{A.9})$$

The NSS model is originally presented in a cross-sectional framework, without time index, but is generally estimated daily (or monthly) as if its parameters  $\beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}, \lambda_{1t}, \lambda_{2t}$  were time dependent. With this practice, the NSS model becomes a complicated factor model, which is not affine in  $\lambda_{1t}, \lambda_{2t}$ . It has been shown in [Filipovic (1999), Theorem 4.1] that a necessary condition for the dynamic Nelson-Siegel (NS) model (A.8), and *a fortiori* for the NSS model (A.9), to be arbitrage free is that  $\lambda_{1t}$  (and  $\lambda_{2t}$ ) is (are) time-independent. Then, for constant  $\lambda_{1t}$  and  $\lambda_{2t}$ , the NSS model becomes an affine term structure model with factors  $\beta_{0t}, \beta_{1t}, \beta_{2t}, \beta_{3t}$  usually called level, slope, curvature, and curvature factors (respectively), because of the shape of the factors loadings  $a'(h)$ <sup>74</sup>. However it is still not compatible with no-arbitrage [Bjork, Christensen (1999), Filipovic (1999)], due to a misleading adjustment for risk in the standard NS model.

---

<sup>74</sup> $a'(h) = \left( 1 \quad \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} \quad \frac{1 - \exp[-\lambda_1 h]}{\lambda_1 h} - \exp[-\lambda_1 h] \quad \frac{1 - \exp[-\lambda_2 h]}{\lambda_2 h} - \exp[-\lambda_2 h] \right)$ .

## Appendix 3.5

### Bonds pricing formula

The price at  $t$  of the zero-coupon bond of term  $h$  is given by :

$$B(t, h) = \mathbb{E}_t^{\mathbb{Q}}[\exp\{-r(t, 1)\}B(t + 1, h - 1)],$$

where  $\mathbb{E}_t^{\mathbb{Q}}$  denotes the risk-neutral conditional distribution given  $I_t$ . The associated geometric rate at time-to-maturity  $h$  is :

$$r(t, h) = -\frac{1}{h} \log B(t, h).$$

Besides, we assume that the rates are affine functions of the factors ( $y_t$ ) :

$$r(t, h) = a'(h)y_t + d(h), \text{ with } a(1) = \alpha \text{ and } d(1) = \beta$$

and that the factors satisfies a vector autoregressive model :

$$y_t = \Phi y_{t-1} + \eta_t,$$

where the shocks  $\eta_t$  are IID, either multivariate Gaussian or a combination of Gaussian and stable shocks.

Thus:

$$\begin{aligned} r(t, h) &= -\frac{1}{h} \log \mathbb{E}_t^{\mathbb{Q}}[\exp\{-\alpha' y_t - \beta\} \exp\{-(h-1)a'(h-1)y_{t+1} - (h-1)d(h-1)\}] \\ &= \frac{1}{h} [(\alpha' + (h-1)a'(h-1)\Phi) y_t + \beta + (h-1)d(h-1)] \\ &\quad - \frac{1}{h} \log \mathbb{E}_t^{\mathbb{Q}}[\exp\{-(h-1)a'(h-1)\eta_{t+1}\}] \end{aligned}$$

which leads to :

$$\begin{aligned} ha(h) &= \alpha' + (h-1)a'(h-1)\Phi \\ &= \alpha' + \alpha' \sum_{k=1}^{h-1} \Phi^k \\ hd(h) &= \beta + (h-1)d(h-1) - \psi((h-1)a'(h-1)) \end{aligned}$$

where  $\psi(u)$  stands for the (conditional and marginal) log-Laplace transform of  $\eta$  under the risk-neutral distribution, i.e  $\psi(u) = \log \mathbb{E}_t^{\mathbb{Q}}[\exp -u\eta]$ .

