ABSTRACT

Title of Dissertation: ASSET PRICING AND PORTFOLIO CHOICE WITH HEAVY-TAIL RETURNS DISTRIBUTIONS AND NONLINEAR EXPECTATIONS

Yoshihiro Shirai
Doctor of Philosophy, 2022

Dissertation Directed by: Professor Dilip B. Madan
Department of Business and Management

The fundamental works of Bachelier, Markovitz, Sharpe and Lintner, Black, Scholes and Merton, and others, which laid out the basis of the discipline that today we refer to as Mathematical Finance, rest in general on few main assumptions: returns are Gaussians, prices are unique/linear, markets are arbitrage-free, and investors are expected utility maximizers.

It has long been recognized, on the other hand, that none of these assumptions hold true in practice, so the traditional theoretical results of Mathematical Finance are only approximately true at best, and their applicability is limited. Typical examples are the volatility smile, the leptokurtic feature of returns, trade- and volume-dependence of prices, existence of infinitesimally small arbitrage opportunities and constant violations of the expected utility theorem axioms.

In this work, I continue the exploration, started a few decades ago, of the consequences of relaxing the assumptions of normality of returns, linearity/uniqueness of prices, and certainty equivalent based financial objectives.

Specifically, in Chapter I, I develop a pure jump model for pricing credit index options, that is based on the double gamma dynamics for the default intensity. In Chapter II, I apply several supervised and unsupervised learning techniques to provide additional evidence of investors' behaviors that contradicts Expected Utility Theory. In Chapter III, I show that spectral risk measures, a well known class of nonlinear expectation operators for pure jump semimartingales, admit an integral representation and, based on it, I define a new class of convex risk measures that are not sublinear.

ASSET PRICING AND PORTFOLIO CHOICE WITH HEAVY-TAIL RETURNS DISTRIBUTIONS AND NONLINEAR EXPECTATIONS

by

Yoshihiro Shirai

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Dilip B. Madan, Chair/Advisor
Professor Patrick M. Fitzpatrick,
Professor Maria K. Cameron
Professor Leonid Koralov
Professor Mark V. Loewenstein

© Copyright by
Yoshihiro Shirai
2022

Dedication

To Julie.

Voliamo insieme, spinti dalla brezza dei nostri sogni...

Acknowledgments

Professor Dilip B. Madan, my advisor, believed in, gave a sense to and inspired many of my research ideas. He has provided me with a multitude of the pieces of an infinite puzzle, some of which I was able to utilize in this work, while many open questions still stand. In addition to that, I came to realize in these years how Professor Madan sets an extraordinary benchmark as a scholar and adviser, who is genuinely dedicated to the improvement of the theoretical and practical aspects of his field, and at the same time looks over and generously shares his intuitions with his students.

I would like to thank Professor Patrick Fitzpatrick, Professor Radu Balan, Professor Pete Kyle, Professor Mark Loewenstein, Professor John Benedetto and Professor Maria Cameron. I enjoyed attending their classes and learned a lot from them. In fact, I only wish I had some more time to learn even more from them and to exploit the many learning and research opportunities at the University of Maryland.

I am indebted to Professor Leonid Koralov, and, again, Professors Fitzpatrick, Cameron and Loewenstein for agreeing to serve on my thesis committee and for sparing their invaluable time reviewing this manuscript, as well as to Professor Ali Hirsa, without whom I would not have even started my PhD.

I also would like to thank Professor Tim Leung, Professor Fabio Maccheroni and Professor Massimo Marinacci, who helped shaping my mind as a researcher in Mathematical Finance. Their teachings and approach to research proved repeatedly useful while working on my dissertation.

Yiran Zhang, Shanhawaz Khalid, Zhang Zhang, Ran Li and Sahil Chopra, my colleagues and friends in the Financial Mathematics seminar, have enriched my experience at the University of Maryland and provided useful advice anytime I asked, and they still do. A special mention is for Patty and Lisa, for their friendship, support and, again, their great advice that only a sharp mind and a life of experiences can give.

I owe my deepest thanks to my family - my wife Julie, for her love, patience and support as we walk through our journey together, and my mother and father, to whom I ultimately owe everything.

Table of Contents

Dedication ii

Acknowledgements iii

Table of Contents v

Chapter 0: Executive Summary 1
- 0.1 Valuation of Credit Index Derivatives . . . 1
- 0.2 Acceptable Bilateral Gamma Parameters . . . 3
- 0.3 Spectral Martingale Measures . . . 6

Chapter 1: Valuation of Credit Index Derivatives 10
- 1.1 Introduction . . . 10
- 1.2 Review and Assumptions . . . 13
  - 1.2.1 Hazard Rates and Doubly Stochastic Random Times . . . 13
  - 1.2.2 Basics of Forward CDS and CDX Contracts . . . 14
- 1.3 A Model for Stochastic Rate and Default Intensity . . . 20
  - 1.3.1 Characteristic Exponent, Zero Coupon Bond Prices and Valuation PIDE . . . 22
- 1.4 Numerical Results . . . 31
- 1.5 Comparison with Market Data . . . 37
- 1.6 Conclusions . . . 44
- 1.7 Appendix: The Finite Difference Scheme for the Valuation PIDE . . . 44

Chapter 2: Admissibility of Bilateral Gamma Parameters 47
- 2.1 Introduction . . . 47
- 2.2 Bilateral Gamma Returns . . . 52
  - 2.2.1 From Brownian Motion to Bilateral Gamma Process . . . 52
  - 2.2.2 Bilateral Gamma Returns under Expected Utility Theory . . . 55
- 2.3 The Acceptance Set . . . 59
  - 2.3.1 Learning the Boundaries . . . 59
  - 2.3.2 Implied Boundaries for Performance Measures . . . 63
  - 2.3.3 Uncertainty Quantification . . . 65
  - 2.3.4 Statistical Significance . . . 66
  - 2.3.5 A Simple Modification of a Lucas Tree Economy . . . 67
- 2.4 The Risks-Neutral Acceptance Set . . . 70
  - 2.4.1 Speed Uncertainty . . . 72

2.4.2 An Equation for the Boundaries of Acceptable Risk Neutral Parameters . . 73
2.4.3 Empirical Verifications 75
2.5 Conclusions 81
2.6 Appendix A: Unsupervised Manifold Learning 82
2.7 Appendix B: Assets Tickers 86

Chapter 3: Spectral Martingale Measures 87
3.1 Introduction 87
3.2 Assumptions and Preliminary Results 91
3.3 Spectral Martingale Measures 95
3.3.1 Properties of the set C^g 95
3.3.2 Definition and Existence 96
3.3.3 An Explicit Finite Difference Scheme 104
3.4 Monotone Claims 106
3.4.1 The Spectral Martingale Measures of Monotone Claims 106
3.4.2 Empirical Tests of Nonlinear Asset Pricing Equations 110
3.4.3 The BG2BG measure distortions 117
3.5 Portfolio Choice 120
3.5.1 Optimality Conditions for a Small Investor 120
3.5.2 Numerical Results 121
3.6 From Conic to Convex Finance 125
3.6.1 Dynamic Rebated Spectral Risk Measures 125
3.6.2 Optimality Conditions for a Large Investor 130
3.6.3 Myopic Risk-Sensitive Asset Management 132
3.6.4 Optimal Investment Amount 137
3.7 Conclusions 141
3.8 Appendix 141

Bibliography 148

Chapter 0: Executive Summary

0.1 Valuation of Credit Index Derivatives

This paper proposes a new valuation method for credit index swaptions (henceforth, CDXO), which are options to enter at a predetermined date a credit index swap. The current literature (see [15] and [4], as well as [82] and [31]) focuses on developing a Black-type formula for the purpose of retrieving the CDXO price from its quotation, which is expressed in terms of the underlying spread, and, particularly, on the issue of including the so called front end protection into the CDXO payoff.[1]

Instead, the main contribution of this paper is to specify an underlying Markov process X that ultimately defines both short rate and credit spread dynamics and is such that: (i) a reliable and fast numerical method can be implemented to obtain CDX and CDXO prices; (ii) the model parameters can be calibrated to fit sufficiently well the option market price surface; and (iii) the model implied statistical properties of the credit spread fit those implied by the market.

We assume in particular that X is the two dimensional process (r, λ), where r is the short rate, and λ the default intensity process of each entity in the underlying index. The default time for entity i is then modeled as the first time the default intensity integrated process Λ reaches a

[1] Applying the conversion formula requires several inputs, such as the CDXO annuity (also referred to as a hypothetical bond), which are typically unavailable to outsiders. Testing the model here proposed with real market data is made possible thanks to time series of CDXO prices provided by Morgan Stanley.

threshold ε_i, where $\varepsilon_1, ..., \varepsilon_n$ are independent copies of an exponential random variable and n is the numbers of entities in the index. We take mean reverting processes for r and λ. The uncertainty for the rate r is a gamma process g_r, whereas for λ it is the sum of a double gamma process $g_\lambda \circ g_\tau$ and a scalar multiple ρ of g_r itself. Multiple gamma processes were first investigated in [73], for the purpose of randomizing the speed at which jumps occur.

To our knowledge, ours is the first application of a pure jump process with infinite arrival rate in credit risk modeling. Default times defined as above, known as doubly stochastic random times, are commonly used in credit risk modeling (see [10] and [77]) and their development goes back to the works of Duffie and Singleton ([35]), Lando ([63]), Jarrow Lando and Turnbull ([54]) and Madan and Unal ([75]). Common specifications for rate and intensity processes are the affine models developed by Duffie and Kan ([32]) for diffusion models and Duffie, Pan and Singleton ([34]) and Duffie and Garlenau ([33]) for jump-diffusions. We mention that reduced form model with non doubly stochastic random times are also possible, although such a direction was not investigated here. For their development see, e.g., Kusuoka ([62]) and Elliott, Jeanblanc and Yor ([38]).

We calculate the Levy measure of the process (r, λ), from which analytical formulas for prices of zero coupon bonds are derived. Moments, stationary distribution and characteristic exponent of (r, λ) are also computed analytically, and level curves of its bivariate density for different parameters are plotted using a 2D-version of the FFT algorithm (similar to the one in [53]). We then derive analytical formulas for discounted payoff of credit index swaps, and a linear partial integro differential equation (PIDE) for credit index swaptions prices, together with a finite difference numerical scheme for its solution. Calibration is performed for each maturity to all traded

strikes of options on the IG CDX index as of 2 January 2020. We do not perform a stability analysis, but we show that the numerical error for a given set of parameters (obtained from calibration to market prices on 2 January 2020) is close enough to the prices obtained via Montecarlo simulation.

Finally, we compare market and model implied summary statistics of credit spreads for a specific maturity. As shown in [18], variance, skewness and kurtosis of an equity position under the risk neutral measure can be replicated with a continuum of option contracts. Here it is shown that variance, skewness and kurtosis of the spread of a credit index can be replicated with a continuum of credit index swaptions under the measure $\mathbb{Q}^A$ corresponding to choosing as numeraire the annuity of the index. Our model is then calibrated to market prices for all strikes and for a specific maturity for the period between 2 January 2020 through 5 June 2020, and market and model implied variance, skewness and kurtosis of the credit spreads are compared. The closer these are, the better the model approximates the market implied densities. The results of our analysis show that our model is generally able to capture positive skewness and leptokurtic features of CDX spreads under the measure $\mathbb{Q}^A$, and the model and market implied moments are of the same magnitude.

We observe, in particular, that the replication of credits spreads with option contracts under the measure $\mathbb{Q}^A$ represents a new way to extract model-free statistical properties of credit spreads from market prices of options, and can be used to validate any model of credit spreads.

0.2 Acceptable Bilateral Gamma Parameters

The 2-fund separation theorem and the capital asset pricing model (CAPM) of Sharpe and Lintzner ([95], [64]) are based on the assumption that investors' preferences are of mean-variance type, so that, in particular, investors chose portfolios on the mean-variance efficient frontier. As-

suming only risk averse VonNeumann-Morgenstern preferences, Ross proved that k-fund separation holds in a competitive equilibrium if and only if the distribution of assets returns is k-separating ([87]). Ross also proved that the absence of arbitrages implies the existence of statistical factors (arbitrage pricing theory), and in a limit economy the two theories are indistinguishable ([23]). Based on these theoretical results, several factors have been investigated in the empirical finance literature to improve the results of the CAPM. The list includes, in particular, value and growth ([42]) and momentum ([55], [16]). See [21] for other examples.

The covariance between a portfolio of stocks and any of such factors represents a risk that cannot be diversified away, and, theoretically, a rational risk-averse investor should require within a given time frame a larger expected return for a riskier position. This is supported by the evidence reported, to cite a few, in French et al. ([44]), in which it is shown that unexpected changes in volatility are negatively related with unexpected changes in returns, and Bakshi and Kapadia ([7]), in which it is shown that the volatility risk premium is negative. Bollerslev et al. ([14]) and Carr and Wu ([19]) show also that the volatility risk premium is a significant predictor of future stock returns. Glosten et al. ([48]) provide evidence on the other hand that, across time, investors do not require a larger risk premium when volatility is higher than when it is lower, arguing instead that during such periods investors may want to save more, or better instruments to hedge risk may be available to them.

In this paper, the problem of estimating empirically the relationship between risk and return is addressed by representing risk as a vector rather than a single, one-dimensional, parameter. In fact, the empirical evidence that assets returns are leptokurtic and skewed suggests that moments beyond variance should be taken into account in the determination of their statistical distribution,

thus implying that representing risk by a single parameter is an excessive simplification. The bilateral gamma distribution introduced by Kuchler and Tappe ([61]) provides a natural modeling framework for our task, as it is the difference of two independent gamma random variables, which can interpreted as gains and losses, and it is completely specified by the vector $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ of expected value and standard deviation of gains and losses respectively, which determine its first four moments. Furthermore, we show that the expected utility of an asset with bilateral gamma return X is a function $F : (\mu_p, \sigma_p, \mu_n, \sigma_n) \rightarrow \mathbb{E}[u(X)]$, increasing in μ_p, and decreasing in σ_p, μ_n and σ_n, so that, under expected utility theory, risk aversion implies that variations in (σ_p, σ_n) should be compensated by variations of equal sign in $\mu_p - \mu_n$. This is a four dimensional analog to the case of mean and variance for normal returns distribution.

Based on these theoretical results, we represent risk by (σ_p, σ_n), and we estimate the variation in the asset expected log return $\mu = \mu_p - \mu_n$ as σ_p or σ_n changes. Our findings show that there is a positive relationship between μ and σ_p, but a negative one between μ and σ_n. We interpret this result based on the behavioral experiments that constitute the backbone of the prospect theory of Kahneman and Tverski, and, in particular, that humans tend to be risk seekers in losses and risk averse in gains. Such intuition is formally explained through a simple variation of Lucas model ([65]) that incorporates a preference specification dependent both on terminal wealth and the investment strategy outcome.[2]

A second contribution of this paper is to go beyond the law of one price, another fundamental assumption of beta and arbitrage pricing theories, and estimate upper and lower compensations for a given vector of risks $(\sigma_p, \mu_n, \sigma_n)$. Asset pricing theories under uncertainty have been inves-

[2]An interesting task, which is left for future research, is then to construct portfolios that replicate the variance of gains, through which one could better hedge market downturns.

tigated by Epstein and Ji ([40]), who propose a Consumption CAPM in which return follows a G-Brownian motion ([84]) and investors have multiple priors preferences ([47]). To our knowledge, however, a theoretical and empirical investigation of volatility risk premia under uncertainty is still absent from the literature, and this work represents a first study in this direction.

Finally, we investigate the related, but so far unexplored, issue of boundaries within which the volatility skew typically lies.[3] We do so by estimating boundaries and lower dimensional embeddings for risk neutral parameters (b_p, c_p, b_n, c_n) of bilateral gamma returns. In particular, an equation for such boundaries is derived assuming that the set of possible risk neutral measures is in the investors' cone of acceptability, and we test it based on market data.

0.3 Spectral Martingale Measures

This work contributes to the development of the theory of continuous time dynamic spectral risk measures (DSRM), offering in addition illustrations of their applications to nonlinear valuations of financial assets and to portfolio choice theory. Recall that spectral risk measures are sublinear Choquet expectations defined by a concave distortion function, i.e. they are expectations under a measure that lifts the negative tail of a distribution and lowers the positive one. Typically, conditional Choquet expectations are not time-consistent, but one can define a time-consistent discrete time DSRM as iterated Choquet expectations, which is also the solution of a backward stochastic difference equation with a Choquet integral as driver function. Passing to the limit, one obtains a continuous time DSRM, which is then the solution of a backward stochastic differential equation. The construction is carried over in Madan et al. ([71]) for a pure jump process X, and

[3] Another possible extension of this work is to estimate acceptable shapes of the entire implied volatility surface.

the DSRM so generated is identified as the infima of expectations over a set of measures $\mathcal{M}$ that correspond to misspecifications in the jump measure of the process X.

Our theoretical contributions are twofold. First we show that the continuous time DSRM of a square integrable claim $C = f(X_T)$ is a martingale under a measure $\mathbb{Q}^L$, which in general depends on C, and we identify the jump measure of X under $\mathbb{Q}^L$. This result is interesting from a financial point of view, as it shows that lower (and also upper) valuations are martingales under an appropriate measure, and confirms the intuition that bid and ask price processes may not be characterized by utter unfairness. It is also interesting mathematically, as $\mathcal{M}$ is multiplicative stable but not weakly compact ([26]). Furthermore, to our knowledge, this is the first result of this kind that identifies the law of the process X under the measure $\mathbb{Q}^L$. An analogous result is Proposition 2 in [66] which provides a similar formula for *conditional* Choquet integrals defined by S-shaped (and thus not concave) distortions, but under the assumptions that f is monotonic and X is a one-dimensional process, which would greatly simplify our proof. In fact, the main purpose of [66] is to prove the existence of a flow $\Phi(s,t,x,p)$, where s, t and x represent, respectively, the time horizon, the current time and the current value of the process X, such that, given (s,t,x), $\Phi(s,t,x,\cdot)$ is an S-shaped distortion that models behaviors illustrated by prospects theory ([58], [51]), and the associated *conditional* Choquet integrals for monotonic payoffs are time consistent. The results of [66] would not apply, in particular, to portfolios of assets, whereas among our applications we show that maximization of DSRM generates portfolios with higher Sharpe than the market portfolio, and could be thought as a better response to the decisions represented by S-shaped distortions.

The second main theoretical contribution of this work is to define a convex, but not coherent,

dynamic risk measure, which is the continuous time analog of the risk measure constructed in [67]. There is a large literature on convex risk measures, starting with [43]. In the static case, they are represented by infima over the set of probability measures of penalized expectations. The penalty is, in particular, a map over the set of probability measures. A widely employed example is given by the entropic risk measure, for which the penalty is defined as the relative entropy with respect to a reference measure $\mathbb{P}$. However, utility based risk measures suffer from various drawbacks. One is theoretical: the VonNeumann-Morgenstern utility axioms are often violated in practice, and they may not correspond to the financial objectives that large financial institutions maximize or even should maximize. Another one is practical: utility functions are typically estimated on consumption data, which may describe realities that are far from those of large financial institutions. A third one is technical: the conditional version of the entropic risk measure is time consistent if and only if the risk aversion parameter is constant, whereas practically it depends on the investment's time horizon. To our knowledge, no convex risk measures other than the utility based ones have been widely studied or even proposed, also because of the inherent difficulty in specifying a reasonable penalty map. To overcome this issue, we assume given a family of risk measures $\{\rho_c\}_{c\geq 0}$ whose acceptance set is monotonic in c. One can then assume that lower valuations take into account all such risk measures, but a rebate must be offered for those that are most uncertainty averse (i.e. corresponding to higher c). Therefore, the resulting static risk measure is the sum of a Choquet integral and a rebate. In the continuous time setting, we similarly define a driver function as the sum of a Choquet integral and a rebate, and we show that it is well defined and concave.

For applications, we begin by estimating the parameters of a well known pair of measure

distortions Γ_+ and Γ_- to fit upper and lower valuation, defined as 5-day high and low prices. This is typically done through the generalized method of moments (GMM), but our identification of the jump measure of X under $\mathbb{Q}^L$ allows us to perform our estimation by matching the tails of the empirical distribution of the upper and lower prices. This method is the one typically applied to estimate (non distorted) distribution of asset price processes, and our results show that it produces substantially different estimates than the GMM. In particular, the resulting distortions are not symmetric, as those obtained with the GMM. Our knowledge of the jump measure of X under $\mathbb{Q}^L$ also allows us to calibrate the parameters of Γ_+ and Γ_- to bid and ask prices of options via Fast Fourier Transform. For the purpose of obtaining a better fit to the relative bid-ask spread, we also propose a new measure distortion, under which the law of X remains of the same class under $\mathbb{Q}^L$.

Finally, we solve numerically two portfolio choice problems. One maximizes the DSRM, and the other one the Rebated DSRM. In both cases the resulting portfolio outperforms the S&P500 during market downturns. The overall Sharpe ratio during the period 2015-2016 of the DSRM maximizing portfolio is higher in general than that of the S&P500. For the Rebated DSRM portfolio we show that it is also possible to select the optimal amount to invest, and that such amount is typically lower during market downturns.

Chapter 1: Valuation of Credit Index Derivatives

1.1 Introduction

The last few decades saw a spectacular rise in trading volumes of credit derivatives, such as credit default swaps, credit index swaps, single tranche credit default obligations, etc. One reason for this is that the main contract's features of credit default swaps, which form the basic asset class in credit markets, have been standardized thanks to the efforts of bodies such as the International Swaps and Derivatives Association, thus allowing a relatively easy implementation of hedging and speculative strategies and making the credit default swap market more liquid than that of corporate bonds. However, the details of credit derivatives contracts remain complex and satisfactory valuation methods for credit index forwards and swaptions are yet to be determined.

To introduce the mathematical problem, recall that a credit default swap (CDS) is an over the counter contract between two counterparties - the protection buyer and seller - in which protection against the risk of default of an underlying entity (usually a company issuing bonds in the debt market) is provided by the seller to the buyer. The latter pays the former a predetermined premium K (defined as a credit spread multiplied by the contract's notional) at regular intervals until the contract expires and obtains a contingent payment from the seller triggered by any credit event (such as default, restructuring, downgrade, etc.) concerning the underlying entity.

A credit index swap (CDX) can be thought of as a portfolio of credit default swaps. There are two families of credit indices, the CDX and the iTraxx. CDX indices refer to American companies and iTraxx indices refer either to European firms or to Asian and Australian ones. Each family is composed of different indices, each of which representing a different class of credit quality. A summary of the main credit indices is shown in table 1.1. It is important to observe that, in order to reflect changes in the credit quality of the constituents, the composition of most credit indices changes every six months on March 20 and September 20. Each series of an index corresponds to a specific roll date, and older series continue to trade, but their market is far less liquid (see [77]).

Name	Pool size	Region	Credit Quality
CDX.NA.IG	125	North America	Investment Grade
CDX.NA.IG.HVOL	30	North America	Low-quality Investment Grade
CDX.NA.HY	100	North America	Speculative Grade
iTraxx Europe	125	Europe	Investment Grade
iTraxx Europe	30	Europe	Low-quality Investment Grade

Table 1.1: Major credit indices and their characteristics (source: [77]).

Similarly to a CDS, the cash flow associated to a credit index swap consists again of a premium payment leg (with payments made by the protection buyer) and a default payment leg (with payments made by the protection seller). Premium payments, which are defined as a credit spread multiplied by the index annuity (a measure of the number of underlying issuers for which a credit event has not occurred yet) are due at deterministic dates $T_0 < T_1 < ... < T_M$, where T_M is the maturity of the contract and T_0 the inception date (for forward-start contracts $T_0 > 0$). A credit event concerning any of the underlying entities triggers a payment by the seller. Standardized credit index swaps have quarterly premium payments and maturity at issuance is three, five, seven and ten years, with five years being the most liquid traded maturity.

There are two main differences between a CDX and a (portfolio of) CDS: (1) the contingent payment of a CDX is the same for each underlying entity and (2) it does not become an empty contract after a single credit event occurs, so the expected discounted value of the cumulated losses before the inception date (called "front end protection") is included in the price.

In [15], the authors develop a Black-like formula for the forward price under the measure $\mathbb{Q}^A$ that takes as numeraire the index annuity. Apart from this formulation, which assumes a log-normal distribution for the underlying spread and whose purpose is to retrieve prices from spread quotations, there are no generally accepted and/or standard valuation methods for the pricing of credit index derivatives, and, to the author's knowledge, a pure-jump modelling framework for the underlying Markov process driving the credit indexes market is absent from the literature. The aim of this paper is to present a valuable remedy to such deficiency.

Our focus on pure-jump models is also motivated by the possibility that such framework offers to apply the theory of dynamic spectral risk measures (see [71]), thus introducing nonlinearity in the valuation of credit index products. Because of this, and although the exploration of the applications of nonlinear valuations of credit index derivatives is left to future research, we ignore here the relatively small accounting issues related to the front end protection, and assume that no defaults can occur before the time T_0 at which the forward/swaption contract expires.

The rest of the chapter is organized as follows. In section 2 we review basic fundamental concepts of credit markets and credit derivative products. in section 3 we specify the pure-jump dynamics of short rate and default intensity, derive the characteristic exponent of the underlying Markov process and the valuation PIDE. A simple finite difference scheme is tested in section 4, and a comparison of model and market results is shown in section 5. Section 6 concludes.

1.2 Review and Assumptions

1.2.1 Hazard Rates and Doubly Stochastic Random Times

Definition 1.2.1. *Suppose that:*

i. $(\Omega, \mathcal{F}, \mathbb{Q})$ *is a filtered probability space;*

ii. $\{\mathcal{F}_t\}_{t\geq 0}$ *a filtration on* $(\Omega, \mathcal{F}, \mathbb{Q})$*;*[1]

iii. $\tau : \Omega \to [0, \infty]$ *is* $\mathcal{F}$*-measurable and* $\{H_t\}_{t\geq 0} := \sigma\left(\{1_{\{\tau>t\}}\}_{t\geq 0}\right)$*, so that* τ *is an* $\mathcal{H}_t$*-stopping time;*

iv. $\Lambda(t) = \log\left(\mathbb{Q}(\tau > t|\mathcal{F}_\infty)\right)$ *is strictly increasing, finite (i.e.* $\mathbb{Q}(\tau > t|\mathcal{F}_\infty) > 0$ *a.s. for every* $t > 0$*),* $\mathcal{F}_t$*-adapted and absolutely continuous, with* $\Lambda(t) = \int_0^t \lambda(s)ds$*.*

Then, τ *is called a doubly stochastic random time with* $\mathcal{F}_t$*-conditional hazard rate process* λ*.*

Remark 1.2.2. *Since* $\Lambda(t)$ *is* $\mathcal{F}_t$*-adapted, we have* $\mathbb{Q}(\tau \leq t|\mathcal{F}_t) = \mathbb{Q}(\tau \leq t|\mathcal{F}_\infty)\ \forall t \geq 0$*.*

Remark 1.2.3. *The term doubly stochastic refers to the fact that both* τ *and* $\lambda(t)$ *are stochastic.*

Lemma 1.2.4. *Suppose* X *is a standard exponentially distributed random variable on* $(\Omega, \mathcal{F}, \mathbb{Q})$ *independent of* $\mathcal{F}_\infty$*, i.e.* $\mathbb{Q}(X \leq t|\mathcal{F}_\infty) = 1 - e^{-t}$ *for every* $t \geq 0$*. Let* $\lambda(t)$ *be a positive* $\mathcal{F}_t$*-adapted stochastic process such that* $\Lambda(t) = \int_0^t \lambda(s)ds$ *is increasing and finite for every* $t > 0$*. Let* $\tau := \inf\{t \geq 0 : \Lambda(t) \geq X\}$*. Then* τ *is a doubly stochastic random time with hazard process* $\lambda(t)$*.*

Proof. By definition $\{\tau > t\} = \{\Lambda(t) < X\}$. Since $\Lambda(t)$ is $\mathcal{F}_\infty$-measurable and X is independent of $\mathcal{F}_\infty$, we have $\mathbb{Q}(\tau > t|\mathcal{F}_\infty) = \mathbb{Q}(\Lambda(t) < X|\mathcal{F}_\infty) = e^{-\Lambda(t)}$, which proves the result. □

[1] In credit risk modelling, $\{\mathcal{F}_t\}_{t\geq 0}$ is typically generated by some random process Ψ representing some measure of economic activity.

Proposition 1.2.5 (Dellacherie Formulas). *Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a filtered probability space, τ a doubly stochastic random time with $\{\mathcal{F}_t\}_{t\geq 0}$-conditional hazard rate process $\lambda(t)$ and $\{r_t\}_{t\geq 0}$ an $\mathcal{F}_t$-adapted random process. Suppose that, for some $T > 0$, X is $\mathcal{F}_T$-measurable, $\{\nu(t)\}_{0\leq t\leq T}$ and $\{Z(t)\}_{t\geq 0}$ are $\mathcal{F}_t$-adapted.*[2] *If the random variables*

$$|X|e^{-\int_t^T r(s)ds},\ \int_t^T \nu(s)e^{-\int_t^s r(u)du}ds,\ \int_t^T |Z(s)\lambda(s)|e^{-\int_t^s r(u)+\lambda(u)du}ds$$

are all integrable with respect to $\mathbb{Q}$, then

$$\begin{aligned}
&\mathbb{E}\left[e^{-\int_t^T r(s)ds}1_{\{\tau>T\}}X\middle|\ \mathcal{F}_t\vee\mathcal{H}_t\right] = 1_{\{\tau>t\}}\mathbb{E}\left[e^{-\int_t^T r(s)+\lambda(s)ds}X\middle|\ \mathcal{F}_t\right],\\
&\mathbb{E}\left[\int_t^T \nu(s)e^{-\int_t^s r(u)du}1_{\{\tau>s\}}ds\middle|\ \mathcal{F}_t\vee\mathcal{H}_t\right] = 1_{\{\tau>t\}}\mathbb{E}\left[\int_t^T \nu(s)e^{-\int_t^s r(u)+\lambda(u)du}ds\middle|\ \mathcal{F}_t\right],\\
&\mathbb{E}\left[e^{-\int_t^T r(s)ds}1_{\{t<\tau\leq T\}}Z(\tau)\middle|\ \mathcal{F}_t\vee\mathcal{H}_t\right] = 1_{\{\tau>t\}}\mathbb{E}\left[\int_t^T Z(s)\lambda(s)e^{-\int_t^s r(u)+\lambda(u)du}ds\middle|\ \mathcal{F}_t\right],
\end{aligned}$$

where $\mathcal{H}_t = \sigma\left(\{1_{\{\tau>t\}}\}\right)$.

Proof. See [77], proposition 10.19. □

1.2.2 Basics of Forward CDS and CDX Contracts

Consider a forward-start CDS with inception date T_0, tenor structure $T_0 < ... < T_M$, CDS spread c and for a notional of 1 U.S. dollar. Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a probability space, $\{\mathcal{F}_t\}_{t\geq 0}$ a filtration on it, $\{r(t)\}_{t\geq 0}$ an $\mathcal{F}_t$-adapted random process, and $\tau : \Omega \to [0, \infty]$ a doubly stochastic random time with hazard rate $\lambda(t)$. Assuming that τ represents the time of the credit event, the

[2] Typically, X is a *survival claim*, i.e. a promised payment if there is no default, ν is a risky stream of payments that stops when default occurs, and Z is a payment made at default.

payments made by the protection seller (protection leg) discounted at time $t \leq T_0$ are given by

$$\Phi(t) = \delta(\tau) e^{-\int_t^\tau r(s)ds} 1_{\{T_0 < \tau \leq T_M\}},$$

where $\delta(t)$ is the $\mathcal{F}_t$-adapted process representing loss given default. Similarly, the premium leg is given by

$$\Psi(t) = c \sum_{j=1}^{M} e^{-\int_t^{T_j} r(s)ds} 1_{\{\tau > T_j\}} [T_j - T_{j-1}]$$

Using Dellacherie formulas, we have

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[\Phi(t)|\mathcal{F}_t \vee \mathcal{H}_t] &= \mathbb{E}^{\mathbb{Q}}\left[\delta(s) e^{-\int_t^\tau r(s)ds} \left(1_{\{t<\tau\leq T_M\}} - 1_{\{t<\tau<T_0\}}\right) |\mathcal{F}_t \vee \mathcal{H}_t\right] \\
&= 1_{\{\tau>t\}} \mathbb{E}^{\mathbb{Q}}\left[\int_{T_0}^{T_M} \lambda(s)\delta(s) e^{-\int_t^s r(u)+\lambda(u)du} ds) |\mathcal{F}_t\right]
\end{aligned}$$

The present value of the protection buyer's cash flow is then given by

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\Phi(t) - \Psi(t)|\mathcal{F}_t\right] &= 1_{\{\tau>t\}} \mathbb{E}^{\mathbb{Q}}\left[\int_{T_0}^{T_M} \lambda(s)\delta(s) e^{-\int_t^s r(u)+\lambda(u)du} ds)\right] \\
&\quad - 1_{\{\tau>t\}} c \sum_{j=1}^{M} (T_j - T_{j-1}) \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_j} r(u)+\lambda(u)du}\right].
\end{aligned}$$

Since the CDS spread $c(t, T_0, T_M)$ is chosen such that the current value of the contract is zero, we then have

$$c(t, T_0, T_M) = \frac{\mathbb{E}^{\mathbb{Q}}\left[\int_{T_0}^{T_M} \lambda(s)\delta(s) e^{-\int_t^s r(u)+\lambda(u)du} ds)\right]}{\sum_{j=1}^{M} (T_j - T_{j-1}) \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_j} r(u)+\lambda(u)du}\right]}.$$

We next provide the relevant definitions for forward contracts on a credit index. Suppose that the premium payments occur at $T_0 < T_1 < ... < T_M$, where T_M is the maturity of the contract and T_0 is the inception date. Define the following quantities:

(i) **Cumulated losses**: $L(t) = \frac{\delta}{n} \sum_{j=1}^{n} 1_{\{\tau_j < t\}}$, where δ is the loss given default (typically common for each name and nonrandom) and τ_j is the time of default of entity j;

(ii) **Outstanding notional**: $N(t) = 1 - \frac{L(t)}{\delta}$;

(iii) **Index Annuity**:

$$A(t) = \sum_{j=1}^{M} e^{-\int_t^{T_j} r(u)du} \int_{T_{j-1}}^{T_j} N(s)ds \approx \sum_{j=1}^{M} e^{-\int_t^{T_j} r(u)du} N(T_j)(T_j - T_{j-1});$$

(iv) **Premium leg**: $\Psi(t, c) = cA(t)$;

(v) **Protection leg**:

$$\Phi(t) = \int_{T_0}^{T_M} e^{-\int_t^s r(u)du} dL(s) \approx \sum_{j=1}^{M} e^{-\int_t^{T_j} r(u)du} [L(T_j) - L(T_{j-1})]$$

(vi) **Front End Protection**: $F(t) = e^{-\int_t^{T_0} r(s)ds} L(T_0)$, $t \leq T_0$, is the discounted value at time t of cumulated losses at time T_0.

The discounted payoff of a credit index swap is then given by

$$e^{-\int_t^{T_0} r(s)ds} [\Phi(T_0) - \Psi(T_0) + F(T_0)] = \Phi(t) - \Psi(t) + F(t). \tag{1.2.1}$$

Thus, any time a default event is triggered for any of the names composing the index, the name that

defaulted is removed from the index and a payment of size δ/n is made by the protection seller, provided the default event happens after the inception date of the credit index swap. If the event happens before the inception date, then the name that defaulted is again removed from the index, and the loss is paid at the inception of the swap.

As mentioned, we avoid technical complications related to the front end protection, and assume no defaults occur before inception. The discounted payoff at $0 \leq t \leq T_0$ is then

$$e^{-\int_t^{T_0} r(s)ds} \left[\Phi(T_0) - \Psi(T_0)\right] = \Phi(t) - \Psi(t). \tag{1.2.2}$$

To model different default rates, we consider a common intensity process $\{\lambda(t)\}_{t\geq 0}$ for each underlying name (i.e. the pool is "homogeneous"), and, for $i = 1, ..., n$, we define the default time

$$\tau^i = \inf\left\{t > 0 : \Lambda(t) > \varepsilon_i\right\},$$

where $\varepsilon_1, ..., \varepsilon_n$ are independent exponential random variables. Then ([77], lemma 9.33), the default times are conditionally independent doubly stochastic random times, i.e. each τ^i is a doubly stochastic random time with respect to $\mathcal{F}_t$ and

$$\mathbb{Q}(\tau^1 > t, ..., \tau^n > t|\mathcal{F}_\infty) = \prod_{i=1}^{n} \mathbb{Q}(\tau^i > t|\mathcal{F}_\infty).$$

In this case, we obtain for $i = 1, ..., n$,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du} 1_{\{\tau^i > T_{\ell-1}\}} | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du} 1_{\{\tau^i > T_{\ell-1}\}} | \mathcal{F}_{T_0} \vee \mathcal{H}^i_{T_0}\right],$$

where for every $T \geq t \geq 0$, $\mathcal{H}_t = \vee_{i=1}^n \mathcal{H}_t^i$ and $\mathcal{H}_t^i = \sigma(\{1_{\{\tau^i > t\}}, t \geq 0\})$.

Hence, if $1 \leq i \leq n$, $1 \leq \ell \leq M$, and using the Dellacherie formulas 1.2.5 and the tower property of conditional expectation, there is a function $g_\ell : \mathbb{R}_+^3 \to \mathbb{R}$ such that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du} 1_{\{\tau^i > T_{\ell-1}\}} | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_{\ell-1}} r(u)du} 1_{\{\tau^i > T_{\ell-1}\}} P(T_{\ell-1}, T_\ell) | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] \\ &= 1_{\{\tau^i > T_0\}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_{\ell-1}} r(u)+\lambda(u)du} P(T_{\ell-1}, T_\ell) | \mathcal{F}_{T_0}\right] \\ &= 1_{\{\tau^i > T_0\}} g_\ell(T_0, r(T_0), \lambda(T_0)),\end{aligned}$$

where, for every $T \geq t \geq 0$,

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(u)du} | \mathcal{F}_t\right].$$

Similarly, there is a function $h_\ell : \mathbb{R}_+^3 \to \mathbb{R}$ such that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du} 1_{\{\tau^i > T_\ell\}} | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] &= 1_{\{\tau^i > T_0\}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)+\lambda(u)du} | \mathcal{F}_{T_0}\right] \\ &= 1_{\{\tau^i > T_0\}} h_\ell(T_0, r(T_0), \lambda(T_0)).\end{aligned}$$

Therefore, setting $g = \sum_{\ell=1}^M g_\ell$, $h = \sum_{\ell=1}^M h_\ell$, and using the tower property of conditional expectation, we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\Phi(T_0) | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] &= \sum_{\ell=1}^M \frac{\delta}{n} \sum_{i=1}^n \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du} (1_{\{\tau^i > T_{\ell-1}\}} - 1_{\{\tau^i > T_\ell\}}) | \mathcal{F}_{T_0} \vee \mathcal{H}_{T_0}\right] \\ &= \frac{\delta}{n} \sum_{i=1}^n 1_{\{\tau^i > T_0\}} \left[(g(T_0, r(T_0), \lambda(T_0)) - h(T_0, r(T_0), \lambda(T_0)))\right].\end{aligned}$$

As for the premium leg, similar calculations give

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[A(T_0)|\mathcal{F}_{T_0}\vee\mathcal{H}_{T_0}\right] &= \sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)du}1_{\{\tau^i>T_\ell\}}|\mathcal{F}_{T_0}\vee\mathcal{H}_{T_0}\right]\\ &= \sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})\frac{1}{n}\sum_{i=1}^{n}1_{\{\tau^i>T_0\}}h_\ell(T_0, r(T_0), \lambda(T_0))\end{aligned}$$

The payoff $\pi(T_0, r(T_0), \lambda(T_0))$ of a receiver credit index swaption[3] is then

$$\begin{aligned}&\pi(T_0, r(T_0), \lambda(T_0))\\ &= \left[\kappa\sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})h_\ell(T_0, r(T_0), \lambda(T_0)) - \delta\left(g(T_0, r(T_0), \lambda(T_0)) - h(T_0, r(T_0), \lambda(T_0))\right)\right]^+.\end{aligned}$$

This allows one to define a PIDE for the price of a credit index swaption at time t in terms of the vector $(t, r(t), \lambda(t))$. [4]

For the approximated payoff π of the credit index swaption, and given the short rate $r(t)$ at time t, the spread $c(t, T_0, T_M)$ is a function of $\lambda(t)$ only:

$$c(t, \lambda(t), T_0, T_M) := \delta\frac{\mathbb{E}^{\mathbb{Q}}\left[\Phi(t)|\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[A(t)\mathcal{F}_t\right]} = \delta\frac{g(t, r(t), \lambda(t)) - h(t, r(t), \lambda(t))}{\sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})h_\ell(t, r(t), \lambda(t))}. \tag{1.2.3}$$

Remark 1.2.6. *For the case of semiannual payments, the credit spread is given by*

$$\kappa(t, \lambda(t), T_0, T_M) = \delta\frac{\mathbb{E}^{\mathbb{Q}}\left[\Phi(t)|\mathcal{F}_t\right]}{\mathbb{E}^{\mathbb{Q}}\left[A(t)\mathcal{F}_t\right]} = 2\delta\left[\frac{g(t, r(t), \lambda(t))}{h(t, r(t), \lambda(t))} - 1\right].$$

[3] An option contract on a CDX index is of receiver type if the holder has the right, not the obligation, to sell protection, and it is of payer type if the holder has the right, not the obligation, to buy protection.

[4] The value of $\lambda(t)$ is not directly observable, but can be retrieved by equating to zero the value of the forward contract at the strike for which put-call parity holds.

If we assume that there is only one payment, then we obtain the following familiar approximated relationship between credit spread and hazard rate:

$$\kappa(t,\lambda(t),T_0,T_M) = 2\delta\left[K_1(t,T_0)e^{\lambda(t)\left(\frac{1-e^{-\theta_\lambda/2}}{\theta_\lambda}\right)} - 1\right] \approx \delta K_1(t,T_0)\lambda(t) + K_2(t,T_0),$$

where $K_1(t,T_0)$ and $K_2(t,T_0)$ are constants that do not depend on $\lambda(t)$ nor $r(t)$. In particular, the credit spread is, in first order approximation, an affine linear function of the hazard rate.

1.3 A Model for Stochastic Rate and Default Intensity

In this section, we propose a specification for the dynamics of the vector process $X = (r,\lambda)$ for the purpose of pricing credit index derivatives. Inspired by [37], a simple choice for the dynamics of X is

$$\begin{cases} dr(t) = \theta_r(\mu_r - r(t))dt + dg_r(t) \\ d\lambda(t) = \theta_\lambda(\mu_\lambda - \lambda(t))dt + (dg_\lambda(t) + \rho dg_r(t)), \end{cases} \tag{1.3.1}$$

where g_r and g_λ are two independent gamma processes with scale parameters c_r and c_λ, and shape parameters γ_r and γ_λ respectively,[5] Parameters θ_r and θ_λ are positive and measure the speed of mean reversion toward the long term average μ_r and μ_λ respectively. The magnitude of the impact of jumps in the short rate on the default intensity is modeled by the variable $\rho \in \mathbb{R}$. To better fit the option price surface,[6] and recognizing that economic activity in the highly liquid Treasury market often evolves at a different pace than in the debt markets, we further subordinate g_λ to a

[5]Equivalently, for instance, $g_r(1)$ is a gamma random variable with mean γ_r/c_r and variance γ_r/c_r^2. Furthermore, $\log\left(\mathbb{E}^{\mathbb{Q}}\left[e^{iug_r(t)}\right]\right) = \gamma_r t \log\left(\frac{c_r}{c_r - iu}\right)$, and its Levy density φ_r is given by $\varphi_r(x) = \gamma_r \frac{e^{-c_r x}}{x}$.

