University of Sussex

School of
Business, Management and Economics

Department of Finance

Discretisation-Invariant Swaps and Higher-Moment Risk Premia

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Author:
Johannes Rauch

Second Supervisor:
Dr Andreas Kaeck

First Supervisor:
Prof Carol Alexander

External Examiner:
Prof Markus Leippold

June 7, 2016

Declaration of original authorship

I hereby declare that this dissertation is my original work and that all sources have been properly and fully acknowledged.

Johannes Rauch

June 7, 2016

Summary

This thesis introduces a general framework for model-free discretisation-invariant swaps. In the first main chapter a novel design for swap contracts is developed where the realised leg is modified such that the fair value is independent of the monitoring partition. An exact swap rate can then be derived from the price a portfolio of vanilla out-of-the-money options without any discrete-monitoring or jump errors. In the second main chapter the P&Ls on discretisation-invariant swaps associated with the variance, skewness and kurtosis of the log return distribution are used as estimators for the corresponding higher-moment risk premia. An empirical study on the S&P 500 investigates the factors determining these risk premia for different sampling frequencies and contract maturities. In the third main chapter the dynamics of conventional and discretisation-invariant variance swaps and variance risk premia are compared in an affine jump-diffusion setting. The ideas presented in this thesis set the ground for many interesting and practically relevant applications.

Acknowledgements

First of all I would like to thank Carol Alexander for her dedication and patience all the way through this PhD journey. It is both a pleasure and a unique experience to work alongside with such a brilliant researcher and passionate academic teacher. My thanks also go to Rudi Zagst who encouraged me to do a PhD with Carol. I am further indebted to Andreas Kaeck for his constant support and guidance, and to Tobias Kuna for his mathematical advice whenever I asked for it. This thesis would not be the same without the original ideas of Anthony Neuberger, who kindly provided me with helpful comments and encouragement. The success of this project is not least due to numerous on and off topic discussions with my friends and colleagues: Casey Chen, Michael Coulon, Julia Kapraun, Mikhail Krayzler, Leonardo Nogueira, Yuri Saporito, Johanna Scheller, Anannit Sumawong and Aanand Venkatramanan. I would like to dedicate this thesis to those who are most important to me and have given me strength and joy at any time: Amanda, Jakob, Elisabeth, Oda and Nikolaus.

Contents

List of Figures

List of Tables

Acronyms

AP aggregation property.

CAPM capital asset pricing model.

CBOE Chicago Board Options Exchange.

CEV constant elasticity of volatility.

CLT central limit theorem.

DI discretisation-invariant.

ERP equity risk premium.

ETF exchange traded fund.

ETN exchange traded note.

GARCH generalised autoregressive conditional heteroscedasticity.

OTM out-of-the-money.

P&L profit and loss.

PCA principal component analysis.

S&P 500 Standard & Poor's 500 Stock Market Index.

SDE stochastic differential equation.

SDI strike-discretisation invariant.

SV stochastic volatility.

SVCJ stochastic volatility with contemporaneous jumps.

TED T-Bills – Eurodollar.

VIX CBOE Volatility Index.

VRP variance risk premium.

Chapter 1

Introduction

Global financial markets are subject to continuous change and evolution as market participants look out for new and profitable investment opportunities. In order to improve the risk and return properties of their portfolios investors constantly scan the economy for risk premiums and diversification potential, accessing a wide range of asset classes and derivative instruments. Simultaneously, as the market environment matures investors develop more sophisticated risk preference profiles, and more powerful tools become necessary to match demand and supply of financial assets in an effective and robust manner.

In the base model for most financial markets, returns are assumed to be normally distributed and independent over non-overlapping investment periods. Naturally, neither of these assumptions holds in practice and the academic literature has brought forward and analysed a plethora of stochastic jump-diffusion models that match more closely frequently observed real world phenomena, such as volatility clusters or the familiar volatility skew. While jump components allow one to model non-normal short-term return distributions, non-normality in long-term distributions can result from stochastic volatility and more complex patterns of serial dependence.

Those models have been very useful for the pricing of a variety of financial derivatives and also contribute to the understanding of basic market mechanisms, for example the leverage effect.[1] However, they all impose assumptions on the underlying price process and any kind of time series analysis may be biased as soon as the assumptions do not hold. In the literature the uncertainty about these assumptions is referred to as model risk, and the pricing and hedging of path-dependent derivatives is particularly affected. It appears that for the purpose of statistical estimation a model-free approach, based only on the assumption of no arbitrage, is more suitable.[2]

Sound theoretical prices for complex financial instruments help to preclude arbitrage opportunities, and they are especially important during turbulent periods when liquidity dries up and there is no reliable market price. The events leading up to the great financial crisis of 2008-9 illustrate the importance of accurate theoretical prices when, following a turning point in the escalating demand for collateralised debt obligations, the failure of market participants to agree on a

[1] The leverage effect refers to the inverse relationship between asset returns and asset volatility in stock markets. It is a stylysed fact that negative returns trigger an increase in volatility, which reflects the investment behavior of risk-averse market participants.

[2] Arbitrage is a risk-less opportunity to make profit, and in its strong form a profit is guaranteed. The no-arbitrage principle is fundamental to financial and economic theory, and to asset and derivatives pricing in particular.

fair value for these products precipitated a credit crunch. More recently it is not so much credit derivatives as volatility derivatives related to variance swaps that are raising concerns in the financial press.

Variance swaps were introduced over-the-counter in the 1990's and are popular instruments for trading variance risk premia by exchanging a floating realised variance with a fixed swap rate, based on some notional amount. A risk-neutral market participant can offer this premium to speculators or risk-averse investors who hedge their exposure to realised variance. When a bank issues a variance swap that pays realised variance, with payment settled at maturity, the rate it charges should be determined so that it expects a small profit after hedging its exposure to realised variance. A theoretical, fair-value variance swap rate provides an indicative quote for the rate actually charged. However, most theoretical swap rates rely on model assumptions that may not hold in practice, and particularly under changing market conditions.

Swap rates have been available from broker dealers for many years and the fair-value rates are normally within the bid-ask spread of market rates, indicating an active market where banks may not be hedging all their exposures in order to charge competitive rates. However, during times of financial distress market rates can be significantly greater than their fair-values. Nowadays, variance swaps and their futures, options, notes, funds and other derivatives are being actively traded on exchanges. Currently, data from the Chicago Board Options Exchange (CBOE) show that about \$3-\$6 bn notional is being traded daily on CBOE Volatility Index (VIX) futures contracts alone and on stock exchanges around the world even small investors can buy and sell over a hundred listed products linked to volatility futures such as exchange traded funds (ETFs) and exchange traded notes (ETNs).[3] The most popular of these is Barclay's VXX note (iPath S&P 500 VIX Short-Term Futures ETN), with a market cap of around $1 trillion as of 31 December 2013. More recently investors have also developed interest in investment opportunities and diversification potential linked to the higher-moments of the return distribution. A variety of indices – such as the CBOE SKEW index – has been developed to meet this demand. These indices are based on options prices and represent higher-moments of the implied probability distribution.

Non-normality can be observed under the risk-neutral measure, where option prices commonly imply a pronounced volatility smile, but also under the physical measure, e.g. in terms of the well-studied leverage effect. In order to illustrate the latter empirically, Table 1.1 displays the standard deviation, skewness and kurtosis of daily, weekly and monthly log returns on S&P 500 futures over two separate 9-year periods. In each case the weekly and monthly observed moment is compared with the moment derived from daily returns under the i.i.d. assumption, where the standard deviation scales with the square root of the horizon, skewness scales with the inverse square root and kurtosis scales with the inverse of the horizon, in accordance with the derivations for the aggregation over time of standardised moments in the Appendix.

The observed values for standard deviation are roughly the same as those extrapolated using the i.i.d. assumption; however, the observed skewness is greater in magnitude than expected, especially during the second period which is influ-

[3] The VIX is an index that captures the 30-days implied volatility of the Standard & Poor's 500 Stock Market Index (S&P 500). It is seen as an important fear barometer for the equity markets. Futures and options written on the VIX have been traded for about a decade.

Return (trading days)		daily (1)	weekly (5)		monthly (20)	
		obs	obs	i.i.d.	obs	i.i.d.
Stdev	1996-2004	0.012	0.025	0.028	0.047	0.055
	2005-2013	0.013	0.028	0.030	0.056	0.060
Skewness	1996-2004	-0.120	-0.073	-0.054	-0.053	-0.027
	2005-2013	0.002	-1.275	0.001	-2.021	0.000
Kurtosis	1996-2004	5.7	4.6	1.136	3.4	0.284
	2005-2013	16.7	16.8	3.334	14.1	0.834

Table 1.1: Moments of S&P 500 futures: observed vs. predicted under i.i.d.

enced by the positive autocorrelation in negative returns during the financial crisis of 2008–9. This finding agrees with Neuberger [2012], although it is much less pronounced during the earlier period. Now, heteroscedasticity is a more ubiquitous feature of financial returns than autocorrelation and, just as autocorrelation increases skewness, volatility clustering increases kurtosis. Indeed, the observed kurtosis is much greater than expected in the i.i.d case during both periods.

For the pre-crisis time period, and looking at the option implied distribution rather than realised returns, Carr and Wu [2003] observe a similar effect. The volatility smile does not flatten out for long maturities, as would be expected from the central limit theorem (CLT) under the assumption of independent returns, indicating serial dependence under the risk-neutral measure. In fact, before the crisis this effect is stronger under the risk-neutral measure than under the objective measure. Wu [2006] reconciles these observations by modeling the tails of a distribution using what he terms the 'exponentially dampened power law'.

Standard definitions of the realised third and fourth moment as well as their normalised versions – realised skewness and kurtosis – are based solely on a single underlying price or return process and disregard the presence of autocorrelation or any other kind of serial dependence. In the idealised case of continuous monitoring, these realised characteristics only capture moments of the jump distribution. However, under the assumption of i.i.d. period returns, the CLT implies that the long-term distribution is approximately normal. In other words, the short-term jump distribution is mostly irrelevant for long-term investors and serial dependence becomes the predominant feature.