### 3-factor Triangular Term Structure Model

Let us first consider a 3-factor term structure model, with triangular autoregressive matrix  $\Phi$  :

$$\Phi = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \text{ and } \eta_t \sim IIN(0, \Sigma).$$

Thus :

$$\begin{aligned} ha'(h) &= \alpha' + \alpha'(Id - \Phi)^{-1}(Id - \Phi^h) \\ hd(h) &= \beta + (h-1)d(h-1) - \frac{1}{2}(h-1)^2 a'(h-1) \Sigma a(h-1), \end{aligned}$$

and :

$$\begin{aligned} (Id - \Phi)^{-1}(Id - \Phi^h) &= \begin{pmatrix} \frac{1}{1-\lambda} & \frac{1}{(1-\lambda)^2} & \frac{1}{(1-\lambda)^3} \\ 0 & \frac{1}{1-\lambda} & \frac{1}{(1-\lambda)^2} \\ 0 & 0 & \frac{1}{1-\lambda} \end{pmatrix} \begin{pmatrix} \lambda^h & h\lambda^{h-1} & h(h-1)\lambda^{h-1}/2 \\ 0 & \lambda^h & h\lambda^{h-1} \\ 0 & 0 & \lambda^h \end{pmatrix} \\ &= \begin{pmatrix} \frac{1-\lambda^h}{1-\lambda} & \frac{1-\lambda^h}{(1-\lambda)^2} + \frac{h\lambda^{h-1}}{1-\lambda} & \frac{h(h-1)\lambda^{h-1}}{2(1-\lambda)} + \frac{h\lambda^{h-1}}{(1-\lambda)^2} + \frac{1-\lambda^h}{(1-\lambda)^3} \\ 0 & \frac{1-\lambda^h}{1-\lambda} & \frac{1-\lambda^h}{(1-\lambda)^2} + \frac{h\lambda^{h-1}}{1-\lambda} \\ 0 & 0 & \frac{1-\lambda^h}{1-\lambda} \end{pmatrix}. \end{aligned}$$

### Combination of two 2-factor Triangular Term Structure Models with a Gaussian random walk

Let us now derive the pricing formula for the fully Gaussian arbitrage-free Nelson Siegel model. This specification combines two 2-factor Triangular Term Structure Models with a Gaussian random walk, and as such, constitutes the discrete time counterpart of the so-called Arbitrage-Free Generalized Nelson-Siegel model (AFGNS) put forward by Christensen, Diebold, and Rudebush (2009). With respect to the previous model, the autoregressive matrix is slightly changed :

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \text{ and } \eta_t \sim IIN(0, \Sigma).$$

in that case :

$$\sum_{k=0}^{h-1} \Phi^k = \begin{pmatrix} h-1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1-\lambda_1^h}{1-\lambda_1} & \frac{1-\lambda_1^h}{(1-\lambda_1)^2} + \frac{h\lambda_1^{h-1}}{1-\lambda_1} & 0 & 0 \\ 0 & 0 & \frac{1-\lambda_1^h}{1-\lambda_1} & 0 & 0 \\ 0 & 0 & 0 & \frac{1-\lambda_2^h}{1-\lambda_2} & \frac{1-\lambda_2^h}{(1-\lambda_2)^2} + \frac{h\lambda_2^{h-1}}{1-\lambda_2} \\ 0 & 0 & 0 & 0 & \frac{1-\lambda_2^h}{1-\lambda_2} \end{pmatrix},$$

which gives the pricing formulas in 4.2 i).

### Combination of two 2-factor Triangular Term Structure Models with an almost flat Stable Term Structure Model: TSM3

Finally, we combine two 2-factor Triangular Term Structure Models with the almost flat Stable Term Structure presented in 3.2. In this specification, the autoregressive matrix, and thus the factor loadings, are the same as in the



AFGNS model. However, the residuals distribution is modified :

$$\eta_t = \begin{pmatrix} \exp(-\gamma t)\eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \eta_{4,t} \\ \eta_{5,t} \end{pmatrix} \quad \text{with} \quad \eta_{1,t} \sim IIS(\alpha, c) \quad \text{and} \quad \begin{pmatrix} \eta_{2,t} \\ \eta_{3,t} \\ \eta_{4,t} \\ \eta_{5,t} \end{pmatrix} \sim IIN(0, \Sigma).$$