[6]Although not reported here, we experimented with real market data and found that model 1.3.1 is not rich enough to fit, in particular, the prices of out of the money options. Note also that other modeling choices for λ that are not investigated here are possible, e.g. one could consider an integrated truncated bilateral gamma process.

gamma time change. Specifically, given a third gamma process g_τ with parameters c_τ and γ_τ, we set $\tilde{g}_\lambda(t) = g_\lambda \circ g_\tau(t)$, and consider

$$\begin{cases} dr(t) = \theta_r[\mu_r - r(t)]dt + dg_r(t) \\ d\lambda(t) = \theta_\lambda[\mu_\lambda - \lambda(t)]dt + (d\tilde{g}_\lambda(t) + \rho dg_r(t))) \,, \end{cases} \tag{1.3.2}$$

Standard calculations (see e.g. [9]) give the following solution for the system 1.3.2 given the initial condition $(r(0), \lambda(0))$:

$$\begin{cases} r(t) = r(0)e^{-\theta_r t} + \mu_r(1 - e^{-\theta_r t}) + \int_0^t e^{-\theta_r(t-u)} dg_r(u) \\ \lambda(t) = \lambda(0)e^{-\theta_\lambda t} + \mu_\lambda(1 - e^{-\theta_\lambda t}) + \int_0^t e^{-\theta_\lambda(t-u)} d(\rho g_r(u) + \tilde{g}_\lambda(u)] \end{cases} \tag{1.3.3}$$

In general, given a gamma process g with parameters c and γ and conditioning on $g_\tau(t)$, the characteristic function of the process $\tilde{g} = g \circ g_\tau$ is

$$\begin{aligned} \mathbb{E}\left[e^{i\theta\tilde{g}(t)}\right] = \mathbb{E}\left[\psi_g(\theta)^{g_\tau(t)}\right] = \mathbb{E}\left[e^{g_\tau(t)\log(\psi_g(\theta))}\right] &= \exp\left(\gamma_\tau t \int_0^\infty \left(e^{x\log(\psi_g(\theta))} - 1\right)\varphi_\tau(x)dx\right) \\ &= \exp\left(\gamma_\tau t \int_0^\infty \left(\psi_g(\theta)^x - 1\right)\varphi_\tau(x)dx\right) \\ &= \exp\left(\gamma_\tau t \int_0^\infty \left(\int_0^\infty e^{i\theta y} p_x(y)dy - 1\right)\varphi_\tau(x)dx\right) \\ &= \exp\left(\gamma_\tau t \int_0^\infty \int_0^\infty \left(e^{i\theta y} - 1\right) p_x(y)\varphi_\tau(x)dydx\right) \\ &= \exp\left(\gamma_\tau t \int_0^\infty \left(e^{i\theta y} - 1\right)\left(\int_0^\infty p_x(y)\varphi_\tau(x)dx\right)dy\right). \end{aligned}$$

where $\psi_g(\theta) := \mathbb{E}\left[e^{i\theta g(1)}\right]$ and where p_x is the density of a gamma distribution with parameters c and γx. Therefore, the Levy density of the process $\tilde{g}$ is the weighted gamma Levy density:

$$\varphi_{\tilde{g}}(y) = \int_0^\infty p_x(y)\varphi_\tau(x)dx = \gamma_\tau \frac{e^{-c_\tau y}}{y} \int_0^\infty \frac{(cy)^{\gamma x}}{\Gamma(\gamma x)} \frac{e^{-cx}}{x} dx,\ y > 0, \tag{1.3.4}$$

where Γ is the gamma function. Note that the use of Fubini-Tonelli in the above derivation follows from the fact that the Levy density $\varphi_{\tilde{g}}$ is well defined, which follows from Stirling approximation and an integration by parts. Finally, from 1.3.4, $\tilde{g}$ has infinite arrival rate, and finite variation.

1.3.1 Characteristic Exponent, Zero Coupon Bond Prices and Valuation PIDE

In this section we compute the joint characteristic exponent of X and its integrated process. To simplify notation, we henceforth drop the tilde and thus assume that the process g_λ is subordinated to g_τ. Note that the integrated Ornstein-Uhlenbeck ("OU") process $Y_r(t)$ associated to $r(t)$ is given by

$$\begin{aligned} Y_r(t) &= \int_0^t r(s)ds = \int_0^t r(0)e^{-\theta_r s} + \mu_r\left(1 - e^{-\theta_r s}\right) ds + \int_0^t \int_0^s e^{-\theta_r(s-u)} dg_r(u)ds \\ &= r(0)\frac{1 - e^{-\theta_r t}}{\theta_r} + \mu_r\left(t - \frac{1 - e^{-\theta_r t}}{\theta_r}\right) + \int_0^t \int_u^t e^{-\theta_r(s-u)} ds dg_r(u) \\ &= \mu_r t + (r(0) - \mu_r)\left(\frac{1 - e^{-\theta_r t}}{\theta_r}\right) + \int_0^t \frac{1 - e^{-\theta_r(t-u)}}{\theta_r} dg_r(u) \end{aligned}$$

Similarly, the integrated OU process $Y_\lambda(t)$ associated to $\lambda(t)$ is given by

$$Y_\lambda(t) = \mu_\lambda t + (\lambda(0) - \mu_\lambda)\left(\frac{1 - e^{-\theta_\lambda t}}{\theta_\lambda}\right) + \int_0^t \frac{1 - e^{-\theta_\lambda(t-u)}}{\theta_\lambda} d(\rho g_r + g_\lambda)(u)$$

Therefore, for every $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$, we have

$$
\begin{aligned}
&\alpha_1 Y_r(t) + \alpha_2 Y_\lambda(t) + \alpha_3 r(t) + \alpha_4 \lambda(t) \\
&= \alpha_1 \left[\mu_r t + (r(0) - \mu_r) \left(\frac{1 - e^{-\theta_r t}}{\theta_r} \right) \right] + \alpha_2 \left[\mu_\lambda t + (\lambda(0) - \mu_\lambda) \left(\frac{1 - e^{-\theta_\lambda t}}{\theta_\lambda} \right) \right] \\
&+ \alpha_3 \left[r(0) e^{-\theta_r t} + \mu_r (1 - e^{-\theta_r t}) \right] + \alpha_4 \left[\lambda(0) e^{-\theta_\lambda t} + \mu_\lambda (1 - e^{-\theta_\lambda t}) \right] \\
&+ \int_0^t \left[\alpha_1 \left(\frac{1 - e^{-\theta_r (t-u)}}{\theta_r} \right) + \alpha_2 \rho \left(\frac{1 - e^{-\theta_\lambda (t-u)}}{\theta_\lambda} \right) + \alpha_3 \left(e^{-\theta_r (t-u)} \right) + \alpha_4 \rho \left(e^{-\theta_\lambda (t-u)} \right) \right] dg_r(u) \\
&+ \int_0^t \left[\alpha_2 \left(\frac{1 - e^{-\theta_\lambda (t-u)}}{\theta_\lambda} \right) + \alpha_4 \left(e^{-\theta_\lambda (t-u)} \right) \right] dg_\lambda(u).
\end{aligned}
$$

Next, set

$$
\begin{aligned}
\xi_r(t, r, \alpha_1, \alpha_3) &= \alpha_1 \left[\mu_r t + (r - \mu_r) \left(\frac{1 - e^{-\theta_r t}}{\theta_r} \right) \right] + \alpha_3 \left[r e^{-\theta_r t} + \mu_r (1 - e^{-\theta_r t}) \right] \\
\xi_\lambda(t, \lambda, \alpha_2, \alpha_4) &= \alpha_2 \left[\mu_\lambda t + (\lambda - \mu_\lambda) \left(\frac{1 - e^{-\theta_\lambda t}}{\theta_\lambda} \right) \right] + \alpha_4 \left[\lambda e^{-\theta_\lambda t} + \mu_\lambda (1 - e^{-\theta_\lambda t}) \right] \\
\psi_r(t, u, \alpha_1, \alpha_2, \alpha_3, \alpha_4) &= \alpha_1 \left(\frac{1 - e^{-\theta_r (t-u)}}{\theta_r} \right) + \alpha_2 \rho \left(\frac{1 - e^{-\theta_\lambda (t-u)}}{\theta_\lambda} \right) \\
&\quad + \alpha_3 \left(e^{-\theta_r (t-u)} \right) + \alpha_4 \rho e^{-\theta_\lambda (t-u)} \\
\psi_\lambda(t, u, \alpha_2, \alpha_4) &= \alpha_2 \left(\frac{1 - e^{-\theta_\lambda (t-u)}}{\theta_\lambda} \right) + \alpha_4 \left(e^{-\theta_\lambda (t-u)} \right).
\end{aligned}
$$

Since $|\psi_r(t, u, \alpha_1, \alpha_2, \alpha_3, \alpha_4)| \leq |\alpha_1| + |\rho \alpha_2| + |\alpha_3| + \rho |\alpha_4|$, we have

$$
\begin{aligned}
&\mathbb{E}_0^{\mathbb{Q}} \left[\exp \left(\int_0^t i \psi_r(t, u, \alpha_1, \alpha_2, \alpha_3, \alpha_4) dg_r(u) \right) \right] \\
&= \lim_{n \to \infty} \mathbb{E}_0^{\mathbb{Q}} \left[\exp \left(\sum_{k=0}^{n-1} i \psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4) \left(g_r((k+1)t/n) - g_r(kt/n) \right) \right) \right],
\end{aligned}
$$

where $\mathcal{F}_t$ is the filtration generated by X and $\mathbb{E}_0^{\mathbb{Q}}[\cdot]$ denotes expectation conditional on $\mathcal{F}_0$. By the stationariness and independence of the increments of the process g_r, we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}_0^{\mathbb{Q}}\left[\exp\left(\sum_{k=0}^{n-1} i\psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4)\left(g_r((k+1)t/n) - g_r(kt/n)\right)\right)\right] \\
&= \lim_{n\to\infty} \prod_{k=0}^{n-1} \mathbb{E}_0^{\mathbb{Q}}\left[\exp\left(i\psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4) g_r(t/n)\right)\right].
\end{aligned}$$

By the properties of the Levy density φ_r of the process g_r, we have, for every $\psi \in \mathbb{C}$,

$$\begin{aligned}
&\frac{1}{t}\log\left(\mathbb{E}_0^{\mathbb{Q}}\left[e^{i\psi g(t)}\right]\right) = \int_{\mathbb{R}} \left(e^{i\psi y} - 1\right)\varphi_r(y)dy = -\gamma_r \log\left(1 - \frac{i\psi}{c_r}\right), \\
&\Rightarrow \lim_{n\to\infty} \prod_{k=0}^{n-1} \mathbb{E}_0^{\mathbb{Q}}\left[\exp\left(i\psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4) g_r(t/n)\right)\right] \\
&\quad = \lim_{n\to\infty} \prod_{k=0}^{n-1} \exp\left(\frac{t}{n}\int_0^{\infty} \left(e^{i\psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4)y} - 1\right)\varphi_r(y)dy\right) \\
&\quad = \lim_{n\to\infty} \exp\left(\sum_{k=0}^{n-1} \frac{t}{n}\int_0^{\infty} \left(e^{i\psi_r(t, kt/n, \alpha_1, \alpha_2, \alpha_3, \alpha_4)y} - 1\right)\varphi_r(y)dy\right) \\
&\quad = \exp\left(\int_0^t \int_0^{\infty} \left(e^{i\psi_r(t, u, \alpha_1, \alpha_2, \alpha_3, \alpha_4)y} - 1\right)\varphi_r(y)dydu\right) \\
&\quad = \exp\left(-\gamma_r \int_0^t \log\left(1 - \frac{i\psi_r(t, u, \alpha_1, \alpha_2, \alpha_3, \alpha_4)}{c_r}\right)du\right).
\end{aligned}$$

Similar calculations can be done with respect to $\psi_\lambda(t, u, \alpha_2, \alpha_4)$, yielding

$$\begin{aligned}
&\mathbb{E}_0^{\mathbb{Q}}\left[\exp\left(\int_0^t i\psi_\lambda(t, u, \alpha_2, \alpha_4)dg_\lambda(u)\right)\right] \\
&\quad = \exp\left(\int_0^t \int_0^{\infty} \left(e^{i\psi_\lambda(t, u, \alpha_2, \alpha_4)y} - 1\right)\varphi_\lambda(y)dydu\right) \\
&\quad = \exp\left(-\gamma_\tau \int_0^t \log\left(1 + \frac{\gamma_\lambda}{c_\tau}\log\left(1 - \frac{i}{c_\lambda}\psi_\lambda(t, u, \alpha_2, \alpha_4)\right)\right)du\right)
\end{aligned}$$

Therefore, the characteristic exponent of the vector process $(Y_r(t), Y_\lambda(t), r(t), \lambda(t))$ is given by

$$\begin{aligned}
&\log\left(\mathbb{E}^{\mathbb{Q}}\left[e^{i\alpha_1 Y_r(T)+i\alpha_2 Y_\lambda(T)+i\alpha_3 r(T)+i\alpha_4\lambda(T)}|\mathcal{F}_t\right]\right)\\
&= \int_t^T\int_0^\infty \left(e^{i\psi_r(T,u,\alpha_1,\alpha_2,\alpha_3,\alpha_4)y}-1\right)\varphi_r(y)dydu + \int_t^T\int_0^\infty \left(e^{i\psi_\lambda(T,u,\alpha_2,\alpha_4)y}-1\right)\varphi_\lambda(y)dydu\\
&\quad + i\xi_r(T-t, r(t), \alpha_1, \alpha_3) + i\xi_\lambda(T-t, \lambda(t), \alpha_2, \alpha_4). \qquad (1.3.5)
\end{aligned}$$

This implies immediately that the risk neutral price $P(t,T)$ at time t of a zero coupon bond with maturity T and no default risk is given by

$$\begin{aligned}
P(t,T) = \exp\Bigg[&-\mu_r(T-t) - (r(t)-\mu_r)\left(\frac{1-e^{-\theta_r(T-t)}}{\theta_r}\right)\\
&+\int_t^T\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu\Bigg].
\end{aligned}$$

1.3.1.1 Density, Summary Statistics and Stationary Distribution

To gain further insights into the dynamics of the vector $(r(t), \lambda(t))$, we will now derive its density $f_{r,\lambda}$ conditional on $(r(0), \lambda(0))$. Denoting by ϕ the (joint) Fourier transform, and assuming $\mu_r = \mu_\lambda = 0$, we have, for every $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$,

$$\begin{aligned}
\log\left(\phi(\alpha_1,\alpha_2)\right) &= \log\left(\iint_{\mathbb{R}^2} e^{-2\pi i(\alpha_1 r+\alpha_2\lambda)} f_{r,\lambda}(r,\lambda)drd\lambda\right) = \log\left(\mathbb{E}_0^{\mathbb{Q}}\left[e^{-2\pi i(\alpha_1 r(t)+\alpha_2\lambda(t))}\right]\right)\\
&= -2\pi i\alpha_1 r(0)e^{-\theta_r t} - 2\pi i\alpha_2\lambda(0)e^{-\theta_\lambda t}\\
&\quad - \gamma_r\int_0^t \log\left(1+\frac{2\pi i}{c_r}\left(\alpha_1 e^{-\theta_r(t-u)} + \rho\alpha_2 e^{-\theta_\lambda(t-u)}\right)\right)du\\
&\quad - \gamma_\tau\int_0^t \log\left(1+\frac{\gamma_\lambda}{c_\tau}\log\left(1+\frac{2\pi i}{c_\lambda}\alpha_2 e^{-\theta_\lambda(t-u)}\right)\right)du \qquad (1.3.6)
\end{aligned}$$

By Fourier inversion and a change of variable, the joint density of $(r(t), \lambda(t))$ is then given by

$$f_{r,\lambda}(r,\lambda) = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} e^{i(\alpha_1 r + \alpha_2 \lambda)} \phi\left(\frac{\alpha_1}{2\pi}, \frac{\alpha_2}{2\pi}\right) d\alpha_1 d\alpha_2 \tag{1.3.7}$$

As shown in [53], this double integral can be computed using a two dimensional fast Fourier transform. Specifically, we set $N = 2^{13}$, $B = 10^6$, $\eta = \frac{2B}{N}$, $\lambda = \frac{2\pi}{N\eta} = \frac{\pi}{B}$, $b = \frac{N\lambda}{2} = \frac{\pi}{\eta}$, and approximate 1.3.7 by a double sum over the grid in the frequency domain

$$F = \left\{\alpha_k = (\alpha_{k_1}, \alpha_{k_2}) : k = (k_1, k_2) \in \{0, 1, ..., N-1\}^2\right\},\ \alpha_{k_i} = -B + k_i\eta,\ i = 1, 2$$

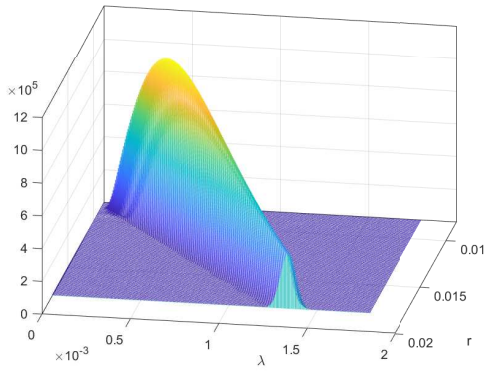
with corresponding grid in the space domain given by

$$S = \left\{x_\ell = (x_{\ell_1}, x_{\ell_2}) : \ell = (\ell_1, \ell_2) \in \{0, 1, ..., N-1\}^2\right\},\ x_{\ell_i} = -b + \ell_i\eta,\ i = 1, 2.$$

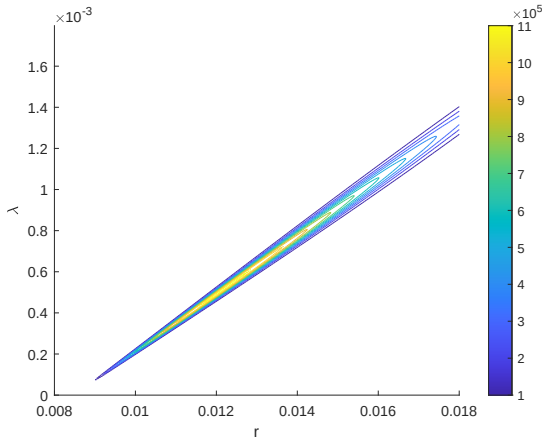
Thus, we have the approximation

$$\begin{aligned} f_{r,\lambda}(r,\lambda) &\approx \frac{\eta^2}{4\pi^2} \sum_{k_1,k_2=0}^{N-1} e^{i\alpha_k x'_\ell} \phi\left(\frac{\alpha_{k_1}}{2\pi}, \frac{\alpha_{k_2}}{2\pi}\right) \\ &= (-1)^{\ell_1+\ell_2} \left(\frac{\eta N}{2\pi}\right)^2 \frac{1}{N^2} \sum_{k_1,k_2=0}^{N-1} e^{2\pi i k\ell'/N} (-1)^{k_1+k_2} \phi\left(\frac{\alpha_{k_1}}{2\pi}, \frac{\alpha_{k_2}}{2\pi}\right), \end{aligned}$$

where the last double sum can be computed for instance in Matlab with the command `ifft2`. Figures 1.1 and 1.2 show the bivariate density and the marginals of the vector $(r(t), \lambda(t)$ for $t = 1$ year, and for $r(0) = 0.0146,\ \theta_r = 0.5500\ c_r = 400.0005,\ \gamma_r = 3.9475,\ \rho = 0.1548,\ \lambda(0) = 0;\ \theta_\lambda = 3.3533,\ c_\lambda = 4.3178,\ \gamma_\lambda = 6.0617,\ c_\tau = 3.5298,\ \gamma_\tau = 190.0001.$

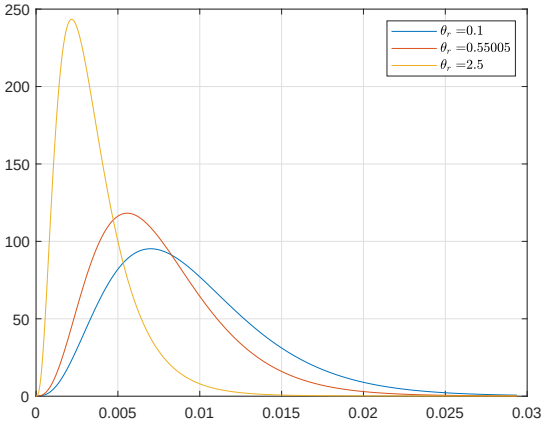


(a) Bivariate density of the random vector $(r(t), \lambda(t))$ for $t = 1$ year.

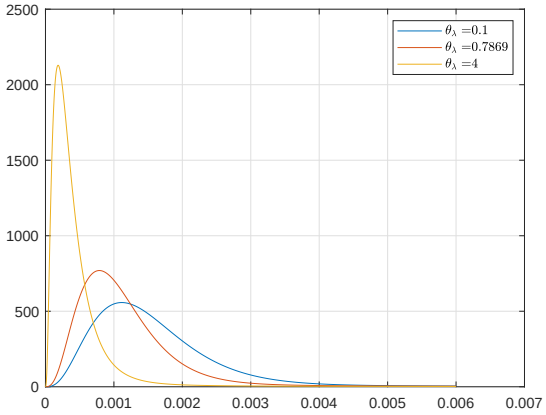


(b) Bivariate density contour of the random vector $(r(t), \lambda(t))$ for $t = 1$ year.

Figure 1.1

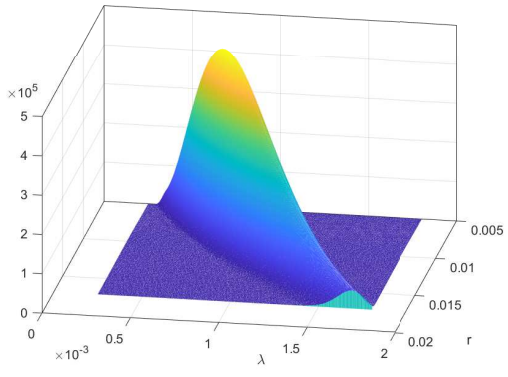


(a) Marginal densities for the short rate for $\theta_r \in \{0.16, 1, 5\}$ and $t = 1$ year.

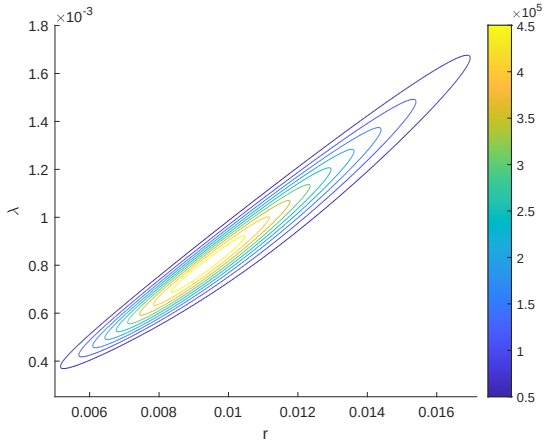


(b) Default intensity marginal for $\theta_\lambda \in \{0.5, 2.06, 4\}$ and $t = 1$ year.

Figure 1.2



(a) Stationary/limiting bivariate density of the random vector $(r(t), \lambda(t))$.



(b) Stationary/limiting bivariate density contour of the random vector $(r(t), \lambda(t))$.

Figure 1.3

The characteristic function allows one to compute the basic summary statistics for $r(t)$ and $\lambda(t)$:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[r(t)] &= r(0)e^{-\theta_r t} + \frac{1}{i}\frac{\partial}{\partial \alpha}\left[-\gamma_r \int_0^t \log\left(1 - \frac{i\alpha e^{-\theta_r(t-u)}}{c_r}\right) du\right]_{\alpha=0} \\
&= r(0)e^{-\theta_r t} + \frac{\gamma_r}{c_r}\frac{1-e^{-\theta_r t}}{\theta_r}, \\
\mathbb{V}^{\mathbb{Q}}[r(t)] &= -\frac{\partial^2}{\partial \alpha^2}\left[-\gamma_r \int_0^t \log\left(1 - \frac{i\alpha e^{-\theta_r(t-u)}}{c_r}\right) du\right]_{\alpha=0} = \frac{\gamma_r}{c_r^2}\frac{1-e^{-2\theta_r t}}{2\theta_r} \\
\mathbb{E}_0^{\mathbb{Q}}[\lambda(t)] &= \lambda(0)e^{-\theta_\lambda t} + \left(\frac{\rho\gamma_r}{c_r} + \frac{\gamma_\tau}{c_\tau}\frac{\gamma_\lambda}{c_\lambda}\right)\frac{1-e^{-\theta_\lambda t}}{\theta_\lambda} \\
\mathbb{V}_0^{\mathbb{Q}}[\lambda(t)] &= \left[\frac{\rho^2\gamma_r}{c_r^2} + \frac{\gamma_\tau\gamma_\lambda}{c_\tau^2 c_\lambda^2}(\gamma_\lambda + c_\tau)\right]\frac{1-e^{-2\theta_\lambda t}}{2\theta_\lambda}.
\end{aligned}$$

These calculations show that higher values of the parameters θ_r and θ_λ imply smaller short rate and default intensity, whereas the smaller θ_r and θ_λ, the higher the variance and expected value of $r(t)$ and $\lambda(t)$ respectively (see also figure 1.2).

Finally, we observe that the change of variable $-(t-u) = v$ gives

$$\begin{aligned}
r(t) &= r(0)e^{-\theta_r t} + \int_0^t e^{-\theta_r(t-u)} dg_r(u) \\
&= r(0)e^{-\theta_r t} + \int_{-t}^0 e^{\theta_r v} dg_r(v) \longrightarrow \int_{-\infty}^0 e^{\theta_r v} dg_r(v), \\
\lambda(t) &= \lambda(0)e^{-\theta_\lambda t} + \int_0^t e^{-\theta_\lambda(t-u)}(dg_\lambda(u) + \rho dg_r(u)) \\
&= \lambda(0)e^{-\theta_\lambda t} + \int_{-t}^0 e^{\theta_\lambda v}(dg_\lambda(u) + \rho dg_r(u)) \longrightarrow \int_{-\infty}^0 e^{-\theta_\lambda v}(dg_\lambda(v) + \rho dg_r(v))
\end{aligned}$$

and also that if the initial condition satisfies

$$(r(0), \lambda(0)) = \left(\int_{-\infty}^0 e^{\theta_r v} dg_r(v), \int_{-\infty}^0 e^{\theta_\lambda v}(dg_\lambda(v) + \rho d_r(v))\right),$$

then

$$r(t) = \int_{-\infty}^{t} e^{-\theta_r(t-u)} dg_r(u) = \int_{-\infty}^{0} e^{\theta_r v} dg_r(v),$$
$$\lambda(t) = \int_{-\infty}^{t} e^{\theta_{-\lambda}(t-u)} (dg_\lambda(u) + \rho dg_r(u)) = \int_{-\infty}^{0} e^{\theta_\lambda v} (dg_\lambda(u) + \rho dg_r(u)).$$

Thus, the limiting/stationary distribution of the process $(r(t), \lambda(t))$ is that of the random vector

$$\left(\int_{-\infty}^{0} e^{\theta_r v} dg_r(v), \int_{-\infty}^{0} e^{\theta_\lambda v} (dg_\lambda(v) + \rho d_r(v)) \right).$$

Similar calculations as above give the Fourier transform of the stationary bivariate density:

$$\phi(\alpha_1, \alpha_2) = \exp\left\{ -\gamma_r \int_{-\infty}^{0} \log\left(1 + \frac{2\pi i}{c_r} \left(\alpha_1 e^{\theta_r v} + \rho \alpha_2 e^{\theta_\lambda v} \right) \right) dv \right.$$
$$\left. -\gamma_\tau \int_{-\infty}^{0} \log\left(1 + \frac{\gamma_\lambda}{c_\tau} \log\left(1 + \frac{2\pi i}{c_\lambda} \alpha_2 e^{\theta_\lambda v} \right) \right) dv \right\}. \tag{1.3.8}$$

Figure 1.3 show the limiting density of the bivariate process $(r(t), \lambda(t))$, obtained by Fourier inversion and approximating the improper integrals in 1.3.8 with the corresponding proper integrals on $[-100, 0]$.

1.3.1.2 Linear PIDE for CDX Swaption prices

By the first fundamental theorem of asset pricing, we know that the price $u(t, r, \lambda)$ of a CDX swpation is given by

$$u(t, r(t), \lambda(t)) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_0} r(u) du} \pi(r(T_0), \lambda(T_0)) \middle| \mathcal{F}_t \right].$$

Therefore, setting for every $t \in [0,T]$

$$M(t) = e^{-\int_0^t r(u)du} u(t, r(t), \lambda(t))$$

we have

$$M(t) = \mathbb{E}^{\mathbb{Q}}\left[M(T) \middle| \mathcal{F}_t \right]$$

i.e. $M(t)$ is a martingale. Furthermore, by Ito's lemma for semimartingales, we have

$$\begin{aligned} M(t) = M(0) &- \int_0^t e^{-\int_0^s r(u)du} r(s) u(s, r(s), \lambda(s)) ds \\ &+ \int_0^t e^{-\int_0^s r(u)du} \left[u_t + u_r \left(\theta_r(\mu_r - r(s)) + u_\lambda \left(\theta_\lambda(\mu_\lambda - \lambda(s)) \right] ds \right. \right. \\ &+ \iint\limits_{(0,t]\times\mathbb{R}^2_+\setminus\{0\}} e^{-\int_0^s r(u)du} \mathcal{D}^{t,r,\lambda}_u(\mathbf{y}) N(ds, d\mathbf{y}). \end{aligned}$$

where N is the Poisson random measure associated to the process $(g_r(t), g_\lambda(t) + \rho g_r(t))$.[7] Therefore, if we denote by φ the Levy density of $(g_r(t), g_\lambda(t) + \rho g_r(t))$, and we add and subtract the

[7] It is easy to see that jumps in r and λ correspond to jumps in g_r and $g_r + \rho g_\lambda$, and their magnitude is also the same. In fact,

$$\begin{aligned} \lim_{s\to 0} r(t+s) - r(t) &= \lim_{s\to 0} \int_t^{t+s} e^{-\theta_r(t-u)} dg_r(u) = \lim_{s\to 0} \int_0^s e^{\theta_r u} dg_r(u+t) \\ &= \lim_{s\to 0} \lim_{n\to\infty} \sum_{k=0}^{n} e^{\theta_r \frac{ks}{n}} \left[g_r(t + (k+1)s/n) - g_r(t + ks/n) \right] \\ &\leq \lim_{s\to 0} \lim_{n\to 0} \sum_{k=0}^{n} e^{\theta_r s} \left[g_r(t + (k+1)s/n) - g_r(t + ks/n) \right] \\ &= \lim_{s\to 0} e^{\theta_r s} \left[g_r(t+s) - g_r(t) \right] = \lim_{s\to 0} \left[g_r(t+s) - g_r(t) \right]. \end{aligned}$$

On the other hand, since $\theta_r > 0$,

$$\lim_{s\to 0} \left[g_r(t+s) - g_r(t) \right] = \lim_{s\to 0} \int_0^s dg_r(u+t) \leq \lim_{s\to 0} \int_0^s e^{\theta_r u} dg_r(u+t) = \lim_{s\to 0} r(t+s) - r(t).$$

Similar considerations hold for jumps in the default intensity. Therefore, the Poisson random measure associated to $(r(t), \lambda(t))$ must be the same as the one associated to $(g_r(t), g_\lambda(t) + \rho g_r(t))$.

compensator to $M(t)$, we obtain, after reversing time, the following PIDE for $u(t, r(t), \lambda(t))$:

$$\begin{cases} u_t + ru - \nabla u \cdot \alpha(r(t), \lambda(t)) - \int_{\mathbb{R}^2_+ \setminus \{0\}} \mathcal{D}_u^{t,r,\lambda}(\mathbf{y})\varphi(\mathbf{y})d\mathbf{y} = 0 \\ u(0, r, \lambda) = \pi(T_0, r, \lambda) \\ + \textit{boundary conditions} \end{cases} \tag{1.3.9}$$

where ∇u is the partial gradient of u with respect to r and λ (a two dimensional row vector) and $\alpha(r, \lambda) := (\theta_r(\mu_r - r), \theta_\lambda(\mu_\lambda - \lambda))^T$. Finally, it is easy to see that, for every $(u_r, u_\lambda) \in \mathbb{R}^2$,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[e^{iu_r g_r(t) + iu_\lambda(\rho g_r(t) + g_\lambda(t)}\right] &= \mathbb{E}^{\mathbb{Q}}\left[e^{i(u_r + \rho u_\lambda)g_r(t)}\right]\mathbb{E}^{\mathbb{Q}}\left[e^{iu_\lambda g_\lambda(t)}\right] \\ &= e^{\iint_{\{y_\lambda = \rho y_r\}}\left(e^{iu_r y_r + u_\lambda y_\lambda} - 1\right)\varphi_r(y_r)dy_r dy_\lambda} e^{\int_0^\infty \left(e^{iu_\lambda y_\lambda} - 1\right)\varphi_\lambda(y_\lambda)dy_\lambda}, \end{aligned}$$

which implies

$$\int_{\mathbb{R}^2_+ \setminus \{0\}} \mathcal{D}_u^{t,r,\lambda}(\mathbf{y})\varphi(\mathbf{y})d\mathbf{y} = \int_0^\infty \mathcal{D}_u^{t,r,\lambda}(y_r, \rho y_r)\varphi_r(y_r)dy_r + \int_0^\infty \mathcal{D}_u^{t,r,\lambda}(0, y_\lambda)\varphi_\lambda(y_\lambda)dy_\lambda.$$

1.4 Numerical Results

We implemented a finite difference scheme for the valuation PIDE, whose construction is reported in the appendix. The scheme was then tested taking as final condition the payoff of a forward start CDX swap, whose current value admits an integral representation. The proof is based on calculations that are similar to those performed in section 1.3.1.1. In particular, we obtain

that, for every $t \in [0, T_0]$, $\ell = 1, ..., M$,

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_0} r(u)du}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)+\lambda(u)du}\big|\mathcal{F}_{T_0}\right]\Big|\,\mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_\ell} r(u)du}e^{\xi_r(T_\ell-T_0,r(T_0),-1,0)+\xi_\lambda(T_\ell-T_0,\lambda(T_0),-1,0)}\big|\mathcal{F}_t\right] \\
&\quad\times e^{\int_{T_0}^{T_\ell}\int_0^\infty\left(e^{\psi_r(T_\ell,u,-1,-1,0,0)y}-1\right)\varphi_r(y)dydu+\int_{T_0}^{T_\ell}\int_0^\infty\left(e^{\psi_\lambda(T_\ell,u,-1,0)y}-1\right)\varphi_\lambda(y)dydu} \\
&= e^{\xi_r(T_0-t,r(t),-1,b_3)+\xi_\lambda(T_0-t,\lambda(t),0,b_4)} \\
&\quad\times e^{\int_t^{T_0}\int_0^\infty\left(e^{\psi_r(T_0,u,-1,0,b_3,b_4)y}-1\right)\varphi_r(y)dydu+\int_t^{T_0}\int_0^\infty\left(e^{\psi_\lambda(T_0,u,0,b_4)y}-1\right)\varphi_\lambda(y)dydu} \\
&\quad\times e^{\int_{T_0}^{T_\ell}\int_0^\infty\left(e^{\psi_r(T_\ell,u,-1,-1,0)y}-1\right)\varphi_r(y)dydu+\int_{T_0}^{T_\ell}\int_0^\infty\left(e^{\psi_\lambda(T_\ell,u,-1,0)y}-1\right)\varphi_\lambda(y)dydu},
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_0} r(u)du}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_{\ell-1}} r(u)+\lambda(u)du}P(T_{\ell-1},T_\ell)|\mathcal{F}_{T_0}\right]\Big|\,\mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_0} r(u)du}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_{\ell-1}} r(u)+\lambda(u)du}e^{r(T_{\ell-1})\alpha_3}\big|\mathcal{F}_{T_0}\right]\Big|\,\mathcal{F}_t\right] \\
&\quad\times e^{\int_{T_{\ell-1}}^{T_\ell}\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T_\ell-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu} \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_0} r(u)du}e^{\xi_r(T_{\ell-1}-T_0,r(T_0),-1,\alpha_3)+\xi_\lambda(T_{\ell-1}-T_0,\lambda(T_0),-1,0)}\Big|\,\mathcal{F}_t\right] \\
&\quad\times e^{\int_{T_0}^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_r(T_{\ell-1},u,-1,-1,\alpha_3,0)y}-1\right)\varphi_r(y)dydu+\int_{T_0}^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_\lambda(T_{\ell-1},u,-1,0)y}-1\right)\varphi_\lambda(y)dydu} \\
&\quad\times e^{\int_{T_{\ell-1}}^{T_\ell}\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T_\ell-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu} \\
&= e^{\xi_r(T_0-t,r(t),-1,a_3)+\xi_\lambda(T_0-t,\lambda(t),0,a_4)} \\
&\quad\times e^{\int_t^{T_0}\int_0^\infty\left(e^{\psi_r(T_0,u,-1,0,a_3,a_4)y}-1\right)\varphi_r(y)dydu+\int_t^{T_0}\int_0^\infty\left(e^{\psi_\lambda(T_0,u,0,a_4)y}-1\right)\varphi_\lambda(y)dydu}. \\
&\quad\times e^{\int_{T_0}^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_r(T_{\ell-1},u,-1,-1,\alpha_3,0)y}-1\right)\varphi_r(y)dydu+\int_{T_0}^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_\lambda(T_{\ell-1},u,-1,0)y}-1\right)\varphi_\lambda(y)dydu} \\
&\quad\times e^{\int_{T_{\ell-1}}^{T_\ell}\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T_\ell-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu},
\end{aligned}
$$

where

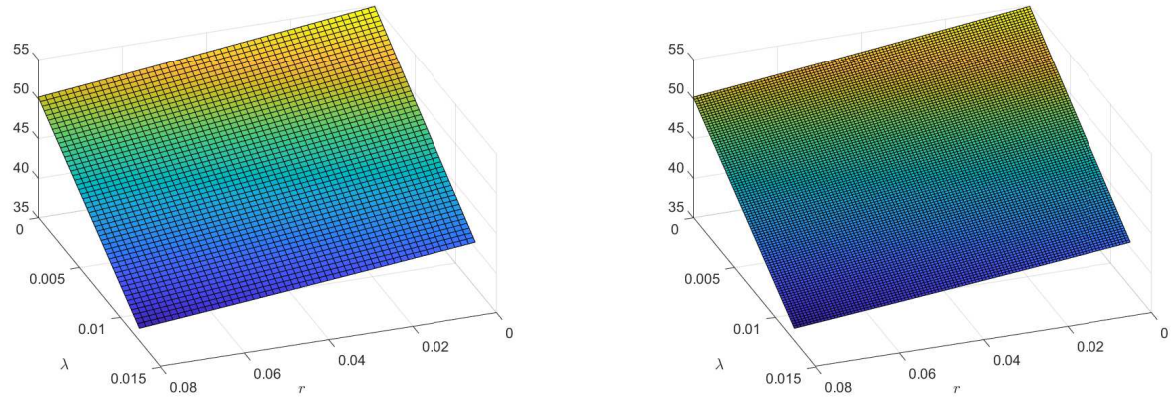
$$\alpha_3 := -\frac{1-e^{-\theta_r(T_\ell - T_{\ell-1})}}{\theta_r},\ a_3 := -\frac{1-e^{-\theta_r(T_{\ell-1}-T_0)}}{\theta_r} + \alpha_3 e^{-\theta_r(T_{\ell-1}-T_0)},$$
$$a_4 := -\frac{1-e^{-\theta_\lambda(T_{\ell-1}-T_0)}}{\theta_\lambda},\ b_3 := -\frac{1-e^{-\theta_r(T_\ell - T_0)}}{\theta_r},\ b_4 := -\frac{1-e^{-\theta_\lambda(T_\ell - T_0)}}{\theta_\lambda}.$$

We considered as before the following set of parameters,

$$r(0) = 0.0146,\ \theta_r = 0.5500\ \ c_r = 400.0005,\ \gamma_r = 3.9475,\ \rho = 0.1548;$$

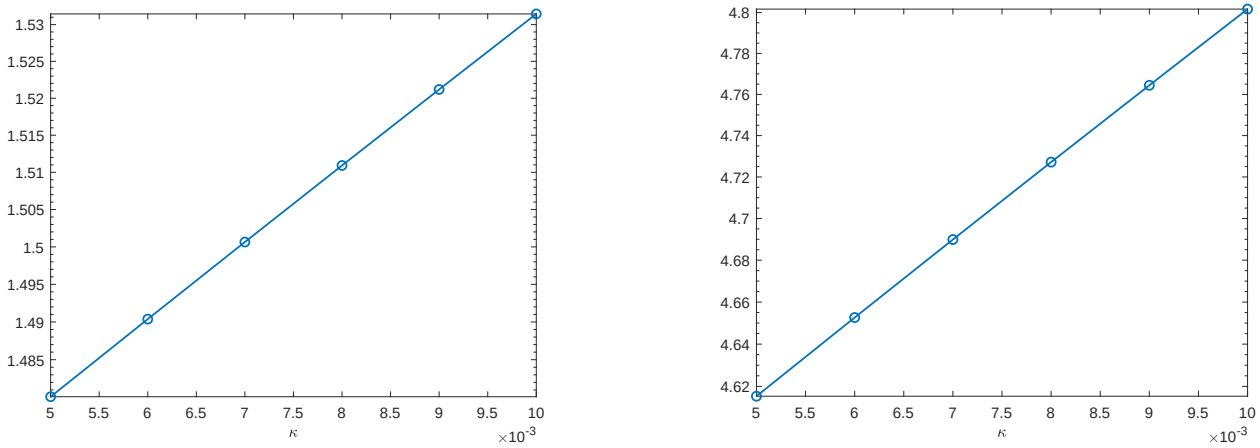
$$\lambda(0) = 0,\ \theta_\lambda = 3.3533,\ c_\lambda = 4.3178,\ \gamma_\lambda = 6.0617,\ c_\tau = 3.5298,\ \gamma_\tau = 190.0001,$$

and we also assumed that the forward contract matures in 15 days, while the underlying asset is a 5-year receiver swap with recovery rate of 0.4, strike $\kappa = 60$ bps and semiannual payments. Figure 1.4 shows the price surface for the forward start credit index swap generated by solving 1.3.9 assuming $N = 50$ (left) and $N = 100$ (right), and with initial condition given by the payoff of the swap at maturity of the forward contract. The ℓ^∞ absolute error is plotted in figure 1.5(a) for strikes $\kappa = 50 : 10 : 100$ bps. Note that even to compute the analytical solutions certain integrations were performed numerically. Figure 1.5 (b) shows the ℓ^∞ absolute error of the solution computed via Montecarlo simulation. For both the PIDE and Montecarlo cases the error is relatively high, and although one could reduce the error for instance by more accurately computing the gamma time changed gamma Levy density 1.3.4 (at the price of higher computational costs), we observe that the error is less than or at least comparable with the bid-ask spread observed in the option market, which, in the period considered, is at least 2 bps.



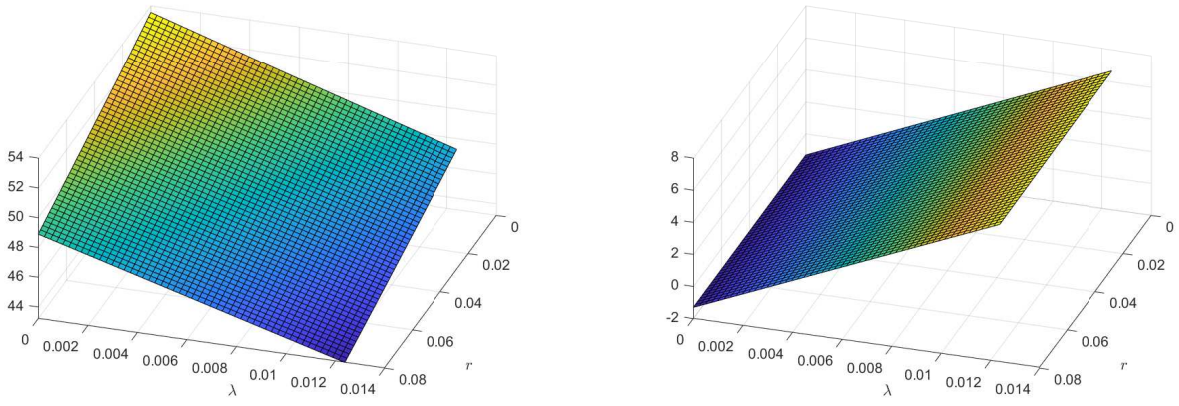
(a) (b)

Figure 1.4: Numerical price surface (in bps) for a forward-start credit index swap, assuming $N = 50$ (a) and $N = 100$ (b), and for $M = 100$ and $N_{sim} = 100$.



(a) (b)

Figure 1.5: ℓ^∞ absolute error (a) and ℓ^∞ difference (in bps) with Montecarlo generated price surface (b) for strikes $\kappa = 50 : 10 : 100$ bps.



(a) (b)

Figure 1.6: Price surface (in bps) of a forward-start CDX including FEP on a 50×50 grid (a) and comparison with the numerical solution (excluding FEP) of 1.7.1 for strike $\kappa = 100$ bps.