The above analysis clearly shows the presence of serial dependence and illustrates that short-term higher-moments are not suitable for forecasting higher-moments of long-term returns. This should also be reflected in the definition of realised moments. Hence, a new realised kurtosis based on aggregating higher-moment characteristics which makes feasible the accurate measurement of long-horizon kurtosis from short-horizon returns is important. The phenomena described above also indicate that the analysis of higher-moment risk premia, which explain the shift between the physical and risk-neutral probability measure, is an important and promising field of study. Notably, the joint estimation of $\mathbb{P}$ and $\mathbb{Q}$ parameters for asset pricing models has attracted attention in recent literature (see e.g. Bardgett et al. [2015]).

This thesis introduces a comprehensive theory for discretisation-invariant (DI) swap contracts written on multiple assets that have exact fair-values, provided only that the no-arbitrage assumption for forward prices holds. That is, swap rates are model-free in a strong sense, and they do not depend on the monitoring frequency of realised cash flows. The definitions of the realised legs take serial dependence into account and link long-term risk with short-term returns on futures and option portfolios. Our theory encompasses a wide variety of DI pay-offs, including those corresponding to higher-moments of the log return distribution and bi-linear functions of vanilla options prices. Based on these new definitions, we expect market swap rates to be closer to their theoretical values and within the no-arbitrage range, particularly in times of financial distress.

In the three main chapters of this thesis we take complementary perspectives on DI swaps. First, we discuss the design and pricing of swap contracts that can be associated with the higher-moments of a log return distribution, and that can be perfectly hedged in discrete time. Second, we conduct a model-free empirical analysis of higher-moment risk premia in the US equity market using these contracts. Our favourable DI design results in exact fair-values for swaps and consequently unbiased risk premium estimates, even when monitoring is performed along a discrete partition. Finally, we compare conventional and DI variance swaps in the context of affine stochastic volatility models with and without jumps in order to illustrate the structure and composition of risk premia as well as their relationship with latent variables. The following paragraphs provide an overview of the focal papers, while secondary papers are reviewed in detail in the beginning of each chapter.

The second chapter builds on recent ideas from Neuberger [2012] and Bondarenko [2014], who suggest an alternative definition of realised variance that results in an exact and model-free swap rate. Neuberger [2012] further includes option portfolios in the definition of the realised leg to define a realised third moment. Starting from the analysis of both model-dependent and discretisation errors that arise when pricing a conventional variance swap, we discuss a variety of alternative DI definitions for the realised leg that are related to higher-moments of the log return distribution of a single asset and allow for exact pricing via the replication theorem of Carr and Madan [2001]. Dynamic trading strategies in a small number of vanilla-style contingent claims, such as the volatility, cubic and quartic contracts discussed in Bakshi et al. [2003], allow to hedge DI swaps in a model-free manner and make higher-moment risk accessible to investors. Our framework also applies to a multivariate setting where the realised leg incorporates the prices of a large number of tradable assets, facilitating the design of covariance swaps.

The third chapter follows the empirical studies by Carr and Wu [2009] and Kozhan et al. [2013], who analyse the variance risk premium in the US equity market, using the profit and loss (P&L) on a variance swap as an estimator for the premium. While Carr and Wu [2009] calculate the floating leg using a model-dependent definition of realised variance, the study of Kozhan et al. [2013] applies the DI definition from Neuberger [2012] in order to avoid the propagation of a pricing bias into the estimator. Kozhan et al. [2013] also analyse the skewness risk premium, using the P&L on a swap based on the DI third-moment characteristic from Neuberger [2012] as an estimator, and find that it is highly correlated with the variance risk premium.

According to their results, it is not possible to make profits on a skewness swap once the variance risk is hedged away. However, this conclusion relies on the specific choice of the third moment characteristic. When we apply our framework to a large sample of S&P 500 options prices, we find that DI higher-moment swaps exhibit relatively low correlation with the negative variance risk premium, particularly for higher monitoring frequencies. In contrast with previous research by Kozhan et al. [2013] our empirical results suggest that significant new risk premia become tradable via the use of DI third and fourth moment swaps as well as frequency swaps that exchange two floating legs at different monitoring frequencies. They also point towards interesting new investment opportunities and diversification potential associated with higher-moments. We conclude our empirical analysis by relating higher-moment risk premia to the standard risk factors introduced by Fama and French [1993] and Carhart [1997] and discover interesting new patterns.

The fourth chapter provides derivations for explicit swap dynamics in affine stochastic volatility models. Specifically, we compare the fair-value price process of a conventional variance swap with Neuberger's variance swap as well as our DI variance swap in the Heston [1993] model and study the impact of jumps in the price and variance process under the stochastic volatility with contemporaneous jumps (SVCJ) model proposed by Duffie et al. [2000]. The chapter also discusses the market price of risk as well as the change of measure from risk-neutral to physical. These explicit solutions may be helpful for the specification analysis and estimation of affine asset pricing models, as e.g. performed in Egloff et al. [2010] and Bardgett et al. [2015].

Chapter five concludes by summarising the mains results and providing an overview of topics for further research. The Appendix contains a range of tools that are used throughout this thesis: Itô's formula for jump-diffusion processes, Girsanov's change of measure for one and more dimensions, the replication theorem by Carr and Madan [2001] and finally a review of the aggregation of standardised and non-standardised moments of a distribution over time.

Chapter 2

Discretisation-Invariant Swaps

Demand for volatility derivatives as a diversifier, a hedge or purely for speculation has increased very significantly during the last few years. However, there is no exact theoretical value for the underlying (the variance swap rate) and for this reason market rates can deviate significantly beyond the approximation used for the no-arbitrage range. For instance, during the turbulent year surrounding the Lehman Brothers collapse in September 2008, market rates for variance swaps written on the Standard & Poor's 500 Stock Market Index (S&P 500) were very often 5% or more above the CBOE Volatility Index (VIX) formula which is commonly used as an approximation for their theoretical value.

In the first main part of this thesis, we illustrate that most pricing problems are already rooted in the definition of the realised leg of conventional variance swaps. We examine the variety of error terms that common practice is facing, discuss existing approaches to modify realised variance in order to improve the pricing accuracy and then present a fundamental condition for swaps that assures perfect replicability in discrete time. We solve this condition for arbitrary higher-moments, providing a state-of-the-art framework for trading higher-moment swaps and measuring higher-moment risk and the associated risk premia. Our framework facilitates the design of swap contracts where the definition of the floating leg encompasses information about serial dependence and hence allows to link short-term returns to moments of the long-term return distribution. Swap rates are model-free and can be derived exactly from vanilla out-of-the-money (OTM) option prices.

For this purpose we present an exhaustive literature review on the pricing of variance and higher-moment swaps, with particular emphasis on the recent work of Anthony Neuberger. We then generalise Neuberger's aggregation property (AP) and introduce the notion of a discretisation-invariant (DI) swap. We also provide dynamic trading strategies that can be used to hedge DI swaps perfectly in discrete time.

2.1 Literature Review

The terms and conditions of a conventional variance swap define the floating leg, realised variance, as the average squared daily log-return on some underlying, commonly an equity index, over the life of the swap. With this definition of realised variance the conventional fair-value variance swap rate calculation proceeds under the assumptions: (i) the forward price of the underlying follows a pure martingale diffusion under the risk-neutral measure; (ii) the floating leg is monitored continuously; and (iii) vanilla options on the underlying with the same maturity as the swap are traded at a continuum of strikes. As Demeterfi et al. [1999] show, the fair-value swap rate – which under assumption (ii) becomes the expected quadratic variation of the log-price – can then be derived from the market prices of these vanilla options.[1] However, in practice none of these assumptions hold and there is a large literature analysing the biases caused by making these false assumptions.

2.1.1 Pricing Conventional Variance Swaps

In an arbitrage-free market, as introduced by Harrison and Kreps [1979], the bank issuing a variance swap to a representative investor will compute this expected pay-off under a risk-neutral measure. According to Breeden and Litzenberger [1978] the risk-neutral measure for a representative investor corresponds to the market implied measure in a complete market. In this case a unique fair value for the variance swap rate can be derived as the expectation of realised variance under the market implied measure. Yet, markets are usually incomplete in the presence of jumps, when assumption (i) is violated. An underlying price process can still be consistent, which is by definition the case if the model prices of all European options match observable market quotes, although the prices of Exotic derivatives may differ. Britten-Jones and Neuberger [2000] derive a simple condition for a continuous process to be consistent and show that all consistent price processes

[1] An alternative derivation as well as further details on the implementation can be found in Jiang and Tian [2005], who then promote the 'model-free implied volatility' as a direct test of market efficiency.

imply the same, model-free volatility.[2] Unfortunately, expected realised variance as defined in a conventional variance swap is model-dependent.

Carr and Wu [2009] discuss the idealised case where assumption (ii) holds but not assumption (i), i.e. continuous monitoring is possible but the underlying price process need not be continuous. Under the assumption of a generic decomposition of the underlying process into a pure jump and a pure geometric diffusion component they apply the replication theorem of Carr and Madan [2001] to find that the fair-value swap rate is a weighted integral over a continuum of European OTM option prices, corrected for a model-dependent 'jump error' term. The replication theorem allows for a general European style claim on some underlying to be represented as the integral over a continuum of put or call options. This theorem is closely related to the spanning approach to derivative pricing by Bakshi and Madan [2000], who state that the characteristic function of a martingale price process and option prices for a continuum of strike prices are interchangeable representations of the claims they span. Accordingly, since any European claim can be expressed in terms of the characteristic function, it is also spanned by put and call options.