Thus :

$$\begin{aligned} \psi(u) &= \log \mathbb{E}_t^{\mathbb{Q}}[\exp -u\eta_{t+1}] \\ &= \log \mathbb{E}_t^{\mathbb{Q}}[\exp - \begin{pmatrix} u_1 & u_2 & u_3 & u_4 & u_5 \end{pmatrix} \begin{pmatrix} \eta_{1,t+1} \\ \eta_{2,t+1} \\ \eta_{3,t+1} \\ \eta_{4,t+1} \\ \eta_{5,t+1} \end{pmatrix}] \\ &= \log \mathbb{E}_t^{\mathbb{Q}}[\exp -u_1\eta_{1,t+1}] \\ &+ \log \mathbb{E}_t^{\mathbb{Q}}[\exp - \begin{pmatrix} u_2 & u_3 & u_4 & u_5 \end{pmatrix} \begin{pmatrix} \eta_{2,t+1} \\ \eta_{3,t+1} \\ \eta_{4,t+1} \\ \eta_{5,t+1} \end{pmatrix}] \\ &= \psi(u_1) + \psi(\begin{pmatrix} u_2 & u_3 & u_4 & u_5 \end{pmatrix}) \\ &= -cu_1^\alpha + \frac{1}{2}u_{25}'\Sigma u_{25} \end{aligned}$$

and the pricing formulas in 4.2 ii) follows.

## Appendix 3.6

### Tables

$\rho$	0	0	0	0.5	0.5	0.5	0.9	0.9	0.9
$\alpha$	0.5	0.8	0.9	0.5	0.8	0.9	0.5	0.8	0.9
$\hat{\rho}_{10}$	0.006	0.153	0.246	0.506	0.579	0.693	0.901	0.933	1
$\hat{\rho}_{50}$	0.000	0.030	0.093	0.500	0.527	0.583	0.900	0.915	0.941
$\hat{\rho}_{100}$	0.000	0.030	0.052	0.500	0.516	0.549	0.900	0.905	0.928
$\hat{\rho}_{200}$	0.000	0.003	0.052	0.500	0.503	0.549	0.900	0.903	0.927

Table 3.5: **Comparison between  $\rho$  and  $\hat{\rho}_T$ .**

Values of  $\hat{\rho}_T = \min_{1 \leq t \leq t+h \leq T} (y_{t+h}/y_t)^{1/h}$  for several values of parameter  $\alpha$ ,  $\rho$ ,  $c = 1$ , and different lengths of the trajectory  $(y_t)_{t=1, \dots, T}$ .

$\mu$	$\Phi$					$\Sigma$				
0	1	0	0	0	0	$0.35 \times 10^{-4}$	0	0	0	0
0	0	0.992	1	0	0	0	$0.28 \times 10^{-4}$	0	0	0
0	0	0	0.992	0	0	0	0	$0.04 \times 10^{-4}$	0	0
0	0	0	0	0.911	1	0	0	0	$0.34 \times 10^{-4}$	0
0	0	0	0	0	0.911	0	0	0	0	$0.04 \times 10^{-4}$

Table 3.6: Estimates of the parameters governing the factors dynamics under the pricing measure - fully Gaussian model.

$\mu_{\mathbb{P}}$	$\Phi_{\mathbb{P}}$				
$-0.219 \times 10^{-4}$	0.976	$-0.026 \times 10^{-11}$	$-0.418 \times 10^{-11}$	$-0.015 \times 10^{-11}$	$-0.158 \times 10^{-11}$
$-0.025 \times 10^{-4}$	$-0.793 \times 10^{-13}$	0.9997	$0.502 \times 10^{-10}$	$0.016 \times 10^{-10}$	$-0.019 \times 10^{-10}$
$-0.017 \times 10^{-4}$	$-0.193 \times 10^{-10}$	$-0.682 \times 10^{-10}$	0.979	$-0.486 \times 10^{-11}$	$0.069 \times 10^{-11}$
$-0.016 \times 10^{-4}$	$-0.005 \times 10^{-11}$	$0.084 \times 10^{-11}$	$0.675 \times 10^{-11}$	0.982	$-0.184 \times 10^{-11}$
$-0.028 \times 10^{-4}$	$-0.020 \times 10^{-11}$	$-0.174 \times 10^{-11}$	$-0.443 \times 10^{-11}$	$-0.056 \times 10^{-11}$	0.9557

$\Sigma$				
$0.35 \times 10^{-4}$	0	0	0	0
0	$0.28 \times 10^{-4}$	0	0	0
0	0	$0.04 \times 10^{-4}$	0	0
0	0	0	$0.34 \times 10^{-4}$	0
0	0	0	0	$0.04 \times 10^{-4}$

Table 3.7: Estimates of the parameters governing the factors dynamics under the physical measure - fully Gaussian model.