In the case of a forward contract, the actual price of the contract, including the front end protection, can be analytically computed. In particular, we have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_\ell} r(u)du}1_{\{\tau^i>T_{\ell-1}\}}|\mathcal{F}_t\vee\mathcal{H}_t\right]=1_{\{\tau^i>t\}}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_{\ell-1}} r(u)+\lambda(u)du}P(T_{\ell-1},T_\ell)|\mathcal{F}_t\right]\\
&=1_{\{\tau^i>t\}}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_{\ell-1}} r(u)+\lambda(u)du}e^{r(T_{\ell-1})\alpha_3}|\mathcal{F}_t\right]e^{\int_{T_{\ell-1}}^{T_\ell}\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T_\ell-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu}\\
&=1_{\{\tau^i>t\}}e^{\xi_r(T_{\ell-1}-t,r(t),-1,\alpha_3)+\xi_\lambda(T_{\ell-1}-t,\lambda(t),-1,0)}\\
&\times e^{\int_t^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_r(T_{\ell-1},u,-1,-1,\alpha_3,0)y}-1\right)\varphi_r(y)dydu+\int_t^{T_{\ell-1}}\int_0^\infty\left(e^{\psi_\lambda(T_{\ell-1},u,-1,0)y}-1\right)\varphi_\lambda(y)dydu}\\
&\times e^{\int_{T_{\ell-1}}^{T_\ell}\int_0^\infty\left(e^{-\frac{1-e^{-\theta_r(T_\ell-u)}}{\theta_r}y}-1\right)\varphi_r(y)dydu},
\end{aligned}$$

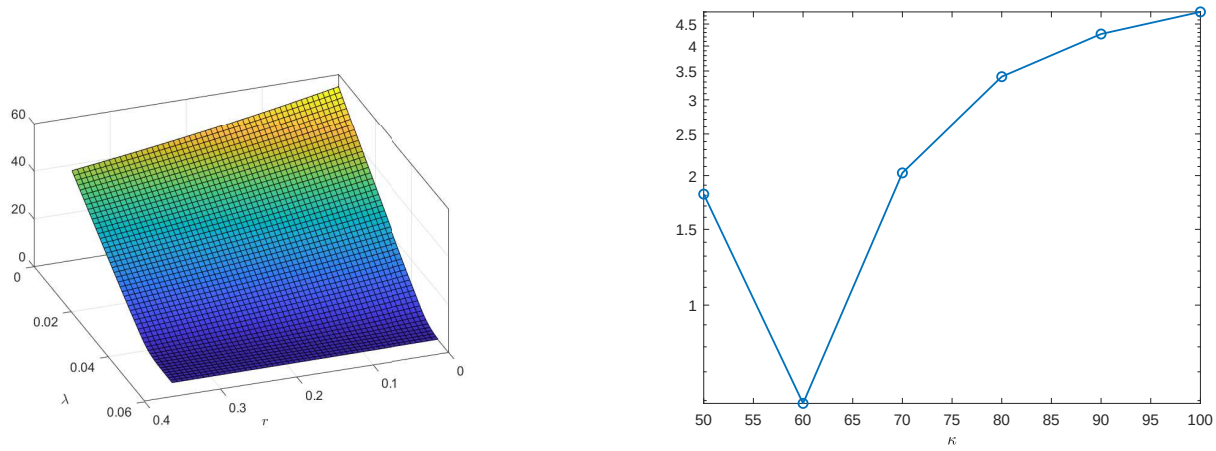
and

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_\ell} r(u)du}1_{\{\tau^i>T_\ell\}}|\mathcal{F}_t\right]=1_{\{\tau^i>t\}}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^{T_\ell} r(u)+\lambda(u)du}|\mathcal{F}_t\vee\mathcal{H}_t\right]\\
&=1_{\{\tau^i>t\}}e^{\xi_r(T_\ell-t,r(t),-1,0)+\xi_\lambda(T_\ell-t,\lambda(t),-1,0)}\\
&\times e^{\int_t^{T_\ell}\int_0^\infty\left(e^{\psi_r(T_\ell,u,-1,-1,0,0)y}-1\right)\varphi_r(y)dydu+\int_t^{T_\ell}\int_0^\infty\left(e^{\psi_\lambda(T_\ell,u,-1,0)y}-1\right)\varphi_\lambda(y)dydu},
\end{aligned}$$

where, as before,

$$\alpha_3=-\frac{1-e^{-\theta_r(T_\ell-T_{\ell-1})}}{\theta_r}.$$

Figures 1.6 shows that the value of the front end protection is relatively small for this set of parameters, although, as noted for instance in [15], its value can be substantial for higher values of $\lambda(0)$.



(a) (b)

Figure 1.7: Receiver CDX option price surface for $\kappa = 60$ bps and $N = 50$ (left) and ℓ^∞ absolute difference with the Montecarlo generated price surface for $\kappa = 50 : 10 : 100$.

We now turn our attention to the option contracts on a CDX index. The numerical price is shown in figure 1.7 (a), while the ℓ^∞ absolute error with respect to the Montecarlo generated surface for $\kappa = 50 : 10 : 100$ and for $r(0) = 146$ bps is shown in figure 1.7 (b).

Finally, the question of convergence of the numerical method for the case of an option payoff is addressed. For $r(0) = 146$ bps, consider CDX spreads κ in the range 60 to 50 bps. The resulting prices for various values of N, reported in table 1.2, show that convergence up to the second decimal (in bps) is obtained already for $N = 50$.

N/κ	60 bps	50 bps	Cpu time
50	53.98734	12.05898	4.656107e+00
100	53.98690	12.05998	1.225736e+01
150	53.98675	12.06023	2.448375e+01
200	53.98669	12.06029	4.205160e+01
250	53.98665	12.06030	6.590890e+01

Table 1.2: CDXO price (in bps) and cpu time for different values of N and strike price κ bps and for $M = N_{sim} = 100$.

1.5 Comparison with Market Data

We calibrated the model to the Treasury yield curve (from www.treasury.gov) and to CDX option prices (provided by Morgan Stanley) as of 2 January 2020 across traded strikes and for each traded maturity. Strike prices are expressed in terms of CDX spreads, and they range from 42.5 bps up to 120 bps. Traded maturities are 13, 43, 76, 104, 139 and 167 business days. The spot CDX spread as of 2 January 2020 was 44 bps. We considered strikes that are up to 30% out of the money (OTM) for receiver and payer contracts for each available maturity.[8] Calibrated parameters for the short rate are the same as those considered above, while those for the default intensity are reported in table 1.3. Figure 1.8 compares the corresponding OTM model and market price.

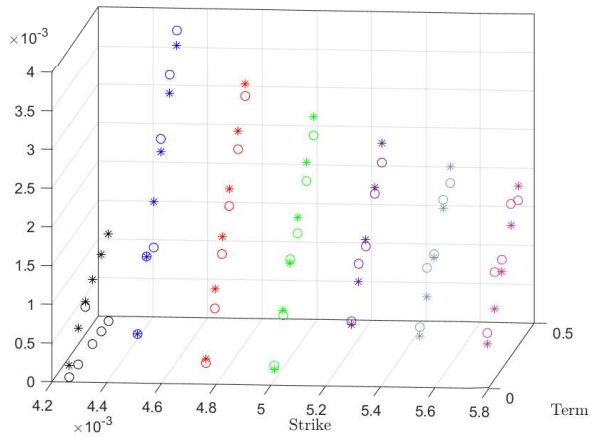
Term (years)	θ_λ	ρ	c_λ	γ_λ	c_τ	γ_τ
0.04	0.1562	0.7869	20.3292	4.1223	604.0000	3.3192
0.13	3.3533	0.1548	4.3178	6.0617	190.0001	3.5298
0.21	2.6789	0.1115	6.1313	2.6983	101.2590	3.6123
0.29	0.0026	0.1280	18.7756	5.1836	312.5091	2.5903
0.39	0.0010	0.1000	10.0981	4.4205	818.1465	4.9855
0.46	0.0010	0.1000	82.2892	1.0241	45.8397	8.4584

Table 1.3

It is also possible to compare market and model implied summary statistics (variance, skewness, kurtosis) of credit spreads for a specific maturity. In particular, as shown above, the payoff of a receiver CDX option maturing at time T_0 with strike spread $\overline{c}$ is given by

$$\pi(T_0) = \left[\overline{c}\mathbb{E}^{\mathbb{Q}}[A(T_0)|\mathcal{F}_{T_0}] - \delta\mathbb{E}^{\mathbb{Q}}[\Phi(T_0)|\mathcal{F}_{T_0}]\right]^+,$$

[8] A receiver option, i.e. an option to sell protection, is OTM if the spot spread is higher than the strike spread, while a payer option is OTM if the spot spread is lower than the strike spread.



1

Figure 1.8: OTM mid price (asterisk) and model price (circle) for maturities and strikes traded on 2 January 2020. Each color corresponds to one of the following strikes: 42.5, 45, 47.5, 50, 52.5, 55, 57.5 (in bps). Maturities are as reported in table 1.3 and model prices are computed using the corresponding parameters also reported in table 1.3.

while the spot credit spread $c(T_0)$ satisfies

$$c(T_0)\mathbb{E}^{\mathbb{Q}}\left[A(T_0)|\mathcal{F}_{T_0}\right] = \delta\mathbb{E}^{\mathbb{Q}}\left[\Phi(T_0)|\mathcal{F}_{T_0}\right].$$

Therefore, we have

$$\pi(T_0) = \mathbb{E}^{\mathbb{Q}}[A(T_0)|\mathcal{F}_{T_0}]\left(\overline{c} - c(T_0)\right)^+.$$

Taking the index annuity as numeraire,[9] and letting $\mathbb{Q}^A$ denote the associated probability measure, the price $u_r(t,c)$ and $u_p(t,c)$ of a receiver/payer option at time t is given by

$$u_r(t,\overline{c}) = \mathbb{E}^{\mathbb{Q}}[A(t)|\mathcal{F}_t]\mathbb{E}^{\mathbb{Q}^A}[(\overline{c} - c(T_0))^+|\mathcal{F}_t],\ u_p(t,\overline{c}) = \mathbb{E}^{\mathbb{Q}}[A(t)|\mathcal{F}_t]\mathbb{E}^{\mathbb{Q}^A}[(c(T_0) - \overline{c})^+|\mathcal{F}_t].$$

[9]Technically, the index annuity may be null on a set of positive measure. This happens in the case of an armageddon event, i.e. all the entities in the index default prior to the option expiration. Our assumption that such event has approximately zero probability (which is more likely the shorter the maturity of the option) ensures that the approximation error in taking the index annuity as numeraire is small enough.

We can then use a result due to Madan and Carr (see [18]), according to which a twice continuously differentiable payoff function $H(c) \in \mathcal{C}^2(\mathbb{R})$ can be written as

$$H(c) = H(\hat{c}) + (c - \hat{c})H'(\hat{c}) + \int_{\hat{c}}^{\infty} H''(\overline{c})(c - \overline{c})^+ d\overline{c} + \int_0^{\hat{c}} H''(\overline{c})(\overline{c} - c)^+ d\overline{c}, \tag{1.5.1}$$

where $\hat{c} \geq 0$ is arbitrary. This allows one to compute model-free summary statistics of the spot credit spread $c(T_0)$ under $\mathbb{Q}^A$. To do so, define the volatility, cubic and quartic contracts as $(c - c_f)^2,\ (c - c_f)^3,\ (c - c_f)^4$, where $c_f = \mathbb{E}^{Q^A}[c(T_0)]$ is the forward credit spread. Setting $\hat{c} = c_f$, 1.5.1 implies

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^2] = 2E_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}\left[\int_{c_f}^{\infty}(c(T_0) - \overline{c})^+ d\overline{c} + \int_0^{c_f}(\overline{c} - c(T_0))^+ d\overline{c}\right],$$

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^3]$$
$$= 6E_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}\left[\int_{c_f}^{\infty}(\overline{c} - c_f)(c(T_0) - \overline{c})^+ d\overline{c} + \int_0^{c_f}(\overline{c} - c_f)(\overline{c} - c(T_0))^+ d\overline{c}\right],$$

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^4]$$
$$= 12E_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}\left[\int_{c_f}^{\infty}(\overline{c} - c_f)^2(c(T_0) - \overline{c})^+ d\overline{c} + \int_0^{c_f}(\overline{c} - c_f)^2(\overline{c} - c(T_0))^+ d\overline{c}\right],$$

which imply, under reasonable assumptions on $c(T_0)$,

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^2] = 2\int_{c_f}^{\infty} u_p(0, \overline{c})d\overline{c} + 2\int_0^{c_f} u_r(0, \overline{c})d\overline{c}, \tag{1.5.2}$$

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^3] = 6\int_{c_f}^{\infty}(\overline{c} - c_f)u_p(0, \overline{c})d\overline{c} + 6\int_0^{c_f}(\overline{c} - c_f)u_r(0, \overline{c})d\overline{c}, \tag{1.5.3}$$

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[(c(T_0) - c_f)^4] = 12\int_{c_f}^{\infty}(\overline{c} - c_f)^2 u_p(0, \overline{c})d\overline{c} + 12\int_0^{c_f}(\overline{c} - c_f)^2 u_r(0, \overline{c})d\overline{c}, \tag{1.5.4}$$

Assuming that non traded deep OTM otions have zero price, one can think of 1.5.2, 1.5.3 and 1.5.4 as reasonable approximations of the first three moments of $c(T_0)$, multiplied by the current value of the index annuity. Note also that calculation of c_f is straightforward since, following the standard put-call parity argument, the price $f_p(t,c)$ at time t of a payer forward with spread c is

$$\begin{aligned} f_p(T_0,c_f) &= \left[\delta\mathbb{E}^{\mathbb{Q}}[\Phi(T_0)|\mathcal{F}_{T_0}] - c_f\mathbb{E}^{\mathbb{Q}}[A(T_0)|\mathcal{F}_{T_0}]\right]^+ - \left[c_f\mathbb{E}^{\mathbb{Q}}[A(T_0)|\mathcal{F}_{T_0}] - \delta\mathbb{E}^{\mathbb{Q}}[\Phi(T_0)|\mathcal{F}_{T_0}]\right]^+ \\ &= u_p(T_0,c_f) - u_r(T_0,c_f) \end{aligned}$$

and so $f_p(0,c_f) = u_p(0,c_f) - u_r(0,c_f)$. Since $f_p(0,0) = \mathbb{E}_0^{\mathbb{Q}}[A(0)]\mathbb{E}_0^{\mathbb{Q}^A}[c(T_0)] = \mathbb{E}_0^{\mathbb{Q}}[A(0)]c_f$, the no arbitrage model-free value of the annuity is

$$\mathbb{E}_0^{\mathbb{Q}}[A(0)] = \frac{f_p(0,0)}{c_f} \approx \frac{u_p(0,0)}{c_f}. \tag{1.5.5}$$

Then, the market implied spread's variance, μ_2, skewness μ_3 and kurtosis μ_4 under $\mathbb{Q}^A$ are:

$$\mu_2 := \mathbb{E}^{\mathbb{Q}^A}\left[(c(T_0)-c_f)^2\right] \tag{1.5.6}$$

$$\approx \frac{2c_f}{f_p(0,0)}\left(\int_{c_f}^{\infty} u_p(0,\overline{c})d\overline{c} + \int_0^{c_f} u_r(0,\overline{c})d\overline{c}\right) \tag{1.5.7}$$

$$\mu_3 := \frac{\mathbb{E}^{\mathbb{Q}^A}[(c(T_0)-c_f)^3]}{\mathbb{E}^{\mathbb{Q}^A}\left[(c(T_0)-c_f)^2\right]^{3/2}} \tag{1.5.8}$$

$$\approx \frac{6c_f}{\mu_2^{3/2}f_p(0,0)}\left(\int_{c_f}^{\infty}(\overline{c}-c_f)u_p(0,\overline{c})d\overline{c} + \int_0^{c_f}(\overline{c}-c_f)u_r(0,\overline{c})d\overline{c}\right) \tag{1.5.9}$$

$$\mu_4 := \frac{\mathbb{E}^{\mathbb{Q}^A}[(c(T_0)-c_f)^4]}{\mathbb{E}^{\mathbb{Q}^A}\left[(c(T_0)-c_f)^2\right]^4} \tag{1.5.10}$$

$$\approx \frac{12c_f}{\mu_2^4 f_p(0,0)}\left(\int_{c_f}^{\infty}(\overline{c}-c_f)^2u_p(0,\overline{c})d\overline{c} + 12\int_0^{c_f}(\overline{c}-c_f)^2u_r(0,\overline{c})d\overline{c}\right) \tag{1.5.11}$$

As shown in table 1.4, model and market implied spread statistics under $\mathbb{Q}^A$ are relatively close, evidencing accuracy of model 1.3.2 in explaining short rate and default intensity dynamics. Note in particular that model 1.3.2 is able to capture the positive skewness and leptokurtic feature of CDX spreads under the measure $\mathbb{Q}_A$.

Variance

Term	Market Implied	Model Implied
0.04	1.054514e+01	7.459700e+00
0.13	5.865457e+01	6.210797e+01
0.21	1.100104e+02	2.152011e+02
0.29	1.812481e+02	2.623242e+02
0.39	2.722237e+02	5.018296e+02
0.46	3.496845e+02	2.280879e+02

Skewness

Term	Market Implied	Model Implied
0.04	8.599558e-01	2.520823e-01
0.13	2.463146e+00	2.974330e+00
0.21	2.660094e+00	3.507166e+00
0.29	3.221129e+00	3.992535e+00
0.39	3.144701e+00	3.809120e+00
0.46	2.896963e+00	2.457448e+00

Kurtosis

Term	Market Implied	Model Implied
0.04	2.308373e+00	2.732897e+00
0.13	9.452965e+00	1.140996e+01
0.21	1.066158e+01	1.230041e+01
0.29	1.566200e+01	1.788540e+01
0.39	1.458586e+01	1.487553e+01
0.46	1.212721e+01	1.028359e+01

Table 1.4: Market and model implied CDX spread statistics for the maturities traded on 2 January 2020.

It is also worth noting that, under 1.3.2,

$$\mathbb{E}_0^{\mathbb{Q}^A}[H(c(T_0))] = \frac{\mathbb{E}_0^{\mathbb{Q}}\left[e^{-\int_0^{T_0} r(u)du}\sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_{T_0}^{T_\ell} r(u)+\lambda(u)du}|\mathcal{F}_{T_0}\right]H(c(T_0))\right]}{\sum_{\ell=1}^{M}(T_\ell - T_{\ell-1})\mathbb{E}_0^{\mathbb{Q}}\left[e^{-\int_0^{T_\ell} r(u)+\lambda(u)du}\right]}, \tag{1.5.12}$$

which can be computed via montecarlo simulation. We performed such computation, but the resulting model implied variance, skewness and kurtosis are not in line with those computed above, ultimately because the statistics in table 1.4 assume that the prices of deep OTM options are zero.

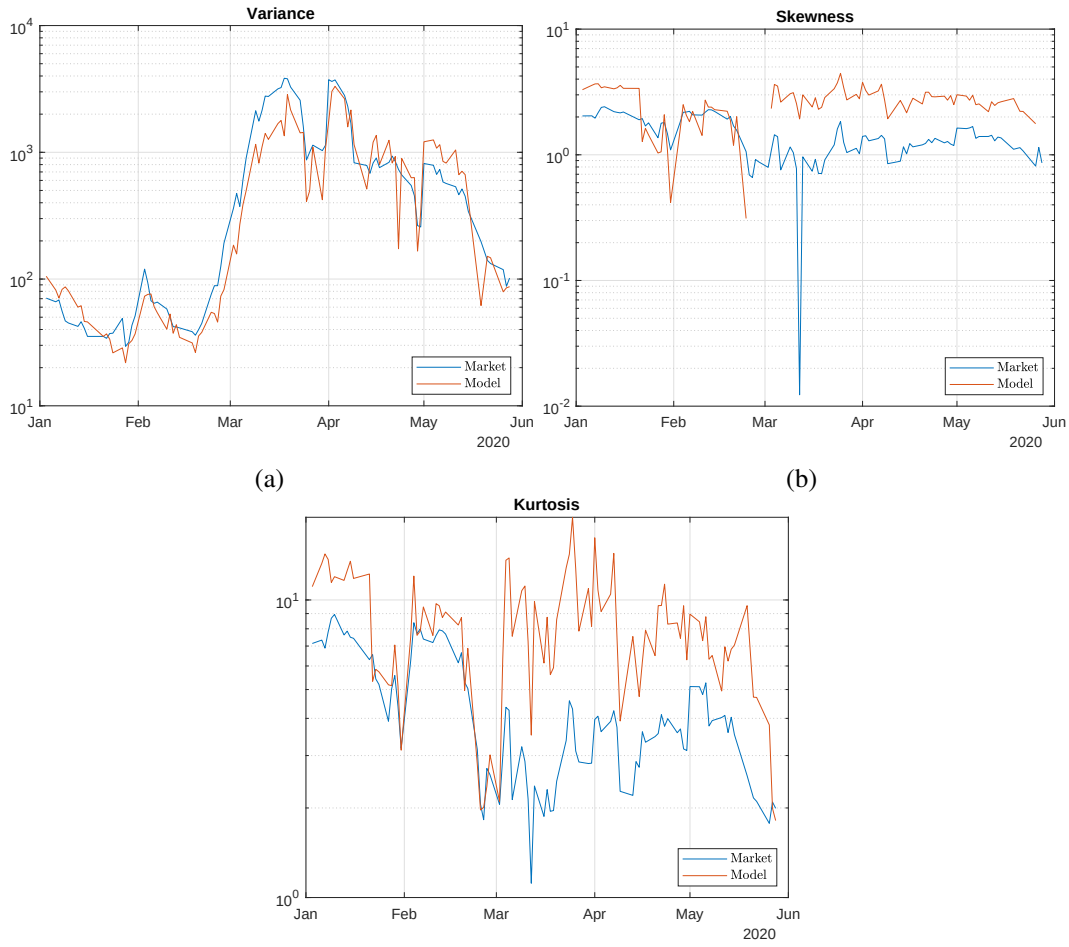


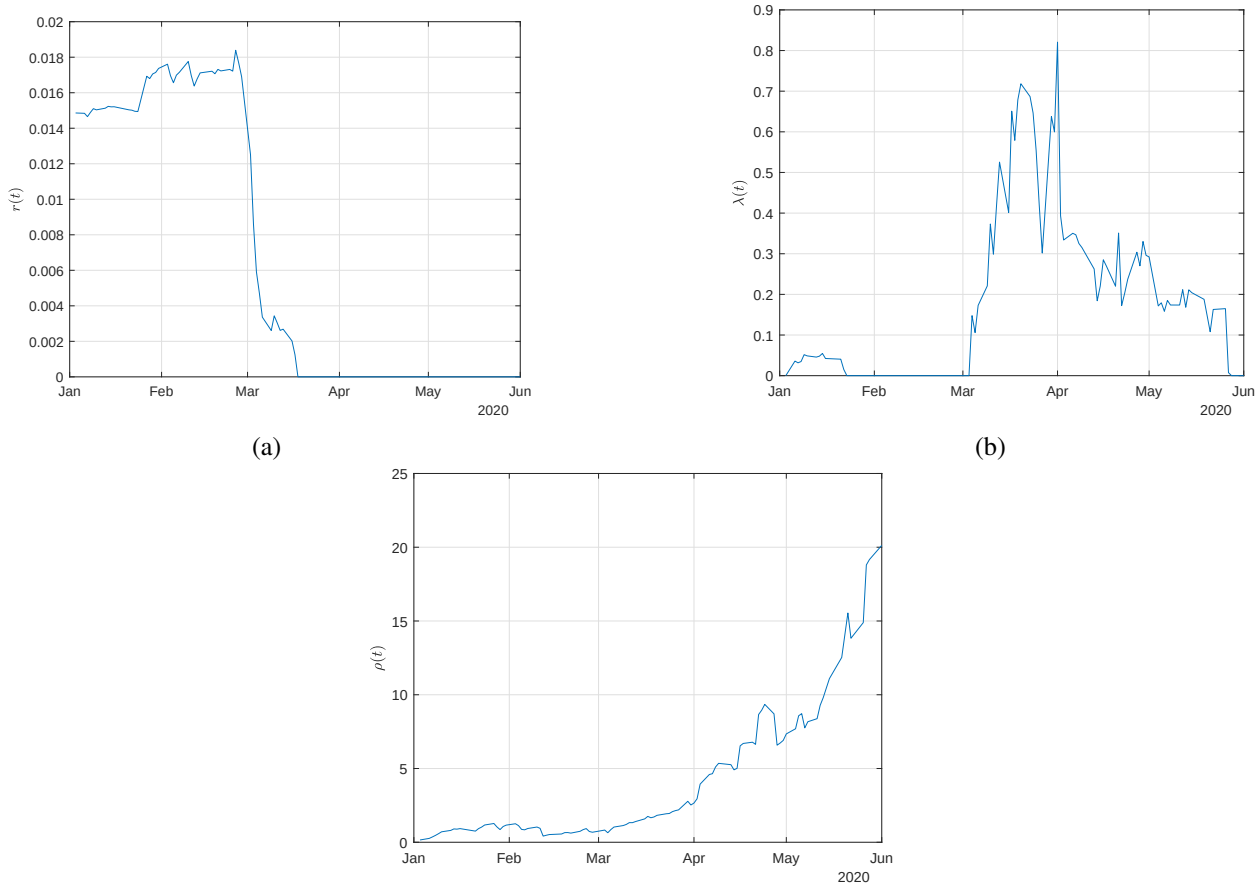
Figure 1.9: Daily market implied and model statistics under $\mathbb{Q}^A$.

Table 1.5 shows model and market implied moments significant correlation between 2 January 2020 and 5 June 2020. Calibration was performed each day using Nelder-Mead algorithm with starting point the optimal parameters for the previous day and maximum 100 iterations.

Statistics	Correlation Coefficient
Variance	0.8892
Skewness	0.3238
Kurtosis	0.4759

Table 1.5: Correlation coefficient for the time series of market and model implied spread statistics.

Finally, figure 1.10 shows the daily realized short rate, default intensity and the parameter ρ.



(a) (b)

(c)

Figure 1.10: Daily realized short rate, hazard rate and parameter ρ.

1.6 Conclusions

We introduced a pure-jump dynamics for the simultaneous modelling of the short rate and default intensity of a pool of entities with similar credit qualities, with the former being a gamma process and the latter also a gamma process but subordinated to another (independent) gamma process. We tested a simple finite difference scheme for the valuation PIDE for forward and option contracts on derivatives determined by short rate and default intensity, and showed that the numerical solution approximates the exact one or a simulated one with reasonable margin of errors. We calibrated the model to the CDX option price surface. For January 2 2020, the calibration error is generally low, but it can be substantial and especially as maturity increases. Finally, we derived a market implied formula for variance, skewness and kurtosis of the credit spread under the Annuity measure, and reported that market and model implied statistics over the year 2020 are of similar magnitude and positively correlated.

1.7 Appendix: The Finite Difference Scheme for the Valuation PIDE

In our finite difference approximation, we treat the integral term fully explicitly. Consider the following mesh on the region $[0,T]\times[0,r_{\max}]\times[0,\lambda_{\max}]$:

$$D=\begin{cases} t_j = j\Delta t;\ \Delta t=\frac{T}{M};\ j=0,1,...,M \\ r_i = i\Delta r;\ \Delta r=\frac{r_{\max}}{N};\ i=0,1,...,N \\ \lambda_k = k\Delta\lambda;\ \Delta\lambda=\frac{\lambda_{\max}}{L};\ k=0,1,...,L \end{cases}$$

We pick $\lambda_{\max} = \rho r_{\max}$ and $L = N$. We denote by $(t_j, r_i, \lambda_k) \in \mathbb{R}^3_+$ the grid points in D, and let $u^j_{i,k} = u(t_j, r_i, \lambda_k)$. Assuming that the $(N+1)^2$ values $u^j_{i,k}$ are known for fixed t_j, we need to construct the difference equation for each point $(t_{j+1}, r_i, \lambda_k)$. Space and time derivatives are approximated using central and forward differences respectively, i.e.

$$u_r(t_{j+1}, r_i, \lambda_k) = \frac{u^{j+1}_{i+1,k} - u^{j+1}_{i-1,k}}{2\Delta r} + O(\Delta r^2),\ u_\lambda(t_{j+1}, r_i, \lambda_k) = \frac{u^{j+1}_{i,k+1} - u^{j+1}_{i,k-1}}{2\Delta\lambda} + O(\Delta\lambda^2),$$
$$u_t(t_{j+1}, r_i, \lambda_k) = \frac{u^{j+1}_{i,k} - u^{j}_{i,k}}{\Delta t} + O(\Delta t^2).$$

We then obtain the following equation for the point $(t_{j+1}, r_i, \lambda_k)$:

$$\begin{aligned}
&\frac{u^{j+1}_{i,k} - u^{j}_{i,k}}{\Delta t} + r_i u^{j+1}_{i,k} - \alpha_1 \frac{u^{j+1}_{i+1,k} - u^{j+1}_{i-1,k}}{2\Delta r} - \alpha_2 \frac{u^{j+1}_{i,k+1} - u^{j+1}_{i,k-1}}{2\Delta\lambda} \\
&\approx \int_0^\infty \mathcal{D}_u^{t_j, r_i, \lambda_k}(y_r, 0)\varphi_r(y_r)dy_r + \int_0^\infty \mathcal{D}_u^{t_j, r_i, \lambda_k}(0, y_\lambda)\varphi_\lambda(y_\lambda)dy_\lambda,
\end{aligned}$$

Equivalently, we have

$$\begin{aligned}
&u^{j+1}_{i,k}(1 + \Delta t r_i) - \frac{\Delta t \alpha_1}{2\Delta r}\left(u^{j+1}_{i+1,k} - u^{j+1}_{i-1,k}\right) - \frac{\Delta t \alpha_2}{2\Delta\lambda}\left(u^{j+1}_{i,k+1} - u^{j+1}_{i,k-1}\right) \\
&\approx u^j_{i,k} + \int_0^\infty \mathcal{D}_u^{t_j, r_i, \lambda_k}(y_r, 0)\varphi_r(y_r)dy_r + \int_0^\infty \mathcal{D}_u^{t_j, r_i, \lambda_k}(0, y_\lambda)\varphi_\lambda(y_\lambda)dy_\lambda. \qquad (1.7.1)
\end{aligned}$$

The integral terms in 1.7.1 can be treated easily via montecarlo simulation. Specifically, at the grid point (t_j, r_i, λ_k), having generated N_{sim} exponentially distributed random variables $\{Y^r_s\}_{s=1,\ldots,N_{sim}}$ with parameter c_r, we have

$$\int_0^\infty \mathcal{D}_u^{t_j, r_i, \lambda_k}(y_r, \rho y_r)\varphi_r(y_r)dy_r \approx \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \mathcal{D}_u^{t_j, r_i, \lambda_k}(Y^r_s, \rho Y^r_s)\frac{\gamma_r}{c_r Y^r_s},$$

and similarly for the second integral. For every $s = 1, ..., N_{sim}$, we compute $\mathcal{D}_u^{t_j,r_i,\lambda_k}(Y_s^r, \rho Y_s^r)$ and $\mathcal{D}_u^{t_j,r_i,\lambda_k}(0, Y_s^\lambda)$ by linearly interpolating u^j, and obtain the following difference equation

$$-S_{i,k}u_{i,k-1}^{j+1} - W_{i,k}u_{i-1,k}^{j+1} + C_{i,k}u_{i,k}^{j+1} - E_{i,k}u_{i+1,k}^{j+1} - N_{i,k}u_{i,k+1}^{j+1} = u_{i,k}^j + \Delta t R_{i,k}^j, \tag{1.7.2}$$

where

$$S_{i,k} = -\frac{\Delta t \alpha_2}{2\Delta\lambda},\ W_{i,k} = -\frac{\Delta t \alpha_1}{2\Delta r},\ E_{i,k} = \frac{\Delta t \alpha_1}{2\Delta r},\ C_{i,k} = 1 + \Delta t r_i,\ N_{i,k} = \frac{\Delta t \alpha_2}{2\Delta\lambda},$$
$$R_{i,k}^j = \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \mathcal{D}_u^{t_j,r_i,\lambda_k}(Y_s^r, \rho Y_s^r) \frac{\gamma_r}{c_r Y_s^r} + \frac{1}{N_{sim}} \sum_{s=1}^{N_{sim}} \mathcal{D}_u^{t_j,r_i,\lambda_k}(Y_s^\lambda, 0) \frac{\gamma_\lambda}{c_\lambda Y_s^\lambda}.$$

Implementation of Boundary Conditions

We impose homogeneous Neumann boundary conditions for each time $t \in [0, T]$:

$$u_{\lambda\lambda}(t, r, \lambda_L) = 0 - u_{\lambda\lambda}(t, r, 0) = 0,\ -u_{rr}(t, 0, \lambda) = 0,\ u_{rr}(t, r_N, \lambda) = 0. \tag{1.7.3}$$

We thus solve

$$\begin{cases} u_t + ru - \nabla u \cdot \alpha = \int_0^\infty \mathcal{D}_u^{t,r,\lambda}(y_r, \rho y_r)\varphi_r(y_r)dy_r + \int_0^\infty \mathcal{D}_u^{t,r,\lambda}(0, y_\lambda)\varphi_\lambda(y_\lambda)dy_\lambda \\ u_{\lambda\lambda}(t, r, \lambda_L) = 0,\ u_{\lambda\lambda}(t, r, 0) = 0,\ u_{rr}(t, 0, \lambda) = 0,\ u_{rr}(t, r_N, \lambda) = 0 \end{cases} \tag{1.7.4}$$

We implement 1.7.3 at the points (t_j, r_1, λ_k), $(t_j, r_{N-1}, \lambda_k)$, (t_j, r_i, λ_1), i.e. we set

$$u_{0,k}^{j+1} = 2u_{1,k}^{j+1} - u_{2,k}^{j+1},\ u_{N,k}^{j+1} = 2u_{N-1,k}^{j+1} - u_{N-2,k}^{j+1},\ u_{i,L+1}^{j+1} = 2u_{i,L}^{j+1} - u_{i,L-1}^{j+1}.u_{i,0}^{j+1} = 2u_{i,1}^{j+1} - u_{i,2}^{j+1}.$$

Chapter 2: Admissibility of Bilateral Gamma Parameters

2.1 Introduction

The definition of *acceptable* risks, based on the axiomatization of the concept of coherent risk measure given in [6] and their convex generalization ([43]), is a major recent advance in mathematical finance, as, among other applications, it provides an operative framework for superhedging in incomplete markets. Starting from a monetary measure, such as Value at Risk, that only satisfies the basic requirements of monotonicity and cash invariance, practical considerations lead one to require that a measure of risk also satisfy subadditivity and positive homogeneity. A measure of risk ρ then defines a set $\mathcal{A}^\rho$ of acceptable risks as those random variables X such that $\rho(X) \geq 0$. Conversely, it is possible to show that given a cone $\mathcal{A}$ of acceptable risks, the functional

$$\rho(X) = \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\} \tag{2.1.1}$$

satisfies monotonicity, cash invariance, subadditivity and positive homogeneity. Based on convex duality, a risk measure is also specified by a set of equivalent probability measures $\mathcal{M}$ as

$$\rho(X) = \inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[X]. \tag{2.1.2}$$

The class $\mathcal{M}$ can be interpreted as the set of possible and credible macroeconomic/financial models, so that 2.1.2 is referred to as the robust representation of ρ, and risk measures become natural tools for the purpose of modeling uncertainty. For convex risk measures, a penalty $\alpha(\mathbb{Q})$ is added to 2.1.2 to take into account that some models $\mathbb{Q} \in \mathcal{M}$ may be more or less plausible than others.

Typical examples of risk measures are those based on certainty equivalent, such as the entropic risk measure, which are known in general as *utility-based shortfall risk measures* and are defined by the acceptance set $\mathcal{A} = \{X : \mathbb{E}[u(X)] \geq u(c)\}$ for a given convex utility u and a threshold c, and those obtained by modifying the tails of the underline statistical measure $\mathbb{P}$, such as the expected shortfall, which are known in general as *spectral risk measures* and are defined by a Choquet integral.

As the above examples confirm, relatively little is known in general about the set $\mathcal{M}$. To gain some empirical understanding of it, we note that a risk measure is essentially an expected value under a worst case scenario measure, and, as such, it defines a minimal current valuation (or maximal bid price) of the future cash flow X, while $-\rho(-X)$ gives a maximal valuation (or the minimal ask price). Then, assuming that log-returns are random variables whose distribution belong to a specific class and is determined by a set $\Theta \in \mathbb{R}^D$ of parameters, observed market prices imply specific boundaries for the set Θ and, in turn, for $\mathcal{M}$. For instance, if $\mathcal{M}$ contains the class of normal random variables parameterized by pairs (μ, σ) of mean and standard deviation of assets returns, one can ask what are maximal and minimal bounds for μ given σ that are implied by historically observed pairs (μ, σ) of traded assets, in turn estimated from market prices. These bounds are then naturally interpreted as structural limits for the reward μ given the risk, measured by σ, that the economic system can offer, in the sense that investors do not trade assets issued

by those less efficient firms that cannot offer sufficiently high compensations, while, on the other hand, the given technological level does not consent exceedingly high returns for lower risk levels.

To fix a reference framework, consider a market composed of one risky asset with log return X and a riskless one in zero net supply with zero risk free rate. Then,

$$1 = \mathbb{E}^{\mathbb{Q}}[e^X] = \mathbb{E}[\eta e^X], \tag{2.1.3}$$

where $\mathbb{Q}$ is a risk neutral measure, η the corresponding stochastic discount factor. If the distribution of X under the statistical measure $\mathbb{P}$ is parameterized by $\theta \in \mathbb{R}^D$, and assuming the existence of a representative investor with utility U defined by a set of parameters $\xi \in \mathbb{R}^m$, there is a function $V : \mathbb{R}^D \times \mathbb{R}^m \to \mathbb{R}$ that evaluates to 1 at (θ, ξ). Specifically (see e.g. [73]), if f_θ is the statistical density of X, the risk neutral density (with respect to the log return) is given by

$$h(x, \theta, \xi) = \frac{U'_\xi(e^x) f_\theta(x)}{\int_{\mathbb{R}} U'_\xi(e^s) f_\theta(s) ds}. \tag{2.1.4}$$

In case of $\log$ utility and log-normal returns with $\theta = (\mu, \sigma)$, equation 2.1.3 then yields

$$e^{\mu - \frac{\sigma^2}{2}} = 1, \tag{2.1.5}$$

and higher σ requires higher μ (assuming fixed risk aversion and risk free rate). In fact, it is natural to think that risks, as measured by σ, move together with their compensation, as measured by μ, as, for instance, new, risky projects yield potentially higher returns, or an adverse event might be compensated by a capital injection. Uncertainty then leads to the existence of a more or less wide

range $[\underline{\mu}, \overline{\mu}]$ of values for μ that are acceptable given σ. With multiple assets the range might be wider, as one typically allocates a fraction of wealth to less rewarding assets for diversification purposes, but observed performance measures imply that the range of values for the reward given the risk(s) are tight ([36]).

As it is well known from [76], [97], [95], [64], the efficient frontier provides the upper limit for the reward μ given σ, and also the lower one, as this is the upper limit for a short position. In general, however, these results lie on the assumption that investors preferences be of the mean-variance type and, in particular, that the optimal portfolio selection is found by maximizing expected utility of terminal wealth. Empirical observations, on the other hand, have shown in many occasions that investors' behavior are not compatible with the expected utility theory axioms.[1]

An alternative to expected utility theory, termed "prospects theory", is based on a series of experiments conducted by psychologists D. Kahneman and A. Tverski ([58]). One of their results is in particular that humans tend to be risk seekers rather than risk averse in the case of pure losses prospects. For instance, the prospect of winning 1000 dollars with probability $1/2$ and winning zero otherwise is generally dominated by the prospect of winning 500 dollars with probability 1, but the prospect of losing 1000 dollars with probability $1/2$ and losing zero otherwise dominates the prospect of losing 500 dollars with probability 1, independently of initial wealth. Based on such evidence, one is then led to interpret an asset's return as the sum of two prospects, one consisting of pure gains, and the other one of pure losses, and investors rank different assets' returns based on the expectations and standard deviations $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ of gains and losses. In particular, higher variance of losses is compensated, ceteris paribus, by lower expectation μ_p of the gains.

[1]For instance, a well known example is the equity premium puzzle, according to which U.S. equity risk premia over Treasury Bills rates reflect an implausible level of aversion to risk under expected utility theory ([78]).

The bilateral gamma distribution ([61]) and its multivariate version ([69]) provide a natural modeling framework for such a preference specification for several reasons. Firstly, it is the difference of two independent gamma variates, interpretable as gains and losses, and it is completely specified by the vector $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ of their expected values and standard deviations. Secondly, even in a continuous time setting, the bilateral gamma process is the difference of two independent gamma processes, while, for instance, path realizations of diffusion processes have infinite variation. Thirdly, the bilateral gamma distribution provides a very good fit to the (log) returns distribution implied by time series of returns and also by options prices ([61]), which shows that it is more suitable than, e.g., the normal distribution for the purpose of modeling asset returns. Finally, as shown below, the expected utility of an asset with bilateral gamma return X is a function $F : (\mu_p, \sigma_p, \mu_n, \sigma_n) \to \mathbb{E}[u(X)]$, increasing in μ_p, and decreasing in σ_p, μ_n and σ_n, so that under expected utility theory variations in $(\sigma_p, \mu_n, \sigma_n)$ are compensated by variations of equal sign in μ_p.

Based on this considerations, we assume in this paper that the set of credible models $\mathcal{M}$ includes the set of bilateral gamma distributions, and we learn bounds $f_M, f_m : (\sigma_p, \mu_p, \sigma_n) \to \mu_p$ for μ_p given risks $(\sigma_p^2, \mu_n, \sigma_n^2)$ via quantile and/or distorted linear and/or Gaussian process regression.

An interesting result obtained is that both boundaries are generally increasing in (σ_p, μ_p), but decreasing in σ_n, suggesting that investors, independently of their wealth, seek for lower (resp. higher) risk when it comes to purely positive (resp. negative) processes. We test the boundaries computed by assessing how well their implied performance measures (Sharpe ratio and acceptability index) compare with those typically observed in the financial markets. Furthermore, we investigate the linearity of f_M and f_m by comparing the results of a linear lower dimensional embedding and a nonlinear one, and we show through a simple variation of a Lucas tree economy

[65] that the behaviors observed are indeed consistent with prospects theory.

Finally, we move our attention to the risk neutral world, based on the suggestive interpretation given in [73] that, for bilateral gamma returns, the scale parameters (b_p, b_n) determine the structure of limit orders, while the speed parameters (c_p, c_n) determine that of market orders. It is then natural to assume that a relationship exists between the two pairs of parameters, in the sense that for given (c_p, c_n), the scale parameters (b_p, b_n) are bounded to a specific range, as the structure of market orders cannot be too independent from that of limit orders and viceversa. As done for the statistical moments, the boundaries of such range are learned through quantile and distorted regression. In this case, we determine theoretical boundaries as well based on the well known robust representation of spectral risk measures ([72]), and evidence is offered of their comparability with the empirically estimated ones.

The rest of the paper is organized as follows. First we show that for bilateral gamma returns, risks and compensations are identified by the vector $(\sigma_p, \mu_n, \sigma_n)$ and μ_p respectively. Empirical observations are reported in section 3, and the variation on Lucas Tree model is presented in section 4. Risk neutral parameters are analyzed in 5. Section 6 concludes.

2.2 Bilateral Gamma Returns

2.2.1 From Brownian Motion to Bilateral Gamma Process

Given its central role in this paper, the construction and properties of the bilateral gamma process are reviewed in this section. In [12], F. Black and M. Scholes proposed to model the dynamics of log-returns as a Brownian motion (GBM), as prices exhibit exponential growth and based on an entropy maximization argument ([73]).

However, returns exhibit heavier tails than those implied by the normal distribution ([41]) and frequent discontinuities in their path trajectories. In addition, risk aversion results in periods of intense trading, determined by widespread selling in securities, alternating with lower activity ones, thus implying that returns' quadratic variation is not linear in time. It also results in higher demand for out of the money (OTM) puts than for the OTM calls, generating a volatility smile.