Carr and Lee [2009] prove that relaxing assumption (ii) leads to a 'discretisation bias' that is related to the third moment of returns so it can be very large during excessively volatile periods; Jarrow et al. [2013] investigate the convergence of the discretely-monitored swap rate to its continuously-monitored counterpart and derive discretisation error bounds that get tighter as the monitoring

[2]The theory of consistent processes extends the idea of local volatility models where the instantaneous volatility is only a function of time and the current underlying price – introduced simultaneously by Dupire [1994] and Derman and Kani [1994] and developed further by Gatheral [2006] – to the more general case where volatility can be driven by idiosyncratic risk factors.

frequency increases; Bernard et al. [2014] generalise these results and provide conditions for determining the sign of the discretisation bias; and Hobson and Klimmek [2012] derive model-free discretisation error bounds and super- and sub-replication strategies for hedging variance swaps. Under a variety of stochastic volatility diffusion and jump models, Broadie and Jain [2008b] derive fair-value swap rates for discretely monitored variance swaps, claiming that for most realistic contract specifications the discretisation error is smaller than the error due to violation of assumption (i). Bernard and Cui [2014] extend this analysis to include a much wider variety of processes by considering the asymptotic expansion of the discretisation bias. Rompolis and Tzavalis [2013] derive bounds for the so-called 'jump error' and demonstrate, via simulations and an empirical study, that price jumps induce a systematic negative bias which is particularly apparent when there are large downward jumps. Thus, when the term of the swap includes a particularly turbulent period the jump bias and the discretisation bias work in the same direction to substantially under-estimate the fair-value swap rate.

The uncertainty about discretely-monitored realised variance is discussed in Barndorff-Nielsen and Shephard [2002] for the general case when the underlying process follows a semimartingale. The authors derive rates of convergence for continuous monitoring as well as asymptotic distributions for a general class of stochastic volatility models. They further remark that realised variance is quite an accurate estimator when volatility is low while the measurement error can become very large during periods of high volatility (see p.472). Hence, they consider it crucial to control this measurement error.

Also, in practice the integral over a continuum of European OTM options, which is well defined according to (iii), must be estimated using the prices of vanilla options that are actually traded. So there is a third bias arising from the numerical computation of the fair-value rate, which is typically based on a restricted range of quoted strikes because deep-OTM (and deep-in-the-money) options lack sufficient liquidity for reliable prices. In fact, beyond a certain moneyness level there are no price quotes at all and the only ways forward are extrapolation – which constitutes an implicit model assumption – or truncation. Jiang and Tian [2005] address the problems attendant to assumption (iii) and derive upper bounds for the 'truncation error'. Also based on a finite number of traded strikes, Davis et al. [2014] derive model-free arbitrage bounds for continuously-monitored variance swap rates and claim that market rates are surprisingly close to the lower bound.

In a recent working paper, Le and Yang [2015] analyse the impact of truncation on replication portfolios for the implied higher-moments introduced by Bakshi et al. [2003]. The authors state that the truncation error increases with the order of the estimated moment, i.e. the impact of truncation is stronger for skewness than it is for variance and it is even stronger for kurtosis. They detect weaknesses in the linear extrapolation approach taken by Jiang and Tian [2005] and the domain symmetrisation approach by Dennis and Mayhew [2002] and argue that the implied volatility, the truncation level (i.e. the range of available strikes) and the strike domain asymmetry are important factors that influence the estimation procedure. Le and Yang [2015] show how this information can be used to successfully stabilise moment estimators.

Alexander and Leontsinis [2011] further show that variance swap rates are subject to systematic biases which depend on the model assumptions underlying the theoretical fair-value formula, the representation chosen for implementing this formula as well as the numerical integration technique. While jump and discretisation biases are usually negative, the error introduced by the common use of Riemann sums is positive. Consequently, it becomes difficult to disentangle the contribution of potentially erroneous model assumptions from biases that relate to technical details of the implementation. The authors develop an analytical integration technique, based on spline interpolation, which alleviates the drawbacks of the standard approach significantly. Indeed, there is a whole plethora of methods for generating a fine grid of European option prices from the finite sample of traded options available for the purpose of numerical integration across the strike dimension. An easy-to-implement spline-smoothing algorithm is described by Fengler [2009], who also demonstrates how the estimation method precludes arbitrage opportunities across both the strike and time-to-maturity dimension. Excluding calendar arbitrage is particularly important for the estimation of risk premia, since the effects of both phenonema accrue over time and may therefore distort one another.

Based on a no-arbitrage argument Broadie and Jain [2008a] develop dynamic hedging strategies that allow to replicate other volatility derivatives using variance swaps and a discrete set of European options, where the optimal number of hedging instruments is determined numerically. Sensitivities as well as replication errors are provided. More recently, Keller-Ressel and Muhle-Karbe [2013] present exact and approximate methods for pricing options on both discretely and continuously monitored realised variance in different Lévy models, laying the ground for similar approaches using affine models.

The errors associated with assumptions (i), (ii) and (iii) have distorted market prices and posed challenging questions to academics for many years. However, it has now been recognised that the model dependence of variance swaps is essentially rooted in the conventional definition of the realised leg as the sum of squared log-returns. In particular, it turns out that the straight-forward generalisation to higher-moment swaps – using higher powers of the log-return for the realised leg, see e.g. Schoutens [2005] – amplifies the effects of model-dependence, making such products unattractive for investors aware of model risk.

2.1.2 Modifying the Floating Leg

More recently, alternative definitions for the realised variance have been explored which result in swap contracts that are easier to price and hedge than standard variance swaps. Martin [2013] advocates the use of a sum of squared 'simple' returns, rather than log-returns, arguing that with this modification both jump and discretisation biases are minimised. His results are explained and extended by Bondarenko [2014]. Likewise, the gamma swaps described by Lee [2010a] weight the realised variance in such a way that replication and valuation are relatively straightforward under the continuous semimartingale assumption. Bondarenko [2014] and Lee [2010b] derive generalised variance swap pay-offs that are also based on weighting functions.

The concept of swap contracts that are based on a generalisation of quadratic variation has recently been presented by Carr and Lee [2013]. Their definition of a share-weighted G-variation swap encompasses conventional variance swaps and gamma swaps (see Lee [2010a]) as special cases, which can be priced by multiples of a log and an entropy contract, respectively, under very general model assumptions. Since the weighting function of a gamma swap is proportional to the asset price, it puts more emphasis on upside- than on downside-variance. Choe and Lee [2014] define a third and fourth realised moment based on the quadratic (co-)variation of the log-price and squared log price processes. However, neither do these processes represent investible assets, nor do the authors justify their assumption that the log-price process has no drift. Also, the authors illustrate (see p.8) that the risk-neutral expectation of either of these realised legs is subject to a model-dependent jump error term of third order.

Based on the very general definition of divergence by Bregman [1967], Schneider and Trojani [2015a] propose a comprehensive, model-free framework for divergence swaps which can be statically hedged by means of synthetic options portfolios, and independent from a specific trading frequency. The authors also promote the notion of 'divergence indices' as a benchmark for testing asset pricing models. Schneider and Trojani [2015b] build on these ideas when they propose divergence swaps as a means of trading risk premia associated with investors' fear in an incomplete market setting. A common feature in all these approaches is that they only re-define the realised leg of a swap based on changes in the underlying, ignoring the possibility to use implied characteristics from option prices to define higher realised moments.

In his recent path-breaking research Neuberger [2012] re-defines the realised variance and introduces a new skewness characteristic in such a way that exact, model-free fair-value variance and skewness swap rates can be calculated. In addition, the same rate applies whether the floating leg is based on intra-day, daily, weekly or monthly returns – in fact, the monitoring does not even have to be regular. Neuberger introduces the AP as a fundamental condition which must be satisfied by the characteristic used to calculate the floating leg. Following the intuition that long-term skewness is mainly caused by the correlation between changes in the underlying and changes in (implied) variance, Neuberger takes the original step of including implied characteristics into the definition of a realised third moment characteristic which satisfies the AP and therefore has a fair-value swap rate that can be priced and hedged exactly, independently of its monitoring frequency, under the minimal assumption of no arbitrage. In this context the log contract, originally introduced by Neuberger [1994] as a model-free volatility trading instrument, gains new importance as an implied characteristic.

A step into a similar direction is taken by Torricelli [2013] who, based on Fourier techniques in a stochastic volatility setting, shows how joint claims on the underlying and its realised variance can be priced, providing a general partial differential equation that the price process of such a derivative must satisfy. Example claims are target volatility options, double digital European options and volatility-capped or -struck options.

The motivation for Neuberger [2012] is to propose a definition of the realised third moment that is computed from high-frequency returns and vanilla option prices which provides an unbiased estimate of the true third moment of long-horizon returns. He demonstrates that, far from diminishing with horizon as would be the case if returns are independent and identically distributed (i.i.d.), "....skewness actually increases with horizon up to one year, and its magnitude is economically important." We reproduce this finding in the introduction to this thesis and extend the story to comprise a similar argument for the fourth standardised moment. Neuberger concludes his work on realised skewness by stating that "[...] it would also be nice to be able to extend the analysis to higher-order moments. This would not be straightforward; [...] the set of functions that possess the aggregation property is quite limited; the way forward here may be to include other traded claims, in addition to those on the variance of the distribution."

2.2 Theoretical Results

Pursuing these ideas, we here extend Neuberger's theoretical results to fourth- and higher-order moments. Indeed, we provide a general theory for DI swaps, using the term to refer to any swap of some realised characteristic for a corresponding implied characteristic which has a fair value that does not depend on the monitoring frequency. We develop a holistic framework, based only on the assumption that the forward price of the underlying is a martingale, which allows the theoretical fair-value rate of a DI swap for any moment of the log-return (or price) distribution to be derived exactly from vanilla option prices. To this end we introduce a canonical set of option-implied 'fundamental contracts' of similar ilk to the log contract introduced by Neuberger [1994] and the higher-moment contracts in Bakshi et al. [2003]. Using these contracts DI swaps can also be hedged perfectly under any partition for monitoring and rebalancing. Our theory extends to DI swaps which are not even linked to moments and with fair values that are simple bi-linear functions of traded vanilla option prices (without the need for integration), swaps based on forward moment characteristics and swaps trading on systematic differences in risk premia.