$\mu$	$\Phi$				
0	0	0	0	0	0
0	0	0.9994	1	0	0
0	0	0	0.9994	0	0
0	0	0	0	0.9364	1
0	0	0	0	0	0.9364

$\alpha$	$\gamma$	$c$
0.099	12.56	$4.46 \times 10^{-26}$

$\Sigma$			
$0.214 \times 10^{-4}$	0	0	0
0	$0.002 \times 10^{-4}$	0	0
0	0	$0.27 \times 10^{-4}$	0
0	0	0	$0.038 \times 10^{-4}$

Table 3.8: Estimates of the parameters governing the factors dynamics under the pricing measure - stable factor model.

$\mu_{\mathbb{P}}$	$\Phi_{\mathbb{P}}$				
$0.006 \times 10^{-4}$	1	0	0	0	0
$-0.141 \times 10^{-4}$	-0.0002	0.9808	-0.0001	-0.0015	-0.0001
$-0.001 \times 10^{-4}$	0.0002	$-0.4927 \times 10^{-5}$	1.0142	$0.4558 \times 10^{-5}$	0.0001
$-0.012 \times 10^{-4}$	0.0001	$-0.4442 \times 10^{-5}$	0.0002	1.0033	0.0001
$0.023 \times 10^{-4}$	0.0001	-0.0001	0.0001	0.0029	0.912

$\gamma_{\mathbb{P}}$	$c_{\mathbb{P}}$
9.01	$6.118 \times 10^{-31}$

$\Sigma$				
$0.214 \times 10^{-4}$	0	0	0	0
0	$0.002 \times 10^{-4}$	0	0	0
0	0	$0.27 \times 10^{-4}$	0	0
0	0	0	$0.038 \times 10^{-4}$	0

Table 3.9: Estimates of the parameters governing the factors dynamics under the physical measure - stable factor model.



## List of Tables

---

3.1	Mean of excess gains on 2001-2007. All numbers must be divided by 1000. . . . .	205
3.2	Mean of excess gains on 2007-2011. All numbers must be divided by 1000. . . . .	205
3.3	Covariance matrix of excess gains on 2001-2007. All numbers must be divided by 1000. . . . .	205
3.4	Covariance matrix of excess gains on 2007-2011. All numbers must be divided by 1000. . . . .	205
3.5	<b>Comparison between <math>\rho</math> and <math>\hat{\rho}_T</math>.</b> . . . . .	231
3.6	Estimates of the parameters governing the factors dynamics under the pricing measure - fully Gaussian model. . . . .	231
3.7	Estimates of the parameters governing the factors dynamics under the physical measure - fully Gaussian model. . . . .	232
3.8	Estimates of the parameters governing the factors dynamics under the pricing measure - stable factor model. . . . .	232
3.9	Estimates of the parameters governing the factors dynamics under the physical measure - stable factor model. . . . .	233



## List of Figures

---

1	Structures par terme de base, stochastique $[\alpha(h)]$ et déterministe $[\beta(h)]$ , pour un modèle affine Gaussien à un facteur, où $h$ (en abscisse) désigne la maturité résiduelle des taux. . . . .	37
2	Structure par terme de la volatilité des taux d'intérêt des obligations du Trésor des Etats-Unis. La maturité des taux, en années, se trouve en abscisse, la volatilité des taux (en pourcentage) en ordonnées. . . . .	38
3	Structures par terme de base $\alpha(h)$ pour la formule (0.0.42) de Nelson, Siegel (1987), où $h$ (en abscisse) désigne la maturité résiduelle des taux. . . . .	39
4	Simulation d'un processus autorégressif gaussien d'ordre 1 sur 100 périodes. Le graphique du haut présente la trajectoire du processus simulé, tandis que celui du bas décrit les chocs affectant le processus à chaque période $t$ (en abscisse). . . . .	41
5	Densité de la distribution de la valeur future du facteur $X_h$ , conditionnellement à la connaissance de sa valeur initiale $X_0$ . Les courbes pleine, hachurée, et pointillée désignent respectivement les densités conditionnelles à 1, 5, et 50 périodes. . . . .	42
1	Stochastic $[\alpha(h)]$ and deterministic $[\beta(h)]$ baseline term structures for a 1-factor affine Gaussian term structure model, where $h$ indicates the interest rates' residual maturity. . . . .	63
2	Term structure of volatility of Treasury bonds' yield, for different periods. The x-axis indicates the residual maturity of interest rates in years. The y-axis stands for the yields' volatility, in percent. . . . .	65
3	Baseline term structure $\alpha(h)$ in Nelson-Siegel (1987) term structure model (0.0.42). $h$ stands for the rates' residual maturities. . . . .	66
4	Simulation of an autoregressive Gaussian process of order 1 over 100 periods. The top graph present the factor's trajectory, while the shock's realizations are plotted in the bottom panel, at each date $t$ (in the x-axis). . . . .	68