Another entropy maximization argument then suggests modeling economic time as a gamma process, and stock market log returns as Brownian motion evaluated at such gamma time. The resulting process, pioneered by D. Madan and E. Seneta ([74]) and termed the variance gamma process, is a pure jump Levy process with infinite activity and finite variation. In fact, such process is the difference of two i.i.d. gamma processes, which naturally correspond to gains and losses. Finally, as downward jumps in prices are generally higher than upward ones, the bilateral gamma process is defined as the difference of two independent gamma processes ([61]). The gains and losses increments have BG distribution $\beta\Gamma(b_p, c_p, b_n, c_n)$, defined by the convolution

$$\beta\Gamma(b_p, c_p, b_n, c_n) = \Gamma(b_p, c_p) * \Gamma(-b_n, c_n),$$

where $b_p, c_p, b_n, c_n > 0$ and, for $\alpha > 0$, $\lambda \in \mathbb{R}$, a $\Gamma(\lambda, \alpha)$-distributed random variable has density

$$f(x) = \frac{1}{\Gamma(\alpha)|\lambda|^{\alpha}}|x|^{\alpha-1}e^{-|x|/|\lambda|}\left(1_{\{\lambda>0\}}(x)1_{\{x>0\}}(x) + 1_{\{\lambda<0\}}(x)1_{\{x<0\}}(x)\right),\ x \in \mathbb{R}$$

with $\Gamma(\alpha)$ the Gamma function at α. Then, following the notation in the introduction,

$$\mu_p = c_p b_p,\ \sigma_p = \sqrt{c_p} b_p,\ \mu_n = c_n b_n,\ \sigma_n = \sqrt{c_n} b_n.$$

By the convolution theorem, the characteristic function of the increments in t units of time is

$$\varphi_t(u) = (1 - iub_p)^{-tc_p} (1 + iub_n)^{-tc_n} , \tag{2.2.1}$$

and it follows easily from 2.2.1 that BG densities are stable under convolution and are infinitely divisible, and so the BG process is a well defined Levy process. From formula 2.2.1 and the Levy-Khintchine representation we also deduce its Levy density to be

$$k(x) = \left(\frac{c_p}{x} e^{-x/b_p} 1_{\{(0,\infty)\}}(x) + \frac{c_n}{|x|} e^{-|x|/b_n} 1_{\{(-\infty,0)\}}(x) \right), \ x \in \mathbb{R}$$

which shows that a BG process enjoys the self decomposability property.[2] Then (see [17] and the references therein) a BG distributed random variable X is a limit law, i.e. there are centering and scaling constants $\{d_n\}_{n\in\mathbb{N}}$ and $\{e_n\}_{n\in\mathbb{N}}$ and a sequence $\{Z_k\}_{k\in\mathbb{N}}$ of i.i.d. random variables such that the distribution of $e_n S_n + d_n$ converges in distribution to X, where $S_n = \sum_{k=1}^{n} Z_k$. In particular, if returns consist of some average of a large number of independent news or other type of influences, it is reasonable to expect that their distribution should be well approximated by a limit law. In the GBM case such law is the Gaussian, but, as noticed in [17], there is "no compelling economic motivation" for the scaling constants to be $\sqrt{n}$.

Evidence of the goodness of fit of the BG density to returns distributions is presented in [61]. As also proved in [61], prices evolving as BG processes create no arbitrage opportunities, and one typically succeed in fitting the option prices surface, at least for a single fixed maturity, through an exponential BG process.

[2] A random variable X is self decomposable if for any $0 < c < 1$ there is an independent random variable X_C such that $X \stackrel{d}{=} cX + X_c$. A Levy process enjoys the self decomposability property if its increments are self decomposable.

2.2.2 Bilateral Gamma Returns under Expected Utility Theory

The notion and characterizations of second order stochastic dominance (SSD) are recalled below (see [88]).

Definition 2.2.1. *Given random variables X and Y, one says that X first (resp. second) order stochastically dominates Y, i.e. $X \succeq_1 Y$ (resp. $X \succeq_2 Y$) if and only if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for every increasing (resp. increasing and concave) real valued function u.*

Theorem 2.2.2. *Let X and Y be random variables with distribution functions F and G respectively. Then, $X \succeq_1 Y$ if and only if $G(t) \geq F(t)$ for every $t \in \mathbb{R}$.*

Theorem 2.2.3. *Let X and Y be random variables with distribution functions F and G respectively. Then, the following are equivalent*

(i) $X \succeq_2 Y$*;*

(ii) there are random variables Z and ε such that $Y \sim X + Z + \varepsilon$, $Z \leq 0$ and $\mathbb{E}[\varepsilon|X + Z] = 0$;

(iii) $\int_{-\infty}^{t} G(s)ds \geq \int_{-\infty}^{s} F(s)ds$ *for every $t \in \mathbb{R}$.*

In addition, if $\mathbb{E}[X] = \mathbb{E}[Y]$, then the following are equivalent:

(i) $X \succeq_2 Y$*;*

(ii) there is a random variable ε such that $Y \sim X + \varepsilon$ and $\mathbb{E}[\varepsilon|X + Z] = 0$;

(iii) $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ *for every u concave.*

Corollary 2.2.4. *Suppose $X \succeq_2 Y$. Then, $\mathbb{E}[X] \geq \mathbb{E}[Y]$ and if $\mathbb{E}[X] = \mathbb{E}[Y]$ then $V(X) \leq V(Y)$.*

Proof. That $\mathbb{E}[X] \geq \mathbb{E}[Y]$ if $X \succeq_2 Y$ follows immediately from the fact that the identity is non decreasing and concave. If $\mathbb{E}[X] = \mathbb{E}[Y]$, then $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for every u concave, and so, setting $u(x) = -x^2 + \mathbb{E}[X]$, one obtains $V(X) = \mathbb{E}[X^2 - \mathbb{E}[X]] \leq \mathbb{E}[Y^2 - \mathbb{E}[X]] = V[Y]$. □

Thus, for bilateral gamma returns, SSD implies higher expected gains and/or lower expected losses, and, for equal expected gains and losses, lower standard deviation of gains and/or losses. A partial converse of this statement is shown below, and is based on the following results.

Theorem 2.2.5. *Let X and Y be random variables with densities f and g. If the likelihood ratio $\frac{f}{g}$ is monotonically increasing, than $X \succeq_1 Y$. If the likelihood ratio is monotonically increasing on $(-\infty, x_0) \cup (x_1, \infty)$ and decreasing on (x_0, x_1), with $x_0 < x_1 \in \mathbb{R}$, then $X \succeq_2 Y$.*

Proof. See [2] and the references therein. □

Theorem 2.2.6. *Let X and Y be two gamma distributed random variable with scale and shape parameters (b, c) and (b', c') respectively. Then,*

(i) if $b = b'$, then $c > c'$ iff $X \succeq_2 Y$;

(ii) if $c = c'$, then $b > b'$ iff $X \succeq_2 Y$;

(ii) $\frac{c}{c'} \leq \max(1, \frac{b'}{b})$ with strict inequality at least when $\frac{b'}{b} = 1$ iff $X \succeq_2 Y$.

Proof. Based on showing that the assumptions of theorem 2.2.5 are satisfied. See [2]. □

Corollary 2.2.7. *Let X and Y be two gamma distributed random variables with scale and shape parameters (b, c) and (b', c') respectively. Then, $X \succeq_2 Y$ (resp. $-X \succeq_2 -Y$) if $\mathbb{E}[X] \geq \mathbb{E}[Y]$ and $V[X] \leq V[Y]$ (resp. $\mathbb{E}[X] \leq \mathbb{E}[Y]$ and $V[X] \leq V[Y]$).*

Proof. Suppose $\mathbb{E}[X] \geq \mathbb{E}[Y]$ and $V[X] \leq V[Y]$ with at least one strict inequality, i.e. $bc \geq b'c'$ and $b^2c \leq b'^2c'$ with at least one strict inequality. Then, $\frac{c}{c'} \geq \frac{b'}{b}$, and $\frac{b'}{b} = \frac{b'^2c'}{b^2c}\frac{bc}{b'c'} > 1$, so $X \succeq_2 Y$ by Theorem 2.2.6. The result for $-X$ and $-Y$ follows from adapting Theorem 2.2.6 to the case of the negative of gamma distributions. The case $\mathbb{E}[X] = \mathbb{E}[Y]$ and $V[X] = V[Y]$ is obvious. □

Corollary 2.2.8. *Let X^+, X^-, Y^+, Y^- be four gamma distributed random variable with scale and shape parameters (b_p, c_p), (b_n, c_n), (b'_p, c'_p) and (b'_n, c'_n) respectively. Then, $X := X^+ - X^-$ second order stochastically dominates $Y := Y^+ - Y^-$ if $\mathbb{E}[X^+] \geq \mathbb{E}[Y^+]$, $\mathbb{E}[X^-] \leq \mathbb{E}[Y^-]$, $V[X^+] \leq V[Y^+]$, and $V[X^-] \leq V[Y^-]$.*

Proof. If no inequality is strict, the result is obvious. Suppose there is exactly one strict inequality, e.g. $\mathbb{E}[X^+] > \mathbb{E}[Y^+]$. Then, by Corollary 2.2.7, $X^+ \succeq_2 Y^+$, and so, for all $t \in [0, \infty)$

$$\int_0^t F^+(s) - G^+(s)ds \leq 0,$$

where F^+ and G^+ denote the cumulative distribution function of X^+ and Y^+ respectively. Then

$$\int_0^t F(s) - G(s)ds = \int_0^\infty \int_0^t F^+(s-\xi) - G^+(s-\xi)dsdF^-(\xi) \leq 0,$$

where F^- is the (common) distribution of $-X^-$ and $-Y^-$, and the conclusion follows from Theorem 2.2.3. The other cases are similar, and the result follows by transitivity of SSD. □

Based on the last corollary, a positive variation in μ_p compensating a positive variation in any among σ_p, μ_n or σ_n is evidence of investors' risk seeking behaviors.

Note that μ_p is not a "reward" accessible to an investor holding the asset, typically defined, for a time horizon T, by $\mu(T) = \log(\mathbb{E}[S_0 e^{X_T}])$ and the variation

$$\lim_{T \downarrow 0} \frac{\mu(T)}{T} = \int_{\mathbb{R}} (e^x - 1)k(x)dx = (1 - b_p)^{-c_p}(1 + b_n)^{-c_n} \tag{2.2.2}$$

would better serves this purpose. Thus, we refer to μ_p as a "compensation" for the risks $(\sigma_p, \mu_n, \sigma_n)$.

2.2.2.1 Log-Returns and Kelly's Criterion

In the case log returns are assumed to be bilateral gamma variates, these results cannot hold anymore, since, for instance, an increase in σ_p and/or σ_n implies higher expected value of the return, and it cannot imply second order stochastic dominance. However, a traditional assumption in the financial and economics literature, justified by some evidence ([5]), is to assume that investors maximize log-returns. In our context, such an assumption implies that an asset is preferred to another one if and only if the expected log-return is higher. More generally, for asset allocation problems, logarithmic utility yields the best return in the long run, assuming the investor faces a long sequence of investment decisions ([60], [79], [24]), but for an investor with a short/medium term horizon, it will not capture aversion to short term high volatility ([93]), thus leading to consider a utility specification with a coefficient of relative risk aversion (CRRA) bounded below by 1.[3] It then follows from the results of this section and proposition 2.2.2.1 below that, for a reasonable utility specification such as $u(\log(\cdot))$, risks and their compensation are captured by $(\sigma_p, \mu_n, \sigma_n)$ and μ_p respectively even when log-returns belong to the bilateral gamma class.

Proposition 2.2.9. *A strictly increasing and concave function* $v \in C^2((0,\infty))$ *satisfies* $x\frac{v''(x)}{v'(x)} \geq 1$ *if and only if* $\exists\, u \in C^2(\mathbb{R})$ *strictly increasing and concave with* $v(x) = u(\log(x))\ \forall x \in (0,\infty)$.

Proof. Suppose such a u exists. Then, for all $x \in (0,\infty)$, $u''(\log(x)) \geq 0$ and $u'(\log(x)) < 0$

$$x\frac{v''(x)}{v'(x)} = -x\frac{\frac{d^2}{dx^2}u(\log(x))}{\frac{d}{dx}u(\log(x))} = 1 - \frac{u''(\log(x))}{u'(\log(x))} \geq 1.$$

If $x\frac{v''(x)}{v'(x)} \geq 1$, set $u(y) = v(e^y)$, so $u'(y) = v'(e^y)e^y > 0$, $u''(y) = v''(e^y)e^{2y} + v'(e^y)e^y \leq 0$. □

[3] In fact, several empirical studies provide evidence for this to be the case (see e.g. [45]).

2.3 The Acceptance Set

2.3.1 Learning the Boundaries

As mentioned in the introduction, not all quadruples $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ can be traded, or, in other words, there are structural limits to how high and/or low is the level of rewards that can be offered for given risks. In order to determine such limits, moments of gains and losses were estimated for 184 stocks (whose ticker is reported in appendix A) for the period 01/01/2008 to 31/12/2020 using one year of data for each estimate.[4] Assuming the boundaries are defined by functions $f_m, f_M : (\sigma_p, \mu_n, \sigma_n) \to \mu_p$, we find f_M and f_m by solving, respectively,

$$\min_{f\in\mathcal{F}}(1-\tau_M)\sum_i \left[\mu_p(i) - f_M(\sigma_p(i), \mu_n(i), \sigma_n(i))\right]^+ - \tau_M \sum_i \left[\mu_p(i) - f_M(\sigma_p(i), \mu_n(i), \sigma_n(i))\right]^-,$$

$$\min_{f\in\mathcal{F}}(1-\tau_m)\sum_i \left[\mu_p(i) - f_m(\sigma_p(i), \mu_n(i), \sigma_n(i))\right]^+ - \tau_m \sum_i \left[\mu_p(i) - f_m(\sigma_p(i), \mu_n(i), \sigma_n(i))\right]^-,$$

where $\mathcal{F}$ is a suitable class of functions which is here assumed to be the class of linear Gaussian process (GPR) regressors, $\tau_M = 0.95$ and $\tau_m = 0.05$. In our implementation of quantile GPR, the kernel hyperparameters were estimated using the standard loss function, while the regression coefficients are chosen to maximize the quantile loss function. Specifically, recall that GPR assumes

$$\mu_p = \alpha + h(\sigma_p, \mu_n\sigma_n)^T\beta + f(\sigma_p, \mu_n\sigma_n) + \varepsilon, \tag{2.3.1}$$

where ε is noise with variance σ_ε^2, h is the map to features space (here assumed to be the identity), and where any finite number collection $\{f(\sigma_p, \mu_n\sigma_n)\}$ is assumed to have Gaussian distri-

[4] Observations are results of likelihood optimization, so 1% of outliers were excluded.

bution with mean 0 and covariance function $\kappa((\sigma_p, \mu_n\sigma_n), (\sigma_p, \mu_n\sigma_n)')$. The prediction μ_p for $x = (\sigma_p, \mu_n, \sigma_n)$ given n observations $(\mu_p^i, \sigma_p^i, \mu_n^i, \sigma_n^i)$ is then given by (see [86])

$$\mu_p = \begin{bmatrix} \kappa(x^1, x) & \dots & \kappa(x^n, x) \end{bmatrix} \left(\begin{bmatrix} (\kappa(x^1, x^1) & \dots & \kappa(x^1, x^n) \\ & \ddots & \\ (\kappa(x^n, x^1) & \dots & \kappa(x^n, x^n) \end{bmatrix}_{i,j} + \sigma_\varepsilon^2 I \right)^{-1} \begin{bmatrix} \mu_p^1 \\ \vdots \\ \mu_p^n \end{bmatrix},$$

where we let $x^i = (\sigma_p^i, \mu_n^i, \sigma_n^i)$. Here we take κ to be the squared exponential kernel, with parameters estimated based on the standard loss function. The vector β and the intercept α are instead chosen by minimization of the quantile loss function.

The linear estimates obtained for f_m and f_M are

$$\begin{aligned} f_m(\sigma_p, \mu_n, \sigma_n) &= 0.0017 + 0.2029\sigma_p + 0.9951\mu_n - 0.3711\sigma_n, \\ f_M(\sigma_p, \mu_n, \sigma_n) &= 0.0017 + 0.2710\sigma_p + 1.0102\mu_n - 0.2311\sigma_n. \end{aligned} \tag{2.3.2}$$

Note, in particular, the negative relationship between σ_n and μ_p.

Similarly, for quantile GPR, $\frac{\partial f_m}{\partial \sigma_n}$ are always negative, while $\frac{\partial f_M}{\partial \sigma_n}$ are positive at all but two of 16 representative points (table 2.1).

Alternatively, f_m and f_M can be obtained via distorted least square ([72]), i.e. solving

$$\min_{f \in \mathcal{F}} \sum_i r_i^2 \left(\Psi(q_i) - \Psi\left(q_i - \tfrac{1}{n}\right) \right), \tag{2.3.3}$$

where, for every i, $r_i := \mu_p(i) - f(\sigma_p(i), \mu_n(i), \sigma_n(i))$ is the residual corresponding to the i-th observation, $\Psi : [0,1] \to [0,1]$ is a distribution function (called a distortion) concave for f_m and

$\frac{\partial f_M}{\partial \sigma_p}$	$\frac{\partial f_M}{\partial \mu_n}$	$\frac{\partial f_M}{\partial \sigma_n}$	$\frac{\partial f_m}{\partial \sigma_p}$	$\frac{\partial f_m}{\partial \mu_n}$	$\frac{\partial f_m}{\partial \sigma_n}$
0.2667	2.4704	0.7577	-0.0130	2.0042	-0.2421
0.8691	1.9402	-1.3539	1.1140	1.8974	-0.8361
1.5243	1.9553	-1.1134	1.4108	1.9274	-1.2346
1.0459	2.0254	-0.4887	0.5666	1.9927	-1.2635
1.0867	1.9956	-1.0836	0.8823	2.0199	-1.2220
0.4639	2.0065	-1.4194	0.6053	2.0648	-1.1568
1.3013	2.0509	-1.4681	1.2715	2.0128	-1.4149
0.9669	2.0019	-0.2462	0.4477	1.9760	-1.0806
1.4434	2.2522	0.3978	0.5052	2.0026	-0.5761
0.9710	1.9465	-0.9840	0.9472	1.8900	-0.8995
1.0653	1.9423	-1.4702	1.3075	1.9230	-0.9990
0.9307	1.9594	-0.5941	0.6044	1.9087	-1.0416
1.3390	2.0444	-1.8664	1.4529	2.0287	-1.5394
0.8652	1.9872	-1.3001	0.9281	2.0499	-1.0898
1.1957	2.0398	-0.9586	0.9027	1.9967	-1.3931
0.9283	1.9956	-0.0906	0.3913	1.9830	-0.9539

Table 2.1: Boundaries gradients (estimated via Quantile GPR), at 16 representative points.

convex for f_M, and q_i is the i-th quantile of the residual's empirical distribution.

The idea behind 2.3.3 is as follows. First, Ψ defines a distorted expectation $\mathcal{E}^{\Psi}[X]$ of a random variable X with distribution function F, as the Stjielties integral with respect to the distribution function $\Psi \circ F$:

$$\mathcal{E}^{\Psi}[X] := \int_{\mathbb{R}} x d\Psi(F(x)).$$

If Ψ is concave, lower values of X are weighted higher, thus implying $\mathcal{E}^{\Psi}[X] \leq \mathbb{E}[X]$, while the

opposite is true if Ψ is convex. Next, given observations $\{x_i\}_{i=0}^n$ of X, $\mathcal{E}^\Psi[X]$ is estimated by

$$\sum_{i=1}^{n} x_{(i)} \left[\Psi(F(x_{(i)})) - \Psi(F(x_{(i-1)}))\right]$$

where $\{x_i\}_{i=1,\dots,n}$ is the ordered sample. If F is unknown, this estimator can be replaced by

$$= \sum_{i=1}^{n} x_i \left[\Psi\left(q_i\right) - \Psi\left(q_i - \frac{1}{n}\right)\right],$$

so the loss function 2.3.3 corresponds to minimizing the estimated distorted expectation of the squared residual. By the tower property of (nonlinear) conditional expectations, the solution to problem 2.3.3 minimizes the distorted squared distance to the estimate of $\mathcal{E}^\Psi[\mu_p|\sigma_p, \mu_n, \sigma_n]$ and so, for an appropriate distortion, it can be thought as a lower/upper bound to the range of compensation μ_p given the risks $(\sigma_p, \mu_n, \sigma_n)$. Following [72], we set $\gamma = 0.75$ and define, for $u \in [0, 1]$,

$$\Psi(u) = 1 - \left(1 - u^{\frac{1}{1+\gamma}}\right)^{1+\gamma}. \tag{2.3.4}$$

The linear estimates obtained for f_m and f_M obtained via distorted least square are

$$\begin{aligned} f_m(\sigma_p, \mu_n, \sigma_n) &= 0.0024 + 0.1118\sigma_p + 0.9276\mu_n - 0.2596\sigma_n, \\ f_M(\sigma_p, \mu_n, \sigma_n) &= 0.0002 + 0.3604\sigma_p + 1.0196\mu_n - 0.1798\sigma_n, \end{aligned} \tag{2.3.5}$$

while the gradient at 16 representative points of the GPR estimates is shown in table 2.2.

Finally, we show in table 2.3 the percentages of observations represented by each of the 16 quantized points.

$\frac{\partial f_M}{\partial \sigma_p}$	$\frac{\partial f_M}{\partial \mu_n}$	$\frac{\partial f_M}{\partial \sigma_n}$	$\frac{\partial f_m}{\partial \sigma_p}$	$\frac{\partial f_m}{\partial \mu_n}$	$\frac{\partial f_m}{\partial \sigma_n}$
-0.0887	1.9864	0.6572	0.0790	2.0073	-0.1164
1.4815	1.9431	-0.3117	0.5231	1.8825	-1.8317
1.1663	1.9467	-1.8564	1.6447	1.9112	-0.6290
0.8911	2.0018	-0.7190	0.6395	1.9916	-1.1617
1.2837	1.9852	-0.6655	0.6932	2.0128	-1.5889
0.7569	1.9771	-0.9006	0.3673	2.1029	-1.5779
1.6313	2.0300	-0.8589	0.9376	2.0127	-2.0179
0.7577	1.9821	-0.5676	0.5676	1.9736	-0.8992
0.5482	2.0124	-0.3588	0.7854	2.0058	-0.0587
1.2923	1.9270	-0.3932	0.6114	1.8902	-1.5024
1.5750	1.9676	-0.7003	0.8115	1.8921	-1.7373
0.9369	1.9325	-0.5317	0.5394	1.9128	-1.1926
1.8380	2.0261	-0.9785	0.9812	2.0308	-2.3639
1.2369	1.9782	-0.6258	0.6211	2.0585	-1.6346
1.2424	2.0136	-0.8368	0.8070	1.9979	-1.5882
0.6792	1.9826	-0.4806	0.5497	1.9767	-0.7058

Table 2.2: Boundaries gradients via Distorted GPR at 16 representative points.

μ_p	0.0694	0.0208	0.0343	0.0167	0.0685	0.1428	0.0308	0.0119
%	0.70	0.76	3.53	10.49	1.79	0.72	5.66	14.99
μ_p	0.0467	0.0165	0.0260	0.0130	0.0453	0.1002	0.0225	0.0088
%	2.11	11.22	5.52	12.11	2.87	1.19	8.33	11.00

Table 2.3: Percentage of observations represented by quantized point μ_p.

2.3.2 Implied Boundaries for Performance Measures

Table 2.4, 2.5 and 2.6 show boundaries for μ_p, Sharpe ratio and acceptability index at the 16 quantized points.[5] We show only results obtained via quantile GPR, as the others are similar.

[5]The Sharpe ratio is defined as $\frac{\mu\sqrt{t}}{\sigma}$, with $\mu = \mu_p - \mu_n$, $\sigma^2 = \sigma_p^2 + \sigma_n^2$ and $t = 250$ business days. The acceptability index is defined in [70] as the maximal γ such that the distorted expectation $\mathcal{E}^{\Psi_\gamma}[X]$ is nonnegative (or nonpositive for short position).

Upper Boundary	Observation	Lower Boundary	Upper Boundary	Observation	Lower Boundary
0.0856	0.0694	0.0697	0.0536	0.0467	0.0469
0.0224	0.0208	0.0190	0.0184	0.0165	0.0143
0.0348	0.0343	0.0339	0.0269	0.0260	0.0252
0.0180	0.0167	0.0153	0.0147	0.0130	0.0112
0.0706	0.0685	0.0661	0.0473	0.0453	0.0428
0.1440	0.1428	0.1421	0.1024	0.1002	0.0986
0.0329	0.0308	0.0284	0.0243	0.0225	0.0204
0.0127	0.0119	0.0107	0.0092	0.0088	0.0081

Table 2.4: Compensation boundaries (estimated via Quantile GPR), at 16 representative points.

Upper Boundary	Observation	Lower Boundary	Upper Boundary	Observation	Lower Boundary
4.0960	-0.1253	-0.0500	1.7966	-0.2803	-0.2231
1.4038	0.3108	-0.8727	2.2708	0.6168	-1.1672
-0.0615	-0.2441	-0.4023	0.4234	-0.0286	-0.4702
3.2179	1.6361	-0.1370	2.7574	0.9385	-0.9590
2.7685	0.8818	-1.3389	2.1867	0.7031	-1.2120
1.2525	0.5067	0.0483	2.3867	0.6779	-0.5442
2.4903	0.7203	-1.2574	2.9818	1.1174	-0.9577
3.0642	1.8490	0.2862	2.7892	2.1576	0.9789

Table 2.5: Sharpe ratio boundaries (estimated via Quantile GPR), at 16 representative points.

As observed in [70], typically acceptability indexes based on the MINMAXVAR distortion for returns on stocks and indexes are less than 0.15, with median values of 0.04 and more than 5% of observations at 0. These findings are consistent with the boundaries for the acceptability index at the 16 representative points shown in table 2.6 for both short and long positions. Note also that the acceptability index tends to be higher for long positions, which is also to be expected.

Upper Boundary	Observation	Lower Boundary	Upper Boundary	Observation	Lower Boundary
0.151	-0.003	-0.003	0.065	-0.009	-0.009
0.051	0.011	-0.031	0.082	0.022	-0.041
-0.015	-0.008	-0.002	0.015	-0.001	-0.017
0.117	0.059	-0.006	0.100	0.034	-0.034
0.101	0.032	-0.049	0.079	0.026	-0.044
0.045	0.019	0.001	0.087	0.025	-0.020
0.090	0.026	0.046	0.109	0.040	-0.035
0.112	0.067	0.009	0.102	0.078	0.034

Table 2.6: Acceptability index boundaries (estimated via Quantile GPR), at 16 representative points. Negative signs represent acceptability indices of short positions and correspond to a negative Sharpe ratio.

2.3.3 Uncertainty Quantification

A dimensional analysis of the manifold defined by the acceptable bilateral gamma parameters yields a quantification of the feasible spectrum for the response μ_p to given risks $(\sigma_p, \mu_n, \sigma_)$. Furthermore, as quantile linear and GP regressions do not produce significantly different estimates for the boundaries f_m and f_M, one may wonder if such boundaries are indeed linear.[6] Our results, summarized in table 2.7, provide evidence that the variance of the noise in the response is minimal, as well as that f_m and f_M are linear.

	PCA	cumulative weight (in $\%$)	Diffusion Map	cumulative weight (in $\%$)
λ_1	2.7529	68.82	0.0113	70.27
λ_2	1.1778	98.27	0.0045	98.58
λ_3	0.0685	99.98	0.0002	99.64
λ_4	0.0009	100.0	0.0001	100.0

Table 2.7: Eigenvalues's weights for PCA and diffusion map on the quantized dataset.

[6]For the nonlinear embedding we utilize the Diffusion map algorithm, recently introduced in [22] and reviewed in the appendix. The hyperparameters (as defined in [22]) are set to $\epsilon = 3$, $\delta = 0.01$, $t = 5$, and $\alpha = 0$.

2.3.4 Statistical Significance

The results shown in table 2.7, suggest the existence of a linear three dimensional embedding of the acceptable bilateral gamma parameters. Estimating the embedding will then allow one to assess the statistical significance of the predictors $(\sigma_p, \mu_n, \sigma_n)$ and the accuracy of the boundaries f_m and f_M given by 2.3.2 and 2.3.5. In particular, we assume that

$$\mu_p = f(\sigma_p, \mu_n, \sigma_n) + \varepsilon, \tag{2.3.6}$$

where ε represents noise and f is affine linear. The coefficients, t-statistic and p-values for each explanatory variable are reported in table 2.8.

Predictor	Coefficient	t-statistics	p-value
Intercept	0.0019	53.8	< 0.0001
σ_p	0.1723	36.3	< 0.0001
μ_n	1.0029	1417.3	< 0.0001
σ_n	-0.2548	-50.7	< 0.0001

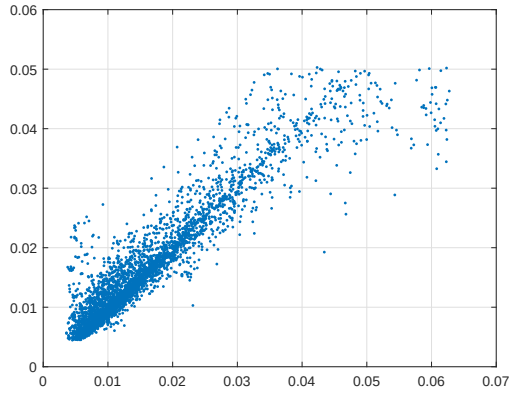
Table 2.8: Coefficients, t-statistics and p-values for the predictors in model 2.3.6.

A deeper analysis of the regressors (figure 2.1) shows a strong linear relationships between σ_p and σ_n, indicating the presence of multicollinearity among predictors. To establish that the same multicollinearity structure holds independently on the size of the sample, we randomly divided the original sample in 50 subsets, and computed quartiles of coefficients and p-values. The results, shown in table 2.9, confirm those of 2.8, and thus provide evidence of the robustness of our results.

Finally, we observe that the Kolmogorov-Smirnov test does not reject the hypothesis that the distribution of the error ε belong to the location-scale family, thus justifying the use of the z-test.

	Coefficient			p-value		
Percentile / Predictor	0.25	0.5	0.75	0.25	0.5	0.75
Intercept	0.0016	0.0019	0.0020	0.0000	0.0000	0.0000
σ_p	0.1481	0.2118	0.2430	0.0000	0.0000	0.0010
μ_n	1.0014	1.0037	1.0080	0.0000	0.0000	0.0000
σ_n	-0.3449	-0.2861	-0.2337	0.0000	0.0000	0.0000

Table 2.9: Percentiles among 50 subsamples of coefficients, t-statistics and p-values for the predictors in model 2.3.6.



(a)

Figure 2.1: Scatter plot of observed pairs (σ_p, σ_n).

2.3.5 A Simple Modification of a Lucas Tree Economy

To formally link the risk-seeking behaviors observed above with those of prospects theory consider the following modification of a Lucas tree economy ([65]). There are two periods, and each agent is endowed with a single risky asset with payoff S_i at the end of period i, $i = 0, 1$. Assume $S_1 = S_0 e^{G-L}$, where G and L are independent gamma variables, there is a risk-free asset in zero net supply with risk-free rate r_f, and that agents decide how much to borrow/lend at time 0. Denoting such amount by ℓ, consumption C_i at period i, $i = 0, 1$, is

$$C_0 = S_0 + \ell, \; C_1 = S_0 e^{G-L} - \ell e^{r_f}.$$

Setting $X = G - L$, $s_0 = \log(S_0)$, $0 < \rho, \beta < 1$, and following [59], preferences are defined by

$$U(C_0, C_1) = \frac{(\log(C_0))^{1-\rho}}{1-\rho} + e^{-\beta}\mathbb{E}\left[\frac{(\log(C_1))^{1-\rho}}{1-\rho}1_{\{s_0+X\geq 0\}} - \frac{(-\log(C_1))^{1-\rho}}{1-\rho}1_{\{s_0+X\leq 0\}}\right].$$

Thus, agents are risk averse if and only if the log-return $G - L$ is above the threshold s_0. Note that such behavior cannot be captured by preferences over terminal wealth.

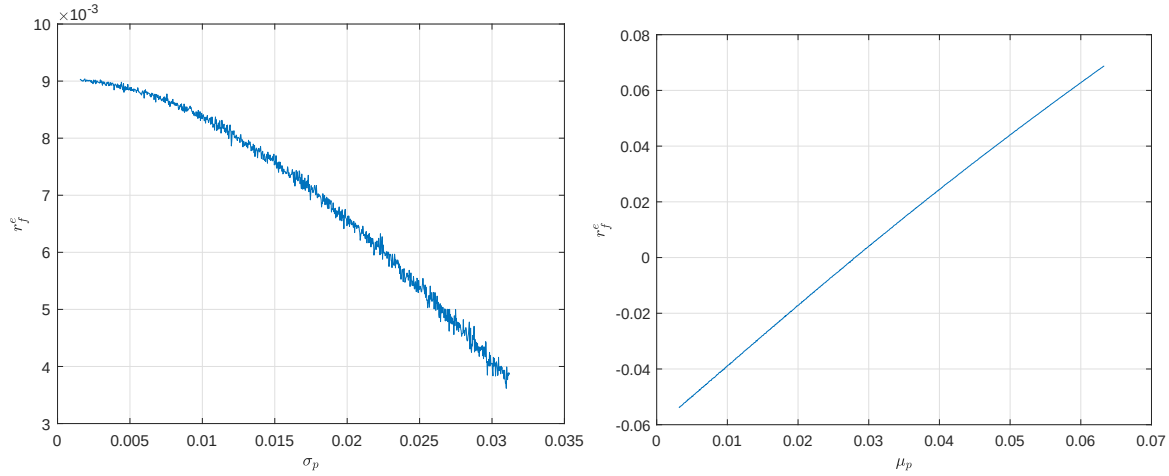
In equilibrium, $\ell = 0$, and so

$$s_0^{-\rho} - e^{r_f}e^{-\beta}\left(\mathbb{E}[(s_0+X)^{-\rho}e^{-X}1_{\{s_0+X\geq 0\}}] - \mathbb{E}[(-s_0-X))^{-\rho}e^{-X}1_{\{s_0+X\leq 0\}}]\right) = 0.$$

Then, the equilibrium interest rate r_f^e satisfies

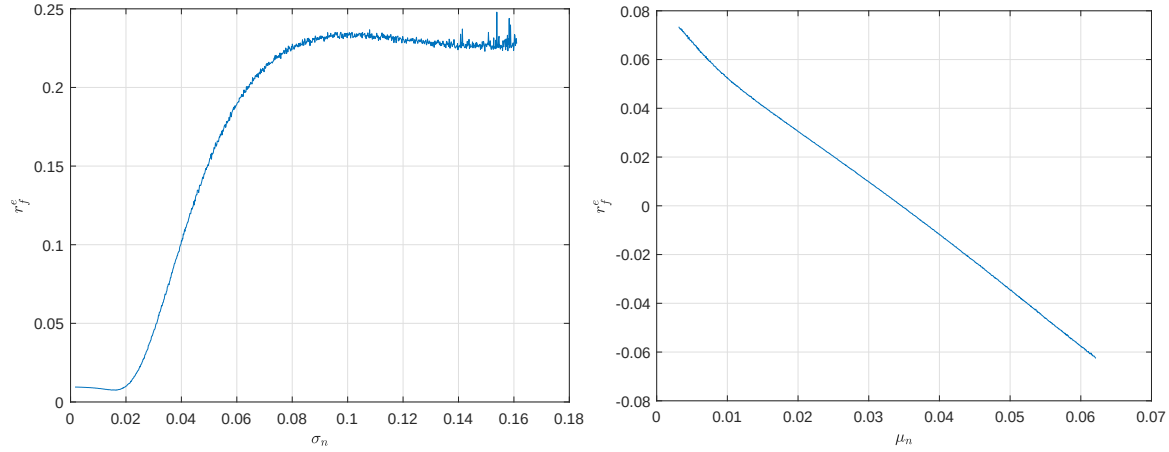
$$r_f^e = \beta - \rho\log(s_0) - \log\left(\mathbb{E}[(s_0+X)^{-\rho}e^{-X}1_{\{s_0+X\geq 0\}}] - \mathbb{E}[(-s_0-X))^{-\rho}e^{-X}1_{\{s_0+X\leq 0\}}]\right).$$

For a risk averse individual, higher risks correspond to lower equilibrium risk free rate, as lending becomes more attractive. Therefore, if the sign of $\partial r_f^e/\partial\sigma_n$ is negative, and that of $\partial r_f^e/\partial\sigma_p$ and $\partial r_f^e/\partial\mu_n$ are positive, the simple setting here described provides an explanation of our empirical findings. In general, it is possible to find values of $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ and of ρ such that this is indeed the case. For instance, setting $(\mu_p, \sigma_p, \mu_n, \sigma_n) = (0.03, 0.01, 0.03, 0.01)$, which are the average values observed in the dataset above described, and setting $\rho = 0.1$ and $\beta = 0.01$, the value of r_f^e computed via Montecarlo simulation as any of the variables $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ changes is shown in figures 2.2 and 2.3. As σ_p, σ_n increase the Montecarlo inegration estimate becomes less accurate, but the patterns in figures 2.2 and 2.3 confirm the behaviors above observed.



(a) (b)

Figure 2.2: Equilibrium rate as a function of σ_p (a) and of μ_p (b).



(a) (b)

Figure 2.3: Equilibrium rate as a function of σ_n (a) and of μ_n (b).

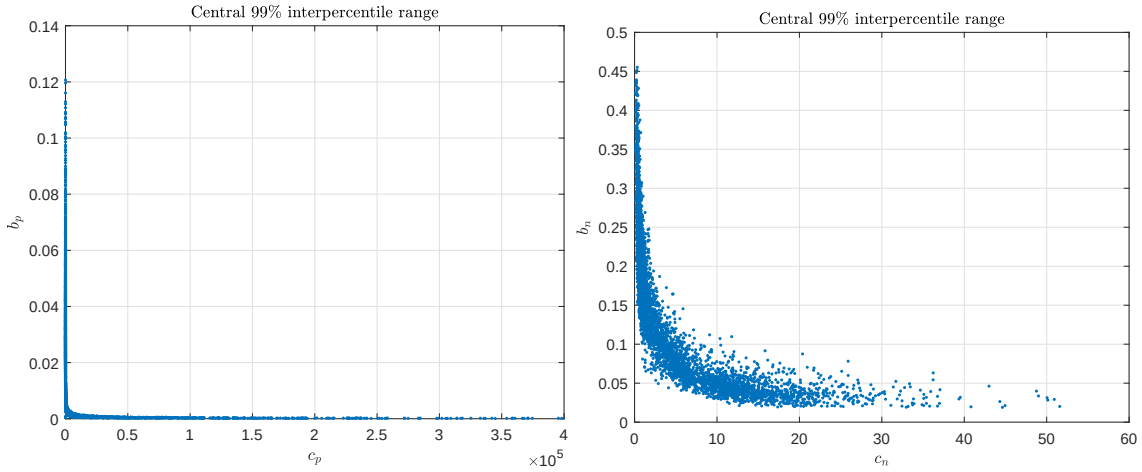
2.4 The Risks-Neutral Acceptance Set

In this section we analyze the "risk-neutral" acceptance set of quadruples (b_p, c_p, b_n, c_n) of BG parameters estimated to option prices. Risk neutral log returns are modeled as $\omega t + X_t$, where X_t is a BG process, $\omega := r + \log\left((1-b_p)^{c_p}(1+b_n)^{c_n}\right)$ and r is the risk free rate. We calibrated the 10 sector ETFs to the mid prices of options for four different maturities[7] and obtained a dataset of 4812 observations. Figure 2.4 shows pairs (b_p, c_p) and (b_n, c_n) excluding 1% of outliers. Boundaries for b_p and b_n in terms of (c_p, b_n, c_n) and (c_p, b_n, c_n) are estimated via quantile and distorted GPR. In both cases, the hyperparameters of the kernel matrix K are estimated using the standard MSE loss function, while $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ are computed so that $\beta + K\alpha$ minimizes the quantile loss function. For quantile GPR, the optimization was performed employing a quasi-Newton method, with the quantile loss function approximated by $S(x) = \tau x + \alpha \log(1 - e^{-x/\alpha})$ as in [102], with $\alpha = 10^{-4}$. [8] Estimates are visualized in figure 2.5 and reported in table 2.10 and 2.11.

Both quantile and distorted regression tend to break down in estimating the boundaries of b_p for large values of c_p, mostly because this parameter ranges between 0 and 10^5, with approximately 60% of the observations concentrated in the range $[0, 30]$ and the remaining ones being sparse (compare figure 2.4.A and 2.5.A) and corresponding to relatively small variations in (b_p, c_n, b_n). To avoid this issue, which does not occur when estimating boundaries of b_n (the range of observations for c_n is $[0, 50]$), the regression algorithms for the boundaries of b_p are only based on observations corresponding to $c_p < 30$.

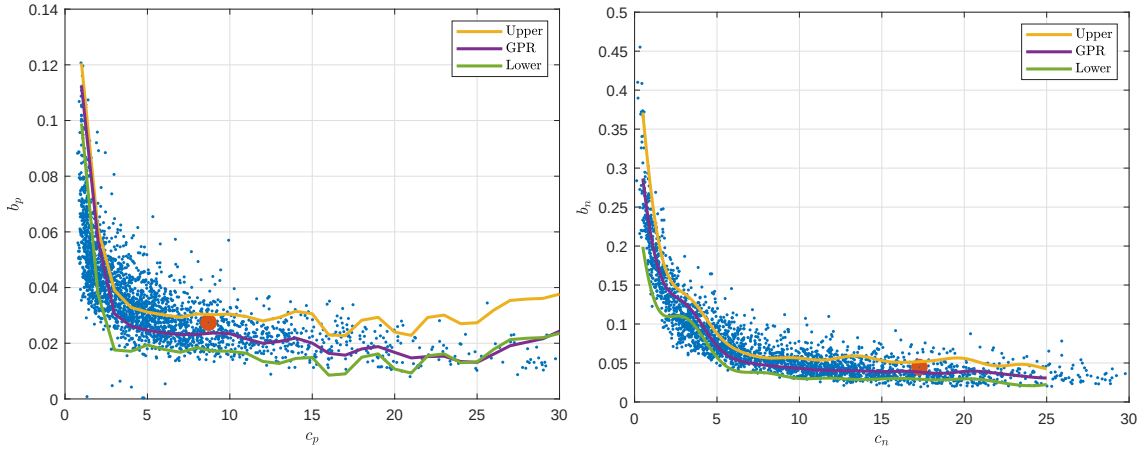
[7] Of all the traded maturities, the middle four were considered. Tickers considered are SPY, XLB, XLE, XLF, XLI, XLK, XLP, XLU, XLV, XLY. Calibration was performed every 10 days between 1/01/2015 through 31/12/2020.

[8] This methodology outperformed both transformation into a linear program and replacement of the constant basis in the GPR with the first 100 trigonometric basis or a feedforward neural network.



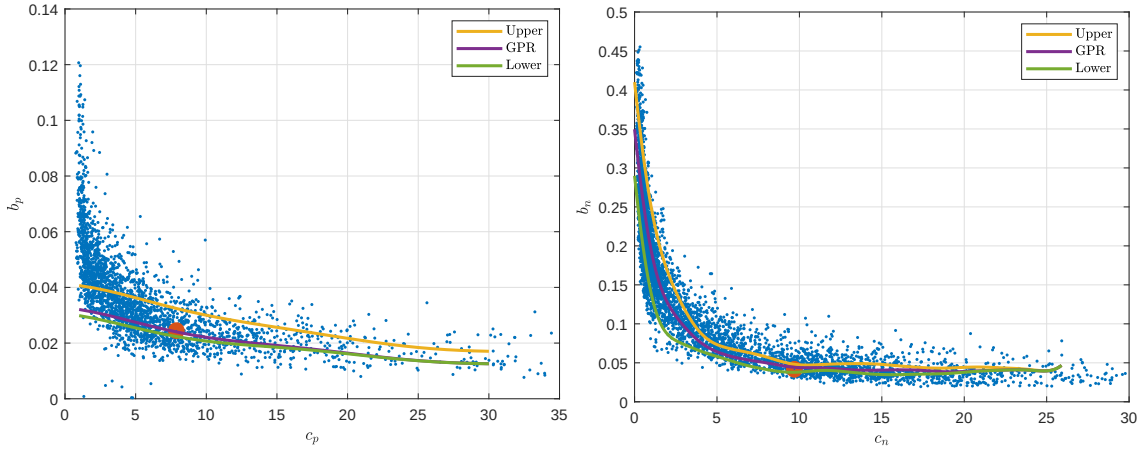
(a) (b)

Figure 2.4: Scatter plot of observed pairs (b_p, c_p) and (b_n, c_n) of risk neutral parameters.



(a) (b)

Figure 2.5: Boundaries around randomly selected point (in red) via quantile GPR with ($\tau = 0.05$).



(a) (b)

Figure 2.6: Boundaries around randomly selected point (in red) via distorted GPR ($\gamma = 0.75$).

Upper Boundary	Observation	Lower Boundary	Upper Boundary	Observation	Lower Boundary
0.0603	0.0451	0.0363	0.0304	0.0220	0.0136
0.0684	0.0557	0.0370	0.0327	0.0265	0.0155
0.0466	0.0382	0.0285	0.0307	0.0216	0.0165
0.0385	0.0320	0.0248	0.0317	0.0215	0.0205
0.0362	0.0296	0.0196	0.0393	0.0313	0.0244
0.0317	0.0243	0.0189	0.0282	0.0202	0.0145
0.0356	0.0276	0.0229	0.0263	0.0192	0.0125
0.0329	0.0248	0.0180	0.0291	0.0197	0.0152

Table 2.10: Boundaries for b_p via quantile GPR at 16 representative points (with $c_p < 30$).