2.2.1 Errors in Variance Swap Rates

The conventional daily, realised variance may be written:

$$\mathrm{RV} := \sum_{t=1}^{T} (x_t - x_{t-1})^2, \tag{2.1}$$

where $x_t := \ln F_t$ and F_t denotes the forward price of the underlying at time t. In practice, the floating leg of a variance swap is set equal to the *average* realised variance during the lifespan of the swap rather than the total variance as in (2.1). However, including this level of detail would only add an unnecessary level of complexity to our analysis. As demonstrated in the preceding literature review, this definition of realised variance entails errors associated with the model assumptions for the underlying price process, the discrete monitoring and the numerical integration over option strikes.

For the idealised case where continuous monitoring is possible and (2.1) can be replaced by the quadratic variation $\langle x \rangle_T$, Carr and Wu [2009] assume a generic decomposition of the underlying process into a pure jump and a pure geometric diffusion component to then apply the replication theorem of Carr and Madan [2001] and prove that,

$$\mathbb{E}^{\mathbb{Q}}\left[\langle x \rangle_T\right] = 2 \int_{\mathbb{R}^+} k^{-2} q(k) dk + \iota,$$

where $q(k)$ denotes the price of a vanilla OTM option with strike k and maturity T. When $k \leq F_0$ the option is a put and when $k > F_0$ the option is a call. This choice of separation strike is standard in the variance swap literature, e.g. in

Bakshi et al. [2003]. When the underlying price follows a pure diffusion the 'jump error', ι, is zero. The operator $\mathbb{E}^{\mathbb{Q}}$ denotes the risk-neutral expectation.

An important source of error in the theoretical fair-value swap rate stems from the fact that the realised leg of the swap is monitored only at discrete points in time. This 'discretisation error' may be written

$$\varepsilon := \mathbb{E}^{\mathbb{Q}}\left[\mathrm{RV} - \langle x \rangle_T\right]. \tag{2.2}$$

Both the discretisation and jump errors affect the theoretical price of the fixed leg. For instance, with the realised variance (2.1) and the generic jump-diffusion setting of Carr and Wu [2009] the fair-value variance swap rate may be written

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{RV}\right] = 2\int_{\mathbb{R}^+} k^{-2} q(k) dk + \iota + \varepsilon. \tag{2.3}$$

By ignoring these errors the risk-neutral expectation of the pay-off becomes $\iota + \varepsilon$ rather than zero and therefore, under the standard variance swap pricing formula, the estimator for the variance risk premium (VRP) is biased.

In addition to this model-dependent bias, the way in which the Chicago Board Options Exchange [2009] and other exchanges implement the integral in (2.3) is subject to an approximation error due to the numerical integration over actually traded strikes:

$$\delta := 2\int_{\mathbb{R}^+} k^{-2} q(k) dk - 2\sum_i k_i^{-2} q\left(k_i\right) \Delta k_i, \tag{2.4}$$

where k_i, $i = 0, \ldots, n$ denote the traded strikes for maturity T and $\Delta k_i := \left(k_{i+1} + k_{i-1}\right)/2$ for $i = 1, \ldots, n-1$ as well as $\Delta k_0 := k_1 - k_0$ and $\Delta k_n := k_n - k_{n-1}$.

We have

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{RV}\right] = 2\sum_{i} k_i^{-2} q\left(k_i\right)\Delta k_i + \iota + \varepsilon + \delta. \tag{2.5}$$

The integration error δ stems from two practical restrictions. Firstly, and most importantly, deep-OTM options lack sufficient liquidity for reliable prices (in other words $k_0 >> 0$ and $k_n << \infty$) so truncation or extrapolation of the integral become necessary, the latter making further model assumptions necessary. Secondly, there is only a small number of strikes available within the traded range. This problem is often tackled by using linear or cubic spline interpolation over strikes.

2.2.2 The Aggregation Property

We consider only one maturity date, T, but various partitions of the interval $\mathbf{\Pi} := [0, T]$, e.g. the 'daily' partition $\mathbf{\Pi}_D := \{0, 1, \ldots, T\}$. The increments along a partition are denoted using a 'carat'. We use $\mathbb{E}_t^{\mathbb{Q}}[.] := \mathbb{E}^{\mathbb{Q}}[.|\mathcal{F}_t]$ to denote the expectation conditional on a filtration $\mathcal{F}_t$ at time t, under the risk-neutral measure and write $\mathbb{E}^{\mathbb{Q}}[.] := \mathbb{E}_0^{\mathbb{Q}}[.]$. Univariate martingale processes are denoted in upper-case and non-martingales in lower-case: e.g. $s := \{s_t\}_{t\in\mathbf{\Pi}}$ denotes the price process underlying the variance swap and $F := \{F_t\}_{t\in\mathbf{\Pi}}$ denotes the fair-value price process of a forward contract on s, i.e. $F_t := \mathbb{E}_t^{\mathbb{Q}}\left[s_T\right]$. As before $x := \{x_t\}_{t\in\mathbf{\Pi}}$ denotes the log forward price process, i.e. $x_t := \ln F_t$.

In the following we relate realised characteristics to an n-dimensional stochastic process $\mathbf{z} := \{\mathbf{z}_t\}_{t\in\mathbf{\Pi}} \in \mathbb{R}^n$. Given some function $\phi : \mathbb{R}^n \to \mathbb{R}$, the 'realised ϕ-characteristic' of $\mathbf{z}$ w.r.t. a partition $\mathbf{\Pi}_N = \{t_i\}_{i=0,\ldots,N}$ over the swap's lifespan $\mathbf{\Pi}$

is defined as

$$\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right) := \sum_{i=1}^{N} \phi\left(\mathbf{z}_{t_i} - \mathbf{z}_{t_{i-1}}\right). \tag{2.6}$$

Let $\{\mathbf{\Pi}_N\}_{N=1,2,\ldots}$ denote a sequence of partitions such that $0 = t_0 \leq t_1 \leq \ldots \leq t_N = T$. If $\max_{i\in\{1,\ldots,N\}}[t_i - t_{i-1}] \to 0$ as $N \to \infty$ we write written $\mathbf{\Pi}_N \to \mathbf{\Pi}$. If it exists we define the 'ϕ-variation' of $\mathbf{z}$ as the continuously monitored limit of the realised characteristic, i.e.

$$\langle \mathbf{z} \rangle_T^{\phi} := \lim_{\mathbf{\Pi}_N \to \mathbf{\Pi}} \sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right). \tag{2.7}$$

In the following we only consider characteristics ϕ with $\phi(\mathbf{0}) = 0$, so the limit (2.7) can be finite. The ϕ-variation is a theoretical construct that, if it exists, can be used to derive a fair-value swap rate by taking its expected value based on some assumed process for the underlying. This is the approach taken by Jarrow et al. [2013] and several other papers that analyse the discrete monitoring error for variance swaps.

But we do not need to assume that the ϕ-variation exists because it does not preclude the definition of a 'ϕ-swap' as a financial contract that exchanges a realised ϕ-characteristic (2.6) with a fixed value, called the 'ϕ-swap rate'. As long as the ϕ-variation exists and is finite the discrete monitoring error for a ϕ-swap under the partition $\mathbf{\Pi}_N$ may be written

$$\varepsilon_N(\phi, \mathbf{z}) := \mathbb{E}^{\mathbb{Q}}\left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right) - \langle \mathbf{z} \rangle_T^{\phi}\right]. \tag{2.8}$$

For instance, with $\mathbf{z} := x$ and $\phi(\hat{x}) := \hat{x}^2$ the definition (2.7) corresponds to the quadratic variation of the log-price and the discrete monitoring error is given by Equation (2.2).

Our focus is on those combinations $(\phi, \mathbf{z})$ for which the swap satisfies

$$\mathbb{E}^{\mathbb{Q}}\left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right)\right] = \mathbb{E}^{\mathbb{Q}}\left[\langle \mathbf{z} \rangle_T^{\phi}\right], \tag{2.9}$$

for all partitions $\mathbf{\Pi}_N$. If (2.9) holds $\forall\, \mathbf{\Pi}_N$ then it holds for the trivial partition $\mathbf{\Pi}_1 = [0, T]$, for which the above becomes: $\mathbb{E}^{\mathbb{Q}}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_0\right)\right] = \mathbb{E}^{\mathbb{Q}}\left[\langle \mathbf{z} \rangle_T^{\phi}\right]$. But $\phi\left(\mathbf{z}_T - \mathbf{z}_0\right) = \phi\left(\sum_{\mathbf{\Pi}_N} \hat{\mathbf{z}}_i\right)$ hence (2.9) implies:

$$\mathbb{E}^{\mathbb{Q}}\left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right)\right] = \mathbb{E}^{\mathbb{Q}}\left[\phi\left(\sum_{\mathbf{\Pi}_N} \hat{\mathbf{z}}_i\right)\right], \tag{2.10}$$

for all $\mathbf{\Pi}_N$. The right hand side of (2.10) may also be written $\mathbb{E}^{\mathbb{Q}}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_0\right)\right]$. The lack of path-dependence in this 'implied characteristic' shows that that the jump error ι must also be zero. Thus, when the discrete monitoring error is zero under all partitions then, even if investors differ in their views about jump risk in an incomplete market, they would still agree on the fair-value ϕ-swap rate.