1.1	Shift of a distribution. The solid line stands for the baseline univariate distribution. . . . .	101
1.2	Conditional distributions of the shocked variable $h(Y_0, \varepsilon; \delta)$ with respect to $Y_0$ for different $a$ and $b$ functions in (1.5.2). The solid line stands for $h(Y_0, \varepsilon; 0) = Y_0$ , and $\Phi^{-1}(\varepsilon)$ has a standard Gaussian distribution. . . . .	103
1.3	Unconditional distributions of the shocked variable $h(Y_0, \varepsilon; \delta)$ for different $a$ and $b$ functions in (1.5.2), where $Y_0$ and $\Phi^{-1}(\varepsilon)$ have standard independent Gaussian distributions. . . . .	103
1.4	Left tail contaminated distribution. The solid line stands for the baseline distribution, the dashed line for the contaminated distribution. . . . .	104
1.5	Montly excess gains on 10Y European sovereign bonds. . . . .	112
1.6	Historical dynamics and histograms of the first systematic factor on periods 2001-2007, and 2007-2011. . . . .	113
1.7	Empirical c.d.f. and p.d.f. of 1000 simulated contaminated factor values $X_\delta$ in terms of variable for various $\delta$ . The solid line stands for the c.d.f. of the baseline factor $X_0$ , while dashed, dotted, and dash-dotted lines represent the contaminated empirical distribution for $\delta = 0.1, 0.5, 1$ . . . . .	115
1.8	Contaminated mean-variance allocation as a function of $\delta$ . . . . .	117
1.9	Impulse response of the Sharpe ratio of crystallized and Mean-Variance portfolios (contamination in variable). . . . .	118
1.10	Impulse response of the Mean and Variance of crystallized and Mean-Variance portfolios (contamination in variable). . . . .	118
1.11	Impulse response of the VaR of crystallized and Mean-Variance portfolios (contamination in variable). . . . .	118
1.12	Impulse response of the Expected-shortfall of crystallized and Mean-Variance portfolios (contamination in variable). . . . .	118
1.13	Impulse response of the conditional distribution of the crystallized portfolio $Y_\delta Y_0$ (contamination in variable). . . . .	120
1.14	Impulse response of the conditional distribution of the optimally adjusted portfolio $Y_\delta Y_0$ (contamination in variable). . . . .	120
3.1	Trajectories of the SAR(1) process. Simulated trajectories of the SAR(1) process $(y_t)$ with length $t = 1, \dots, 200$ for unitary scale parameter $c = 1$ , different values of parameters $\alpha$ and $\rho$ and a same history of shocks. . . . .	154
3.2	Trajectories of historical means. Simulated trajectories of the historical mean $\bar{y}_t = \frac{1}{t}(y_1 + \dots + y_t), t = 1, \dots, 200$ of the SAR(1) process $y_t$ for unitary scale parameter $c = 1$ , different values of the parameters $\alpha$ and $\rho$ . Four different trajectories are plotted for each graph. . . . .	155