Upper Boundary	Observation	Lower Boundary	Upper Boundary	Observation	Lower Boundary
0.0728	0.0593	0.0430	0.2394	0.1862	0.1266
0.0541	0.0346	0.0261	0.2422	0.1885	0.1284
0.2576	0.1992	0.1347	0.2378	0.1852	0.1268
0.2392	0.1857	0.1261	0.2198	0.1701	0.1193
0.2597	0.2012	0.1356	0.2175	0.1674	0.1194
0.2452	0.1906	0.1291	0.2259	0.1756	0.1215
0.2703	0.2089	0.1406	0.2113	0.1679	0.1214
0.2638	0.2041	0.1374	0.2302	0.1764	0.1280

Table 2.11: Boundaries for b_n via quantile GPR at 16 representative points.

2.4.1 Speed Uncertainty

Scale and shape parameters represent, respectively, limit and market orders. Typically, professional traders place limit orders based on stable patterns, so the scale parameters reflect the phase of the economic cycle, while the speed parameters are noisy responses to it.[9] However, "speed uncertainty" cannot be achieved by change of measure, but would require a G-bilateral

[9]Diffusion map showed that more than 95% of the dataset variance is explained by two eigenvectors.

gamma process (similar to a G-Brownian Motion as in [84]). This can be implemented via nonlinear Levy processes ([80]), with the ask price of a claim $C = f(X_T)$, where X is a bilateral gamma process, being the unique viscosity solutions of the equation

$$ru(t,x) + \sup_{c_p,c_n\in\Theta}\left\{\int_{\mathbb{R}\setminus\{0\}}[u(t,x+y)-u(t,x)]k(y)dy\right\} = u_t(t,x),$$

with initial condition $u(0,x) = f(x)$ and $\Theta \subset \mathbb{R}^2$ is compact. However, the empirical analysis in the next section is based on the large literature on the magnitude of the spread between upper and lower valuations arising from spectral risk measure, so this approach is left for future research.

2.4.2 An Equation for the Boundaries of Acceptable Risk Neutral Parameters

The boundaries found are naturally linked to acceptance sets implied by risk measures. In particular, given an asset's bid and ask price processes $\{B_t\}_{t\geq 0}$ and $\{A_t\}_{t\geq 0}$ on $(\Omega,\mathcal{F},\mathbb{P})$, absence of arbitrages implies the existence of a pair $(\mathbb{Q},S)$ such that $\mathbb{Q}\sim\mathbb{P}$, $B_t \leq S_t \leq A_t$ for every $t\geq 0$ and $\{e^{-rt}S_t\}_{t\geq 0}$ is a martingale under $\mathbb{Q}$.[10] One can then assume that

$$B_0 = \inf_{Q\in\mathcal{M}}\mathbb{E}^Q[S_0e^{-r+\omega+X_1}],\ A_0 = \sup_{Q\in\mathcal{M}}\mathbb{E}^Q[S_0e^{-r+\omega+X_1}],$$

where $\mathcal{M}$ is a collection of probability measures that are equivalent to the statistical measure $\mathbb{P}$. Such collection is a financial primitive of the economy that depends on regulator's requirements for financial stability as well as trading, costs and incentives of market operators, and a risk Z is deemed acceptable if $\mathbb{E}^Q[Z]\geq 0, \forall Q\in\mathcal{M}$. If $\mathcal{M}$ is defined by the spectral risk measure arising

[10] For the existence of $\mathbb{Q}$ and the associated processes $\{S_t\}_{t\geq 0}$ see [56] and [94].

from a distortion Ψ, bid and ask prices are integrals of distorted probabilities of tail events (see [72]), and the higher their distortion the higher size of the set $\mathcal{M}$ and the bid-ask spread.

If for given $(\hat{c}_p, \hat{b}_n, \hat{c}_n)$, b_p lies in the interval $[\underline{b}_p, \overline{b}_p]$, it is natural to assume

$$\{\mathbb{Q}_{b_p}\}_{b_p \in [\underline{b}_p, \overline{b}_p]} \subset \mathcal{M}. \tag{2.4.1}$$

where, under $\mathbb{Q}_{b_p}$, $\{X_t\}_{t \geq 0}$ is a BG process with parameters $(b_p, \hat{c}_p, \hat{b}_n, \hat{c}_n)$.[11] Since

$$\inf_{b_p \in [\underline{b}_p, \overline{b}_p]} \mathbb{E}^{\mathbb{Q}_{b_p}}[e^{-r+\omega+X_1}] = \frac{(1-\hat{b}_p)^{\hat{c}_p}}{(1-\underline{b}_p)^{\hat{c}_p}}, \quad \sup_{b_p \in [\underline{b}_p, \overline{b}_p]} \mathbb{E}^{\mathbb{Q}_{b_p}}[e^{-r+\omega+X_1}] = \frac{(1-\hat{b}_p)^{\hat{c}_p}}{(1-\overline{b}_p)^{\hat{c}_p}}, \tag{2.4.2}$$

where $\omega = r + \log((1-\hat{b}_p)^{\hat{c}_p}(1+\hat{b}_n)^{\hat{c}_n})$, 2.4.1 implies

$$\frac{B_0}{S_0} \leq \frac{(1-\hat{b}_p)^{\hat{c}_p}}{(1-\underline{b}_p)^{\hat{c}_p}} \leq \frac{(1-\hat{b}_p)^{\hat{c}_p}}{(1-\overline{b}_p)^{\hat{c}_p}} \leq \frac{A_0}{S_0}. \tag{2.4.3}$$

Further pushing 2.4.1 to be satisfied with an equality, one obtains:

$$\frac{B_0}{A_0} = \frac{(1-\overline{b}_p)^{\hat{c}_p}}{(1-\underline{b}_p)^{\hat{c}_p}} \tag{2.4.4}$$

Note that, in 2.4.4, B_0 and A_0 are functions of $\hat{c}_p, \hat{b}_n, \hat{c}_n$. In other words, the parameters $\hat{c}_p, \hat{b}_n, \hat{c}_n$ are measures of economic activity and thus, together with the structural limits $\underline{b}_p, \overline{b}_n$, determine bid and ask prices. Similarly, if $\hat{c}_p$ and $\hat{c}_n$ determine boundaries for b_p and b_n,

$$\frac{B_0(\hat{c}_p, \hat{c}_n)}{A_0(\hat{c}_p, \hat{c}_n)} = \frac{(1-\overline{b}_p)^{\hat{c}_p}}{(1-\underline{b}_p)^{\hat{c}_p}} \frac{(1+\underline{b}_n)^{\hat{c}_n}}{(1+\overline{b}_n)^{\hat{c}_n}}. \tag{2.4.5}$$

[11] By proposition 6.1 in [61], such a measure exists and is equivalent to $\mathbb{Q}$.

2.4.3 Empirical Verifications

Typically, equations 2.4.4 and/or 2.4.5 are not satisfied, at least with respect to daily closing bid-ask ratios. However, since large orders are executed over several days, one can consider other distorted valuations, such as 5-days high/low prices. In general, one can compare the size of $\mathcal{M}$ required for 2.4.1 to hold with typically observed acceptability indexes. To do so, we let ν_{b_p} denote the BG Levy measure with parameters $(b_p, \hat{c}_p, \hat{b}_n, \hat{c}_n)$, and replace 2.4.1 with

$$\{\nu_{b_p}\}_{b_p \in [\underline{b}_p, \bar{b}_p]} \subset \mathcal{N}. \tag{2.4.6}$$

The collection $\mathcal{N}$ is such that distorted rewards are defined by

$$\begin{aligned}
\underline{\mu} &= \omega - \int_0^\infty G^+(\nu(e^x - 1 < -a))da + \int_0^\infty (G^-(\nu(e^x - 1 > a))da, \\
\overline{\mu} &= \omega - \int_0^\infty G^-(\nu(e^x - 1 < -a))da + \int_0^\infty (G^+(\nu(e^x - 1 > a))da
\end{aligned}$$

where ν is the Levy measure of X under $\mathbb{Q}$ and G^+ and G^- are (see [37])

$$G^+(x) = x + \frac{1}{c}(1 - e^{-cx})^{1/(1+\gamma)},\ G^-(x) = x - \frac{1}{c}(1 - e^{-cx}).$$

Then, as proved in [72], $\tilde{\nu} \in \mathcal{N}$ if and only if $\frac{d\tilde{\nu}}{d\nu}$ satisfies

$$\begin{aligned}
S(\lambda) :&= \int_{\mathbb{R}} \left(\frac{d\tilde{\nu}}{d\nu} - \lambda\right)^+ d\nu(x) \leq \Phi(\lambda),\ \lambda > 1, \\
\tilde{S}(\lambda) :&= \int_{\mathbb{R}} \left(\lambda - \frac{d\tilde{\nu}}{d\nu}\right)^+ d\nu(x) \leq -\tilde{\Phi}(\lambda),\ 0 \leq \lambda < 1,
\end{aligned} \tag{2.4.7}$$

where Φ and $\tilde{\Phi}$ are Fenchel conjugates of G^+ and G^- respectively, and are given by

$$\Phi(\lambda) := \frac{1}{c}\left[-(1-\lambda)\log(u(\lambda)) + (1-u(\lambda))^{1/(1+\gamma)}\right],$$
$$-\tilde{\Phi}(\lambda) := \frac{1}{c}[\lambda + (1-\lambda)\log(1-\lambda)],$$

with $u : (1,\infty) \to (0,1)$ defined as the unique solution of

$$\frac{u}{(1-u)^{\gamma/(1+\gamma)}} = (\lambda - 1)(1+\gamma).$$

Next, we find requirements on c, γ for 2.4.6 to hold. Note that, for $\tilde{\nu} = \nu_{\bar{b}_p}$,

$$S(\lambda) = \int_{L(\lambda)}^{\infty} \frac{c_p}{x}\left(e^{-x/\bar{b}_p} - \lambda e^{-x/\hat{b}_p}\right)dx \; = c_p\left(Ei\left(\frac{L(\lambda)}{\bar{b}_p}\right) - \lambda Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right)\right)$$

and, similarly, for $\tilde{\nu} = \nu_{\underline{b}_p}$,

$$\tilde{S}(\lambda) = \int_{\tilde{L}(\lambda)}^{\infty} \frac{c_p}{x}\left(\lambda e^{-x/\hat{b}_p} - e^{-x/\underline{b}_p}\right)dx \; = c_p\left(\lambda Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right) - Ei\left(\frac{L(\lambda)}{\underline{b}_p}\right)\right),$$

where Ei is the exponential integral function and

$$L(\lambda) = \frac{\log(\lambda)\bar{b}_p\hat{b}_p}{\bar{b}_p - \hat{b}_p},\; \tilde{L}(\lambda) = -\frac{\log(\lambda)\hat{b}_p\underline{b}_p}{\hat{b}_p - \underline{b}_p}.$$

Lemma 2.4.1. *Suppose $\underline{b}_p > 0.55\hat{b}_p$. Then, $\nu_{b_p} \in \mathcal{N}$ holds for every $b_p \in [\underline{b}_p, \hat{b}_p]$ if and only if*

$$c \le \lim_{\lambda \to 1^-} \frac{1}{\tilde{S}(\lambda)} \tag{2.4.8}$$

Proof. Note that for every $0 < \lambda < 1$

$$
\begin{aligned}
-\tilde{\Phi}'(\lambda) - \tilde{S}'(\lambda) &= \frac{-\log(1-\lambda)}{c} - c_p \left(Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right) - \lambda e^{-L(\lambda)/\hat{b}_p} \frac{L'(\lambda)}{L(\lambda)} + e^{-L(\lambda)/\underline{b}_p} \frac{L'(\lambda)}{L(\lambda)} \right) \\
&= \frac{-\log(1-\lambda)}{c} - c_p Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right),
\end{aligned}
$$

so that a stationary point ℓ of $-\tilde{\Phi} - \tilde{S}$ must satisfy $cc_p = -\log(1-\ell)/Ei\left(\frac{L(\ell)}{\hat{b}_p}\right)$. Since

$$
\begin{aligned}
\frac{d}{d\lambda} \frac{-\log(1-\lambda)}{Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right)} &= \frac{\frac{Ei\left(\frac{L(\ell)}{\hat{b}_p}\right)}{1-\lambda} - \frac{e^{-L(\lambda)/\hat{b}_p}\log(1-\lambda)}{\lambda\log(\lambda)}}{Ei\left(\frac{L(\ell)}{\hat{b}_p}\right)^2} \\
&\leq e^{-L(\lambda)/\hat{b}_p} \frac{\frac{\log\left(1-\frac{\underline{b}_p}{(\hat{b}_p-\underline{b}_p)\log(\lambda)}\right)}{1-\lambda} - \frac{\log(1-\lambda)}{\lambda\log(\lambda)}}{Ei\left(\frac{L(\ell)}{\hat{b}_p}\right)^2},
\end{aligned}
$$

the function $-\tilde{\Phi} - \tilde{S}$ admits at most one stationary point in $(0,1)$ if

$$
\frac{\log\left(1-\frac{\underline{b}_p}{(\hat{b}_p-\underline{b}_p)\log(\lambda)}\right)}{1-\lambda} - \frac{\log(1-\lambda)}{\lambda\log(\lambda)} < 0, \Leftrightarrow \log(\lambda)\left[1-(1-\lambda)^{\frac{1-\lambda}{\lambda\log(\lambda)}}\right] < \frac{\underline{b}_p}{(\hat{b}_p-\underline{b}_p)}. \quad (2.4.9)
$$

Since, for $0 < \lambda < 1$, $\log(\lambda)\left[1-(1-\lambda)^{\frac{1-\lambda}{\lambda\log(\lambda)}}\right] < 1.2$, condition 2.4.9 holds if $0.55\hat{b}_p < \underline{b}_p$. Since

$$
\begin{aligned}
\lim_{\lambda\to 0^+} -\tilde{\Phi}(\lambda) = \lim_{\lambda\to 0^+} \tilde{S}(\lambda) = 0, \ \lim_{\lambda\to 0^+} \frac{-\tilde{\Phi}(\lambda)}{\tilde{S}(\lambda)} = \infty \\
\lim_{\lambda\to 1^-} -\tilde{\Phi}'(\lambda) = \lim_{\lambda\to 1^-} \tilde{S}'(\lambda) = \infty, \ \lim_{\lambda\to 1^-} \frac{-\tilde{\Phi}'(\lambda)}{\tilde{S}'(\lambda)} = 0,
\end{aligned}
$$

it must be the case that if $\lim_{\lambda\to 1^-} -\tilde{\Phi}(\lambda) \geq \lim_{\lambda\to 1^-} \tilde{S}(\lambda)$, which is 2.4.8, then $-\tilde{\Phi} - \tilde{S}$ admits

a positive maximum in $(0, \lambda)$. Therefore, if $0.55\hat{b}_p < \underline{b}_p$ and 2.4.8 are satisfied, $-\tilde{\Phi} - \tilde{S}$ must be nonnegative on $(0, 1)$, since it would otherwise admit two stationary points. □

We note that $0.55\hat{b}_p < \underline{b}_p$ for all the 16 representative points. The next two lemmas identify necessary and sufficient conditions for the case $b_p \geq \hat{b}_p$.

Lemma 2.4.2. *Suppose $b_p \in [\hat{b}_p, \overline{b}_p]$. Then, it is necessary for $\nu_{b_p} \in \mathcal{N}$ to hold that*

$$\gamma > \frac{\overline{b}_p - \hat{b}_p}{\hat{b}_p} := \tilde{\gamma} \tag{2.4.10}$$

$$c \leq \frac{1}{c_p}. \tag{2.4.11}$$

Proof. Note that

$$\lim_{\lambda \to \infty} \Phi(\lambda) = \lim_{\lambda \to \infty} S(\lambda) = 0,$$

so, by l'Hopital's theorem, $\Phi(\lambda) \geq S(\lambda)$ implies

$$\frac{S''(\lambda)}{\Phi''(\lambda)} = O(1)$$

as $\lambda \to \infty$. Furthermore, using the implicit definition of u,

$$\Phi'(\lambda) = \frac{1}{c}\log(u(\lambda)),\ \Phi''(\lambda) = \frac{u'(\lambda)}{cu(\lambda)},$$
$$S'(\lambda) = -c_p Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right),\ S''(\lambda) = c_p \frac{1}{\lambda^{\overline{b}_p/(\overline{b}_p - \hat{b}_p)+1}\log(\lambda)},$$

and, using implicit differentiation,

$$u'(\lambda) = \left(\frac{u^{2+1/\gamma}}{(1-u+\gamma)(1+\gamma)^{1/\gamma}}\right)\left(\frac{1}{(\lambda-1)}\right)^{2+1/\gamma} \sim \left(\frac{1}{(\gamma)(1+\gamma)^{1/\gamma}}\right)\left(\frac{1}{\lambda}\right)^{2+1/\gamma}.$$

We thus need

$$2+\frac{1}{\gamma} < \frac{\overline{b}_p}{\overline{b}_p - \hat{b}_p} + 1 \Rightarrow \gamma > \tilde{\gamma}.$$

For 2.4.11 simply note that

$$\lim_{\lambda \to 1^+} \frac{\Phi(\lambda)}{S(\lambda)} = \frac{1}{cc_p}.$$

□

Lemma 2.4.3. *There is a function $\kappa_p : (\tilde{\gamma}, \infty) \to (0, \tilde{c}]$, where*

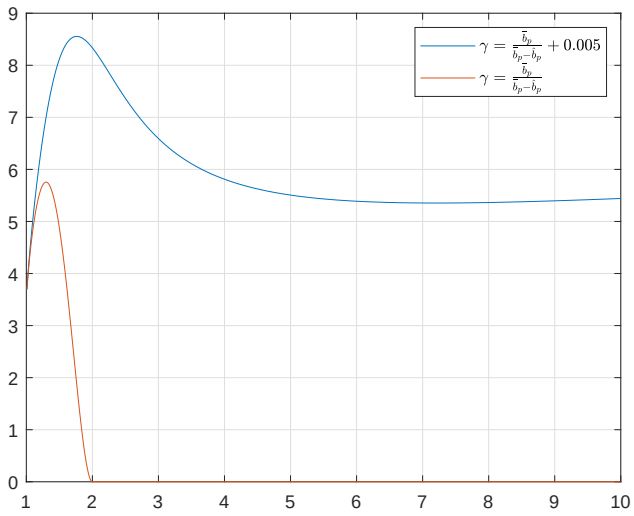
$$\tilde{c} := \min\left\{\lim_{\lambda \to 1^-} \frac{1}{\tilde{S}(\lambda)}, \frac{1}{c_p}\right\},$$

such that, for every γ that satisfies 2.4.10, $\nu_{b_p} \in \mathcal{N}$ if $c < \kappa_p(\gamma)$ and $\nu_{b_p} \notin \mathcal{N}$ if $c > \kappa_p(\gamma)$.

Proof. Fix $\gamma > \tilde{\gamma}$. As in the previous lemma, and since u does not depend on c, there is $\ell > 1$ independent of c such that for every $\lambda > \ell$,

$$\frac{-(1-\lambda)\log(u(\lambda)) + (1-u(\lambda))^{1/(1+\gamma)}}{Ei\left(\frac{L(\lambda)}{\overline{b}_p}\right) - \lambda Ei\left(\frac{L(\lambda)}{\hat{b}_p}\right)} \geq 1.$$

Hence, if $c < \tilde{c}$, $\Phi(\lambda) > S(\lambda)$ for every $\lambda > \ell$. Since $\Phi - S$ is continuous and decreasing in c for



(a)

Figure 2.7: The function $\Phi(\lambda)/S(\lambda)$ for the first of the 16 representative points, with $c = \tilde{c}$ and assuming $\gamma = \tilde{\gamma} + 0.005$ (blue) and $\gamma = \tilde{\gamma}$ (red).

every $\lambda \in [1, \ell]$, with $\lim_{c\to 0} \Phi - S = \infty$, $\lim_{c\to\infty} \Phi - S = -S < 0$, there is a bounded set of values $c > 0$ such that $\Phi(\lambda) - S(\lambda) \geq 0$ for every $\lambda \in [1, \ell]$. Letting $\bar{c}$ denote the supremum of such values, one can set $\kappa(\gamma) = \min\{\bar{c}, \tilde{c}\}$. □

The function κ_p typically grows very fast, so that $\kappa_p(\gamma) = \tilde{c}$ for values of γ that are slightly larger than $\tilde{\gamma}$. For instance, figure 2.7 depicts the function $\Phi(\lambda/S(\lambda)$ for $\gamma = \tilde{\gamma}$ and $\gamma' = \tilde{\gamma} + 0.005$, and for $c = \tilde{c}$ and the bilateral gamma parameters set as in the most representative of the quantized points. In fact, this is the case for each of the 16 quantized points, as shown in table 2.12.

c	γ	$\tilde{\gamma}$	c	γ	$\tilde{\gamma}$	c	γ	$\tilde{\gamma}$	c	γ	$\tilde{\gamma}$
0.462	0.357	0.347	0.212	0.259	0.219	0.075	0.400	0.380	0.138	0.288	0.258
0.666	0.263	0.233	0.130	0.336	0.296	0.169	0.256	0.216	0.061	0.397	0.377
0.262	0.252	0.232	0.126	0.327	0.287	0.069	0.484	0.424	0.039	0.387	0.357
0.219	0.209	0.199	0.096	0.358	0.328	0.056	0.529	0.469	0.040	0.527	0.467

Table 2.12: Triples $(\tilde{c}, \gamma, \tilde{\gamma})$ where γ is the minimal value ensuring 2.4.6 with $c = \tilde{c}$.

Remark 2.4.4. *Recalling that γ is similar to the acceptability index for probability distortions,*

while $\frac{10}{c}$ *roughly corresponds to the maximum distorted frequencies (so higher* c *corresponds to smaller* $\mathcal{N}$*), we note that the values reported in 2.12 have the same magnitude and are consistent in general with those typically seen in the literature (see for instance [37], [39] and [72]). We observe, in particular, that*

i. the three most representative points (top left corner of table 2.12), are consistent with the pair $(0.25, 0.25)$ *used in [72] to estimate capital requirements (chapter 15.5.2), and that the relatively high values of* γ *is compensated by high values of* c*;*

ii. for each triple, $\tilde{c}$ *is higher, and in the less frequent cases close to, the value* 0.01*, which, as shown in [39], generates higher returns compared to* $c = 1$ *and* $c = 0.25$ *for a portfolio constructed by maximization of the lower valuation.*

2.5 Conclusions

For an asset with (log) returns in the bilateral gamma class, a justification is provided, based on expected utility theory, that risks from holding the asset can be decomposed into a three dimensional vector of expected losses, variance of gains and variance of losses, while compensation for the risks is given by expected gains. Evidence is then provided that moments of bilateral gamma returns lie on a manifold with boundaries, and such boundaries are estimated via quantile and distorted linear and nonlinear regressions. It is observed that they imply a positive relationship between expected gains and variance of gains/expected losses, but a negative one between compensation and variance of losses thus implying market's operators being risk seekers in pure loss prospects. The claim that such finding are compatible with the experimental evidence that constitute prospects theory is then justified through a simple modification of Lucas Tree model. The

analysis is corroborated by performing a similar one to the case of risk neutral parameters, assuming a separate drift to satisfy the martingale condition. An inverse relationship between shape and scale parameters of loss and gain process is observed and a theoretical boundary for scale parameters, in line with certain empirical observations, is described based on the theory of Conic finance. Finally, we observed that our estimates of the boundaries are generally larger than those implied by regulatory capital requirements.

2.6 Appendix A: Unsupervised Manifold Learning

Because of its linearity, one can expect PCA to overestimate the dimension of a dataset that lives on a nonlinear submanifold of $\mathbb{R}^D$.[12] Among the existent nonlinear embeddings (e.g. [96], [90], [52], [98]), we chose to use the diffusion map, particularly for its known robustness with respect to noise perturbation, i.e. the case in which X is a realization of a random variable $\eta + \nu$, where $\mathbb{E}[\nu] = 0$. The construction of the diffusion mapping is as follows:

(i) define the diffusion kernel matrix $K = \{a_{j,k}\}_{j,k=1}^{n} \in \mathbb{R}^{n\times n}$ by $a_{j,k} := e^{-\|x_j - x_k\|^2/\varepsilon}$ for every $j, k = 1, ..., n$, where $\|\cdot\|$ is any norm;

(ii) convert the kernel matrix K into a stochastic matrix $P = \{p_{j,k}\}_{j,k=1}^{n}$ by setting $P = Q^{-1}K$, where Q is a diagonal matrix with elements $\{q_{j,j}\}_{j=1}^{n}$ on the diagonal given by $q_{j,j} = \sum_{k=1}^{n} a_{j,k}$, $j = 1, ..., n$, and interpret $p_{j,k}$ as the one step transition probability from x_j to x_k of a Markov chain on the graph X;

(iii) note that the diagonal elements $q = \{q_{j,j}\}_{j=1}^{n}$ of Q define an invariant measure for the

[12] If $\mathcal{M}$ is not a linear submanifold of $\mathbb{R}^D$, multidimensional scaling, not presented here, generalizes PCA by allowing for distances among data points to be the geodesic distances on the manifold. Such algorithm assumes, however, knowledge of geodesic distances, which in turn presumes knowledge of the manifold itself.

Markov chain defined by P; define an invariant probability distribution by setting $\pi = \frac{q}{\sum_{j=1}^{n} q_{j,j}}$;

(iv) for every $t \in \mathbb{N}$, let $p_{j,k}^t$ denote the t-step transition probability from j to k, i.e. $P^t = \{p_{j,k}^t\}_{j,k=1}^n$, and define the discrete family of diffusion distance

$$D_t(x_j, x_k)^2 = \sum_{m=1}^{n} \frac{1}{\pi_m} \|p_{j,m}^t - p_{k,m}^t\|^2;$$

(v) the family of diffusion maps $\Psi_t : \mathbb{R}^D \to \mathbb{R}^{n-1}$ is defined so that $D_t(x_j, x_k) = \|\Psi_t(x_j) - \Psi_t(x_k)\|$; dimensionality reduction is achieved by keeping the first d components of Ψ_t.

In particular, it is possible to show that

$$D_t(x_j, x_k)^2 = \sum_{m=1}^{n} \lambda_m^{2t} |r_m(j) - r_m(k)|^2, \tag{2.6.1}$$

where, given the svd decomposition of P, $P = R\Lambda L$, with ordered eigenvalues $1 = \lambda_0 \geq |\lambda_1| \geq ... \geq |\lambda_{n-1}|$, $r_m(j)$ denotes the j-th entry of the m-th column r_m of the matrix R. Therefore, the diffusion map for given $t \in \mathbb{N}$ is defined by $\Psi_t(x_j) = \begin{bmatrix} \lambda_1^t r_1(j) & \dots & \lambda_{n-1}^t r_{n-1}(j) \end{bmatrix}^T$. Note that the first eigenvector r_0 is removed as the diffusion distances do not change by including it.

Remark. Given an accuracy parameter $\delta \in (0, 1)$, and setting the dimension of the embedding to

$$s(\delta, t) = \max\{m \in \mathbb{N} : |\lambda_m|^t > \delta |\lambda_1|^t\},$$

then, up to relative precision δ,

$$D_t(x_j, x_k)^2 = \sum_{m=1}^{s(\delta,t)} \lambda_m^{2t} |r_m(j) - r_m(k)|^2,$$

which justifies the embedding. In particular, for a d-dimensional embedding, relative precision δ is reached for

$$\left(\frac{|\lambda_1|}{|\lambda_d|}\right)^t \leq \frac{1}{\delta}, \Rightarrow t = \left\lceil \frac{\log(\delta)}{\log(|\lambda_d|) - \log(|\lambda_1|)} \right\rceil \tag{2.6.2}$$

Limit as $\varepsilon \to 0$ and anisotropic densities.

The diffusion map construction is given in [22] in a continuous setting. Specifically, given a density q on $\mathcal{M}$ and a map $k : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ symmetric and nonnegative, the function

$$p(x, y) = \frac{k(x, y)}{\int_{\mathcal{M}} k(x, y) q(y) dy}$$

defines a Markov transition function P by

$$P(x, C) = \int_C p(x, y) q(y) dy$$

for every $x \in \mathcal{M}$ and $C \subset \mathcal{M}$ Borel measurable.[13] Its n-th transition probability is

$$P^n(x, C) = \int_C \int_{\mathcal{M}} \dots \int_{\mathcal{M}} p(x, y_1) \dots p(y_{n-2}, y_{n-1}) q(y_1) dy_1 \dots q(y_{n-1}) dy_{n-1},$$

[13] Assuming $\mathcal{M}$ is, e.g., compact, integration over measurable sets can be defined via Riesz-Markov Theorem.

and its n-th power Markov operator P^n is given for f bounded measurable by

$$P^n(f)(x) = \int_{\mathcal{M}} f(y) P^n(x, dy). \tag{2.6.3}$$

The summation formulas in the diffusion map algorithm above presented are the Montecarlo integration formulas of the corresponding quantities in the continuous time setting. The advantage of considering such setting is that a relationship between the Laplace Beltrami operator and the graph Laplacian results from it. In addition, by extending the eigenvectors r_i to eigenfunctions of P, one can extend the embedding from the dataset to the whole manifold $\mathcal{M}$, i.e. its parameterization.

Theorem 2.6.1. *Suppose $\mathcal{M}$ is a compact smooth submanifold of $\mathbb{R}^D$, and q a density on $\mathcal{M}$. Let $\alpha \in \mathbb{R}$, $k_\varepsilon(x, y) = h\left(\frac{\|x-y\|^2}{\varepsilon}\right)$ a rotational invariant kernel (i.e. symmetric and nonnegative). Let*

$$q_\varepsilon(x) = \int_{\mathcal{M}} k_\varepsilon(x, y) q(y) dy,\ k_\varepsilon^{(\alpha)}(x, y) = \frac{k_\varepsilon(x, y)}{q_\varepsilon^{(\alpha)}(x) q_\varepsilon^{(\alpha)}(y)},\ p_{\varepsilon,\alpha}(x, y) = \frac{k_\varepsilon^{(\alpha)}(x, y)}{\int_{\mathcal{M}} k_\varepsilon^{(\alpha)}(x, y) q(y) dy},$$

and let $P_{\alpha,\varepsilon}$ denote the Markov operator associated to $p_{\varepsilon,\alpha}$. Let Δ denote the Laplace-Beltrami operator on $\mathcal{M}$, whose eigenfunctions (with Neumann boundary conditions) $\{\phi_l\}_{l\in\mathbb{N}}$ form a Hilbert basis for $L^2(\mathcal{M})$. Then for every $K \in \mathbb{N}$, $f \in E_K := span\{\phi_l\}_{0\le l\le K}$,

$$\lim_{\varepsilon\to 0} \frac{I - P_{\varepsilon,\alpha}}{\varepsilon}(f) = \frac{\Delta(f q^{1-\alpha})}{q^{1-\alpha}} - \frac{\Delta(q^{1-\alpha})}{q^{1-\alpha}} f.$$

Proof. See [22]. □

When q is uniform and k_ε is the Gaussian kernel, the result is well known in $\mathbb{R}^n$, i.e. the Gaussian is the fundamental solution of the heat equation. Furthermore, in this case, for $\alpha = 0$

the diffusion map construction is effectively equivalent to that of the Laplacian eigenmap, in turn constructed based on the graph Laplacian on $\mathcal{M}$. Hence, when q is uniform, theorem 2.6.1 yields the known connection between the graph Laplacian and the Laplace-Beltrami operator on $\mathcal{M}$.

When q is nonuniform, one can set $\alpha = 1$, and obtain once again the convergence to the Laplace-Beltrami operator. The theoretical importance of such result is explained by the known relationship between the geometric characteristics of a smooth manifold $\mathcal{M}$ and the spectrum of the Laplace-Beltrami operator on $L^2(\mathcal{M})$.[14]

2.7 Appendix B: Assets Tickers

a	aapl	abc	abt	adbe	adm	aep	afl	akam	all	amat	amp
amt	amzn	antm	aon	apa	apd	axp	ba	bac	bax	bby	bdx
ben	biib	bk	bmy	c	cah	cat	ccl	cf	chrw	cl	cma
cmcsa	cmi	cms	cof	cop	cost	crm	csco	ctsh	ctxs	cvs	cvx
d	de	dgx	dhr	dis	dov	duk	ebay	ecl	el	eog	eqt
etn	f	fcx	fdx	fitb	flr	fls	fslr	gd	ge	gild	gis
glw	gs	hal	hd	hes	hog	hon	hp	hpq	hum	ibm	ice
intc	isrg	itw	ivz	jci	jnj	jnpr	jpm	jwn	k	kim	klac
kmb	ko	kr	kss	lmt	lnc	low	m	ma	mcd	mck	mdt
met	mmc	mmm	mo	mrk	mro	ms	msft	mtb	mur	nem	nke
nov	nsc	ntap	nvda	nyt	orcl	oxy	payx	pcar	pfe	pg	ph
pnc	ppg	pru	pxd	rf	rhi	rl	rok	rrc	sbux	schw	slb
so	spg	spx	spy	stt	stz	syk	syy	t	tgt	tjx	tmo
trv	txn	txt	unh	unp	ups	usb	vix	vlo	vno	vz	wfc
whr	wmb	wmt	wy	x	xlb	xle	xlf	xli	xlk	xlp	xlu
xlv	xly	xom	xrx								

Table 2.13

[14] An example of such relationship is the famous puzzle "can one hear the shape of a drum?". The answer to this question is negative, but something can be inferred. Weyl's formula, for instance, remarkably connects the area of a bounded domain with its dimension and the rate of decay of the (Neumann, assuming the drum has a clamped boundary) spectrum of the Laplacian (i.e. the solutions of the Helmotz problem for the wave equation).

Chapter 3: Spectral Martingale Measures

3.1 Introduction

The traditional valuation of a derivative claim, sometimes referred to as (linear) martingale pricing theory and developed starting from the seminal work of F. Black and M. Scholes ([12], [49], [50], [28]), assumes the law of one price and the absence of arbitrage opportunities (or, more generally, of "free lunches with vanishing risk"), which, together, imply linearity of prices. In a complete market, such as the one originally considered by Black and Scholes, martingale pricing can be justified relatively easily with a replication argument, which also provides the basics for dynamic hedging. In general, the existence of a so called martingale measure results from a separation argument, and valuations can be performed by specifying a model directly in the risk neutral world and calibrating it to market prices of liquid instruments. However, not all risks can be replicated, and certain positions can only be super-hedged (or sub-hedged depending on the direction of the trade).

To better address this and other problems arising from market incompleteness, attention has more recently moved towards removing the assumption of the law of one price and/or that of linear pricing. In general, this can be achieved by specifying a nonlinear expectation operator (as defined in [85]) and, in time-dependent models, a nonlinear conditional expectation satisfying a

time-consistency property. In continuous time, two types of such operators have been utilized in practical implementations: g-expectations and G-expectation, both of which are due to S. Peng ([83], [84]). Specifically, the g-expectation $\mathcal{E}_g^{t,X_t}[f(X_T)]$, conditional at time t, of a claim $f(X_T)$ that depends on the time T value of a semimartingale X, is defined as the time t value of the solution of a backward stochastic differential equation (BSDE) with driver function g (generally nonlinear), and terminal value given by the claim $f(X_T)$ itself. As it is well known, the value function $\mathcal{E}_g^{t,x}[f(X_T)]$ for $t \in [0,T]$ and $x \in \mathbb{R}^n$ is the (viscosity) solution of a semilinear PIDE (see [8] and [81]). In other words, the nonlinearity arises by introducing nonlinear drift and integral terms into an otherwise linear integro-differential equation. Well known examples, rooted in the g-expectation framework, are, for X a diffusion process, the entropic risk measure, in which the driver function is quadratic and it arises from indifference pricing with exponential utility ([89]), and, for X a pure jump process, the spectral risk measures, characterized by a BSDE driven by a Choquet integral [71]. In the G-expectation approach, instead, the drift and diffusion components of the linear PIDE are transformed through a nonlinear function G, so that a fully nonlinear valuation equation is obtained. For instance, the heat generator $\frac{1}{2}\Delta u$ transformed into $\frac{1}{2}G(\Delta u)$, with $G(a) = \overline{\sigma}a^+ - \underline{\sigma}a^-$, yields the nonlinear expectation that corresponds to the G-Brownian motion ([84]). Stochastic integration and BSDEs based on G-Brownian motion are constructed in [85], while the construction of general nonlinear Levy processes is investigated in [80]. The applications include modeling of volatility uncertainty.

Here the focus is on the g-expectation approach and on spectral risk measures in particular. In fact, theoretical and empirical arguments ([74], [73], [61]) show that, although not completely satisfactory, finite variation pure jump Levy processes fit well the statistical and the risk neutral

distributions of stocks returns, at least for a specific maturity, and if the underlying semimartingale X is such a process, the g-expectation operator is arguably capable of modeling valuations nonlinearities. Furthermore, solutions of BSDEs are essentially dynamic time-consistent risk measures ([46], [71]), and so they can be represented as suprema/infima of (penalized) standard expectation operators over a set of probability measures ([6], [1], [27]). In fact, an alternative to g-expectations could be to start directly with a set of measures and define a risk measure operator accordingly. However, it is time-consistency that provides the connection, fundamental for implementation purposes, with PIDEs, so one would need to prove that the underlying set of measures satisfies the condition known in the literature as m-stability ([26]).

When a nonlinear expectation $\mathcal{E}$ is sublinear (or superlinear for lower valuations), the Hahn-Banach theorem implies that for a claim $f(X_T)$ with finite second moment there is a probability measure $\mathbb{Q}^U$, referred to as an *extreme measure* in [20], which, in general, depends on f, such that $\mathcal{E}[f(X_T)] = \mathbb{E}^{\mathbb{Q}^U}[f(X_T)]$. For a dynamic risk measure, one can similarly find measures $\{\mathbb{Q}^U_t\}_{t\in[0,T]}$ such that $\mathcal{E}_t[f(X_T)]$ is an expectation under $\mathbb{Q}^U_t$ for each $0 \leq t \leq T$. We show that, when X is a pure jump process and $\mathcal{E}$ is a spectral risk measure, there is, under mild conditions, a probability measure $\mathbb{Q}^U$, which we refer to as a (upper) *spectral martingale measure* for the claim $f(X_T)$, such that $\{\mathcal{E}_t[f(X_T)]\}_{0\leq t\leq T}$ is a (local) martingale under $\mathbb{Q}^U$. This is a non trivial generalization with respect to the static case, as it requires a consistent choice of the measures $\{\mathbb{Q}^U_t\}_{t\in[0,T]}$. From a technical perspective, this result allows to link spectral risk measures with the viscosity solutions of the corresponding semilinear PIDE in a standard fashion (as done e.g. in theorem 3.1 in [8]). Furthermore, as a byproduct of the proof of their existence, we also obtain a simple general formula for the Levy density of the process X under $\mathbb{Q}^U$ in terms of the deterministic function u

defined by $u(t,x) = E^{\mathbb{Q}^U}[f(X_T)]$, which, in turn, can be used to develop explicit finite difference schemes for the valuation PIDE.

If X has dimension 1 and f is monotone (e.g. an option), one can determine explicitly the Levy density of the process X under $\mathbb{Q}^U$. Then, the corresponding probability densities and prices of one dimensional monotonic claims can be performed with standard Fourier inversion techniques. From an empirical point of view, knowledge of the Levy density under $\mathbb{Q}^U$ allows one to assess how well such Levy densities fit market data. This is done here in two alternative ways. The first one is a standard application of the generalized method of moments (GMM), and the second one (originally proposed in [68] for non-distorted returns) attempts to match the tails of the empirical distribution of upper and lower returns. We also assess the ability of a specific measure distortion to capture the daily observed bid-ask spreads of options for different level of moneyness, and we obtain conditions under which the nonlinear valuation is an exponential bilateral gamma (BG) process under $\mathbb{Q}^U$, assuming the non-distorted one is an exponential BG process under $\mathbb{P}$.

The Levy density of X under $\mathbb{Q}^U$ can also be determined if f is a portfolio of exponential Levy processes. In this case, the optimal weights for the portfolio choice problem under uncertainty can be obtained, based on the Hamilton Jacobi Bellman equation, by solving a finite dimensional optimization problem. The optimal portfolio weights are constructed for the 10 sector ETFs, assuming a multivariate Bilateral Gamma (MBG) dynamics ([69], and the resulting Sharpe ratio is compared with SPY as benchmark.

A final contribution of this paper is to show the existence and the applications to portfolio choice theory of a convex, but not coherent, spectral risk measure, which is the dynamic version of the risk measure introduced in [67]. The importance of giving up not just to linearity, but also to

sublinearity of valuations emerges when one recognizes the diminishing returns to scale of capital. This can and is usually achieved assuming diminishing marginal utility of returns, which is however not consistent with goals and incentives of large investors. Spectral martingale measures are crucial in the proof of the existence of this new dynamic risk measure. The corresponding portfolio choice problem for the myopic investor is numerically solved to determine optimal investment amount and weights for a portfolio constituted of the 10 sector ETFs, and we provide a comparison with common formulations of the risk-sensitive portfolio choice problem as well as with the static measure introduced in [67].

The rest of the paper is organized as follows. Spectral risk measures are reviewed in Section 2, and an existence theorem for $\mathbb{Q}^U$ is given in section 3, together with an implicit formula for the Levy density of X under $\mathbb{Q}^U$. The empirical analyses for monotone claims, both in the statistical and risk neutral worlds, are performed in section 4. Portfolio choice problems and the definition and the applications of a new dynamic convex spectral risk measure are the subjects of, respectively, sections 5 and 6. Section 7 concludes.

3.2 Assumptions and Preliminary Results

Unless otherwise specified, it is assumed throughout this paper that:

(i) $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space where $\mathcal{F} = \mathcal{F}_T$ and $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the right-continuous and completed filtration generated by a Poisson random measure N on $[0,T] \times \mathbb{R}^D \setminus \{0\}$, whose compensator ν has no atoms and satisfies for some $p_0 > 2$

$$\int_{\mathbb{R}^n\setminus\{0\}} |x|^{p_0}\nu(dy) \in \mathbb{R}^D \setminus \{0\};$$

(ii) X is defined, for $a \in \mathbb{R}^D$ and $t \geq 0$, by[1]

$$X_t = at + \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} y\tilde{N}(dy, dt),$$

where $\tilde{N}(dy, dt) = N(dy, dt) - \nu(dy)dt$;

(iii) $\Gamma_+, \Gamma_- : [0, \nu(\mathbb{R})) \to \mathbb{R}_+$ are bounded, concave, satisfy $\Gamma_-(x) \leq x$ and

$$\int_{(0,\nu(\mathbb{R}))} \frac{\Gamma_\pm(y)}{2y^{3/2}} dy < \infty; \tag{3.2.1}$$

(iv) $g : L^2(\nu) \to \mathbb{R}$ where $L^2(\nu) := L^2(\mathbb{R}\setminus\{0\}, \mathcal{B}(\mathbb{R}\setminus\{0\}), \nu)$,[2] is specified $\forall z \in L^2(\nu)$ by

$$g(z) := \int_0^\infty \Gamma_+(\nu(z^+ > w))dw + \int_0^\infty \Gamma_-(\nu(z^- > w))dw. \tag{3.2.2}$$

Our main results are stated in terms of any pair of measure distortions Γ_+ and Γ_- that satisfy the above assumptions.

Example. Measure distortions satisfying the above assumptions are defined (see [37]) by

$$\Gamma_+(x) = a\left(1 - e^{-cx}\right)^{1/(1+\gamma)},\ \Gamma_-(x) = \frac{b}{c}\left(1 - e^{-cx}\right), \tag{3.2.3}$$

with $0 < \gamma < 1$, $0 < b \leq 1$ and $a, c > 0$.[3] These are obtained by composing the MAXVAR/MINVAR probability distortions with the change of variable $x \to 1 - e^{-cx}$. If an event A satisfies $\nu(A) \geq \frac{10}{c}$, then $\exp(-c(\nu(A)) \approx 0$, and, essentially, $\nu(A)$ is not reweighted.

[1] In applications, X is the (eventually stochastic) logarithm of the value of an asset or a portfolio of assets.

[2] Given a topological space (X, τ), the Borel σ-algebra on X is denoted by $\mathcal{B}(X)$.

[3] Note that if $\gamma \geq 1$ then assumption 3.2.1 does not hold, and if $b > 1$ there is $x > 0$ such that $\Gamma_-(x) > x$. The requirement $\gamma > 0$ ensures that the associated MAXVAR probability distortion is strictly concave.

Spectral Risk Measures

Suppose a claim $C \in L^2 := L^2(\Omega, \mathcal{F}, \mathbb{P})$ is given. Let $\mathcal{U}$ denote the norm topology on $L^2(\nu)$.