Neuberger [2012] calls (2.10) the aggregation property (AP).[3] The AP does not hold for (ϕ, x) when $\phi\left(\hat{x}\right) := \hat{x}^2$ is the realised characteristic for a conventional variance swap, but Neuberger finds two alternative generalised variance

[3]Another variation of Neuberger's property, namely

$$\mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} \tilde{\phi}\left(F_{t_i}, F_{t_{i-1}}\right)\right] = \mathbb{E}^{\mathbb{Q}}\left[\tilde{\phi}\left(F_T, F_0\right)\right],$$

is discussed in Bondarenko [2014]. This property is less general than ours in that only one dimension is considered, but more general in that the function need not be defined on increments only, i.e. $\tilde{\phi}\left(F_{t_i}, F_{t_{i-1}}\right) \neq \phi\left(F_{t_i} - F_{t_{i-1}}\right)$ necessarily.

characteristics: The log characteristic

$$\lambda(\hat{x}) := 2\left(\mathrm{e}^{\hat{x}} - 1 - \hat{x}\right) \tag{2.11}$$

for which the AP holds when x is the log of any martingale; and the entropy characteristic

$$\eta(\hat{x}) := 2\left(\hat{x}\mathrm{e}^{\hat{x}} - \mathrm{e}^{\hat{x}} + 1\right) \tag{2.12}$$

which satisfies the AP under the additional assumption of independent increments. Using Taylor expansion about the origin, one can see that both λ and η may be associated with the second moment of the distribution of $\hat{x}$, because they satisfy $\lim_{\hat{x}\to 0} \psi(\hat{x}) / \hat{x}^2 = 1$ for $\psi \equiv \lambda$ and $\psi \equiv \eta$. Then $v_t^{\psi} := v_t^{\lambda} = \mathbb{E}_t^{\mathbb{Q}}[\lambda(x_T - x_t)]$ and $v_t^{\psi} := v_t^{\eta} = \mathbb{E}_t^{\mathbb{Q}}[\eta(x_T - x_t)]$, called the log and entropy variance processes because they are closely related to the log contract, which pays x_T, and the entropy contract, which pays $s_T x_T$ at maturity, respectively.

Neuberger [2012] finds more characteristics by including the conditional variance process v_t or other conditional processes v_t^{ψ} in $\mathbf{z}$. He first examines the two-dimensional stochastic process $\mathbf{z} := (F, v)'$ with $v := \{v_t\}_{t\in\mathbf{\Pi}}$ being the conditional variance process

$$v_t := \mathbb{E}_t^{\mathbb{Q}}\left[(F_T - F_t)^2\right],$$

and proves that the set $\mathbb{A}$ of all functions ϕ which satisfy (2.10) for $\mathbf{z}$ is given by

$$\mathbb{A} := \left\{(\alpha, \gamma)\,\hat{\mathbf{z}} + \Omega\hat{F}^2 + h\left(\hat{F}^3 + 3\hat{F}\hat{v}\right)\right\}, \tag{2.13}$$

for $\alpha, \gamma, \Omega, h \in \mathbb{R}$. He calls this the arithmetic world of price changes and then turns towards what he calls the geometric world of log-returns. There we have $\mathbf{z} := \left(x, v^{\psi}\right)'$ with $v^{\psi} := \left\{v_t^{\psi}\right\}_{t\in\mathbf{\Pi}}$ being a generalised conditional variance process of the form

$$v_t^{\psi} := \mathbb{E}_t^{\mathbb{Q}}\left[\psi\left(x_T - x_t\right)\right],$$

which must satisfy the condition $\lim_{x\to 0}\psi(x)/x^2 = 1$ in order for v^{ψ} to be associated with (implied) variance. The set $\mathbb{G}$ of all functions which satisfy (2.10) for this $\mathbf{z}$ is given by

$$\mathbb{G} := \left\{\boldsymbol{\gamma}'\hat{\mathbf{z}} + \beta\left(\mathrm{e}^{\hat{x}} - 1\right) + \Omega\left(2\hat{x} - \hat{v}^{\psi}\right)^2 + h\left(2\hat{x} + \hat{v}^{\psi}\right)\mathrm{e}^{\hat{x}}\right\}, \tag{2.14}$$

for $\beta, \Omega, h \in \mathbb{R}$ and $\boldsymbol{\gamma} \in \mathbb{R}^2$, subject to the following constraints:

$$\begin{array}{ll} \text{if } \Omega \neq 0, & h = 0 \text{ and } \psi = \lambda \text{ as defined in (2.11)}, \\ \text{if } h \neq 0, & \Omega = 0 \text{ and } \psi = \eta \text{ as defined in (2.12)}, \\ \text{if } \Omega = h = 0, & \psi \text{ can be any generalised variance.} \end{array} \tag{2.15}$$

The log characteristic corresponds to the parameterisation $\gamma_1 = -\beta = -2$ and $\gamma_2 = \Omega = h = 0$ of $\mathbb{G}$. Since the log characteristic is an AP-characteristic w.r.t. the log of any martingale, x, one can change the definition of the floating leg of a variance swap from (2.1) to $\sum_{\mathbf{\Pi}_D}\lambda\left(\hat{x}_t\right)$, and the result will be a log variance swap whose fair-value swap rate can be derived from the replication theorem of Carr and Madan [2001] without any discrete monitoring error or jump error terms:[4]

$$\mathbb{E}^{\mathbb{Q}}\left[\lambda\left(\sum_{\mathbf{\Pi}_D}\hat{x}_t\right)\right] = 2\int_{\mathbb{R}^+} k^{-2}q(k)dk.$$

Within the set $\mathbb{G}$ of pay-off functions Neuberger further identifies the characteristic

$$3\hat{v}^{\psi}\left(\mathrm{e}^{\hat{x}}-1\right)+\tau\left(\hat{x}\right),$$

with $\tau\left(\hat{x}\right) := 6\left(\hat{x}\mathrm{e}^{\hat{x}}-2\mathrm{e}^{\hat{x}}+\hat{x}+2\right)$, which corresponds to the parameterisation $\gamma_1 = 6$, $\beta = -12$, $\gamma_2 = -3$, $\Omega = 0$ and $h = 3$, and argues that it approximates the third moment of log returns since $\lim_{\hat{x}\to 0}\tau\left(\hat{x}\right)/\hat{x}^3 = 1$. However, the first term does not vanish under expectation for partial increments even if F follows a martingale. In fact it measures the covariance between returns and changes in implied variance. For the fair-value swap rate we have

$$\mathbb{E}^{\mathbb{Q}}\left[3\left(v^{\psi}_{T}-v^{\psi}_{0}\right)\left(\mathrm{e}^{x_T-x_0}-1\right)+\tau\left(x_T-x_0\right)\right] = \mathbb{E}\left[\tau\left(x_T-x_0\right)\right],$$

which is dominated by the higher-order terms of τ for sufficiently large $x_T - x_0$. Therefore the association of either the floating or the fixed leg of this swap with the third moment is questionable.

The subsequent empirical study by Kozhan et al. [2013] shows that the risk premium associated with this swap is strongly correlated with the VRP. The

[4]See Neuberger [2012], p.7: “If the measure is a pricing measure, it says that the fair price of a one-month variance swap computed daily (a swap that pays the realised daily variance over a month) is the same as the price of a contingent claim that pays $(S_T - S_0)^2$. Indeed, because the relationship holds under any pricing measure (because the process is a martingale under any pricing measure), it also implies that a variance swap can be perfectly replicated if the contingent claim exists (or can be synthesised from other contingent claims) and the underlying asset is traded.”

flexibility to define a great variety of swaps with potentially diverse risk premia motivates our more general class of $(\phi, \mathbf{z})$ for which (2.10) holds, and from henceforth we refer to (2.10) as the AP.

2.2.3 Discretisation Invariance

Consider a multivariate stochastic process $\mathbf{z} \in \mathbb{R}^n$ which contains only deterministic functions of the forward prices $\mathbf{F} := \{\mathbf{F}_t\}_{t \in \mathbf{\Pi}} \in \mathbb{R}^d$ of d tradable assets, or derivatives on these assets, in an arbitrage-free market. For instance, the process $\mathbf{z}$ may contain forward prices and/or the logs of these prices. We make the minimal no-arbitrage assumption only to ensure that the forward prices follow a multivariate $\mathbb{Q}$-martingale. We define a DI swap to be any ϕ-swap on $\mathbf{z}$ for which the AP (2.10) holds. Two trivial DI swaps are: (a) if ϕ is linear, say $\phi(\hat{\mathbf{z}}) = \boldsymbol{\alpha}'\hat{\mathbf{z}}$ for some $\boldsymbol{\alpha} \in \mathbb{R}^n$, then (2.10) holds for any process $\mathbf{z}$ because $\sum_{\mathbf{\Pi}_N} \hat{\mathbf{z}}_i = \mathbf{z}_T - \mathbf{z}_0$; (b) if $\mathbf{z}$ contains only constant processes then $\hat{\mathbf{z}}_i = \mathbf{0}\ \forall i \in \{1, \ldots, N\}$, so (2.10) holds for any function with $\phi(\mathbf{0}) = 0$. Note that (2.9) also holds in both cases: in (a) because $\langle \mathbf{z} \rangle_T^\phi = \mathbf{z}_T - \mathbf{z}_0$ and in case (b) because $\langle \mathbf{z} \rangle_T^\phi = 0$, provided $\phi(\mathbf{0}) = 0$.

In what follows we only consider characteristics $\phi \in C^2$ (and for which $\phi(\mathbf{0}) = 0$). Let $\mathbf{\Delta} \in \mathbb{R}^{n \times d}$ and $\mathbf{\Gamma} \in \mathbb{R}^{n \times d \times d}$ denote the first and second partial derivatives of $\mathbf{z}$ w.r.t. $\mathbf{F}$ and denote by $\mathbf{J}(\hat{\mathbf{z}}) \in \mathbb{R}^n$ the Jacobian vector and $\mathbf{H}(\hat{\mathbf{z}}) \in \mathbb{R}^{n \times n}$ the Hessian matrix of first and second partial derivatives of ϕ w.r.t. $\hat{\mathbf{z}}$.

Theorem 1: Equivalence of the Aggregation Property

If $(\phi, \mathbf{z})$ is such that either the AP (2.10) holds, or the ϕ-variation of $\mathbf{z}$ exists and (2.9) holds, then the following second-order system of partial differential equations holds:

$$[\mathbf{J}(\hat{\mathbf{z}}) - \mathbf{J}(\mathbf{0})]' \mathbf{\Gamma} + \mathbf{\Delta}' [\mathbf{H}(\hat{\mathbf{z}}) - \mathbf{H}(\mathbf{0})] \mathbf{\Delta} = \mathbf{0}. \tag{2.16}$$

Further, when $\mathbf{F}$ follows a diffusion with finite ϕ-variation then (2.9), (2.10) and (2.16) are equivalent.