3.3	Histogram of the historical mean [SAR(1)]. Histogram of the historical mean $\bar{y}_{200} - y_1 \frac{1-\rho^{200}}{200(1-\rho)}$ based on 200 different trajectories of the SAR(1) process $y_t$ , for $c = 1$ , and different values of parameters $\alpha$ and $\rho$ . . . . .	156
3.4	Histogram of the historical mean [ARG(1)]. Histogram of the centered historical means $\bar{y}_{200}^* - E(y^*)$ based on 200 different trajectories of the ARG(1) process $y_t^*$ , for different values of parameters $d$ and $b$ , with $c = 1$ . . . . .	157
3.5	Pattern of the SARTSM. Pattern of SARTSM $r(t, h)$ with $h = 1, \dots, 100$ , for different values of $r(t, \infty)$ , $\rho$ and $\alpha$ , when the short-term rate $r(t, 1)$ is set at 3%. The solid, dashed and dotted lines stand for $\alpha = 0.5, 0.8, 0.9$ , respectively. . . . .	160
3.6	Times Series of the unsmoothed Fama-Bliss zero-coupon US T-bond yields from January 1987 to December 2002. The yields shown have 1 year (dash-dot line), 5 years (dashed line), 10 years (dotted line), 30 years (solid line). . . . .	168
3.7	Time series of the estimated pricing errors for the model with stable level factor (solid line), and the Fully Gaussian model (dashed line), from January 1987 to December 2002. The models are estimated on the level of yields. The reported values are in basis points. . . . .	170
3.8	Time series of the estimated Root Mean Squared Errors for the model with stable level factor (solid line), and the Fully Gaussian model (dashed line), from January 1987 to December 2002. The models are estimated on the level of yields. The reported values are in basis points. . . . .	170
3.9	Time average of the pricing errors and Root Mean Squared Errors on the level of yields, for all maturities (from 3 months to 30 years) for the model with stable level factor (solid line), and fully Gaussian model (dashed line). The mean of pricing errors and Root Mean Squared Errors obtained by Christensen, Diebold, Rudebush (2009) are reported as black circles. The x-axis stands for the yield maturity in years. The reported values are in basis points. . . . .	170
3.10	Fitted yield curve for 4 specific dates. The observed yields are indicated as black circles, and the fitted yield curve are from the model with stable factor (solid line) and the fully Gaussian model (dashed line). The reported values are in percentage points. . . . .	171
3.11	Normalized Factor loadings $a(h)$ for the second to fifth factor. For each factor, the factor loading function $a(h)$ is normalized by its average value over the whole set of maturities. The x-axis stands for the yield's maturity in years. The reported values are in percentage points. The solid, and dashed lines, indicate the factor loading functions for the model with stable level factor, and for the Fully Gaussian model, respectively. . . . .	171

3.12	Time series of the estimated factors from January 1987 to December 2002 for the fully Gaussian model (left plot) and for the model with stable level factor (right plot). All factors are scaled by the average value over all maturities of the corresponding factor loading function. The reported values are in percentage points. For each model, the solid line with black cross, the solid line, the dashed line, the dotted line, and the dash-dot line, indicate the time series of the first to fifth factors, respectively. . . . .	172
3.13	Time series of the estimated factors from January 1987 to December 2002. All factors are scaled by the average value over all maturities of the corresponding factor loading function. The reported values are in percentage points. The solid, and dashed lines, indicate the factor loading functions for the model with stable level factor, and for the Fully Gaussian model, respectively. The graph on the south-east corner presents the deterministic baseline term structure $d_2(h, \theta)$ and $d(h, \theta)$ for the model with stable level factor (solid line) and the fully Gaussian model (dashed line), respectively. . . . .	172
3.14	Time series of the estimated stable factor's variation from February 1987 to December 2002. The variations are scaled by the average value over all maturities of the level factor loading function. The reported values are in percentage points. . . . .	173
3.15	Eigenvalues and factor loadings from the principal components analysis of the 2001-2011 excess gains covariance matrix	206
3.16	Empirical c.d.f and p.d.f. of 1000 simulated factor $X_\delta$ , contaminated in terms of distribution for various $\delta$ . The solid line stands for the c.d.f. of the baseline factor $X_0$ , while dashed, dotted, and dash-dotted lines represent the contaminated empirical distribution for $\delta = 0.1, 0.5, 1$ . . . . .	207
3.17	Contaminated mean-variance allocation as a function of $\delta$ . . . . .	208
3.18	Impulse response of the Sharpe ratio of crystallized and Mean-Variance portfolios (contamination in distribution). . . . .	208
3.19	Impulse response of the Mean and Variance of crystallized and Mean-Variance portfolios (contamination in distribution). . . . .	208
3.20	Impulse response of the VaR of crystallized and Mean-Variance portfolios (contamination in distribution). . . . .	209
3.21	Impulse response of the Expected-shortfall of crystallized and Mean-Variance portfolios (contamination in distribution). . . . .	209