Definition 3.2.1. *For $t \in [0,T]$, the upper and lower valuations of C are respectively defined as*

$$U_t = \mathcal{U}[C|\mathcal{F}_t] = \operatorname*{ess\,sup}_{\mathbb{Q}^\xi \in \mathcal{M}^g} \mathbb{E}^{\mathbb{Q}^\xi}[C|\mathcal{F}_t],\ L_t = \mathcal{L}[C|\mathcal{F}_t] = \operatorname*{ess\,inf}_{\mathbb{Q}^\xi \in \mathcal{M}^g} \mathbb{E}^{\mathbb{Q}^\xi}[C|\mathcal{F}_t],$$

where $\mathcal{M}^g$ is the set of measures absolutely continuous with respect to $\mathbb{P}$ such that their Radon-Nikodym derivative ξ is in L^2 and $\mathbb{E}[\xi|\mathcal{F}_t] = \mathcal{E}(M^\xi)_t$, with $\{M_t^\xi\}_{t\in[0,T]}$ represented by

$$M_t^\xi = \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^\xi(\omega, s, y)\tilde{N}(ds, dy),$$

for some $\mathcal{P}$-measurable[4] $\psi^\xi : [0,T]\times\Omega \to L^2(\nu)$ such that, for every Borel set A with $\nu(A) < \infty$,

$$-\Gamma_-(\nu(A)) \le \int_A \psi^\xi(\omega, t, y)\nu(dy) \le \Gamma_+(\nu(A)). \tag{3.2.4}$$

Theorem 3.2.2. *There is a square integrable[5] $L^2(\nu)$-valued process Z such that, $\forall z \in L^2(\nu)$,*

$$U_t(\omega) = C + \int_t^T g(Z_s(\omega,\cdot))ds - \int_{(0,T]\times\mathbb{R}^D\setminus\{0\}} Z_s(\omega, y)\tilde{N}(ds, dy), \tag{3.2.5}$$

$$g(z) = \sup_{\psi\in C^g} \int_{\mathbb{R}^D\setminus\{0\}} \psi(x)z(x)\nu(dx), \tag{3.2.6}$$

where C^g is the set of Borel measurable functions $\psi \in L^2(\nu)$ that satisfy 3.2.4. Similarly for L.

Proof. See theorem 4.3 in [71]. □

[4] $\mathcal{P}$ denotes the predictable σ-algebra on $[0,T]\times\Omega$.

[5] That is, $\mathbb{E}^{\mathbb{P}}[\sup_{t\in[0,T]} \|Z_t\|_{2,\nu}] < \infty$.

Theorem 3.2.3. *Let $f : \mathbb{R}^D \to \mathbb{R}$, and suppose $u \in C^{1,1}([0,T]) \times \mathbb{R})$ solves*

$$\begin{cases} u_t + \mathcal{G}u + g(\mathcal{D}_u^{t,x}) = 0 \\ u(T,x) = f(x) \end{cases} \tag{3.2.7}$$

where $\mathcal{D}_u^{t,x}(y) = u(t,x+y) - u(t,x)$ and

$$\mathcal{G}(u)(t,x) = a^T \nabla u(t,x) + \int_{\mathbb{R}^D \setminus \{0\}} \left(\mathcal{D}_u^{t,x}(y) - \nabla u(t,x)^T y \right) \nu(dy).$$

If ∇u is uniformly bounded, a solution of the BSDE 3.2.5 with $C = f(X(T))$ is

$$U_t = u(t, X_t), \ Z_t(\omega, y) = \mathcal{D}_u^{t, X_{t-}(\omega)}(y).$$

Proof. The result is an application of Ito's lemma for semimartingales. □

Theorem 3.2.4. *Let $z \in L^2(\nu)$. Then $\exists \psi \in C^g$ such that*

$$-g(-\tilde{z}) \leq \int_{\mathbb{R}^D \setminus \{0\}} \psi(y)\tilde{z}(y)\nu(dy) \leq g(\tilde{z}) \ \forall \tilde{z} \in L^p(\nu), \text{ and } \int_{\mathbb{R}^D \setminus \{0\}} \psi(y) z(y) \nu(dy) = g(z).$$

Proof. Define $I : \{az\}_{a \in \mathbb{R}} \to \mathbb{R}$ by $I[az] := ag(z)$ and extend I by Hahn-Banach theorem. □

Remark 3.2.5. *For u as in theorem 3.2.3, $\exists \psi^\xi : [0,T] \times \Omega \to C^g$ such that, $\forall (t,\omega) \in [0,T] \times \Omega$*

$$\int_{\mathbb{R}^D \setminus \{0\}} \psi^\xi(\omega, t, y) \mathcal{D}_u^{t, X_t(\omega)}(y) \nu(dy) = g(\mathcal{D}_u^{t, X_t(\omega)}(y)).$$

Note, however, that ψ^ξ is not $\mathcal{P}$-measurable in general.

3.3 Spectral Martingale Measures

3.3.1 Properties of the set C^g

Lemma 3.3.1. *The set C^g is convex.*

Proof. See [71] and the references therein. □

Lemma 3.3.2. *The set C^g is closed and bounded in $L^p(\nu)$ for $1 \leq p \leq 2$.*

Proof. Fix $\psi \in C^g$ nonnegative. For $p = 1$, setting $A_n = (-\infty, -1/n) \cup (1/n, \infty)$, one obtains

$$\int_{\mathbb{R}\setminus\{0\}} |\psi(y)|\nu(dy) = \lim_{n\to\infty} \int_{A_n} |\psi(y)|\nu(dy) \leq \lim_{n\to\infty} \Gamma_+(\nu(A_n)) + \Gamma_-(\nu(A_n)) < \infty,$$

where the first equality follows from the Monotone Convergence theorem. Thus, C^g is bounded in $L^1(\nu)$. For $1 < p \leq \frac{3}{2} := p_1$, let $\alpha = \frac{1}{p-1}$. Then, $\alpha \geq 2$ and Tchebicheff inequality yields

$$\|\psi\|_{p,\nu}^p \leq \int_0^\infty \Gamma_+(\mu(\psi^{p-1} > x))dx \leq \int_0^\infty \Gamma_+\left(\frac{\|\psi\|_{1,\nu}}{x^{1+1/\alpha}}\right)dx = \|\psi\|_{1,\nu} \int_0^\infty \frac{\Gamma_+(y)}{\alpha y^{1+1/\alpha}}dx < \infty.$$

Next, set $\alpha = \frac{p_1}{p-1}$, and Tchebicheff inequality and boundedness in L^{p_1} give boundedness in L^p for $p_1 < p \leq p_1 + \frac{1}{4} =: p_2$. Repeating this process one can cover all the cases for $p < \sum p_i = 2$. For $p = 2$, Tchebycheff inequality yields

$$\|\psi\|_{2,\nu}^2 \leq \|\psi\|_{2,\nu} \int_0^\infty \frac{\Gamma_+(y)}{2y^{3/2}}dx.$$

Therefore, if $s_n \uparrow \psi$ are nonnegative, simple and integrable, it must be the case that $\{s_n\}_{n\in\mathbb{N}} \subset C^g$ and $\|s_n\|_{2,\nu}$ is bounded in $L^2(\nu)$, so the same bound also holds for $\|\psi\|_{2,\nu}$. Similarly for the case

ψ nonpositive, and the boundedness result follows. It is obvious that C^g is closed in $L^1(\nu)$. If $\psi_n \subset C^g$ converges in $L^p(\nu)$ to ψ, $1 < p \leq 2$, then for every Borel set A of finite measure

$$\int_A \psi(y)\nu(dy) = \lim_{n\to\infty} \int_A \psi(y)\nu(dy) \leq \Gamma_+(\nu(A),$$

and similarly for the lower bound, as requested. □

Corollary 3.3.3. *The driver function g is Lipschitz-continuous for the $L^p(\nu)$-norm for $p \geq 2$.*

Proof. Suppose $z \in L^p(\nu)$. Then, by Tchebycheff's inequality,

$$g(z^+) = \int_{\mathbb{R}^D\setminus\{0\}} \Gamma_+(\nu(z > x))dx \leq \|z\|_{p,\nu} \int_{\mathbb{R}^D\setminus\{0\}} \frac{\Gamma_+(y)}{py^{1+1/p}}dy.$$

and the last integral is finite if $p \geq 2$. Similarly for z^-. □

Based on the technical properties of the set C^g, the following results follow easily.

Theorem 3.3.4. *The set C^g is weakly compact in $L^p(\nu)$ for $1 < p \leq 2$.*

Proof. Being closed, convex and bounded, C^g is weakly compact for $1 < p \leq 2$. □

3.3.2 Definition and Existence

Definition 3.3.5. *Let $C \in L^p$, $p \geq 2$. Measures $\mathbb{Q}^U, \mathbb{Q}^L \in \mathcal{M}^g$ are called, respectively, upper and lower spectral martingale measures for C if, for all $t \in [0, T]$,*

$$\mathcal{U}_t[C] = \mathbb{E}_t^{\mathbb{Q}^U}[C],\ \mathcal{L}_t[C] = \mathbb{E}_t^{\mathbb{Q}^L}[C].$$

Definition 3.3.6. *Let Θ be any set. Two functions $f, g : \Theta \to \mathbb{R}$ are called comonotone if there are no pairs $\theta_1, \theta_2 \in \Theta$ such that $f(\theta_1) < f(\theta_2)$ and $g(\theta_1) > g(\theta_2)$.*

We will show existence of spectral martingale measures under the mild assumption that Γ^+, Γ^- are differentiable. Definition 3.3.6 is based on proposition 4.5 in [30]. Recall also that Choquet integrals are additive over comonotone functions (see [30] proposition 5.1). In particular, by definition 3.3.6, if f and h are comonotone and g and h are comonotone, than $f + g$ and h are comonotone, so Choquet integrals are additive over finite sets of pairwise comonotone functions.

Example. Suppose $D = 1$. Fix $a > 0$, and set $C(\omega) = 1_{\{X_T(\omega) \in [a,\infty)\}}(\omega)$. Let M denote the class of claims $C_\alpha(\omega) = 1_{\{X_T(\omega) \in [\alpha,\infty)\}}(\omega)$ with $\alpha \geq a$. All such claims are comonotonic with C. On the other hand, the claim $C'(\omega) = 1_{\{X_T(\omega) \in [a+1,a+2)\}}(\omega)$ and C are comonotonic, but $C' \notin M$. Thus, the function ψ^ξ in remark 3.2.5 can be specified at $(t, x) = (T, 0)$ for C by ψ_1, ψ_2 and ψ_3 defined by

$$\psi_1(y) = -\Gamma'_+(\nu(y,\infty)1_{\{y>a\}},$$

$$\psi_2(y) = -\Gamma'_+(\nu(y,\infty)1_{\{a+1>y>a\}}$$

$$- \left(\Gamma'_+(\nu(y,a+2)1_{\{a+1>y>a\}} - \Gamma'_+(\nu(y,\infty)1_{\{a+2>y>a+1\}}\right)/2 + \Gamma'_+(\nu(y,\infty)1_{\{y>a+2\}},$$

$$\psi_3(y) = \frac{\Gamma_+(\nu(a,\infty))}{\nu(a,\infty)} 1_{\{y>a\}}$$

However, ψ_1 is consistent with the claims C_α for every $\alpha > 0$ but not with C', ψ_2 is consistent with C' but not with the claims in M, ψ_3 is not consistent with C_α for $\alpha > a$ nor with C'. ■

Theorem 3.3.7. *Suppose ν is σ-finite and Γ_+, Γ_- are differentiable on $(0, \infty)$. Then, a claim*

$C \in L^2$ admits spectral martingale measures $\mathbb{Q}^U$ and $\mathbb{Q}^L$. In addition, for every $t \geq 0$,

$$\mathbb{E}\left[\frac{d\mathbb{Q}^U}{d\mathbb{P}}\middle|\mathcal{F}_t\right] = \mathcal{E}(M^U)_t,\ \mathbb{E}\left[\frac{d\mathbb{Q}^L}{d\mathbb{P}}\middle|\mathcal{F}_t\right] = \mathcal{E}(M^L)_t,$$

where

$$M_t^U = \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^U(s, X_t(\omega), y)\tilde{N}(ds, dy)$$
$$M_t^L = \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^L(s, X_t(\omega), y)\tilde{N}(ds, dy)$$

and where

$$\begin{aligned}\psi^U(t,x,y) &= \Gamma'_+(\nu(A^{z(t,x,y)})1_{\{z(t,x,y)>0\}} - \Gamma'_-(\nu(A^{z(t,x,y)})1_{\{z(t,x,y)<0\}},\\ \psi^L(t,x,y) &= \Gamma'_-(\nu(A^{z(t,x,y)})1_{\{z(t,x,y)>0\}} - \Gamma'_+(\nu(A^{z(t,x,y)})1_{\{z(t,x,y)<0\}},\end{aligned} \tag{3.3.1}$$

with $\Gamma'_+ := \frac{d}{dx}\Gamma_+$, $\Gamma'_- := \frac{d}{dx}\Gamma_-$, and z denoting the deterministic function corresponding to the process Z in theorem 3.2.5.

In the proof below, for $\alpha \in \mathbb{R}\setminus\{0\}$, $(t,x,y) \in [0,T]\times\mathbb{R}^D\times\mathbb{R}^D\setminus\{0\}$, we set

$$z_\alpha(t,x,y) := sign(\alpha)1_{\{A^\alpha_{t,x}\}}(y),$$

where $A^\alpha_{t,x}$ denotes the α-upper (lower) level set of $z(t,x,\cdot)$ if $\alpha > 0$ (if $\alpha < 0$). Furthermore, for $(t,x) \in [0,T]\times\mathbb{R}^D$, $\Sigma_{t,x}$ denotes the completed σ-algebra generated by $z(t,x,\cdot)$, and $\nu_{t,x}$ denotes the restriction of ν to $\Sigma_{t,x}$.

Proof. Step 1. *For every* $(t,x) \in [0,T] \times \mathbb{R}^D$*, there is a* $\nu_{t,x}$*-a.e. unique function* $\psi^U(t,x,\cdot) : \mathbb{R}^D \setminus \{0\} \to \mathbb{R}$ *such that* $\psi^U(t,x,\cdot)$ *is* $\Sigma_{t,x}$ *measurable and, for every* $\alpha > 0$*,*

$$\int_{A^\alpha} \psi^U(t,x,y)\nu(dy) = g(z_\alpha(t,x,\cdot)),\ \int_{A^{-\alpha}} \psi^U(t,x,y)\nu(dy) = g(z_{-\alpha}(t,x,\cdot)). \tag{3.3.2}$$

Furthermore, for every $(t,x) \in [0,T] \times \mathbb{R}^D$*,* $\psi^U(t,x,\cdot)$ *satisfies*

$$g(z(t,x,\cdot)) = \int_{\mathbb{R}^D \setminus \{0\}} \psi^U(t,x,y) z(t,x,y)\nu(dy). \tag{3.3.3}$$

Proof of step 1. Based on definition 3.3.6, it is easy to see that if $I \subset \mathbb{R} \setminus \{0\}$ is finite, then the collection of functions $\{z(t,x,\cdot), \{z_\alpha(t,x,\cdot)\}_{\alpha \in I}\}$ are pairwise comonotone. Hence,

$$g\left(z(t,x,\cdot) + \sum_{\alpha \in I} z_\alpha(t,x,\cdot)\right) = g(z(t,x,\cdot)) + \sum_{\alpha \in I} g(z_\alpha(t,x,\cdot)),$$

and, by Hahn Banach theorem, there is $\psi_I^U \in C^g$ such that 3.3.3 holds and, for every $\alpha \in I$,

$$g(z_\alpha(t,x,\cdot)) = \int_{\mathbb{R}^D \setminus \{0\}} \psi_I^U(y) z^\alpha(t,x,\cdot)\nu(dy).$$

Therefore, the (closed and convex, and thus) weakly closed sets $\Psi_\alpha(t,x)$ defined for $\alpha \in \mathbb{R} \setminus \{0\}$ by

$$\Psi_\alpha(t,x) := \left\{\psi \in C^g : \text{3.3.3 holds and } g(z_\alpha(t,x,\cdot)) = \int_{\mathbb{R}^D \setminus \{0\}} \psi(y) z_\alpha(t,x,y)\nu(dy)\right\},$$

have nonempty intersection over any finite $I \subset \mathbb{R} \setminus \{0\}$. By the finite intersection property of C^g,

$$\Psi(t,x) = \bigcap_{\alpha>0} \Psi_\alpha(t,x) \neq \emptyset.$$

By the choice axiom, there is $\tilde{\psi}^U$ such that for every $(t,x) \in [0,T] \times \mathbb{R}^D$,

$$\tilde{\psi}^U(t,x,\cdot) \in \Psi(t,x).$$

Define a (signed) measure on $(\mathbb{R}^D \setminus \{0\}, \Sigma_{t,x})$ by setting, for every $A \in \Sigma_{t,x}$,

$$\nu^U_{t,x}(A) = \int_A \tilde{\psi}^U(t,x,y)\nu_{t,x}(dy),$$

and note that $\nu^U_{t,x} \ll \nu_{t,x}$. It is easily shown that, for every $\alpha \in \mathbb{R} \setminus \{0\}$,

$$\begin{aligned}
\nu^U_{t,x}(A^\alpha) &= g(z_\alpha(t,x,\cdot)) \\
&= \int_0^\infty \Gamma_+\left(\nu(z^+_\alpha(t,x,\cdot) > s)\right) ds - \int_0^\infty \Gamma_-\left(\nu(z^-_\alpha(t,x,\cdot) > s)\right) ds \qquad (3.3.4) \\
&= \Gamma_+(\nu_{t,x}(A^\alpha))1_{\{\alpha>0\}} - \Gamma_-(\nu_{t,x}(A^\alpha))1_{\{\alpha<0\}},
\end{aligned}$$

so the monotone class theorem implies $\nu^U_{t,x}$ is independent on the choice of $\tilde{\psi}^U(t,x,\cdot)$. Setting

$$\psi^U(t,x,y) := \frac{d\nu^U_{t,x}}{d\nu_{t,x}}(y),$$

the $\nu_{t,x}$ a.e. uniqueness of the Radon-Nikodym derivative implies that if ψ is $\Sigma_{t,x}$-measurable and

satisfies

$$\int_{A^\alpha} \psi(y)\nu(dy) = g(z_\alpha(t,x,\cdot)),\ \int_{A^{-\alpha}} \psi(y)\nu(dy) = g(z_{-\alpha}(t,x,\cdot)) \tag{3.3.5}$$

for every $\alpha > 0$, then $\psi(y) = \psi^U(t,x,y)$ $\nu_{t,x}$-a.e., and similarly for ψ^L. ■

Step 2. Given $(t,x) \in [0,T]\times\mathbb{R}^D$, if Γ_+, Γ_- are differentiable at $\nu(A^{z(t,x,y)})$ for ν-a.e. y, the maps ψ^U, ψ^L, defined in step 1 satisfy 3.3.1 ν-a.e.

Proof of step 2. Define measures $\tilde{\nu}^U_{t,x}$ and $\tilde{\nu}_{t,x}$ on $\mathcal{B}(\mathbb{R}\setminus\{0\})$ by setting, for every $\alpha_+ > 0$, $\alpha_- < 0$,

$$\tilde{\nu}^U_{t,x}([\alpha_+,\infty)) = \nu^U_{t,x}(A^{\alpha_+}),\ \tilde{\nu}^U_{t,x}((-\infty,\alpha_-]) = \nu^U_{t,x}(A^{\alpha_-}),$$

$$\tilde{\nu}_{t,x}([\alpha_+,\infty)) = \nu_{t,x}(A^{\alpha_+}),\ \tilde{\nu}_{t,x}((-\infty,\alpha_-]) = \nu_{t,x}(A^{\alpha_-}),$$

where $\nu^U_{t,x}$ and ν^U are as in step 1 of the proof of step 1. Then, for $\tilde{\nu}_{t,x}$-almost every $\alpha > 0$,[6]

$$\frac{d\tilde{\nu}^U_{t,x}}{d\tilde{\nu}_{t,x}}(\alpha) = \lim_{\varepsilon\downarrow 0} \frac{\Gamma_+(\nu(A^{\alpha+\varepsilon})) - \Gamma_+(\nu(A^{\alpha-\varepsilon}))}{\nu(A^{\alpha+\varepsilon}) - \nu(A^{\alpha-\varepsilon})} = \Gamma'_+(\nu(A^\alpha)) =: \tilde{\psi}^U(t,x,y)$$

for $y \in z^{-1}(t,x,\{\alpha\})$, where for $B \in \mathcal{B}(\mathbb{R}\setminus\{0\})$, $z^{-1}(t,x,B) = \{y \in \mathbb{R}^D\setminus\{0\} : z(t,x,y) \in B)$. Similarly if $\alpha < 0$. Next note that $\tilde{\psi}^U$ so defined is $\Sigma_{t,x}$-measureable and that, for every $B \subset \mathcal{B}(\mathbb{R}\setminus\{0\})$ and $\mathcal{B}(\mathbb{R}\setminus\{0\})$-measurable function θ, by construction of Lebesgue integral,

$$\int_B \theta(p)\tilde{\nu}_{t,x}(dp) = \int_{z^{-1}(t,x,B)} \theta(z(t,x,y))\nu_{t,x}(dy), \tag{3.3.6}$$

as 3.3.6 holds for indicator functions, hence for simple functions, hence for measurable functions.

[6] For the ball ratio limit representation of the Radon-Nykodim derivative see [13], theorem 5.5.8.

In particular, for every $\alpha > 0$ (and similarly if $\alpha < 0$)

$$\begin{aligned}\int_{A^\alpha} \tilde{\psi}^U(t,x,y)\nu(dy) &= \int_\alpha^\infty \Gamma'_+(\nu(A^p))\tilde{\nu}_{t,x}(dp) \\ &= \int_\alpha^\infty \frac{d\tilde{\nu}^U_{t,x}}{d\tilde{\nu}_{t,x}}(p)\tilde{\nu}_{t,x}(dp) \\ &= \Gamma_+(\nu(A^\alpha)).\end{aligned}$$

Thus, $\tilde{\psi}^U$ satisfies 3.3.5, so $\tilde{\psi}^U = \psi^U$ $\nu_{t,x}$-a.e. (and $\tilde{\psi}^U \in C^g$). The case for ψ^L is similar. ■

Step 3. Conclusion.

Proof of step 3. Since z is measurable, ψ is measurable, and the stochastic integrals

$$\begin{aligned}M^U_t &= \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^U(s, X_s(\omega), y)\tilde{N}(ds,dy) \\ M^L_t &= \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^L(s, X_s(\omega), y)\tilde{N}(ds,dy)\end{aligned}$$

are well defined. These in turn define measures $\mathbb{Q}^U$ and $\mathbb{Q}^L$ that satisfy the required properties. □

Note that Γ_+ and Γ_- are increasing and concave, and therefore differentiable almost everywhere. It could be, however, that for some $(t,x) \in [0,T]\times\mathbb{R}^D \setminus \{0\}$, Γ_+ and/or Γ_- are not differentiable at some point p such that $\nu(\{z(t,x,y) = p\}) > 0$. In this case theorem 3.3.7 does not apply, but it is still possible to show that spectral martingale measures still exist assuming the following regularity condition on z. The proof is in the appendix.

Definition 3.3.8. *A claim $C \in L^2$ is called regular if the deterministic function z defined by the process Z in theorem 3.2.5, is uniformly continuous on $[0,T]\times\mathbb{R}^D$ uniformly in $y \in \mathbb{R}^D \setminus \{0\}$.*

Theorem 3.3.9. *Let ν be σ-finite and suppose that $C \in L^2(\nu)$ is a regular claim. Then, C admits upper and lower spectral martingale measures.*

Remarks.

(i) **Regularity** of C is not needed for step 1 of theorem 3.3.9, nor for theorem 3.3.7. However, there may be a set $B \subset \mathbb{R}^D \setminus \{0\}$ with $\nu(B) > 0$ such that $\{\nu(A^{z(t,x,y)})\}_{y \in B}$ is a subset of the (countable) set of points in which Γ_+ and/or Γ_- are nondifferentiable. In this case, one can apply theorem 3.3.9 if C is regular.

(ii) **Viscosity Solutions**. Suppose that $u : [0,T] \times \mathbb{R}^D \setminus \{0\} \to \mathbb{R}$ is a viscosity solution (as defined e.g. in [8]) of

$$\begin{cases} u_t + \mathcal{G}u + \int_{\mathbb{R}^D \setminus \{0\}} \mathcal{D}_u^{t,x}(y)\psi^U(t,x,y)\nu(dy) = 0 \\ u(T,x) = f(x) \end{cases} \tag{3.3.7}$$

where $\mathcal{G}$ is as in theorem 3.2.3 and ψ^U is constructed from $\mathcal{D}_u^{t,x}$. If u is uniformly continuous and unique, a standard argument as in theorem 3.4 in [8] shows that $u(t, X_t(\omega))$ solves 3.2.5. This observation leads to the finite difference scheme defined in the next section for computing nonlinear valuations. A detailed investigation is left for future research.

(iii) **Extension to Hunt Processes**. It is also possible to allow ν, as well as the distortions Γ_+, Γ_-, to depend on (t,x), as long as ψ^U and ψ^L are weakly measurable (e.g. as in theorem 3.3.9, they are uniformly continuous in (t,x) uniformly in y). For instance, X could be locally a BG process, i.e. the Levy density $\kappa(t,x,\cdot)$ of $\nu_{t,x}$ is a BG Levy density with the property that $(t,x,y) \to \nu_{t,x}(A^{z(t,x,y)})$ is measurable.

(iv) **Uniqueness of SMMs**. When $D = 1$, it could be that $\Sigma_{t,x} = \mathcal{B}(\mathbb{R} \setminus \{0\})$. When this

is true for each $(t,x) \in [0,T] \times \mathbb{R}$, the SMMs are unique. If $\Sigma_{t,x}$ is strictly contained in $\mathcal{B}(\mathbb{R} \setminus \{0\})$, one could define a comonotone class for a regular claim C as any set of pairwise (conditionally, as defined in [57]) comonotonic claims that contains C. Using Zorn's lemma it is then easy to show that any claim C admits a maximal comonotone class $\mathcal{M}_C$, i.e. a comonotone class such that any claim not in $\mathcal{M}_C$ is not comonotone with at least one claim in $\mathcal{M}_C$, and it is easy to see that at least one claim in $\mathcal{M}_C$ generates $\mathcal{B}(\mathbb{R}) \setminus \{0\}$ at each point $(t,x) \in [0,T] \times \mathbb{R}$.

3.3.3 An Explicit Finite Difference Scheme

It is not possible in general to determine analytically the functions ψ^U and ψ^L. One can, however, construct a finite difference scheme for the solution of the time reversed linear equation

$$\begin{cases} u_t + a' \nabla u(t,x) - \int_{\mathbb{R}^D \setminus \{0\}} \mathcal{D}_u^{t,x}(y)[1 + \psi^U(t,x,y)]\nu(dy) = 0 \\ u(0,x) = f(x), \end{cases} \tag{3.3.8}$$

where it is assumed that ν has finite variation and $a \in \mathbb{R}^D$ is defined by

$$a = \int_{\mathbb{R}^D \setminus \{0\}} (e^y - 1)\nu(dy),$$

with

$$e^y = (e^{y_1}, ..., e^{y_D}).$$

Consider M (resp. N) equal subintervals in the variable t (resp. x) on the region $[0,T] \times [x_{\min}, x_{\max}]$:

$$P = \begin{cases} t_j = j\Delta t;\ \Delta t = \frac{T}{M}; j = 0,1,...,M \\ \\ x_i = x_{\min} + i\Delta x;\ \Delta x = \frac{x_{\max} - x_{\min}}{N}; i = 0,1,...,N \end{cases}$$

Let $(t_j, x_i) \in \mathbb{R}_+ \times \mathbb{R}$ denote the grid points in P, and let $u_{j,i} = u(t_j, x_i)$. Assuming that the $N+1$ values $\{u_{j,i}\}_{i=1,...,N}$ are known for fixed t_j, and in order to construct the difference equation for each point (t_{j+1}, x_i), let $z_{ji}(y)$ be the piecewise linear approximation of $z(t_j, x_i, y)$. For $k = 1, ..., K$, $K \in \mathbb{N}$, $i = 1, ..., N+1$, set $y_{k,i} = x_k - x_i$ and define $s_k = z_{j,i}(y_k)$ and

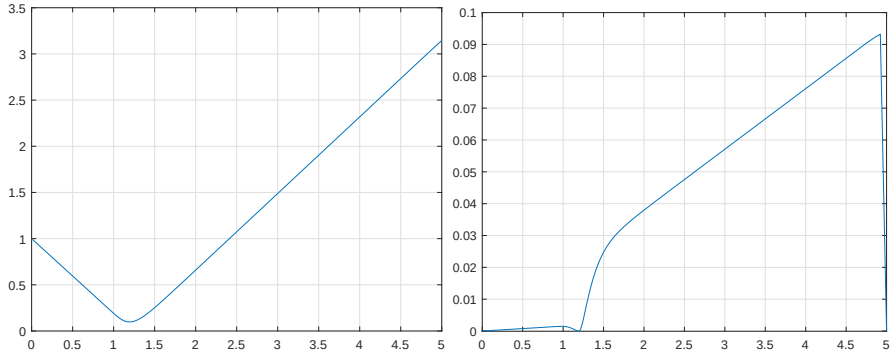
$$A^s_{i,j} = \begin{cases} \{z_{j,i} > s\} \text{ if } s > 0 \\ \\ \{z_{j,i} < s\} \text{ if } s < 0, \end{cases}$$

so that $\psi^U(t_{j+1}, x_i, y_k)$ can be approximated by

$$\psi^U_{j,i}(y_k) = \Gamma'_+(\nu(A^{s_k}_{i,j}))1_{\{s_k > 0\}} - \Gamma'_-(\nu(A^{s_k}_{i,j}))1_{\{s_k < 0\}}.$$

Standard quadratures can then be utilized to compute the risk charge at each point (t_{j+1}, x_i) in the spacetime grid P, and equation 3.3.8 can be solved using a standard explicit finite difference scheme. Figure 3.1 shows contour plot and risk charge for a straddle where, for the numerical scheme, we used $N = 400$, $M = 50$. The underlying is assumed to be an exponential BG process with parameters

$$(b_p, c_p, b_n, c_n) = (0.0075, 1.5592, 0.0181, 0.6308), \tag{3.3.9}$$



(a) (b)

Figure 3.1: Upper valuation and driver function for an ATM straddle with maturity 1 month.

obtained by estimation based on the SPY ETF prices between 2 Januray 2020 and 31 December 2020. Specifically, the parameters (b_p, b_n) are the scale parameters of, respectively, positive and negative jumps, while (c_p, c_n) their shape parameters. The measure distortions are instead as in 3.2.3, with parameters

$$(c, \gamma, a, b) = (0.01, 0.25, 100, 1). \tag{3.3.10}$$

These are the values, as found in [39], that maximize the return for a portfolio choice problem with lower valuation as financial objective.

3.4 Monotone Claims

3.4.1 The Spectral Martingale Measures of Monotone Claims

If $D = 1$ and the claim's payoff is a monotonic function, then ψ^U does not depend on (t, x), and so it can be fully determined. Specifically, the following result follows directly from theorem 3.3.9.

Theorem 3.4.1. *If $D = 1$, $z \in L^p(\nu)$ non-decreasing, $z(0) = 0$, $p \geq p_\Gamma$, Γ_+, Γ_- differentiable, then*

$$g(z) = \int_{\mathbb{R}\setminus\{0\}} \psi^U(y) z(y) \nu(dy)$$

for every $t \in [0,T]$, where, for every $y \in \mathbb{R} \setminus \{0\}$,

$$\psi^U(y) = \Gamma'_+ \left(\nu([y,\infty))\right) 1_{\{y>0\}} - \Gamma'_- \left(\nu((-\infty,y])\right) 1_{\{y<0\}}.$$

Similarly, if $z \in L^p(\nu)$ is non-increasing, then,

$$g(z) = \int_{\mathbb{R}\setminus\{0\}} \psi^L(y) z(y) \nu(dy)$$

for every $t \in [0,T]$, where, for every $y \in \mathbb{R} \setminus \{0\}$,

$$\psi^L(y) = -\Gamma'_- \left(\nu([y,\infty))\right) 1_{\{y>0\}} + \Gamma'_+ \left(\nu((-\infty,y])\right) 1_{\{y<0\}}.$$

Note that ψ^U, ψ^L may not be in $L^2(\nu)$ if 3.2.1 is not satisfied. For instance, if Γ_+ is as in 3.2.3,

$$\psi^U(y) = \frac{ac}{1+\gamma} \left(1 - e^{-c\nu(y,\infty)}\right)^{-\frac{\gamma}{1+\gamma}} e^{-c\nu(y,\infty)},\ y \geq 0,$$

and if $\gamma \geq 1$ and ν an exponential integral (as e.g. in [74] or [61]), $\psi^U \notin L^2(\nu)$. Based on the properties of the set C^g, one could extend the theory of spectral martingale measures to the general

Banach space $L^p(\nu)$ with $p \geq 2$. This would require, however, a definition of stochastic integral for non square integrable processes (as in [92]) and a generalization of the theory of spectral risk measures. We leave this issue for future research, so in this dissertation the $L^2(\nu)$ requirement for ψ^U and ψ^L will be enforced.

Theorem 3.4.2. *Suppose $u \in C^{1,1}([0,T] \times \mathbb{R})$ satisfies the assumptions of theorem 3.2.3 with f increasing. Then, $f(X_T)$ admits upper and lower spectral martingale measures $\mathbb{Q}^U, \mathbb{Q}^L \in \mathcal{M}^g$. The Doob martingales associated to the Radon-Nykodim derivatives of $\mathbb{Q}^U$ and $\mathbb{Q}^L$ satisfy, respectively, $H^U = \mathcal{E}(M^U)_T$ and $H^L = \mathcal{E}(M^L)_T$ with M^U and M^L defined, $\forall t \in [0,T]$, by*

$$M_t^U := \int_{[0,t]\times\mathbb{R}\setminus\{0\}} \psi^U(y)\tilde{N}(ds,dy),\ M_t^L := \int_{[0,t]\times\mathbb{R}\setminus\{0\}} \psi^L(y)\tilde{N}(ds,dy).$$

Thus, X is a Levy process under $\mathbb{Q}^U$ and $\mathbb{Q}^L$ with Levy measure defined for $A \in \mathcal{B}(\mathbb{R}\setminus\{0\})$ by

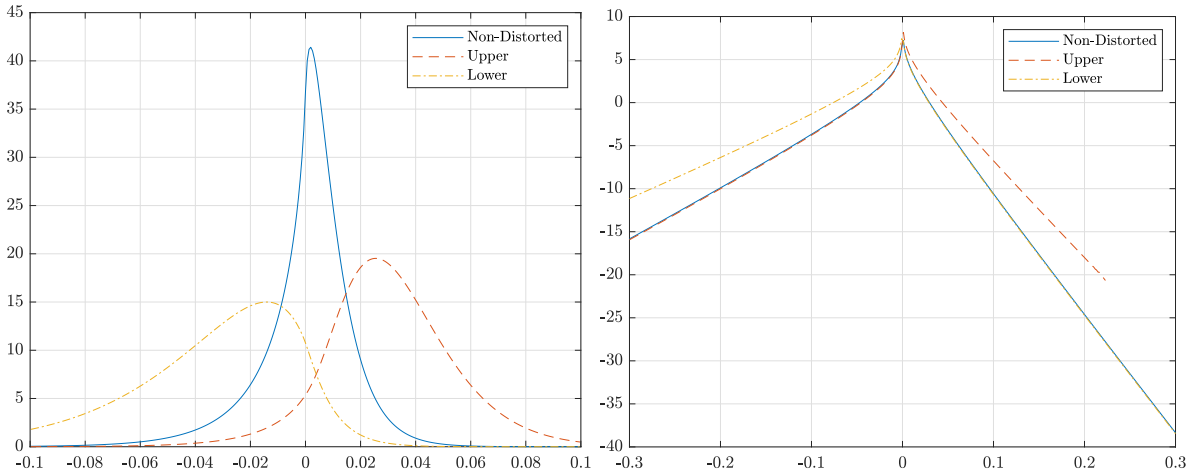
$$\nu^U(A) = \nu(A) + \Gamma_+(\nu(A\cap(0,\infty))) - \Gamma_-(\nu(A\cap(-\infty,0))),$$
$$\nu^L(A) = \nu(A) - \Gamma_-(\nu(A\cap(0,\infty))) + \Gamma_+(\nu(A\cap(-\infty,0))).$$

Proof. By time-consistency, u is increasing if f is ([1]), and the result follows from theorem 3.4.1.

□

The Densities of the Spectral Martingale Measures of Monotone Claims

Given the measures ν^U and ν^L, the characteristic exponents under $\mathbb{Q}^U$ and $\mathbb{Q}^L$ of the process X can be numerically computed for $t \in [0,T]$ based on the Levy-Kintchine formula as, respec-



(a) (b)

Figure 3.2: Plot of the probability densities of X under $\mathbb{Q}^U$ and $\mathbb{Q}^L$ (a) and of the log Levy densities (b).

tively,

$$\mathbb{E}^{\mathbb{Q}^U}\left[e^{i\theta U_t}|X_0 = x\right] = e^{t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta y}-1)(1+\psi^U(y))\nu(dy)}$$

$$\mathbb{E}^{\mathbb{Q}^L}\left[e^{i\theta L_t}|X_0 = x\right] = e^{t\int_{\mathbb{R}\setminus\{0\}}(e^{i\theta y}-1)(1+\psi^L(y))\nu(dy)}.$$

The corresponding probability densities can then be obtained via Fourier inversion. For instance, the non distorted and upper and lower densities of X_t, for $t = 1$, are plotted in figure 3.2a, where, as above, it is assumed that X is a BG process and the measure distortions are as in 3.2.3. Parameters are as in 3.3.9 and 3.3.10. The probability density shown in figure 3.2a is for a specific tenor (here set to 1 month). For a more general understanding of the distortive effects of nonlinear valuations we also plot the log Levy density of X under $\mathbb{P}$, $\mathbb{Q}^U$ and $\mathbb{Q}^L$ in figure 3.2b.

3.4.2 Empirical Tests of Nonlinear Asset Pricing Equations

Consider, for a fixed, distant, investment horizon $T \in (0, \infty)$, a non-dividend paying stock. Suppose that its non-distorted price process Y is given, for $0 \leq t \leq T$, by

$$Y_t = Y_0 + \int_{[0,t]\times\mathbb{R}\setminus\{0\}} Y_{t-}(e^y - 1)N(dy, ds),$$

and consider the claim $C = Y_T$. It was so far assumed that upper and lower valuations for such a claim are given by $U_t = \mathcal{U}[C|\mathcal{F}_t]$ and $L_t = \mathcal{L}[C|\mathcal{F}_t]$. In the formulation with discounting, we let

$$\mu^U = \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\nu^U dy,\ \mu = \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\nu(dy),\ \mu^L = \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\nu^L dy,$$

and we define $U_t = \mathcal{U}[e^{-\mu(T-t)}C|\mathcal{F}_t]$, and $L_t = \mathcal{L}[e^{-\mu(T-t)}C|\mathcal{F}_t]$. The results proved in the previous sections remain valid, provided they are applied to the discounted distorted price processes, i.e. it is the processes $\{e^{-\mu t}U_t\}_{0\leq t\leq T}$ and $\{e^{-\mu t}L_t\}_{0\leq t\leq T}$ that are now martingales under $\mathbb{Q}^U$ and $\mathbb{Q}^L$ respectively. Simple calculations then yield

$$U_t = e^{(\mu^U-\mu)(T-t)}Y_0 e^{X_t},\ L_t = e^{(\mu^L-\mu)(T-t)}Y_0 e^{X_t}.$$

That is,

$$U_t = U_0 + \int_{[t,T]\times\mathbb{R}\setminus\{0\}} U_{t-}(e^y - 1)\psi^U(y)\nu(dy)ds + \int_{[0,t]\times\mathbb{R}\setminus\{0\}} U_{t-}(e^y - 1)N(dy, ds),$$
$$L_t = L_0 + \int_{[t,T]\times\mathbb{R}\setminus\{0\}} L_{t-}(e^y - 1)\psi^L(y)\nu(dy)ds + \int_{[0,t]\times\mathbb{R}\setminus\{0\}} L_{t-}(e^y - 1)N(dy, ds).$$

Hence,

$$\frac{dU_t}{U_{t-}} = \mu dt - \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\psi^U(y)\nu(dy)dt + \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\tilde{N}(dy, dt),$$
$$\frac{dL_t}{L_{t-}} = \mu dt - \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\psi^L(y)\nu(dy)dt + \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\tilde{N}(dy, dt).$$

These relations in turn imply:

$$\mu^U dt = E^{\mathbb{P}}\left[\frac{dU_t}{U_{t-}}\right] = \mu dt - \frac{RC_t^U}{U_{t-}}dt, \tag{3.4.1}$$
$$\mu^L dt = \mathbb{E}^{\mathbb{P}}\left[\frac{dL_t}{L_{t-}}\right] = \mu dt + \frac{RC_t^L}{L_{t-}}dt., \tag{3.4.2}$$

where the risk charges RC^U and RC^L are simply the respective driver functions, i.e.

$$\frac{RC_t^U}{U_{t-}}(\omega) := \int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\psi^U(y)\nu(dy),\ \frac{RC_t^L}{L_{t-}}(\omega) := -\int_{\mathbb{R}\setminus\{0\}} (e^y - 1)\psi^L(y)\nu(dy),$$
$$\Rightarrow \mu^U \leq \mu \leq \mu^L. \tag{3.4.3}$$

3.4.2.1 Numerical Results for SPY

For estimation purposes, suppose that $f(x) = e^x - 1$, and that X represents the log-return process of a non dividend paying stock (e.g. an index) under no transaction costs. The discrete version of the pricing equations 3.4.1 and 3.4.2 is given, for a discrete set of times $t = 1, ..., T$, by

$$\mathbb{E}_t\left[\frac{U_{t+1} - U_t}{U_t} - \int_{\mathbb{R}} (e^y - 1)\nu^U(dy)\right] = 0,\ \mathbb{E}_t\left[\frac{L_{t+1} - L_t}{L_t} - \int_{\mathbb{R}} (e^y - 1)\nu^L(dy)\right] = 0.$$

By the law of iterated expectations, if g is measurable and $\mathbb{E}[|g(U_t)|] < \infty$ and $\mathbb{E}[|g(L_t)|] < \infty$,

$$\mathbb{E}_t\left[\left(\frac{U_{t+1}-U_t}{U_t} - \int_{\mathbb{R}}(e^y-1)\nu^U(dy)\right)g(U_t)\right] = 0, \tag{3.4.4}$$

$$\mathbb{E}_t\left[\left(\frac{L_{t+1}-L_t}{L_t} - \int_{\mathbb{R}}(e^y-1)\nu^L(dy)\right)g(L_t)\right] = 0. \tag{3.4.5}$$

Setting $g(u) = u^h$, $h = 1, 2, ...$ and assuming that 3.4.4 and 3.4.5 hold, at least, locally, the typical approach in estimating daily distortion parameters c_{t_N} and γ_{t_N} at time t_N is to solve

$$\frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{U_{t_i+1}-U_{t_i}}{U_{t_i}} - \int_{\mathbb{R}}(e^y-1)\nu^U(dy)\right)U_{t_i}^h\right] = 0 \tag{3.4.6}$$

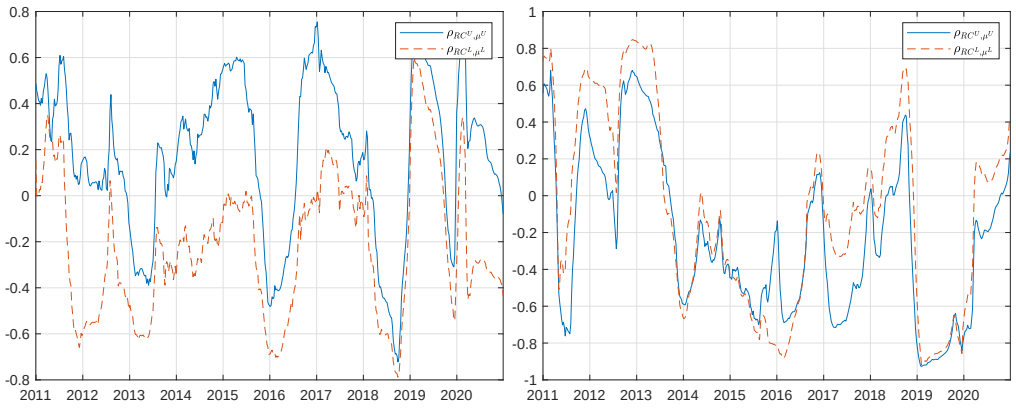
$$\frac{1}{N}\sum_{i=1}^{N}\left[\left(\frac{L_{t_i+1}-L_t}{L_{t_i}} - \int_{\mathbb{R}}(e^y-1)\nu^L(dy)\right)L_{t_i}^h\right] = 0. \tag{3.4.7}$$

These are the generalized method of moments (GMM) estimators. An alternative is to match the tails of the empirical distribution, resulting in the digital moments (DM) estimators ([68]).