Proof: Let the forward price process $\mathbf{F}$ follow the $\mathbb{Q}$-dynamics $d\mathbf{F}_t = \boldsymbol{\sigma}_t d\mathbf{W}_t$ where $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_t\}_{t\in\mathbf{\Pi}} \in \mathbb{R}^{d\times d}$ and $\mathbf{W} = \{\mathbf{W}_t\}_{t\in\mathbf{\Pi}} \in \mathbb{R}^d$ is a multivariate Wiener process with $T^{-1}\langle\mathbf{W}\rangle_t = \mathbf{I}$, the identity matrix. Then $d\langle\mathbf{F}\rangle_t = \boldsymbol{\sigma}_t\boldsymbol{\sigma}_t' dt$ is the quadratic covariation process of $\mathbf{F}$. [5] Let $\mathbf{\Delta} := \nabla_{\mathbf{F}}' \mathbf{z} \in \mathbb{R}^{n\times d}$ and $\mathbf{\Gamma} := \nabla_{\mathbf{F}}'' \mathbf{\Delta} \in \mathbb{R}^{n\times d\times d}$ denote the first and second partial derivatives of $\mathbf{z}$ w.r.t. $\mathbf{F}$ where $\nabla_{\mathbf{F}} := \left(\frac{\partial}{\partial F_1}, \ldots, \frac{\partial}{\partial F_d}\right)'$. Then, applying Itô's Lemma and the cyclic property of the trace operator, we have

$$d\mathbf{z}_t = \mathbf{\Delta}_t d\mathbf{F}_t + \tfrac{1}{2}\mathrm{tr}\left(\mathbf{\Gamma}_t d\langle\mathbf{F}\rangle_t\right), \tag{2.17}$$

and the quadratic covariation process of $\mathbf{z}$ follows the dynamics

$$d\langle\mathbf{z}\rangle_t = \mathbf{\Delta}_t \boldsymbol{\sigma}_t \boldsymbol{\sigma}_t' \mathbf{\Delta}_t' dt. \tag{2.18}$$

Since we want the discrete monitoring error to be zero for all possible forward price processes, it must hold in particular for any specific martingale. We can therefore derive a necessary condition for the functions spanning $\mathbb{F}$ by starting from the assumptions that (2.9) holds w.r.t. $(\phi, \mathbf{z})$ and that $\mathbf{z}$ follows the dynamics specified in (2.17).

[5] The quadratic covariation is a straightforward generalisation of the quadratic variation for multivariate processes and is defined as $\langle\mathbf{z}\rangle_T := \lim_{\mathbf{\Pi}_N\to\mathbf{\Pi}} \sum_{\mathbf{\Pi}_N} \hat{\mathbf{z}}_i \hat{\mathbf{z}}_i' = \int_{\mathbf{\Pi}} d\mathbf{z}_t d\mathbf{z}_t'$. Note that the quadratic covariation $\langle\mathbf{z}\rangle$ is a matrix while the ϕ-variation $\langle\mathbf{z}\rangle^{\phi}$ is a scalar.

Denote the Jacobian vector of first partial derivatives of ϕ by $\mathbf{J}(\hat{\mathbf{z}}) := \nabla_{\mathbf{z}} \phi(\hat{\mathbf{z}}) \in \mathbb{R}^n$ and the Hessian matrix of second partial derivatives of ϕ by $\mathbf{H}(\hat{\mathbf{z}}) := \nabla'_{\mathbf{z}} \mathbf{J}(\hat{\mathbf{z}}) \in \mathbb{R}^{n \times n}$ where $\nabla_{\mathbf{z}} := \left(\frac{\partial}{\partial \hat{z}_1}, \ldots, \frac{\partial}{\partial \hat{z}_n} \right)'$. Then Itô's Lemma yields

$$\phi(\mathbf{z}_T - \mathbf{z}_0) = \int_{\mathbf{\Pi}} \mathbf{J}'(\mathbf{z}_t - \mathbf{z}_0)\, d\mathbf{z}_t + \tfrac{1}{2} \mathrm{tr} \int_{\mathbf{\Pi}} \mathbf{H}(\mathbf{z}_t - \mathbf{z}_0)\, d\langle \mathbf{z} \rangle_t. \tag{2.19}$$

Similarly,

$$\begin{aligned} \sum_{\mathbf{\Pi}_N} \phi(\hat{\mathbf{z}}_i) &= \sum_{i=1}^{N} \left\{ \int_{t_{i-1}}^{t_i} \mathbf{J}'\left(\mathbf{z}_t - \mathbf{z}_{t_{i-1}}\right) d\mathbf{z}_t + \tfrac{1}{2} \mathrm{tr} \int_{t_{i-1}}^{t_i} \mathbf{H}\left(\mathbf{z}_t - \mathbf{z}_{t_{i-1}}\right) d\langle \mathbf{z} \rangle_t \right\} \\ &= \int_{\mathbf{\Pi}} \mathbf{J}'\left(\mathbf{z}_t - \mathbf{z}_{m(t)}\right) d\mathbf{z}_t + \tfrac{1}{2} \mathrm{tr} \int_{\mathbf{\Pi}} \mathbf{H}\left(\mathbf{z}_t - \mathbf{z}_{m(t)}\right) d\langle \mathbf{z} \rangle_t, \end{aligned} \tag{2.20}$$

where $m(t) := \max\{t_i \in \mathbf{\Pi}_N | t_i \leq t\}$. Taking the limit as $\mathbf{\Pi}_N \to \mathbf{\Pi}$ yields the ϕ-variation

$$\langle \mathbf{z} \rangle_T^{\phi} = \int_{\mathbf{\Pi}} \mathbf{J}' d\mathbf{z}_t + \tfrac{1}{2} \mathrm{tr} \int_{\mathbf{\Pi}} \mathbf{H} d\langle \mathbf{z} \rangle_t, \tag{2.21}$$

where $\mathbf{J} := \mathbf{J}(\mathbf{0})$ and $\mathbf{H} := \mathbf{H}(\mathbf{0})$. With (2.19) and (2.21), the condition (2.9) is equivalent to

$$\mathbb{E}^{\mathbb{Q}} \left[\int_{\mathbf{\Pi}} \left[\mathbf{J}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{J}\right]' d\mathbf{z}_t + \tfrac{1}{2} \mathrm{tr} \int_{\mathbf{\Pi}} \left[\mathbf{H}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{H}\right] d\langle \mathbf{z} \rangle_t \right] = 0. \tag{2.22}$$

Substituting (2.17) and (2.18) in (2.22), and using $\mathbb{E}[d\mathbf{F}_t] = 0$ yields that (2.9) is equivalent to

$$\mathrm{tr} \mathbb{E}^{\mathbb{Q}} \left[\int_{\mathbf{\Pi}} \left\{ \left[\mathbf{J}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{J}\right]' \mathbf{\Gamma}_t + \mathbf{\Delta}'_t \left[\mathbf{H}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{H}\right] \mathbf{\Delta}_t \right\} \boldsymbol{\sigma}_t \boldsymbol{\sigma}'_t dt \right] = 0. \tag{2.23}$$

Now consider the spectral decomposition

$$[\mathbf{J}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{J}]' \boldsymbol{\Gamma}_t + \boldsymbol{\Delta}_t' [\mathbf{H}(\mathbf{z}_t - \mathbf{z}_0) - \mathbf{H}] \boldsymbol{\Delta}_t =: \mathbf{E}_t \boldsymbol{\Lambda}_t \mathbf{E}_t', \tag{2.24}$$

where $\boldsymbol{\Lambda}_t = \operatorname{diag}\{\lambda_{1t}, \ldots, \lambda_{dt}\}$ is a diagonal matrix of eigenvalues and $\mathbf{E}_t$ is an orthogonal matrix of eigenvectors. In order to derive a necessary condition for (2.9) we select the particular volatility process:

$$\boldsymbol{\sigma}_t := \exp\left\{\tfrac{1}{2}\xi \mathbf{E}_t \boldsymbol{\Lambda}_t \mathbf{E}_t'\right\},$$

where $\xi \in \mathbb{R}$ is an arbitrary constant. Because $\exp\{\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{-1}\} = \mathbf{E}\exp\{\boldsymbol{\Lambda}\}\mathbf{E}^{-1}$ for $\boldsymbol{\Lambda}, \mathbf{E} \in \mathbb{R}^{d\times d}$ we have

$$\boldsymbol{\sigma}_t \boldsymbol{\sigma}_t' = \mathbf{E}_t \exp\{\xi \boldsymbol{\Lambda}_t\} \mathbf{E}_t'. \tag{2.25}$$

Now inserting (2.24) and (2.25) into (2.23) and differentiating w.r.t. T, then using the cyclic property of the trace yields

$$\mathbb{E}^{\mathbb{Q}}\left[\operatorname{tr}\left(\boldsymbol{\Lambda}_t \exp\{\xi \boldsymbol{\Lambda}_t\}\right)\right] = 0.$$

Differentiating once w.r.t. ξ and evaluating the equation at $\xi = 0$ yields the condition

$$\mathbb{E}^{\mathbb{Q}}\left[\operatorname{tr}\left(\boldsymbol{\Lambda}_t^2\right)\right] = \sum_{i=1}^{d} \mathbb{E}^{\mathbb{Q}}\left[\left(\lambda_t^i\right)^2\right] = 0,$$

which implies that all eigenvalues in $\boldsymbol{\Lambda}_t$ must be equal to zero. Hence we know that both sides in (2.24) are zero and, given that this must hold for all $\mathbf{F}_t$ and $\mathbf{z}_0$, we can write

$$\left[\mathbf{J}\left(\hat{\mathbf{z}}\right)-\mathbf{J}\right]'\boldsymbol{\Gamma}+\boldsymbol{\Delta}'\left[\mathbf{H}\left(\hat{\mathbf{z}}\right)-\mathbf{H}\right]\boldsymbol{\Delta}=\mathbf{0}, \tag{2.26}$$

where $\mathbf{F}$ and $\hat{\mathbf{z}}$ are independent variables. We have derived this $d \times d$ system of partial differential equations based on the assumption that $\mathbf{F}$ follows a particular martingale diffusion, so it represents a necessary condition for the more general case where $\mathbf{F}$ can be any martingale diffusion. However, since (2.26) is also sufficient for (2.23) to hold, the two conditions are equivalent.[6] $\square$