Assuming X is a BG process with parameters (b_p, c_p, b_n, c_n) obtained by matching mid prices digital moments, the GMM and DM estimators for the measure distortion parameters can be obtained based on upper and lower valuations U and L taken as the 5-day maximum and minimum of daily closing prices. Figure 3.5 shows estimated risk charges for the SPY ETF from 2010 through 2020. The distortions are assumed as in 3.2.3.

Remarks.

(i) Figure 3.4 shows estimated and empirical tails as of December 31 2020 for SPY (the magnitude of the mean squared error with Andersen-Darling weights is, on average of the order of $1e03$). Note in particular, that the left empirical tail is fatter than the right one.



(a) (b)

Figure 3.3: Estimated daily correlation between risk charges and returns.

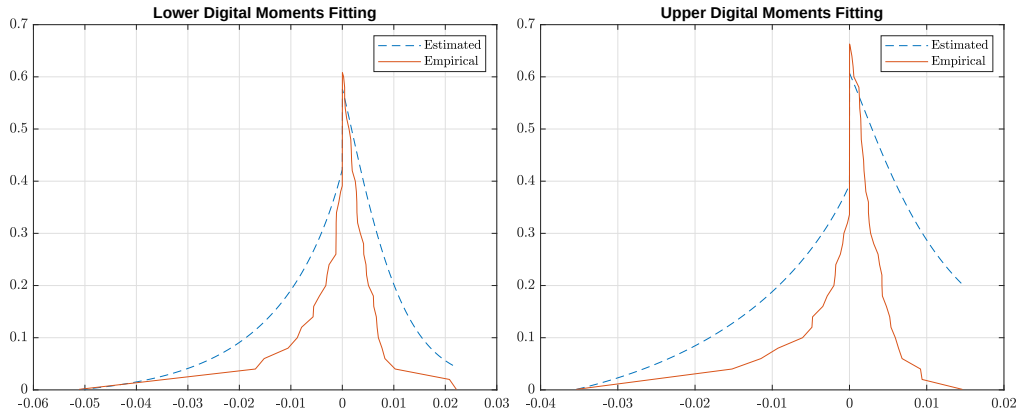
(ii) The DM estimated risk charges are as much as twice the GMM estimated ones, suggesting that replication of the drifts in 3.4.1 and 3.4.2 through 3.4.6 and 3.4.7 is not sufficient to generate measure distortions that adequately capture the tails of the nonlinear valuations.

(iii) Similarly, as one could expect, the correlation between DM estimated risk charges and returns is negative approximately 69% of the days considered for the upper valuation and 57% for the lower one, but for the GMM-estimators, the percentages are 24% and 77% respectively.

(iv) Tables 3.1 and 3.2 report the first five most representative points of 16 quantized points of DM and GMM estimators of measure distortions parameters. With DM estimators, Γ_+ is dominated by Γ_- (figure 3.6), as $\frac{b}{c} > a$. This is not the case for GMM estimators, which explains the positive correlation of GMM's upper valuations with upper returns. On the other hand, b approaches 1 for both estimators.

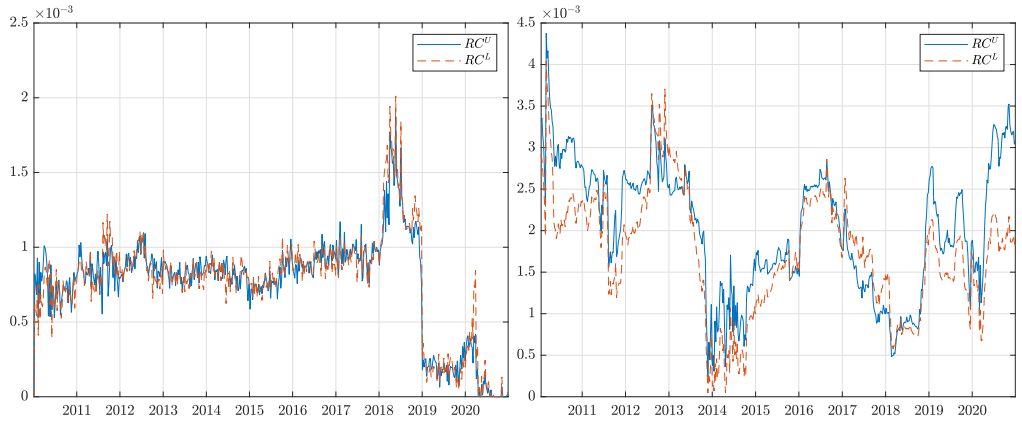
(v) Inequality 3.4.3 holds on average for the upper, non-distorted and lower returns of the SPY and the 10 sector ETFs for the period 2010-2020,[7] as reported in table 3.3.

[7]The non-distorted mean return is computed based on mid prices.



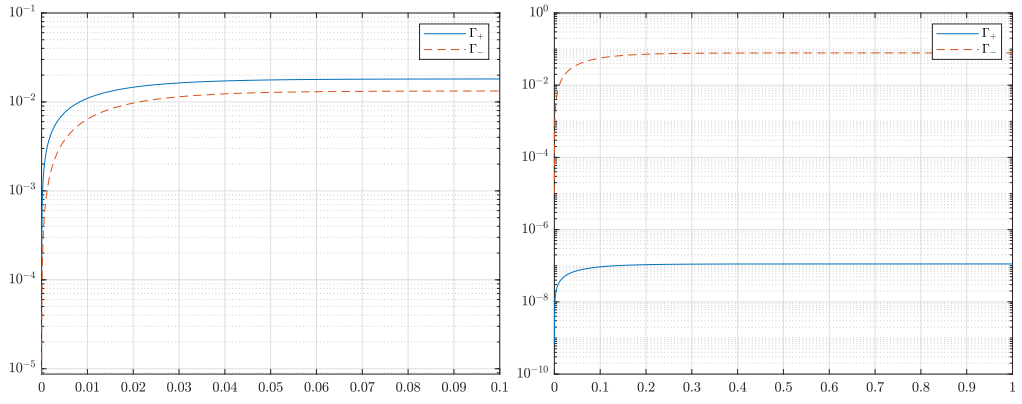
(a) (b)

Figure 3.4: DM-estimated and empirical upper and lower tails as of 31 December 2020 for SPY. The Andersen-Darling statistics evaluates to approximately 718 after 150 iterations of the Nelder-Mead algorithm.



(a) (b)

Figure 3.5: Risk charges for the SPY ETF based on GMM (a) and DM (b) estimators.



(a) (b)

Figure 3.6: Measure distortions for the SPY ETF based on first quantized point of the (a) GMM estimators and (b) DM estimators of the measure distortion parameters. The scale is logarithmic.

c	γ	a	b	$\frac{b}{c}$
12.7092	0.7689	1.1216e-07	0.9949	0.0783
5.8715	0.3860	8.7573e-06	0.9998	0.1705
9.0998	0.4124	5.9994e-06	0.9982	0.1098
3.9814	0.3558	2.0926e-06	1.0000	0.2514
11.3294	0.4826	9.3874e-06	0.9991	0.0883

Table 3.1: First five quantized points of the DM estimators of the measure distortions parameters (c, γ, b, a) for SPY. Parameters were estimated each day during the period 2010-2020.

c	γ	a	b	$\frac{b}{c}$
65.5791	0.4550	0.0181	0.8712	0.0133
78.4547	0.5233	0.0175	0.8705	0.0111
56.0514	0.3189	0.0161	0.8977	0.0160
24.0361	0.5432	0.0236	0.9464	0.0395
42.5793	0.4410	0.0204	0.9376	0.0221

Table 3.2: First five quantized points of the GMM estimators of the measure distortions parameters (c, γ, b, a) for SPY. Parameters were estimated each day during the period 2010-2020.

ETF	μ^U	μ	μ^L	ETF	μ^U	μ	μ^L
XLB	4.16	4.36	4.41	XLP	6.98	7.05	7.08
XLE	-4.90	-4.79	-4.62	XLU	4.38	4.45	4.44
XLF	1.48	1.82	2.04	XLV	10.12	10.20	10.22
XLI	6.94	7.04	7.04	XLY	12.81	13.03	13.12
XLK	13.36	13.51	13.52	SPY	8.11	8.25	8.28

Table 3.3: Averages (in percentage points) over the period 2010-2020 of the annualized upper, non-distorted and lower daily returns for the 10 sector ETFs and SPY. For each day between January 2010 and December 2020, daily returns where computed based on previous 252 days.

3.4.2.2 Calibration to options bid-ask spreads

A different, forward-looking, approach is to calibrate measure distortions and BG parameters to bid and ask prices of option. In fact, as shown above, discounted upper and lower valuations are martingales under the measures $\mathbb{Q}^U$ and $\mathbb{Q}^L$, so assuming that measure distortion parameters remain the same for a few different maturities and/or strikes, standard calibration procedures can be employed to match theoretical and market prices of options. More in details, the upper price

process C^U of a European call and the lower price process P^L of a European put both with strike K and maturity T are given (in the formulation with risk free rate discounting) by

$$C^U(t, X_t) = \mathbb{E}_t^{\mathbb{Q}^U}\left[e^{-r(T-t)}\left(Y_0 e^{\omega T + X_T} - K\right)^+\right], \tag{3.4.8}$$

$$P^L(t, X_t) = \mathbb{E}_t^{\mathbb{Q}^U}\left[e^{-r(T-t)}\left(K - Y_0 e^{\omega T + X_T}\right)^+\right], \tag{3.4.9}$$

where Y represents the price of the underlying with no transaction costs and

$$\omega = r + \log\left[(1-b_p)^{-c_p}(1+b_n)^{-c_n}\right]$$

ensures that Y has drift r under $\mathbb{Q}$. Assuming that Y_0 is simply the mid price, one can then calibrate the BG and measure distortion parameters to option prices.[8]

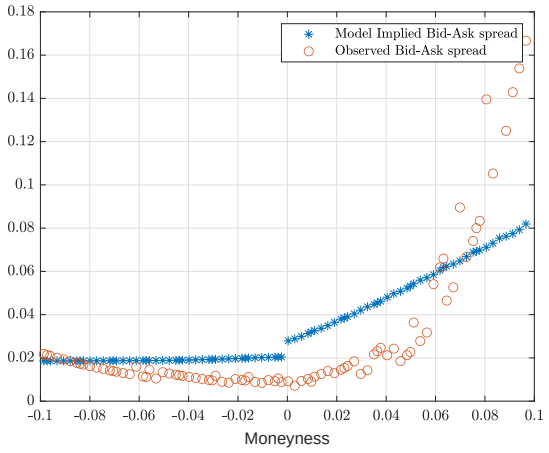
Results of Calibration

We considered again the SPY ETF, and calibrated the measure distortions specified in 3.2.3 and BG parameters to bid and ask prices of call and put options observed on 31 December 2020. Figure 3.7 shows the OTM options relative bid-ask spreads. The distortion parameters are

$$(c, \gamma, a, b) = (0.0021, 0.1996, 0.0011, 0.0067)$$

Not that $\frac{b}{c} > a$ in this case too, but now b is closer to 0.

[8] Alternatively, Y_0 can be assumed to be a parameter to be calibrated.



(a)

Figure 3.7: Model and market relative spreads of OTM options on 31 December 2020. Negative and positive moneyness correspond to puts and calls respectively.

3.4.3 The BG2BG measure distortions

In this section we define measure distortions Υ_+ and Υ_- such that the process X is a BG process under both $\mathbb{Q}^U$ and $\mathbb{Q}^L$, given that it is a BG process under $\mathbb{Q}$. In this case, ψ^U_∞ must be given by

$$\psi^U(y) = \frac{\kappa^U(y)}{\kappa(y)} - 1,$$

where κ^U and κ are the BG Levy densities under $\mathbb{Q}^U$ and $\mathbb{Q}$ respectively. This implies

$$\Upsilon_+(\nu[y,\infty)) = \left[\frac{1}{c_p^U}E_1(y/b_p^U) - \frac{1}{c_p}E_1(y/b_p)\right]1_{\{y>0\}},$$

$$\Upsilon_-(\nu[(-\infty,y]) = -\left[\frac{1}{c_n^U}E_1(-y/b_n^U) - \frac{1}{c_n}E_1(-y/b_n)\right]1_{\{y<0\}},$$

$$\Rightarrow \Upsilon_+(x) = \frac{1}{c_p^U}E_1[E_1^{-1}(xc_p)b_p/b_p^U] - x, \tag{3.4.10}$$

$$\Rightarrow \Upsilon_-(x) = -\frac{1}{c_n^U}E_1[E_1^{-1}(xc_n)b_n/b_n^U] + x, \tag{3.4.11}$$

where, for every $\alpha > 0$, E_α is defined as:

$$E_\alpha(x) = \int_x^\infty \frac{e^{-t}}{t^\alpha} dt.$$

The measure distortions Υ_+ and Υ_- defined by 3.4.10 and 3.4.11 respectively are a dynamic version of the probability distortion arising from the Wang transform ([99], [101], [100]), with the BG distribution replacing the Gaussian. The next result gives necessary and sufficient conditions for them to satisfy the usual assumptions.

Proposition 3.4.3. *The distortions Υ_+ and Υ_- with $b_p^U \geq b_p > b_p^U/2$, $c_p = c_p^U$, $c_n = c_n^U$ and $b_n^U \leq b_n$, are bounded, increasing, concave, satisfy assumption 3.2.1 and $\Upsilon_-(x) \leq x$.*

Proof. It is clear that for Υ_+ and Υ_- to be both bounded it is necessary and sufficient that $c_p = c_p^U$ and $c_n = c_n^U$. To prove concavity, note that

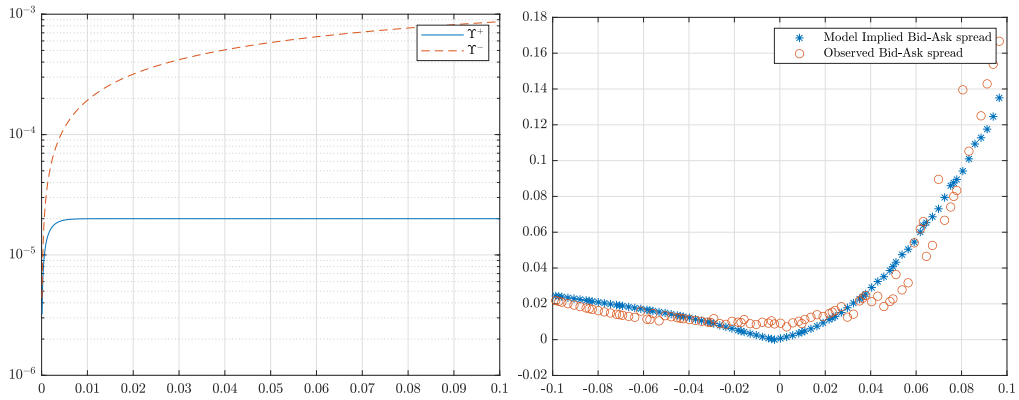
$$\Upsilon'_+(x) = \frac{e^{-E_1^{-1}(xc_p)b_p/b_p^U}}{E_1^{-1}(xc_p)c_p b_p/b_p^U} \frac{E_1^{-1}(xc_p)c_p b_p/b_p^U}{e^{-E_1^{-1}(xc_p)}} - 1 = e^{-E_1^{-1}(xc_p)(b_p/b_p^U - 1)} - 1 > 0$$

$$\Upsilon''_+(x) = c_p e^{-E_1^{-1}(xc_p)(b_p/b_p^U - 1)} \frac{E_1^{-1}(xc_p)}{e^{-E_1^{-1}(xc_p)}} \left(\frac{b_p}{b_p^U} - 1 \right) < 0$$

always hold if $b_p^U > b_p$. As for assumption 3.2.1, setting $0 < \varepsilon < \frac{b_p}{b_p^U} - \frac{1}{2}$ and using the substitution $xc_p = E_1(y)$, we obtain

$$\lim_{x \to 0} \frac{E_1[E_1^{-1}(xc_p)b_p/b_p^U]}{x^{1/2+\varepsilon}} = \lim_{y \to \infty} c_p^{1/2+\varepsilon} \frac{E_1(yb_p/b_p^U)}{E_1(y)^{1/2+\varepsilon}} = \lim_{y \to \infty} c_p^{1/2+\varepsilon} \frac{b_p^U}{b_p} e^{-y\left(\frac{b_p}{b_p^U} - \frac{1}{2} - \varepsilon\right)} y^{\varepsilon - 1/2} = 0,$$

which implies 3.2.1. The proof for Υ_- is similar. □



(a) (b)

Figure 3.8: Figure (a): The BG2BG distortions (in log scale), and, figure (b): BG2BG and market implied relative spreads on options on SPY. Parameters are estimated as of 31 December 2020.

Note that, from theorem 3.4.2, κ^U is a BG density with parameters (b_p^U, c_p, b_n^U, c_n). If one further assumes $b_p = b_n$ and $c_p = c_n = c$, then κ^L is also a BG density and its parameters are (b_n^U, c, b_p^U, c). Alternatively, one can again let upper and lower measure distortions differ, in which case one is led to upper and lower BG processes with parameters (b_p^U, c_p, b_n^U, c_n) and (b_p^L, c_p, b_n^L, c_n) respectively. One must then require $b_p^U > b_p^L$ and $b_n^U < b_p^L$ for there to exist b_p and b_n that satisfy the assumptions of proposition 3.4.3. However, by inverting the roles of Υ_+ and Υ_- for the upper and/or the lower valuations, these conditions are not really needed for X to be a BG process under $\mathbb{Q}^U$ and $\mathbb{Q}^L$. Thus, one only need that the shape parameters be unchanged under $\mathbb{Q}$, $\mathbb{Q}^U$ and $\mathbb{Q}^L$.

The comparison between model and market implied relative spread is shown in figure 3.8(b). Overall, Υ_+ and Υ_- outperform Γ_+ and Γ_-. The calibrated parameters are

$$(b_p, c_p, b_n, c_n) = (0.0038, 614.5676, 0.0979, 3.7175$$

$$(b_p^U, b_n^U, b_p^L, b_n^L) = (0.0039, 0.0972, 0.0038, 0.0985),$$

and note that $b_p^U > b_p^L$ and $b_n^U < b_p^L$. Also, Υ_- dominates Υ_+, as was obtained with DM estimators.

3.5 Portfolio Choice

3.5.1 Optimality Conditions for a Small Investor

In this section, the spectral martingale measure of a portfolio of assets is determined based on theorem 3.3.9, and the wealth allocation maximizing the portfolio's lower valuation is solved for multivariate bilateral gamma log returns. We consider D securities $S = (S^1, ..., S^D)$, satisfying

$$\frac{dS_t^i}{S_{t-}^i} = a_i dt + \int_{\mathbb{R}^D \setminus \{0\}} \left(e^{R^i y} - 1 \right) \tilde{N}(dy, dt).$$

where $R^i \in \mathbb{R}^D$ is a row vector. Given $L_0, ..., L_D \in \mathbb{R}_+$, let

$$B = \left\{ x \in \mathbb{R}^D : x_i \geq -L_i,\ \sum_{i=1}^{D} x_i \leq 1 + L_0 \right\}.$$

Recall that a predictable strategy $\theta : [0,T] \times \Omega \to B$ is called a feedback allocation strategy if there is $\hat{\theta} : [0,T] \times \mathbb{R}_+ \to B$, called a feedback, such that $\theta_t(\omega) = \hat{\theta}\left(t \wedge \tau_{\hat{\theta}}, \Pi_{t \wedge \tau_{\hat{\theta}}}(\omega)\right)$, $\tau_{\hat{\theta}} := \inf\{t \in [0,T] : \Pi_t^{\hat{\theta}} \leq 0\}$ and the SDE

$$\begin{cases} \frac{d\Pi_t^{\hat{\theta}}}{\Pi_{t-}^{\hat{\theta}}} & = \theta_t^T a dt + \int_{\mathbb{R}^D \setminus \{0\}} \theta_t^T (e^{Ry} - 1) \tilde{N}(dy, dt), \\ \Pi_0^{\hat{\theta}} & = \pi,\ \Pi_t^{\hat{\theta}} = \Pi_{t \wedge \tau_{\hat{\theta}}}^{\hat{\theta}},\ t \in [T \wedge \tau_{\hat{\theta}}, T] \end{cases} \tag{3.5.1}$$

admits a unique solution, where R is the matrix whose rows are the vectors R^i. The set of all feedbacks is denoted by $\hat{\Theta}$ and that of feedback allocation strategies by $\mathcal{A}$. We consider a small investor interested in choosing the feedback that delivers the portfolio with maximal lower valuation.

Specifically, the investor solves at each time $t \in [0, T]$

$$\max_{\theta \in \mathcal{A}} \mathcal{L}_t[\Pi_T^\theta]. \tag{3.5.2}$$

Based on theorem 3.3.9, for $\theta \in B$, $y \in \mathbb{R}^D \setminus \{0\}$ and $(t, \omega) \in [0, T] \times \Omega$, set

$$\psi^L(\theta, y) = \Gamma'_-(\nu(\theta^T e^v > \theta^T e^{Ry})1_{\{\theta^T(e^{Ry}-1)>0\}} - \Gamma'_+(\nu(\theta^T e^{Rv} - 1 < \theta^T e^{Ry})1_{\{\theta^T(e^{Ry}-1)<0\}}$$

If $L_0 = L_1 = ... = L_D = 0$ (i.e. no short selling), by theorem 6.5 in [71],[9] the optimal control θ_* is independent of (t, ω) and satisfies

$$\theta_* \in \arg\sup_{\theta \in B} \left\{ \theta^T a - \int_{\mathbb{R}^D \setminus \{0\}} \theta^T (e^{Ry} - 1)\psi^L(\theta, y)\nu(dy) \right\}, \tag{3.5.3}$$

while the value function $V(t, \Pi_t^\theta) = \mathcal{L}_t[\Pi_T^\theta]$ is given by $V(t, \Pi_t^\theta) = C(t)\Pi_t^\theta$, where

$$C(t) = e^{(T-t)\left(\theta_*^T a - \int_{\mathbb{R}^D \setminus \{0\}} \theta_*^T (e^{Ry}-1)\psi^L(\theta_*(t,\omega)^T, y)\nu(dy)\right)}.$$

3.5.2 Numerical Results

We solved problem 3.5.3 assuming log returns have a multivariate BG process (see [69]) and Γ_- and Γ_+ are as in 3.2.3 with $a = 1/c$, $b = 1$. With no dependence,

$$\nu(\theta^T(e^v - 1) > \varpi)1_{\{\varpi>0\}} = \sum_{k=1}^{D} \nu_{pk}(\varpi),\ \nu(\theta^T(e^v - 1) < \varpi)1_{\{\varpi<0\}} = \sum_{k=1}^{D} \nu_{nk}(\varpi),$$

[9] In [71] each asset is a stochastic exponential, but it is assumed that ν is supported on $(-1, \infty)^D$.

where

$$\nu_{pk}(\varpi) = \left[c_{pk} E_1 \left(\frac{\log(\varpi/\theta_k + 1)}{b_{pk}} \right) 1_{\{\theta_k > 0\}} + c_{nk} E_1 \left(\frac{-\log(\varpi/\theta_k + 1)}{b_{nk}} \right) 1_{\{\theta_k < -\varpi\}} \right] 1_{\{\varpi > 0\}}$$
$$\nu_{nk}(\varpi) = \left[c_{nk} E_1 \left(\frac{-\log(\varpi/\theta_k + 1)}{b_{nk}} \right) 1_{\{\theta_k > -\varpi\}} + c_{pk} E_1 \left(\frac{\log(\varpi/\theta_k + 1)}{b_{pk}} \right) 1_{\{\theta_k < 0\}} \right] 1_{\{\varpi < 0\}},$$

E_1 is, as before, the exponential integral function:

$$E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt,$$

and $b_{pj}, c_{pj}, b_{nj}, c_{nj}$, $j = 1, ..., D$ are the BG parameters of asset j. Hence

$$\int_{\mathbb{R}^D \setminus \{0\}} \theta^T (e^y - 1) \psi^L(\theta, y) \nu(dy) = \sum_{j=1}^D \left[\int_0^\infty \theta_j (e^{y_j} - 1) e^{-c \sum_{k=1}^D \nu_{pk}(\theta_j (e^{y_j} - 1))} \kappa_j(y_j) dy_j \right.$$
$$\left. - \int_{-\infty}^0 \frac{\theta_j (e^{y_j} - 1)}{1 + \gamma} \left(1 - e^{-c \sum_{k=1}^D \nu_{nk}(\theta_j (e^{y_j} - 1))} \right)^{-\gamma/(1+\gamma)} e^{-c \sum_{k=1}^D \nu_{nk}(\theta_j (e^{y_j} - 1))} \kappa_j(y_j) dy_j \right]. \tag{3.5.4}$$

Finally, we have

$$\theta^T a = \sum_{j=1}^D \int_{\mathbb{R} \setminus \{0\}} \theta_j (e^{y_j} - 1) \kappa_j(y_j) dy_j = \sum_{j=1}^D \theta_j \log[(1 - b_{pj})^{-c_{pj}} (1 + b_{nj})^{-c_{nj}}]. \tag{3.5.5}$$

When the joint log-price process is a multivariate BG process, the asset S^i is given by

$$S^i = e^{X^i + Y^i},$$

where $Y^1, ..., Y^D$ are independent BG processes with parameters $\{\tilde{b}_{pi}, \tilde{c}_{pi}, \tilde{b}_{ni}, \tilde{c}_{ni}\}_{i=1,...,D}$, and $X = (X^1, ..., X^D)$ is a multivariate variance gamma process, i.e. X is a multivariate Brownian

motion with mean vector $\vartheta = (\vartheta_1, ..., \vartheta_D)$ defined, for a given $\zeta > 0$ and $i = 1, ..., D$, by

$$\vartheta_i = \frac{\tilde{b}_{pi} - \tilde{b}_{ni}}{\zeta},$$

and diffusion matrix $\Sigma = diag(\vec{\sigma})Cdiag(\vec{\sigma})$, where $C \in \mathbb{R}^{D\times D}$ is a given correlation matrix and $\sigma \in \mathbb{R}^D$ is determined, for $i = 1, ..., D$, by

$$\sigma_i^2 = \frac{2\tilde{b}_{pi}\tilde{b}_{ni}}{\zeta}.$$

In this case, S^i is a BG process with parameters $(b_{pi}, c_{pi}, b_{ni}, c_{ni})$, $i = 1, ..., D$, given by

$$b_{pi} = \tilde{b}_{pi},\ c_{pi} = \tilde{c}_{pi} - \frac{1}{\zeta},\ b_{ni} = \tilde{b}_{ni},\ c_{ni} = \tilde{c}_{ni} - \frac{1}{\zeta},$$

and the joint Levy measure ν of the log returns satisfies

$$\nu(A) = \int_A \kappa_{VG}(y)dy + \sum_{i=1}^{D} \int_{y_i:(0,...,0,y_i,0,...,0)\in A} \kappa_i(y_i)dy_i$$

where, for $y \in \mathbb{R}^D \setminus \{0\}$, $y_i \in \mathbb{R} \setminus \{0\}$, $i = 1, ..., D$,

$$\kappa_{VG}(\vec{y}) = \frac{2e^{\vartheta^T\Sigma^{-1}\vec{y}}\left(\sqrt{\left(\vartheta^T\Sigma^{-1}\vartheta + \frac{2}{\zeta}\right)(\vec{y}^T\Sigma^{-1}\vec{y})}\right)^{n/2}}{\zeta(2\pi)^{D/2}\sqrt{|\Sigma|}}$$
$$\times K_{D/2}\left(\sqrt{\left(\vartheta^T\Sigma^{-1}\vartheta + \frac{2}{\zeta}\right)(\vec{y}^T\Sigma^{-1}\vec{y})}\right),$$
$$k_i(y_i) = \frac{c_{ni} - \frac{1}{\zeta}}{|y_i|}e^{-|y_i|/b_{ni}}1_{\{y_i<0\}} + \frac{c_{pi} - \frac{1}{\zeta}}{y_i}e^{-y_i/b_{pi}}1_{\{y_i>0\}},$$

and where $K_{D/2}(\cdot)$ denotes the Bessel function of second kind of order $D/2$. Then,

$$
\begin{aligned}
&\int_{\mathbb{R}^D\setminus\{0\}} \theta^T(e^y-1)\psi^L(\theta,y)\nu(dy)\\
&= \int_{\{\theta^T(e^u-1)>0\}} \theta^T(e^y-1)e^{-c\left[\nu_p(\theta^T(e^y-1))+\sum_{k=1}^D \nu_{pk}(\theta^T(e^y-1))\right]}\nu_{VG}(dy)\\
&\quad - \int_{\theta^T(e^u-1)<0\}} \frac{\theta^T(e^y-1)}{1+\gamma}\left(1-e^{-c\left[\nu_n(\theta^T(e^y-1))+\sum_{k=1}^D \nu_{nk}(\theta^T(e^y-1))\right]}\right)^{-\gamma/(1+\gamma)}\\
&\qquad\qquad e^{-c\left[\nu_n(\theta^T(e^y-1))+\sum_{k=1}^D \nu_{nk}(\theta^T(e^y-1))\right]}\nu_{VG}(y)dy\\
&\quad + \sum_{j=1}^2\left[\int_0^\infty \theta_j(e^{y_j}-1)e^{-c\left[\nu_p(\theta_j(e^{y_j}-1))+\sum_{k=1}^2 \nu_{pk}(\theta_j(e^{y_j}-1))\right]}\kappa_j(y_j)dy_j\right.\\
&\qquad - \int_{-\infty}^0 \frac{\theta_j(e^{y_j}-1)}{1+\gamma}\left(1-e^{-c\sum_{k=1}^2 \nu_{nk}(\theta_j(e^{y_j}-1))}\right)^{-\gamma/(1+\gamma)}\\
&\qquad\qquad \left. e^{-c\left[\nu_n(\theta_j(e^{y_j}-1))\sum_{k=1}^2 \nu_{nk}(\theta_j(e^{y_j}-1))\right]}\kappa_j(y_j)dy_j\right].
\end{aligned}
$$

For simplicity, we assume that only the independent BG processes are distorted, so that we obtain again the same risk charge as in the case of independent BG processes, while the non-distorted variation is now

$$\theta^T a = \int_{\mathbb{R}^D\setminus\{0\}} \theta^T(e^y-1)\nu_{VG}(dy) + \sum_{j=1}^D \theta_j \log[(1-b_{pj})^{-c_{pj}}(1+b_{nj})^{-c_{nj}}]. \tag{3.5.6}$$

To estimate the parameters of the MBG model, we matched tail probabilities with those implied by data. We refer to [72] for the details of how to apply this method to estimating MBG parameters.

We solved the optimization problem over the five year period comprised between January 2016 and December 2020, with the assets being the 10 ETFs with tickers SPY, XLB, XLE, XLF,

XLI, XLK, XLP, XLU, XLV, XLY. The BG parameters were estimated daily based on the previous one year of data, while the correlation parameters were estimated based on data between January 1 and December 31 2015, and kept constant. For the distortion parameters we used as before $(c, \gamma, a, b) = (0.01, 0.25, 100, 0.1)$, leaving the task of finding those that give the best result (e.g. in terms of cumulated returns) for future research. Sharpe ratio quantiles are shown in table 3.4, together with those of SPY and those obtained assuming the ETFs are independent BG processes.

Percentile	MBG	BG	SPY
0.25	0.5484	0.5293	0.4460
0.5	1.0745	0.9831	0.9069
0.75	1.6757	1.6127	1.6982

Table 3.4: Sharpe ratio percentiles.

Even with this rather crude estimate of C and ζ, the maximal lower valuation portfolio outperformed the SPY Sharpe ratio on average, and especially during times of market downturns, and better results were achieved than with the simpler independent BG process assumption.

3.6 From Conic to Convex Finance

3.6.1 Dynamic Rebated Spectral Risk Measures

Attention was so far restricted to the case in which only the direction of the trade affects the price of an asset. On the other hand, however, large trades operated by institutional investors can and do often move the markets, or, in other words, the size of the trade can also affect valuations. Therefore, portfolio construction for a large and/or highly leveraged investor requires to abandon the conic finance framework, as the valuation functional is not just non additive but also non homogeneous, and the set of acceptable risks, although still convex, is no longer a cone.

In a static setting, following [67], for a family of probability distortions $\{\Psi_c\}_{c>0}$, one can set

$$\mathcal{M}_c = \{\mathbb{Q}' \in \Delta(\mathbb{P}) : \mathbb{Q}'(A) \leq \Psi_c(\mathbb{P}(A)) \; \forall A \in \mathcal{F}\}$$

where $\Delta(\mathbb{P})$ is the set of probability measures on $(\Omega, \mathcal{F})$ absolutely continuous w.r.t. $\mathbb{P}$. Given a rebate function $b : \mathbb{R}^+ \to \mathbb{R}^+$, a claim C is said to be acceptable if it satisfies

$$\mathbb{E}^{\mathbb{Q}'}[C] \geq -b(c) \tag{3.6.1}$$

for every $\mathbb{Q}' \in \mathcal{M}_c$ and for every $c > 0$. The lower valuation of the claim C is then defined as

$$\mathcal{L}[C] = \inf_{c>0} \left\{ \inf_{\mathbb{Q}' \in \mathcal{M}_c} \mathbb{E}^{\mathbb{Q}'}[C] + b(c) \right\}. \tag{3.6.2}$$

It is important to note that, based on 3.6.1, given an acceptable claim C, λY is acceptable for every $\lambda \geq 0$ if and only if $\mathbb{E}^{\mathbb{Q}'}[C] \geq 0$ for every $\mathbb{Q}' \in \mathcal{M}_c$ and for every $c > 0$, which obviously does not hold in general. That is, if C is traded, it may not be the case that any multiple of it is also traded.

Consider the family of measure distortions $\{\Gamma_{+,c}, \Gamma_{-,c}\}_{c>0}$ in 3.2.3, for fixed γ, $a = 1/c$, $b = 1$, and note that $-\Gamma_{+,c}$ and $\Gamma_{-,c}$ are decreasing in c. Then, $\exists \psi^{L,c} : L^2(\nu) \to C^g$ such that

$$\begin{aligned}\mathcal{L}_t^c[f(X_T)] = f(X_T) &- \int_t^T \int_{\mathbb{R}^D \setminus \{0\}} \psi^{L,c}((Z_s(\omega, y), y) Z_s(\omega, y) \nu(dy) ds \\ &- \int_t^T \int_{\mathbb{R}^D \setminus \{0\}} Z_s^c(\omega, y) \tilde{N}(ds, dy),\end{aligned}$$

where the superscript c denotes that the underlying measure distortions are $\Gamma_{+,c}$ and $\Gamma_{-,c}$. A natural

generalization of 3.6.2 can then be achieved by setting, for each $z \in \mathcal{L}^2(\nu)$,

$$c(z) := \arg\sup_{c>0} \int_{\mathbb{R}^D \setminus \{0\}} \psi^{L,c}(z,y) z(y) \nu(dy) - b(c).$$

One can then define a driver function g on $L^2(\nu)$ by

$$g(z) := \int_{\mathbb{R}^D \setminus \{0\}} \psi^{L,c(z)}(z,y) z(y) \nu(dy) - b(c(z), \tag{3.6.3}$$

and a lower valuation as the solution, if it exists unique, of

$$\begin{cases} \mathcal{L}_t[f(X_T)] = f(X_T) - \int_t^T g(Z_s(\omega,\cdot))ds - \int_t^T \int_{\mathbb{R}^D \setminus \{0\}} Z_s(\omega,y)\tilde{N}(ds,dy) \\ \mathcal{L}_T[f(X_T)] = f(X_T). \end{cases} \tag{3.6.4}$$

We refer to this class of risk measures as "*dynamic rebated spectral risk measures*".

Existence of Lower Valuations

Theorem 3.6.1. *Equation 3.6.4 admits a unique solution if the rebate b is continuous on an interval $(c_\ell, c_u] \subset \mathbb{R}_+$, $\lim_{c \to c_\ell} b(c) = +\infty$ and $b(c) = +\infty$ for $c \leq c_\ell$ and $b(c) = 0$ for $c \geq c_u$.*

Proof. It is sufficient to show (see e.g. [91]) that the driver function $g : L^2(\nu) \to L^2(\nu)$ defined by 3.6.3 is Lipschitz continuous, where

$$g^c(z) = \int_{\mathbb{R}^D \setminus \{0\}} \psi_z^{L,c} z(y) \nu(dy),$$

$$\psi_z^{L,c}(y) = \Gamma'_{-,c}(\nu(A^{z(y)})) 1_{\{z(y)>0\}} - \Gamma'_{+,c}(\nu(A^{z(y)})) 1_{\{z(y)<0\}}.$$

By the assumptions on the rebate function b, we have

$$g(z) = \sup_{c\in[c_\ell,c_u]} g^c(z) - b(c) \geq g^{c_u}(z) - b(c_u) = g^{c_u}(z) \geq 0.$$

Furthermore, continuity of b and the dominated convergence theorem imply that, for each $z \in L^2(\nu)$, the function $c \to g^c(z) - b(c)$ is continuous on $(c_\ell, c_u]$. Since, in addition, $\lim_{c\to c_\ell} g^c(z) - b(c) = -\infty$, there must be a map $c : L^2(\nu) \to [c_\ell, c_u]$ such that, for each $z \in L^2(\nu)$,

$$g(z) = g^{c(z)}(z) - b(c(z)).$$

Hence, for every $z_1, z_2 \in L^2(\nu)$, such that, without loss of generality, $g(z_1) \geq g(z_2)$, we have

$$\begin{aligned}
|g(z_1) - g(z_2)| = g(z_1) - g(z_2) &= g^{c(z_1)}(z_1) - b(c(z_1)) - g^{c(z_2)}(z_2) + b(c(z_2)) \\
&\leq g^{c(z_1)}(z_1) - b(c(z_1)) - g^{c(z_1)}(z_2) + b(c(z_1)) \\
&= g^{c(z_1)}(z_1) - g^{c(z_1)}(z_2) \\
&= \int_{\mathbb{R}^D\setminus\{0\}} \left[\psi_{z_1}^{L,c(z_1)}(y)z_1(y) - \psi_{z_2}^{L,c(z_1)}(y)z_2(y)\right] \nu(dy) \\
&\leq \int_{\mathbb{R}^D\setminus\{0\}} \left[\psi_{z_2}^{L,c(z_1)}(y)z_1(y) - \psi_{z_2}^{L,c(z_1)}(y)z_2(y)\right] \nu(dy) \\
&\leq \int_{\mathbb{R}^D\setminus\{0\}} |\psi_{z_2}^{L,c(z_1)}(y)||z_2(y) - z_1(y)|\nu(dy) \\
&\leq \|\psi_{z_2}^{L,c(z_1)}\|_{L^2(\nu)}\|z_1 - z_2\|_{L^2(\nu)} \\
&\leq K\|z_1 - z_2\|_{L^2(\nu)},
\end{aligned}$$

where $K = \sup\{\|\psi\|_{L^2(\nu)} : \psi \in C^{g^c}\ \forall c \in [c_\ell, c_u]\} < \infty$ from the proof of lemma 3.3.2. □

Example. An example of rebate satisfying the above assumptions is, for $\chi, \chi_2 > 0$,

$$b(c) = \begin{cases} 0 \text{ if } c \geq c_u, \\ \chi e^{\left(\frac{1}{c-c_\ell}\right)^{\chi_2} - \left(\frac{1}{c_u - c}\right)^{\chi_2}} \text{ if } c_u > c > c_\ell, \\ +\infty \text{ if } c \leq c_\ell. \end{cases} \tag{3.6.5}$$

Concavity of Lower Valuations

Having established the existence of a lower valuation, we now prove that $\mathcal{L}_t[f(X_T)]$ is concave in the time t value $X_t = x$ of the underlying Markov process X. To do so, it is sufficient to show that the driver function g is a convex functional (see [29], proposition 6.2.3).[10]

Proposition 3.6.2. *Suppose b satisfies the assumptions of theorem 3.6.1 and that it is, in addition, strictly convex on $(c_\ell, c_u]$. Then, the driver function g defined by 3.6.3 is convex in z.*

Proof. Fix $z_1, z_2 \in L^2(\nu)$, $\lambda \in [0,1]$, $\overline{\lambda} = 1 - \lambda$. As before, there is $c(\lambda z_1 + \overline{\lambda} z_2)$ such that

$$\begin{aligned} g(\lambda z_1 + \overline{\lambda} z_2) &= g^{c(\lambda z_1 + \overline{\lambda} z_2)}(\lambda z_1 + \overline{\lambda} z_2) - b(c(\lambda z_1 + \overline{\lambda} z_2)) \\ &\leq \lambda g^{c(\lambda z_1 + \overline{\lambda} z_2)}(z_1) + \overline{\lambda} g^{c(\lambda z_1 + \overline{\lambda} z_2)}(z_2) - b(c(\lambda z_1 + \overline{\lambda} z_2)) \\ &= \lambda \left[g^{c(\lambda z_1 + \overline{\lambda} z_2)}(z_1) - b(c(\lambda z_1 + \overline{\lambda} z_2)) \right] + \overline{\lambda} \left[g^{c(\lambda z_1 + \overline{\lambda} z_2)}(z_2) - b(c(\lambda z_1 + \overline{\lambda} z_2)) \right] \\ &\leq \lambda \left[g^{c(z_1)}(z_1) - b(c(z_1)) \right] + \overline{\lambda} \left[g^{c(z_2)}(z_2) - b(c(z_2)) \right] = \lambda g(z_1) + \overline{\lambda} g(z_2). \end{aligned}$$

where the first inequality follows from the fact that g^c is sublinear for every $c > 0$. □

[10] Note in particular from the proof of theorem 3.6.1 that the driver function g satisfies the gradient condition (assumption (ii) in Definition 6.2.3 in [29]), with $\delta^{z,z'} = \psi_{z'}^{L,c(z)}$, which is bounded in $L^2(\nu)$ uniformly in z, z' as in the proof of lemma 3.3.2 since $c(z) \in [c_\ell, c_u]$.

3.6.2 Optimality Conditions for a Large Investor

Based on the results of the previous section and standard arguments, the value function V and the corresponding optimal Markovian rule θ^* for the portfolio with maximal lower valuation are determined by the following HJB equation

$$\begin{aligned} \dot{V} + \sup_{\theta \in B} \left\{ \int_{\mathbb{R}^D \setminus \{0\}} \mathcal{D}^{\theta}_{t,\varpi} V(y) \nu(dy) - g(\mathcal{D}^{\theta}_{t,\varpi} V) \right\} = 0,\ (t, \varpi) \in [0,T] \times \mathbb{R}_+, \\ V(t,\varpi) = \varpi,\ (t,\varpi) \in [0,T] \times \mathbb{R}_-, \\ V(T,\varpi) = \varpi,\ \varpi \in \mathbb{R}. \end{aligned} \tag{3.6.6}$$

where B is a compact subset of the $D+1$ dimensional symplex and

$$\begin{aligned} c(t,\varpi) &= \arg \sup_{c \in [c_\ell, c_u]} \int_{\mathbb{R}^D \setminus \{0\}} \psi^{L,c}(t,\varpi,y) \mathcal{D}^{\theta}_{t,x} V(y) \nu(dy) - b(c), \\ \mathcal{D}^{\theta}_{t,\varpi} V(y) &= V(t, \varpi + \theta^T (e^{Ry} - 1)) - V(t,\varpi). \end{aligned}$$

In particular, the valuation function is concave in the current value ϖ of the portfolio (as for expected utility maximization), and the optimal rule θ^* is determined by

$$\theta_*(t,\varpi) = \arg \sup_{\theta \in B} \left\{ \int_{\mathbb{R}^D \setminus \{0\}} \mathcal{D}^{\theta^*}_{t,\varpi} V(y) \nu(dy) - g(V(t,\varpi,\cdot)) \right\} \tag{3.6.7}$$

The Diminishing Returns to Scale of Concave Lower Valuations

In classical general equilibrium theory, diminishing returns to scale for a given technology model scarcity of physical production factors, such as labor, machinery or raw materials, and/or

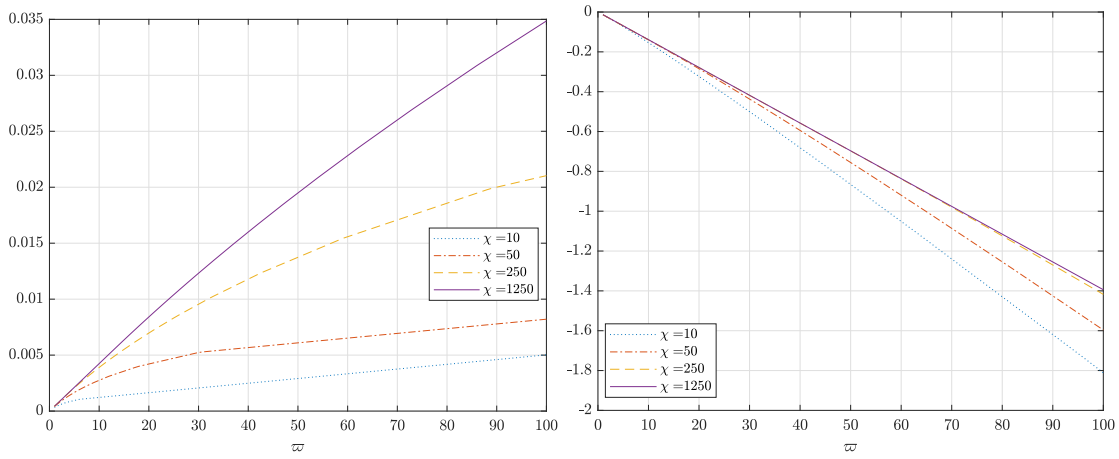
of demand capacity. Here, such diminishing returns are achieved through the introduction of the rebate function, which therefore represents an alternative to the traditional behavioral assumption, often violated in practice, that investors be risk averse in the sense of expected utility theory.