For given $\mathbf{z}$ the system in Theorem 1 may be solved numerically to yield the characteristics that define a DI swap on $\mathbf{z}$. However, in order to define realised characteristics that can be monitored in practice we are only interested in analytic solutions of (2.16). The following Theorem is proved by solving (2.16) for a particular $\mathbf{z}$ and then showing, by straightforward evaluation of (2.9), that the necessary condition is sufficient:

Theorem 2: Discretisation-Invariant Characteristics

Let $\mathbf{F}$ follow any d-dimensional martingale process and set $\mathbf{z} = (\mathbf{F}, \mathbf{x})'$ with $\mathbf{x} := \ln \mathbf{F}$. Then the solutions to (2.16) form a vector space over $\mathbb{R}$, defined by:

$$\mathbb{F} := \left\{ \phi : \mathbb{R}^n \to \mathbb{R} \,\middle|\, \phi\left(\hat{\mathbf{z}}\right) = \boldsymbol{\alpha}'\hat{\mathbf{F}} + \hat{\mathbf{F}}'\boldsymbol{\Omega}\hat{\mathbf{F}} + \boldsymbol{\beta}'\left(\mathrm{e}^{\hat{\mathbf{x}}} - \mathbf{1}\right) + \boldsymbol{\gamma}'\hat{\mathbf{x}} \right\},$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{R}^d$, $\boldsymbol{\Omega}' = \boldsymbol{\Omega} \in \mathbb{R}^{d \times d}$.

[6] The proof can be performed analogously, this time assuming the AP, by substituting (2.19) and (2.20) into condition (2.10) which yields the same solution (2.26). This version does not require the existence of the ϕ-variation. Furthermore, if we relax our assumption that $\mathbf{F}$ follows a diffusion and allow any martingale then (2.26) still represents a necessary condition for (2.23).

Proof: When $\mathbf{z} = (\mathbf{F}, \mathbf{x})'$ we have $\boldsymbol{\Delta}(\mathbf{F}) = (\mathbf{I}, \mathrm{diag}(\mathbf{F})^{-1})' \in \mathbb{R}^{2d\times d}$ and $\boldsymbol{\Gamma}(\mathbf{F}) = (\mathbf{0}, -\mathrm{diag}_3(\mathbf{F})^{-2})' \in \mathbb{R}^{2d\times d\times d}$ where $\mathrm{diag}_3(\mathbf{F})$ denotes a three dimensional tensor with the elements of $\mathbf{F}$ on the diagonal and zeros everywhere else. We shall further use the following decompositions:

$$[\mathbf{J}(\hat{\mathbf{z}}) - \mathbf{J}(\mathbf{0})] = \begin{bmatrix} \mathbf{J}_{\mathbf{F}}(\hat{\mathbf{z}}) \\ \mathbf{J}_{\mathbf{x}}(\hat{\mathbf{z}}) \end{bmatrix} \in \mathbb{R}^{2d}$$

and

$$[\mathbf{H}(\hat{\mathbf{z}}) - \mathbf{H}(\mathbf{0})] = \begin{bmatrix} \mathbf{H}_{\mathbf{FF}}(\hat{\mathbf{z}}) & \mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}}) \\ \mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}})' & \mathbf{H}_{\mathbf{xx}}(\hat{\mathbf{z}}) \end{bmatrix} \in \mathbb{R}^{2d\times 2d}.$$

Then (2.26) may be written:

$$\begin{aligned} &-\mathbf{J}_{\mathbf{x}}(\hat{\mathbf{z}})' \,\mathrm{diag}_3(\mathbf{F})^{-2} + \mathbf{H}_{\mathbf{FF}}(\hat{\mathbf{z}}) + \mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}})\,\mathrm{diag}(\mathbf{F})^{-1} \\ &+\mathrm{diag}(\mathbf{F})^{-1}\mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}})' + \mathrm{diag}(\mathbf{F})^{-1}\mathbf{H}_{\mathbf{xx}}(\hat{\mathbf{z}})\,\mathrm{diag}(\mathbf{F})^{-1} = \mathbf{0} \end{aligned}$$

and multiplying from left and right with $\mathrm{diag}(\mathbf{F})$ yields

$$\begin{aligned} &-\mathrm{diag}(\mathbf{J}_{\mathbf{x}}(\hat{\mathbf{z}})) + \mathrm{diag}(\mathbf{F})\mathbf{H}_{\mathbf{FF}}(\hat{\mathbf{z}})\,\mathrm{diag}(\mathbf{F}) \\ &+\mathrm{diag}(\mathbf{F})\mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}}) + \mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}})'\,\mathrm{diag}(\mathbf{F}) + \mathbf{H}_{\mathbf{xx}}(\hat{\mathbf{z}}) = \mathbf{0}. \end{aligned}$$

Since this condition must be fulfilled for all martingale Itô processes $\mathbf{F}$ (and for $\mathbf{F} = \mathbf{1}$ in particular) this implies $\mathbf{H}_{\mathbf{FF}}(\hat{\mathbf{z}}) = \mathbf{H}_{\mathbf{Fx}}(\hat{\mathbf{z}}) = \mathbf{0}$ as well as $\mathbf{H}_{\mathbf{xx}}(\hat{\mathbf{z}}) = \mathrm{diag}(\mathbf{J}_{\mathbf{x}}(\hat{\mathbf{z}}))$. Therefore the solution must be of the form

$$\phi(\hat{\mathbf{z}}) = \boldsymbol{\alpha}'\hat{\mathbf{F}} + \hat{\mathbf{F}}'\boldsymbol{\Omega}\hat{\mathbf{F}} + \boldsymbol{\beta}'\left(\mathrm{e}^{\hat{\mathbf{x}}} - \mathbf{1}\right) + \boldsymbol{\gamma}'\hat{\mathbf{x}},$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{R}^d$ and $\boldsymbol{\Omega}' = \boldsymbol{\Omega} \in \mathbb{R}^{d\times d}$ is a symmetric matrix.

Swaps associated with $\boldsymbol{\alpha}$ are DI since $\lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \boldsymbol{\alpha}' \hat{\mathbf{F}}_i = \boldsymbol{\alpha}' \left(\mathbf{F}_T - \mathbf{F}_0 \right)$ even without expectation for any process. The same holds for swaps associated with $\boldsymbol{\gamma}$. For the swaps associated with $\boldsymbol{\Omega}$ we can apply

$$
\begin{aligned}
\mathbb{E}\left[\lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \hat{\mathbf{F}}_i' \boldsymbol{\Omega} \hat{\mathbf{F}}_i \right] &= \mathbb{E}\left[\lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \operatorname{tr}\left(\boldsymbol{\Omega} \hat{\mathbf{F}}_i \hat{\mathbf{F}}_i' \right) \right] \\
&= \operatorname{tr}\mathbb{E}\left[\boldsymbol{\Omega} \lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \left(\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}} \right) \left(\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}} \right)' \right] \\
&= \operatorname{tr}\mathbb{E}\left[\boldsymbol{\Omega} \lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \left(\mathbf{F}_{t_i} \mathbf{F}_{t_i}' - \mathbf{F}_{t_{i-1}} \mathbf{F}_{t_{i-1}}' \right) \right] \\
&= \operatorname{tr}\mathbb{E}\left[\boldsymbol{\Omega} \left(\mathbf{F}_T \mathbf{F}_T' - \mathbf{F}_0 \mathbf{F}_0' \right) \right] \\
&= \operatorname{tr}\mathbb{E}\left[\boldsymbol{\Omega} \left(\mathbf{F}_T - \mathbf{F}_0 \right) \left(\mathbf{F}_T - \mathbf{F}_0 \right)' \right] \\
&= \mathbb{E}\left[\left(\mathbf{F}_T - \mathbf{F}_0 \right)' \boldsymbol{\Omega} \left(\mathbf{F}_T - \mathbf{F}_0 \right) \right],
\end{aligned}
$$

where the only requirement is that $\mathbf{F}$ follows a martingale (not necessarily an Itô process). Finally, for all swaps associated with $\boldsymbol{\beta}$ we have

$$
\mathbb{E}\left[\lim_{\boldsymbol{\Pi}_N \to \boldsymbol{\Pi}} \sum_{\boldsymbol{\Pi}_N} \boldsymbol{\gamma}' \left(\mathrm{e}^{\hat{\mathbf{x}}} - \mathbf{1} \right) \right] = \mathbb{E}\left[\boldsymbol{\gamma}' \left(\mathrm{e}^{\mathbf{x}_T - \mathbf{x}_0} - \mathbf{1} \right) \right] = 0.
$$

Therefore, if $\mathbf{z} = (\mathbf{F}, \mathbf{x})'$, the necessary condition (2.26) is sufficient for all martingales. We can assume w.l.o.g. that $\boldsymbol{\Omega}$ is a symmetric matrix because $\hat{\mathbf{F}}' \boldsymbol{\Omega} \hat{\mathbf{F}}$ is a quadratic form. □

Using Theorem 2 we may define realised characteristics for DI swaps based on a wide variety of underlying variables $\mathbf{F}$. For instance, we can include the log contract $X_t := \mathbb{E}_t[x_T]$ or the fair-value price process of any other European pay-off in $\mathbf{F}$. Note that with $\mathbf{z} = (F, X, x)'$, $X = \{X_t\}_{t\in\mathbf{\Pi}}$, we can relate many of the characteristics introduced by Neuberger [2012] to specific characteristics in $\mathbb{F}$. For instance, when we set $\mathbf{F} = F$, the log characteristic can be obtained by choosing $\boldsymbol{\alpha} = \mathbf{0}$, $\boldsymbol{\Omega} = \mathbf{0}$, $\boldsymbol{\beta} = 2$ and $\boldsymbol{\gamma} = -2$.