To illustrate the effect on the lower valuation of a portfolio due to the introduction of the rebate, consider for a fixed time horizon T the measure distorted variation $\mathcal{L}(\varpi)$ of an asset with current value ϖ and BG log-return. Such quantity is given by

$$\begin{aligned}\mathcal{L}(\varpi) = \varpi a - \sup_{c\in[c_\ell,c_u]} \Bigg[&\int_0^\infty \varpi(e^y-1)\Gamma'_{-,c}\left(c_p E_1\left(\tfrac{y}{b_p}\right)\right)\kappa(y)dy \\ &- \int_{-\infty}^0 \varpi(e^y-1)\Gamma'_{+,c}\left(c_n E_1\left(\tfrac{|y|}{b_n}\right)\right)\kappa(y)dy - b(c)\Bigg],\end{aligned} \tag{3.6.8}$$

and it coincides with the lower valuation as the time horizon converges to 0. In a dynamic setting, $\mathcal{L}(\varpi)$ is also the first step of the finite difference scheme for the solution of the semilinear valuation PIDE with concave driver given by the functional g defined in 3.6.3. Figure 3.9 shows the diminishing returns to scale of the capital invested ϖ for different choices of the parameter χ and the bounds c_u and c_ℓ. The BG parameters correspond to those estimated for SPY as of 31 December 2020, i.e. $(b_p, c_p, b_n, c_n) = (0.0075, 1.5592, 0.0181, 0.6308)$. As expected, the resulting plots are concave. It is worth noting, in addition, that, for the rebated distorted variation to be increasing in the amount ϖ of capital invested, the bound c_ℓ needs to be high enough, so that there are no values of $c \in [c_\ell, c_u]$ such that

$$\begin{aligned}\varpi a + b(c) - &\int_0^\infty \varpi(e^y-1)\Gamma'_{-,c}\left(c_p E_1\left(\tfrac{y}{b_p}\right)\right)\kappa(y)dy \\ &+ \int_{-\infty}^0 \varpi(e^y-1)\Gamma'_{+,c}\left(c_n E_1\left(\tfrac{|y|}{b_n}\right)\right)\kappa(y)dy < 0.\end{aligned}$$



(a) (b)

Figure 3.9: The diminishing return to the scale of capital ϖ invested for BG parameters as of 31 December 2020 for several values of χ. The bounds for c are $c_u = 1000$, $c_\ell = 200$ in (a) and $c_u = 1$, $c_\ell = 0.01$ in (b). The other distortion parameters are $\gamma = 0.01$, $a = 1/c$ and $b = 1$.

3.6.3 Myopic Risk-Sensitive Asset Management

With multiple risky assets, maximizing the lower valuation generates an internal solution as the value of the portfolio itself is distorted. In this section we illustrate how the introduction of a rebate can be utilized to risk-sensitize the wealth allocation problem when an investor chooses between a risky asset and a risk free one. Specifically, our goal is to show that maximization of the rebated spectral risk measure is consistent with that of other, well known, risk measures in the numerically simple case of one risky asset. To avoid further numerical difficulties, this analysis is limited to the myopic risk-sensitive portfolio problem. The full solution of the HJB equation 3.6.6 with long term (i.e. infinite) horizon is left for future research.

Denoting by r the risk free rate, an investor maximizing the expected net return over the horizon T solves

$$\max_{\theta\in[0,1]} \left\{ \varpi(1-\theta)(e^{rT}-1) + \varpi\theta \int_{\mathbb{R}} (e^x - 1) F_T(dx) \right\},$$

where F_T is the cdf of a BG distribution with parameters (b_p, c_pT, b_n, c_nT). Thus $\theta^* = 0$ or $\theta^* = 1$, depending on whether the risk free rate is larger than the risky log return or not. In the conic finance framework, and for a given family of probability distortions $\{\psi_\gamma\}_{\gamma>0}$, the investor solves

$$\max_{\theta\in[0,1]} \left\{ \varpi(1-\theta)(e^{rT}-1) + \varpi\theta \int_{\mathbb{R}} (e^x - 1)\psi'_\gamma(F(x))f(x)dx \right\},$$

where f_T is the density of F_T, and again one obtains a bang-bang solution. Introducing a rebate as before, the problem becomes

$$\max_{\theta\in[0,1]} \left\{ \varpi(1-\theta)(e^{rT}-1) + \min_{\gamma>0} \left\{ b(\gamma) + \varpi\theta \int_{\mathbb{R}} (e^x - 1)\psi'_\gamma(F(x))f(x)dx \right\} \right\}, \tag{3.6.9}$$

which, in general, admits an internal solution. Here we use the MINVAR distortion ψ defined by

$$\psi_\gamma(x) = 1 - (1-x)^{1+\gamma}.$$

This approach offers therefore an alternative to the traditional mathematical formulation of the risk sensitive portfolio choice problem based on the maximization of the certainty equivalent of a certain financial objective.

For instance, with exponential utility and log return objective (see [25], Chapter 2, and the references therein), the risk-sensitive portfolio choice problem is given by

$$\begin{aligned} &\max_{\theta\in[0,1]} \left\{ -\frac{1}{\epsilon} \log\left(\mathbb{E}\left[\left(\varpi(1-\theta)e^{rT} + \varpi\theta e^{X_T} \right)^{-\epsilon} \right] \right) \right\} \\ &= \log(\varpi) + \max_{\theta\in[0,1]} \left\{ -\frac{1}{\epsilon} \log\left(\mathbb{E}\left[\left((1-\theta)e^{rT} + \theta e^{X_T} \right)^{-\epsilon} \right] \right) \right\}. \end{aligned} \tag{3.6.10}$$

It is worth noting that, while the rebate function can be specified and calibrated based on an investment firm's capital requirements, the coefficient of absolute risk aversion is more difficult to estimate, and generally requires an analysis of consumption data.

Because of such limitations and given the relatively large literature on the estimate of the relative risk aversion parameter, it is also useful to specify the risk-sensitive allocation problem for a utility function with constant relative risk aversion. In this case it is more natural to assume that the certainty equivalent of the portfolio's return, rather than its log or net return, is maximized, and so such specification is

$$\begin{cases} \max_{\theta\in[0,1]} \left\{ \sqrt[1-\eta]{\mathbb{E}\left[(\varpi(1-\theta)e^{rT}+\varpi\theta e^{X_T})^{1-\eta}\right]} \right\} \text{ if } \eta \neq 1 \\ \max_{\theta\in[0,1]} \left\{ \exp\left\{ \mathbb{E}\left[\log\left(\varpi(1-\theta)e^{rT}+\varpi\theta e^{X_T} \right) \right] \right\} \right\} \text{ if } \eta = 1 \end{cases} \tag{3.6.11}$$

where $\eta > 0$ is the coefficient of relative risk aversion.

Finally, we specify the portfolio choice problem for the dynamic rebated lower valuation by considering the measure distorted variation as a financial objective rather than the probability distorted return (see [72]), i.e. the investor chooses to solve

$$\begin{aligned} \max_{\theta\in[0,1]} \Bigg\{ (1-\theta)\varpi rT + \theta\varpi aT + \sup_{c\in[c_\ell,c_u]} \Bigg[\theta T\varpi \Bigg(\int_0^\infty (e^y-1)\Gamma'_{-,c}\left(c_p E_1\left(\tfrac{y}{b_p}\right)\right)\kappa(y)dy \\ - \int_{-\infty}^0 \varpi(e^y-1)\Gamma'_{+,c}\left(c_n E_1\left(\tfrac{|y|}{b_n}\right)\right)\kappa(y)dy \Bigg) - b(c) \Bigg] \Bigg\}, \end{aligned} \tag{3.6.12}$$

The advantage of this approach, compared to the specification with a probability distortion, is mainly numerical, as the BG cdf can be expressed analytically only in terms of the Whittaker function or a Fourier integral, while its Levy density is available in closed form.

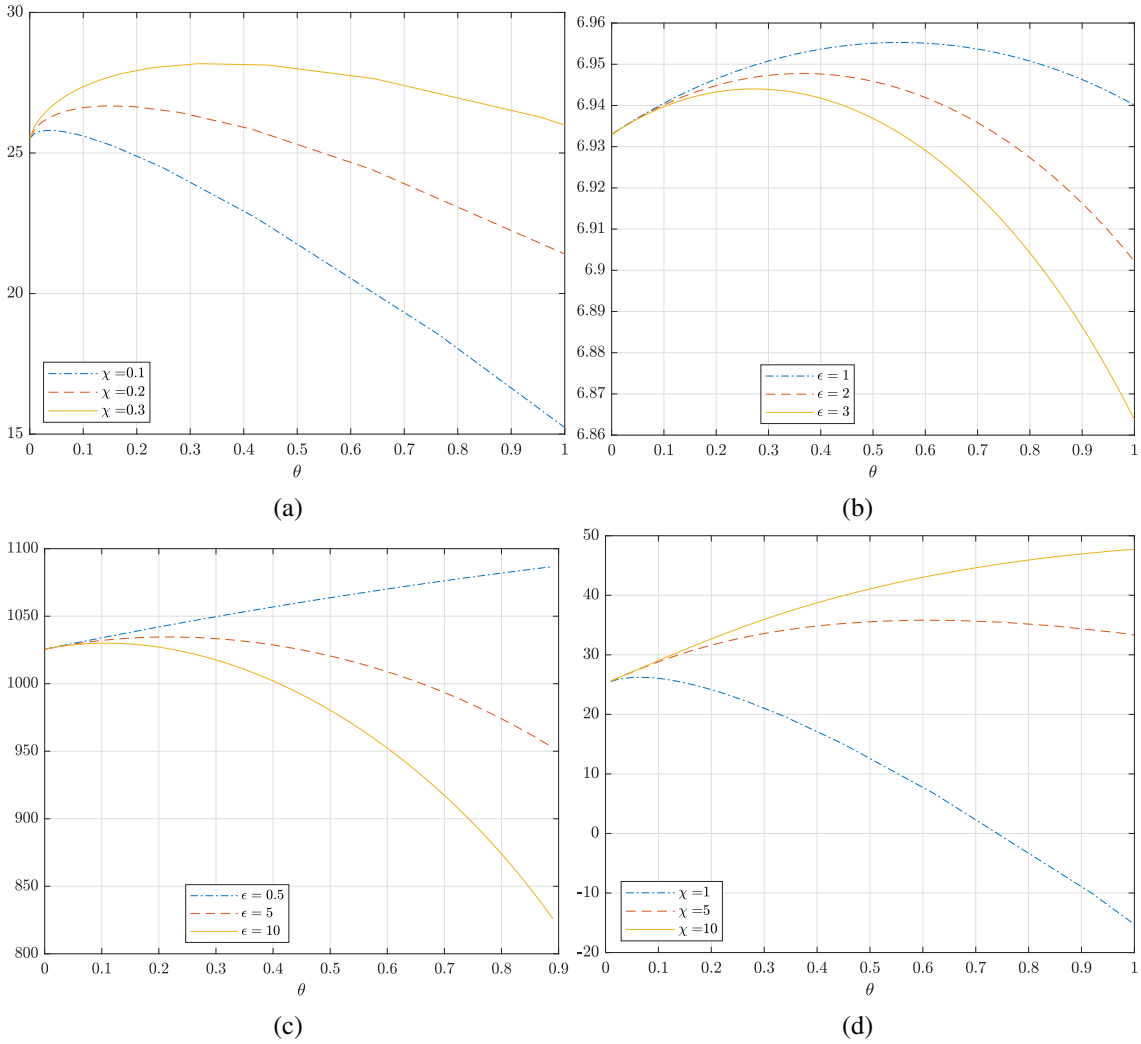


Figure 3.10: The rebated lower net return (a), the certainty equivalent of log return for exponential utility (b) and that of return for CRRA utility (c), and the rebated lower variation (d) for controls $\theta \in [0, 1]$ and for BG parameters calibrated to SPY as of 31 December 2020, and for a time horizon $T = 1$ year.

The value function for the four financial objectives and the SPY as risky asset, with BG parameters calibrated as of 31 December 2020, is plotted in figure 3.10. We assumed $\varpi = 1000$ and $T = 1$ day. The optimal allocation to the risky asset for each objective and for each day between 2 January 2020 and 31 December 2020 is shown in figure 3.11.

Remarks.

(i) The optimal control depends on ϖ with the rebated variation, while it does not with CE.

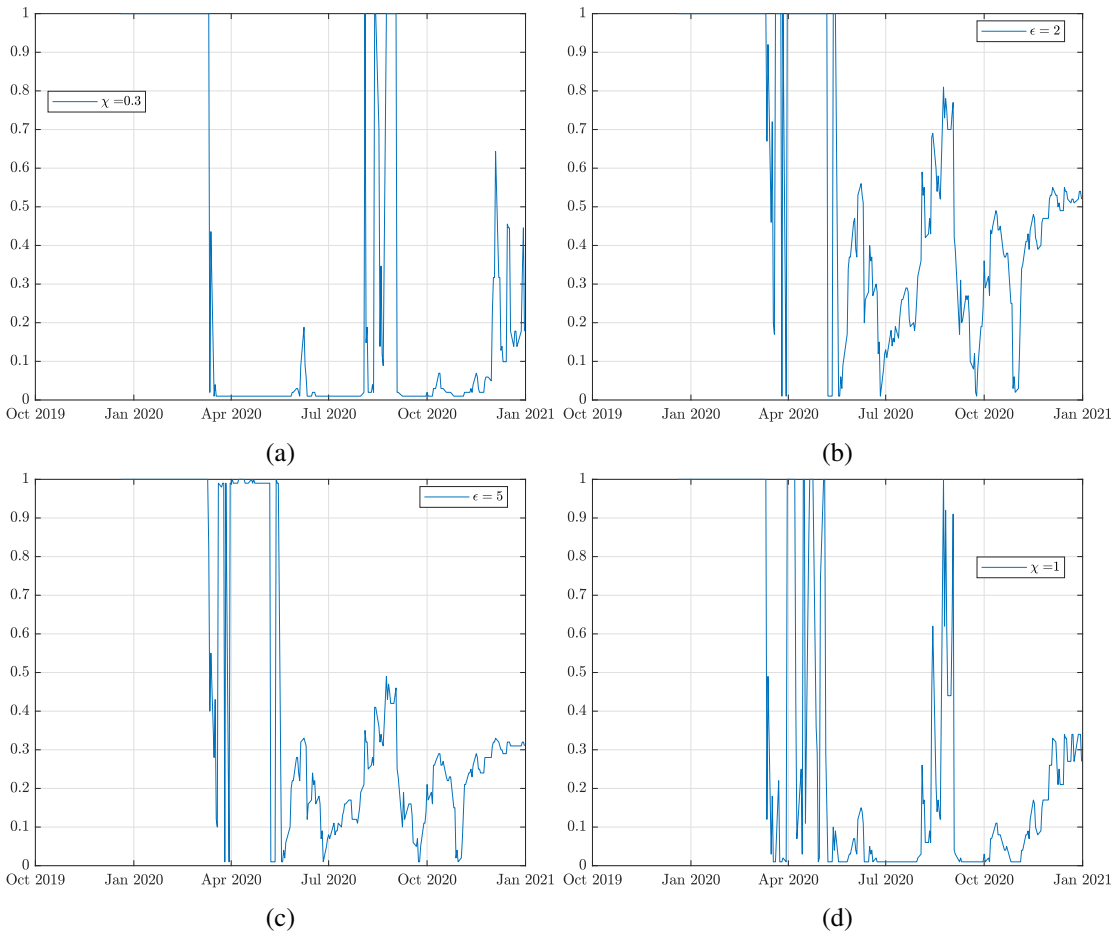


Figure 3.11: Optimal allocation in 2020 for (a) the rebated lower return, (b) the CE of log return for exponential utility and (c) that of return for CRRA utility, (d) the rebated lower variation. The risky asset is the SPY

(ii) For any given utility function, there exists a rebate such that the corresponding objective functions have the same level sets. For instance, with exponential utility, one can set

$$b(c,\theta,\varpi) = \varpi T\left[\theta\int_{\mathbb{R}\setminus\{0\}}(e^y-1)\psi^{L,c}(y)\nu(dy) - a\theta + (1-\theta)r\right] + \\ -\frac{1}{\epsilon}\log\left(\mathbb{E}\left[\left((1-\theta)e^{rT}+\theta e^{X_T}\right)^{-\epsilon}\right]\right).$$

(iii) The correlation between the optimal controls obtained is high, as reported in table 3.5.

	PD	CARA	CRRA	MD
PD	1	0.74	0.74	0.86
CARA		1	0.98	0.87
CRRA			1	0.89
MD				1

Table 3.5: Correlation among optimal controls.

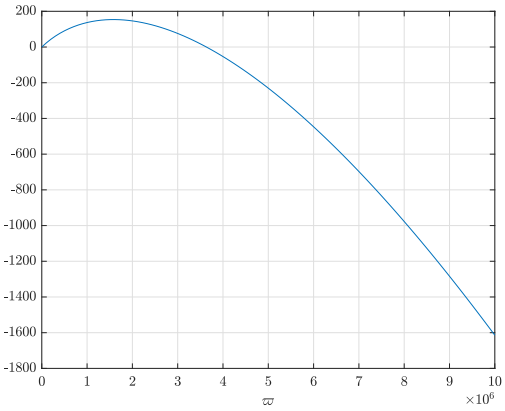
3.6.4 Optimal Investment Amount

In addition to the wealth allocation problem analyzed in previous sections, the diminishing returns to scale of rebated concave lower valuations, illustrated in figure 3.9, allows one to specify an optimization problem also for the amount to invest. We consider again a myopic investor and, based on the results of the previous section, we take the measure distorted variation as financial objective. We thus obtain the following maximization problem:

$$\max_{\theta,\varpi}\left\{\varpi T\theta^T a - \sup_{c\in[c_\ell,c_u]}\left\{\varpi T\int_{\mathbb{R}^D\setminus\{0\}}\theta^T(e^y-1)\psi^{L,c}(y)\nu(dy) - b(c)\right\}\right\}, \tag{3.6.13}$$

where $\theta^T a$ is as in 3.5.6 and the distorted variation as in 3.5.4.

As shown in figure 3.12, the introduction of the rebate and the diminishing return of capital



(a)

Figure 3.12: Rebated lower valuation as a function of total amount invested for a portfolio of the 10 ETFs with uniform weight and rebate parameters $c_u = 100$, $c_\ell = 2$. The underlying model is the multivariate bilatera gamma, with parameters estimated as of Febraury 28 2020.

may imply the existence of an optimal investment amount. In fact, for given θ, as long as there is $c \in [c_\ell, c_u]$ such that the lower valuation is negative for the distortion level at c, then

$$\lim_{\varpi \to \infty} \varpi T \theta^T a - \sup_{c \in [c_\ell, c_u]} \left\{ \varpi T \int_{\mathbb{R}^D \setminus \{0\}} \theta^T (e^y - 1) \psi^{L,c}(y) \nu(dy) - b(c) \right\} = -\infty,$$

and so an optimal investment amount exists if the variation is positive for lower values of ϖ, i.e. if the rebate is sufficiently high. To illustrate the effect of taking the value function in 3.6.13 as a financial objective, we solved problem 3.6.13 for each day between February 2 2020 and April 30 2020 and every 10 days between 2016 and 2020, that is we constructed optimal weights and optimal investment amount for a portfolio composed of the 10 ETFs for such periods. During the February-April 2020 period the pandemic hit the U.S. economy, and a risk sensitive investor is expected to take a more cautious approach and reduce amount invested. We assumed an MBG structure for the underlying assets, and, as before, we only distorted the independent BG component of each of them. For the rebate we took $c_u = 100$, $c_\ell = 2$, $\gamma = 0.01$, $a = 1/c$ and $b = 1$, which are the same parameters used to generate figure 3.12. The optimal investment amounts for the periods

Figure 3.13: SPY, optimal investment amounts and cumulated returns for February-April 2020 are shown in figures (a), (c) and (e), and for 2016-2020 in figures (b), (d), (e).

considered are shown in figure 3.13(b) and (d) , while the daily gross returns are shown in figure 3.13(c) and (e). For each day considered, as initial conditions for the optimization problem we set $\varpi = 2.5$ million and $\theta_i = 1/10$ for each $i = 1, ..., 10$. Given that the optimal investment amount over this period does not exceed 10 million dollars, we compared the daily gross returns from the portfolio so obtained with those of a 5 million dollars investment in the SPY ETF. Figure 3.13(a) depicts the SPY index, while figure 3.13(c) and (e) shows the daily gross return of a 10 million dollar investment in the SPY.

Remarks.

(i) The optimal amount tends to decrease in bear markets and increase in bull markets.

(ii) The return from the portfolio shows substantially less volatility and limited maximal total losses, as clearly illustrated in figures 3.13(c) and (e). In particular, SPY's daily losses can be much more prominent than those of the portfolio.

(iii) Positive returns are also more limited, as confirmed by the statistics for the daily Sharpe ratio, computed each day using returns from the previous 30 days, and shown in table 3.6. Note, however, that the piortfolio tends to outperform the SPY when markets are under stress.

	February-April 2020		2016-2020	
Percentile	MBG	SPY	MBG	SPY
0.05	-2.8746	-5.4765	-2.4120	-1.8424
0.25	-0.7023	-0.9201	-0.7023	-0.0932
0.5	-0.0932	0.0964	0.6232	0.6337
0.75	0.7559	0.8950	1.3001	1.7479
0.95	4.7515	9.7549	5.5594	6.9534

Table 3.6: Sharpe ratio percentiles for the period February to April 2020 and 2016-2020.

3.7 Conclusions

In this paper, we showed the existence of spectral martingale measures, which are defined as probability measures under which spectral risk measures are local martingales. We found an implicit formula for the jump density of the underlying semimartingale X, and analyzed its applications to one dimensional monotone claims and portfolios of exponential Levy processes, in which case our formula becomes explicit. Specifically, for one dimensional monotone claims, we showed that measure distortion parameters can be estimated through Generalized Method of Moments as well as Digital Moments, and presented some evidence that the latter outperform the former. Based on our estimates, the distortion of large losses tends to dominate that of large gains, and risk charges tend to increase in periods of market stress. When the risk neutral measure is distorted, similar results holds and, in addition, the distorted prices can be chosen to be BG processes if the non-distorted one is. We also analyzed and solved the portfolio choice problem with spectral risk measure as financial objective, and found that its solution can outperform the SPY under standard performance metrics. Finally we defined the Dynamic Rebated Spectral Risk Measures, which allow one to determine not just the optimal allocation but also the optimal investment amount. We showed in the case of a myopic risk sensitive investor that substantial losses can be avoided, although the Sharpe ratios are generally lower than SPY's.

3.8 Appendix

In this section, we provide a proof of theorem 3.3.9, which is here recalled together with the relevant definition of regular claim.

Definition. A claim $C \in L^2$ is called regular if the deterministic function z defined by the process Z in theorem 3.2.5, is uniformly continuous on $[0,T] \times \mathbb{R}^D$ uniformly in $y \in \mathbb{R}^D \setminus \{0\}$.

Theorem. *Let ν be σ-finite and suppose that $C \in L^2(\nu)$ is a regular claim. Then, C admits upper and lower spectral martingale measures.*

In the proof below note that differentiability of Γ_+ and Γ_- is not required in the proof of step 1 of theorem 3.3.7, which therefore still defines the maps ψ^U and ψ^L.

Proof. *Step 1. The map $(t,x) \to \psi^U(t,x,\cdot)$ is weakly continuous in $L^2(\nu)$.*

Proof of step 1. Define a map $F : [0,T] \times \mathbb{R}^D \to C^g$ by

$$F(t,x) := \psi^U(t,x,\cdot).$$

Suppose

$$\{(t_m, x_m)\}_{m\in\mathbb{N}} \subset [0,T] \times \mathbb{R}^D,$$

$(t_m, x_m) \to (t,x)$, and let

$$\{(t_{m_k}, x_{m_k})\}_{k\in\mathbb{N}} \subset \{(t_m, x_m)\}_{m\in\mathbb{N}}$$

be any subsequence. Then, since C^g is weakly compact, there is

$$\{(t_{m_{k_j}}, x_{m_{k_j}})\}_{j\in\mathbb{N}} \subset \{(t_{m_k}, x_{m_k})\}_{k\in\mathbb{N}}$$

and $\psi \in C^g$ such that

$$F(t_{m_{k_j}}, x_{m_{k_j}}) \to \psi$$

weakly in $L^2(\nu)$. Fix $\alpha > 0$. Since C is regular, $z_\alpha^+(\cdot, \cdot, y)$ is uniformly continuous and bounded on $[0, T] \times \mathbb{R}^D$ uniformly in y, and so for each $\varepsilon > 0$ there is $j_\varepsilon \in \mathbb{N}$ such that, for every $j > j_\varepsilon$,

$$\begin{aligned}
&\left| \int_{\mathbb{R}^D \setminus \{0\}} F(t_j, x_j)(y) z_\alpha^+(t_j, x_j, y) - \psi(y) z_\alpha^+(t, x, y) \nu(dy) \right| \\
&\leq \int_{\mathbb{R}^D \setminus \{0\}} F(t_j, x_j)(y) \left| z_\alpha^+(t_j, x_j, y) - z_\alpha^+(t, x, y) \right| \nu(dy) \\
&\quad + \left| \int_{\mathbb{R}^D \setminus \{0\}} (\psi(y) - F(t_j, x_j)(y)) z_\alpha^+(t, x, y) \nu(dy) \right| \\
&\leq \varepsilon \left[\Gamma^+(\infty) + \Gamma^-(\infty) \right] + \left| \int_{\mathbb{R}^D \setminus \{0\}} (\psi(y) - F(t_j, x_j)(y)) z_\alpha^+(t, x, y) \nu(dy) \right| \\
&\to \varepsilon \left[\Gamma^+(\infty) + \Gamma^-(\infty) \right],
\end{aligned}$$

where for easy of notation we wrote j for m_{k_j}. Since ε arbitrary, this implies

$$\int_{\mathbb{R}^D \setminus \{0\}} F(t_j, x_j)(y) z_\alpha^+(t_j, x_j, y) \nu(dy) \to \int_{\mathbb{R}^D \setminus \{0\}} \psi(y) z_\alpha^+(t, x, y) \nu(dy). \tag{3.8.1}$$

On the other hand, for every fixed $s > 0$ and $j > j_\varepsilon$,

$$s \leq z_\alpha^+(t_j, x_j, y)\ s + \varepsilon \leq z_\alpha^+(t, x, y),$$

which imply

$$s - \varepsilon \leq z_\alpha^+(t, x, y),\ s \leq z_\alpha^+(t_j, x_j, y)$$

Thus,

$$\nu\left(\{z_\alpha^+(t, x, y) > s + \varepsilon\}\right) \leq \nu\left(\{z_\alpha^+(t_j, x_j, y) > s\right) \leq \nu\left(\{z_\alpha^+(t_j, x_j, y) > s - \varepsilon\}\right)$$

which implies

$$\begin{aligned} &|\nu\left(\{z_\alpha^+(t_j, x_j, y) > s\}\right) - \nu\left(\{z_\alpha^+(t, x, y) > s\}\right)| \\ &\leq \nu\left(\{z_\alpha^+(t, x, y) > s - \varepsilon\}\right) - \nu\left(\{z_\alpha^+(t, x, y) > s + \varepsilon\}\right), \end{aligned} \tag{3.8.2}$$

and the right hand side in the above inequality tends to zero by the monotone convergence theorem as $\varepsilon \downarrow 0$. Then, since Γ^+ is continuous and bounded,

$$\begin{aligned} \int_0^\infty \Gamma^+(\nu(z_\alpha^+(t_j, x_j, y) > s))ds &= \int_0^\alpha \Gamma^+(\nu(z_\alpha^+(t_j, x_j, y) > s))ds \\ &\to \int_0^\infty \Gamma^+(\nu(z_\alpha^+(t, x, y) > s))ds, \end{aligned}$$

where the limit follows from the dominated convergence theorem. Therefore,

$$\begin{aligned} g(z_\alpha^+(t,x,\cdot)) &= \lim_{k\to\infty} g(z_\alpha^+(t_j,x_j,\cdot)) \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^D\setminus\{0\}} F(t_j,x_j)(y) z_\alpha^+(t_j,x_j,y)\nu(dy) \\ &= \int_{\mathbb{R}^D\setminus\{0\}} \psi(y) z_\alpha^+(t,x,y)\nu(dy), \end{aligned}$$

and similarly we obtain

$$g(z_\alpha^-(t,x,\cdot)) = \int_{\mathbb{R}^D\setminus\{0\}} \psi(y) z_\alpha^-(t,x,y)\nu(dy).$$

Thus, ψ satisfies 3.3.5. We next show that ψ is $\Sigma_{t,x}$ measurable. To do so, let $B \in \mathcal{B}(\mathbb{R})$ with $\nu(z^{-1}(t,x,B)) < \infty$. Since ν is σ-finite, for every $\varepsilon > 0$ there are rationals $\{q_{\ell,k}\}_{\ell\in L, k\in K}$ such that

$$\nu\left(z(t,x,\cdot)^{-1}(B) \Delta \bigcup_{\ell\in L} \bigcap_{k\in K} A^{q_{\ell,k}} \right) < \varepsilon,$$

where $A^{q_{\ell,k}}$ is the $q_{\ell,k}$ upper or lower level set of $z(t,x,\cdot)$, and Δ is the set difference operator (see e.g. theorem 11.4 in [11]). A basis for $L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t,x}, \nu_{t,x})$ is then given by functions of the form

$$1_{\{\cup_{\ell\in L}\cap_{k\in K} A^{q_{\ell,k}}\}},$$

where L and K are finite sets of natural numbers and $\{q_{\ell,k}\}$ is an enumeration of the rationals.

Similarly, a countable basis for $L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t_j,x_j}, \nu_{t,x})$, $j \in \mathbb{N}$, is given by functions of the form

$$1_{\{\cup_{\ell\in L}\cap_{k\in K} A_j^{q_{\ell,k}}\}},$$

where $A_j^{q_{\ell,k}}$ is a level set of $z(t_j, x_j, \cdot)$. As shown above, uniform continuity of z implies that

$$\nu_{t_j,x_j}(A_j^{q_{\ell,k}}) = \nu(A_j^{q_{\ell,k}}) \to \nu(A^{q_{\ell,k}}) = \nu_{t,x}(A^{q_{\ell,k}}).$$

A simple calculation shows then that

$$1_{\{\cup_{\ell\in L}\cap_{k\in K} A_j^{q_{\ell,k}}\}} \to 1_{\{\cup_{\ell\in L}\cap_{k\in K} A^{q_{\ell,k}}\}}, \tag{3.8.3}$$

strongly in $L^2(\nu)$, provided that the limit is finite. In the rest of the proof we denote by $1_{\{B_m\}}$ and $1_{\{B_m^j\}}$ the m-th elements of the bases of $L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t,x}, \nu_{t,x})$ and $L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t_j,x_j}, \nu_{t_j,x_j})$ respectively. Now suppose that

$$h \in L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t,x}, \nu_{t,x})^\perp \subset L^2(\nu).$$

Set

$$h_j = h - \sum_{m\in\mathbb{N}} \frac{\langle h, 1_{\{B_m^j\}}\rangle}{\nu(B_m^j)} 1_{\{B_m^j\}} \in L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t_j,x_j}, \nu_{t_j,x_j})^\perp.$$

Then, by 3.8.3, $h_n \to h$ in $L^2(\nu)$, and so

$$0 = \langle h_j, \psi^U(t_j, x_j, \cdot)\rangle \to \langle h, \psi\rangle.$$

Therefore, $\psi \in L^2(\mathbb{R}^D \setminus \{0\}, \Sigma_{t_j,x_j}, \nu)$ and, in particular, it is $\Sigma_{t,x}$ measurable. From step 1, it must then be the case that, ν a.e., $\psi^U(t,x,y) = \psi(y)$, i.e.

$$F(t_j, x_j) \to F(t, x)$$

weakly in $L^2(\nu)$. Since C^g is bounded, the weak topology on it is metrizable, and since every subsequence of $F(t_m, x_m)$ contains a further subsequence converging to $F(t,x)$, it must be the case that F is weakly continuous. ■

Step 2. Conclusion.

Proof of step 2. Since F is weakly continuous, it is weakly measurable, and, since $L^2(\nu)$ with the norm topology is a separable Banach space, by Pettis theorem ([3], lemma 11.37), F is strongly measurable.[11] Hence, F is also Borel measurable, and the stochastic integral

$$M_t^U = \int_{[0,t]\times\mathbb{R}^D\setminus\{0\}} \psi^U(s, X_s(\omega), y)\tilde{N}(ds, dy)$$

is well defined. This in turn defines a measure $\mathbb{Q}^U$ that satisfies the required properties. A similar argument holds for the lower spectral martingale measure. □

[11] That is, F is the $\lambda \otimes \nu$-a.e. pointwise limit of measurable functions, where λ is the Lebesgue measure on $[0, T]$.

Bibliography

[1] B. Acciaio and I. Penner. Dynamic risk measures. In G. Di Nunno and B. Oksendal, editors, *Advanced Mathematical Methods in Finance*. Springer, 2011.

[2] M. Ali. Stochastic dominance and portfolio analysis. *Journal of Financial Economics*, 2:205–229, 1975.

[3] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis*. Springer, 2006.

[4] A. Armstrong and M. Rutkowski. Valuation of credit default index swaps and swaptions. *International Journal of Theoretical and Applied Finance*, 12(7):1027–1053, 2009.

[5] K. Arrow. *Essay in the Theory of Risk-Bearing*. North-Holland, 1971.

[6] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[7] G. Bakshi and N. Kapadia. Delta hedged gains and the negative market volatility risk premium. *The Review of Financial Studies*, 16(2):527–566, 2003.

[8] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral partial differential equations. *Stochastics and Stochastics Reports*, 60:57–83, 1996.

[9] O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2:41–68, 1998.

[10] T. R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance, 2002.

[11] P. Billinglsey. *Probability and Measure*. Wiley, 1995.

[12] F. Black and M. Scholes. The pricing of options and corporates liabilities. *The Journal of Poitical Economy*, 81(3):637–654, 1973.

[13] V. Bogachev. *Measure Theory*. Springer-Verlag, 2007.

[14] T. Bollerslev, M. Gibson, and H. Zhou. Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, 160:235–245, 2011.

[15] D. Brigo and M. Morini. No-armageddon measure for arbitrage-free pricing of index options in a credit crisis. *Mathematical Finance*, 21:583–593, 2011.

[16] M Carhart. On persistence in mutual funds performance. *The Journal of Finance*, 52(1):57–82, 1997.

[17] P. Carr, H. Geman, D. Madan, and M. Yor. Self-decomposabilitt and option pricing. *Mathematical Finance*, 17:31–57, 207.

[18] P. Carr and D. Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 1:19–37, 2001.

[19] P. Carr and L. Wu. Variance risk premiums. *The Review of Financial Studies*, 22:1311–1341, 2008.

[20] A. Cherny. Pricing with coherent risk. *Theory of probability and its applications*, 52:389–495, 2008.

[21] J. Cochrane. Presidential address: Discount rates. *The Journal of Finance*, 61(4):1047–1107, 2011.

[22] R. Coifman and S. Lafon. Diffusion maps. *Appl. Comput. Harmon. Anal.*, 21:5–30, 2006.

[23] G. Connor. A unified beta pricing theory. *Journal of Economic Theory*, 34:13–31, 1984.

[24] T. Cover. Universal portfolios. *Mathematical Finance*, 1:1–29, 1991.

[25] M.H.A. Davis and S. Leo. *Risk-sensitive Investment Management*. World Scientific Publishing Company, Inc., 2014.

[26] F. Delbaen. *The Structure of m–Stable Sets and in Particular of the Set of Risk Neutral Measures*, pages 215–258. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[27] F. Delbaen. Monetary utility functions on $C_b(X)$ spaces. *preprint*, 2021.

[28] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.

[29] L. Delong. *Backward Stochastic Differential Equations with Jumps and their Actuarial and Financial Applications*. Springer Verlag London, 2013.

[30] D. Denneberg. *Non-additive measure and integral*. Springer Netherlands, 1994.

[31] S. Doctor and J. Goulden. An introduction to credit index options and credit volatility. *J.P. Morgan Credit Derivatives Research*, 2007, 2007.

[32] D. Duffee and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 4:921–950, 6.

[33] D. Duffie and N. Garleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41–59, January/February 2001.

[34] D. Duffie, J. Pan, and K. J. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[35] D. Duffie and K. J. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12(4):687–720, 1999.

[36] E. Eberlein and D. Madan. Hedge fund performance: Sources and measures. *International Journal of Theoretical and Applied Finance*, 12(3):267–282, 2009.

[37] E. Eberlein, D. Madan, M. Pistorius, and M. Yor. A simple stochastic rate model for rate equity hybrid products. *Applied Mathematical Finance*, 20(5):461–488, 2013.

[38] R. Elliot, M. Jeanblanc, and M. Yor. On models of default risk. *Mathematical Finance*, 10:179–195, 2000.

[39] R. Elliot, D. Madan, and K. Wang. High dimensional markov trading of a single stock. *SSRN Electronic Journal*, 2022.

[40] L. Epstein and S. J. Ambiguous volatility and asset pricing in continuous time. *The Review of Financial Studies*, 26(7):1740–1786, 2013.

[41] E. Fama. The behavior of stock market prices. *Journal of Business*, 38:34–105, 1965.

[42] E. Fama and K. French. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84, 1996.

[43] H. Follmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[44] K. French, S. Schwert, and R Stambaugh. Expecter stock returns and volatility. *Journal of Financial Economics*, 19:3–29, 1987.

[45] I. Friend and M. Blume. The demand for risky assets. *The American Economic Review*, 65(2):900–922, 1975.

[46] E. Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39:19–34, 2006.

[47] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique priors. *Journal of Mathematical Economics*, pages 141–153, 1989.

[48] L Glosten, R. Jagannathan, and D. Runkle. On the relation between the expected value and the volatility of the nominal excess return. *The Journal of Finance*, 48(5):1779–1801, 1993.

[49] J. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[50] J. Harrison and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11:215–260, 1981.

[51] X. He and X. Zhou. Portfolio choice under cumulative prospect theory: An analytical treatment. *Management Science*, 57:315–331, 2011.

[52] G. Hinton and S. Roweis. Stochastic neighbor embedding. *Neural Information Processing Systems (NIPS'02)*, 15:357–864, 2003.

[53] T. Hurd and Z. Zhou. A Fourier transform method for spread option pricing. *Siam Journal of Financial Mathematics*, 1:142–157, 2010.

[54] R. Jarrow, D. Lando, and S. Turnbull. A Markov model for the term structure of credit risk spreads. *Review of Financial Studies*, 10(2):481–523, 1997.

[55] N. Jegadeesh and S. Titman. Returns to buying winners and selling losers:implications for stock market efficiency. *Journal of Finance*, 48(1):65–91, 1993.

[56] E. Jouini and H. Kallal. Efficient trading strategies in the presence of market frictions. *Review of Financial Studies*, 14(2):343–369, 1995.

[57] E. Jouini and C. Napp. Conditional comonotonicity. *Decisions in Economics and Finance*, 27:153–166, 2004.

[58] D. Kahneman and A. Tverski. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

[59] D. Kahneman and A. Tverski. Advences in prospects theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–293, 1992.

[60] J. Kelly. A new interpretation of information rate. *IRE Transactions on Information Theory*, 2:185–189, 1956.

[61] U. Kuchler and S. Tappe. Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and its Applications*, 118:261–283, 2008.

[62] S. Kusuoka. A remark on default risk models. *Advances in Mathematical Economics*, 1:69–81, 1999.

[63] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, 2(2-3):99–120, 1998.

[64] J. Lintner. Security prices, risk, and mximal gains from diversification. *The Journal of Finance*, 20(4):587–615, 1965.

[65] R. Lucas. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445, 1978.

[66] J. Ma, T. Wong, and J. Zhang. Time-consistent conditional expectation under probability distortion. *Mathematics of Operations Research*, 46(3):1149–180, 2021.

[67] D. Madan. Execution costs and efficient execution frontiers. *working paper*, 2010.

[68] D. Madan. Efficient estimation of expected stock returns. *Finance Research Letters*, 23:31–38, 2015.

[69] D. Madan. Multivariate distributions for financial returns. *International Journal of Theoretical and Applied Finance*, 2020.

[70] D. Madan and E. Eberlein. Hedge fund performance: Sources and measures. *International Journal of Theoretical and Applied Finance*, 12(3):267–282, 2009.

[71] D. Madan, M. Pistorius, and M. Stadje. On dynamic spectral risk measures, a limit theorem and optimal portfolio allocation. *Finance and Stochastics*, 21:1073–1102, 2017.

[72] D. Madan and W. Schoutens. *Nonlinear Valuation and Non-Gaussian Risks in Finance*. Cambridge University Press, 2022.

[73] D. Madan, W. Schoutens, and K. Wang. Bilateral multiple gamma returns: Their risks and rewards. *International Journal of Financial Engineering*, 7(1):1–27, 2020.

[74] D. Madan and E. Seneta. The variance gamma model for share market returns. *The Journal of Business*, 63:511–24, 1990.

[75] D. Madan and H. Unal. Pricing the risks of default. *Review of Derivatives Research*, 2:121–160, 1998.

[76] H. Markovitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[77] A. McNeil, R. Frey, and P. Embrecht. *Quantitative Risk Management-Concepts, Techniques and Tools*. Princeton University Press, 2005.

[78] R. Mehra and E. Prescott. The equity premium: a puzzle. *Journal of Monetary Economics*, 15:145–161, 1985.

[79] R.C. Merton. Lifetime portfolio selection under uncertainty: the continuous time model. *Review of Economic Studies*, 51:247–257, 1969.

[80] A. Neufeld and M. Nutz. Nonlinear Levy processes and their characteristics. *Transwactions of the American Mathematical Society*, 369(1):69–95, 2017.

[81] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In B.L. Rozovskii and R.B. Sowers, editors, *Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control and Information Sciences*, volume 176. Springer, Berlin, Heideleberg, 1992.

[82] C. Pedersen. Valuation of portfolio credit default swaptions. *Lehman Brothers Quantitative Credit Research*, 2003, 2003.

[83] S. Peng. Bsde and related g-expectation. *Backward Stochastic Differential Equations: Pitman Research Notes in Math. Series*, 364:141–159, 1996.

[84] S. Peng. G-Expectation, G-Brownian motion and related stochastic calculus of Ito type. 2006. arXiv:math/0601035v2 [math.PR].

[85] S. Peng. *Nonlinear Expectations and Stochastic Calculus under Uncertainty*. Springer, 2019.

[86] C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[87] S. Ross. Mutual fund separation in financial theory - the separating distributions. *Journal of Economic Theory*, 17:254–286, 1978.

[88] M. Rothschild and J. Stiglitz. Increasing risk: I. a definition. *Journal of Economic Theory*, 2:225–243, 1970.

[89] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.

[90] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.

[91] M. Royer. Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic Processes and their Applications*, 116:1358–1376, 2006.

[92] B. Rudiger. Stochastic integration for compensated poisson measures and the levy-ito formula. In S. Albeverio, A.B. de Monvel, and H. Ouerdiane, editors, *Proceedings of the International Conference on Stochastic Analysis and Applications*, pages 145–167. Kluwer Academic Publishers, 2004.

[93] P. Samuelson. Why we should not make mean log of wealth bit though years to act are long. *Journal of Banking and Finance*, 3:305–307, 1979.

[94] W. Schachermeyer. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Mathematical Finance*, 14(1):19–48, 2004.

[95] W. Sharpe. Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

[96] J. Tenebaum, V. de Silva, and J Langford. A global geometric framwork for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.

[97] J. Tobin. Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25:65–86, 1958.

[98] L. van der Maarten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[99] S.S. Wang. A class of distortions operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, 67:15–36, 2000.

[100] S.S. Wang. Equilibrium pricing transform: new results using Buhlman's 1980 economic model. *Astin Bulletin*, 33:57–73, 2002.

[101] S.S. Wang. A universal framework for pricing financial and insurance risks. *Astin Bulletin*, 32:213–234, 2002.

[102] S. Zheng. Gradient descent algorithms for quantile regression with smooth approximation. *International Journal of Machine Learning and Cybernetics*, (2):191–207, 2011.