We next consider the replication of the value process $V^{\phi} := \left\{V_t^{\phi}\right\}_{t\in\mathbf{\Pi}}$ of a ϕ-swap, i.e. $V_t^{\phi} := \mathbb{E}_t\left[\sum_{\mathbf{\Pi}_N} \phi(\hat{\mathbf{z}}_i)\right] - v_0^{\phi}$, where $v_0^{\phi} := \mathbb{E}[\phi(\mathbf{z}_T - \mathbf{z}_0)]$ denotes the fair-value swap rate at inception. Note that $V_0^{\phi} = 0$ by definition. When hedging the swap we seek to replicate $\hat{V}_t^{\phi} := V_t^{\phi} - V_{t-1}^{\phi}$, for which the following is useful:

Theorem 3: Replicating Discretisation-Invariant Swaps

For $t \in \mathbf{\Pi}_N$ the increments in the value process of any DI swap may be written

$$\hat{V}_t^{\phi} = \phi(\hat{\mathbf{z}}_t) + \hat{v}_t^{\phi}, \tag{2.27}$$

where $\hat{v}_t^{\phi} := v_t^{\phi} - v_{t-1}^{\phi}$ and $v_t^{\phi} := \mathbb{E}_t[\phi(\mathbf{z}_T - \mathbf{z}_t)]$ denotes the fair-value swap rate for the residual time-to-maturity. Further, when $\mathbf{z} = (\mathbf{F}, \mathbf{x})'$ as in Theorem 2 we have

$$\hat{V}_t^{\phi} = \boldsymbol{\alpha}'\hat{\mathbf{F}}_t + \operatorname{tr}\left(\boldsymbol{\Omega}\left[\hat{\boldsymbol{\Sigma}}_t - 2\mathbf{F}_{t-1}\hat{\mathbf{F}}_t'\right]\right) + \boldsymbol{\beta}'\left(\mathrm{e}^{\hat{\mathbf{x}}_t} - \mathbf{1}\right) + \boldsymbol{\gamma}'\hat{\mathbf{X}}_t,$$

where $\hat{\boldsymbol{\Sigma}}_t := \boldsymbol{\Sigma}_t - \boldsymbol{\Sigma}_{t-1}$ with $\boldsymbol{\Sigma}_t := \mathbb{E}_t\left[\mathbf{F}_T\mathbf{F}_T'\right]$ and $\hat{\mathbf{X}}_t := \mathbf{X}_t - \mathbf{X}_{t-1}$ with $\mathbf{X}_t := \mathbb{E}_t[\mathbf{x}_T]$. The corresponding fair-value ϕ-swap rate at inception is $v_0^{\phi} = \operatorname{tr}\left(\boldsymbol{\Omega}\left[\boldsymbol{\Sigma}_0 - \mathbf{F}_0\mathbf{F}_0'\right]\right) + \boldsymbol{\gamma}'(\mathbf{X}_0 - \mathbf{x}_0)$.

Proof: With the value process of a DI swap being defined as

$$V_t^{\phi} := \mathbb{E}_t \left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_i\right) \right] - v_0^{\phi},$$

the increments of the value process along the partition $\mathbf{\Pi}_N$ are given by

$$\begin{aligned}
\hat{V}_i^{\phi} &= V_{t_i}^{\phi} - V_{t_{i-1}}^{\phi} = \mathbb{E}_{t_i} \left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) \right] - \mathbb{E}_{t_{i-1}} \left[\sum_{\mathbf{\Pi}_N} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) \right] \\
&= \sum_{\tilde{i}=1}^{i} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) + \mathbb{E}_{t_i} \left[\sum_{\tilde{i}=i+1}^{N} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) \right] - \sum_{\tilde{i}=1}^{i-1} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) - \mathbb{E}_{t_{i-1}} \left[\sum_{\tilde{i}=i}^{N} \phi\left(\hat{\mathbf{z}}_{\tilde{i}}\right) \right] \\
&= \phi\left(\hat{\mathbf{z}}_i\right) + \mathbb{E}_{t_i}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_{t_i}\right)\right] - \mathbb{E}_{t_{i-1}}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_{t_{i-1}}\right)\right] \\
&= \phi\left(\hat{\mathbf{z}}_i\right) + \hat{v}_i^{\phi}
\end{aligned}$$

where $\hat{v}_i^{\phi} = v_{t_i}^{\phi} - v_{t_{i-1}}^{\phi}$ and $v_t^{\phi} = \mathbb{E}_t\left[\phi\left(\mathbf{z}_T - \mathbf{z}_t\right)\right]$. Combining the above with Theorem 2 yields

$$\begin{aligned}
\hat{v}_i^{\phi} &= \mathbb{E}_{t_i}\Big[\boldsymbol{\alpha}'\left(\mathbf{F}_T - \mathbf{F}_{t_i}\right) + \left(\mathbf{F}_T - \mathbf{F}_{t_i}\right)'\mathbf{\Omega}\left(\mathbf{F}_T - \mathbf{F}_{t_i}\right) \\
&\quad + \boldsymbol{\beta}'\left(\mathrm{e}^{\mathbf{x}_T - \mathbf{x}_{t_i}} - \mathbf{1}\right) + \boldsymbol{\gamma}'\left(\mathbf{x}_T - \mathbf{x}_{t_i}\right)\Big] \\
&\quad - \mathbb{E}_{t_{i-1}}\Big[\boldsymbol{\alpha}'\left(\mathbf{F}_T - \mathbf{F}_{t_{i-1}}\right) + \left(\mathbf{F}_T - \mathbf{F}_{t_{i-1}}\right)'\mathbf{\Omega}\left(\mathbf{F}_T - \mathbf{F}_{t_{i-1}}\right) \\
&\quad + \boldsymbol{\beta}'\left(\mathrm{e}^{\mathbf{x}_T - \mathbf{x}_{t_{i-1}}} - \mathbf{1}\right) + \boldsymbol{\gamma}'\left(\mathbf{x}_T - \mathbf{x}_{t_{i-1}}\right)\Big] \\
&= \mathbb{E}_{t_i}\left[\mathbf{F}_T'\mathbf{\Omega}\mathbf{F}_T + \boldsymbol{\gamma}'\mathbf{x}_T\right] - \mathbf{F}_{t_i}'\mathbf{\Omega}\mathbf{F}_{t_i} - \boldsymbol{\gamma}'\mathbf{x}_{t_i} \\
&\quad - \mathbb{E}_{t_{i-1}}\left[\mathbf{F}_T'\mathbf{\Omega}\mathbf{F}_T + \boldsymbol{\gamma}'\mathbf{x}_T\right] + \mathbf{F}_{t_{i-1}}'\mathbf{\Omega}\mathbf{F}_{t_{i-1}} + \boldsymbol{\gamma}'\mathbf{x}_{t_{i-1}} \\
&= \mathrm{tr}\left(\mathbf{\Omega}\hat{\mathbf{\Sigma}}_i\right) + \boldsymbol{\gamma}'\hat{\mathbf{X}}_i - \mathbf{F}_{t_i}'\mathbf{\Omega}\mathbf{F}_{t_i} + \mathbf{F}_{t_{i-1}}'\mathbf{\Omega}\mathbf{F}_{t_{i-1}} - \boldsymbol{\gamma}'\hat{\mathbf{x}}_i
\end{aligned}$$

where $\hat{\mathbf{\Sigma}}_i = \mathbf{\Sigma}_{t_i} - \mathbf{\Sigma}_{t_{i-1}}$, $\mathbf{\Sigma}_t = \mathbb{E}_t\left[\mathbf{F}_T\mathbf{F}_T'\right]$ as well as $\hat{\mathbf{X}}_i = \mathbf{X}_{t_i} - \mathbf{X}_{t_{i-1}}$, $\mathbf{X}_t = \mathbb{E}_t\left[\mathbf{x}_T\right]$.

Thus

$$
\begin{aligned}
\hat{V}_i^{\phi} &= \boldsymbol{\alpha}'\hat{\mathbf{F}}_i + \left(\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}}\right)'\boldsymbol{\Omega}\left(\mathbf{F}_{t_i} - \mathbf{F}_{t_{i-1}}\right) + \boldsymbol{\beta}'\left(\mathrm{e}^{\hat{\mathbf{x}}_i} - \mathbf{1}\right) + \boldsymbol{\gamma}'\hat{\mathbf{x}}_i + \hat{v}_i^{\phi} \\
&= \boldsymbol{\alpha}'\hat{\mathbf{F}}_i + \mathrm{tr}\left(\boldsymbol{\Omega}\left[\hat{\boldsymbol{\Sigma}}_i - 2\mathbf{F}_{t_{i-1}}\hat{\mathbf{F}}_i'\right]\right) + \boldsymbol{\beta}'\left(\mathrm{e}^{\hat{\mathbf{x}}_i} - \mathbf{1}\right) + \boldsymbol{\gamma}'\hat{\mathbf{X}}_i.
\end{aligned}
$$

The fair-value swap rate becomes

$$
\begin{aligned}
v_0^{\phi} &= \mathbb{E}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_0\right)\right] = \mathbb{E}\left[\boldsymbol{\alpha}'\left(\mathbf{F}_T - \mathbf{F}_0\right) + \left(\mathbf{F}_T - \mathbf{F}_0\right)'\boldsymbol{\Omega}\left(\mathbf{F}_T - \mathbf{F}_0\right)\right. \\
&\quad \left. + \boldsymbol{\beta}'\left(\mathrm{e}^{\mathbf{x}_T - \mathbf{x}_0} - \mathbf{1}\right) + \boldsymbol{\gamma}'\left(\mathbf{x}_T - \mathbf{x}_0\right)\right] \\
&= \mathbb{E}\left[\left(\mathbf{F}_T - \mathbf{F}_0\right)'\boldsymbol{\Omega}\left(\mathbf{F}_T - \mathbf{F}_0\right) + \boldsymbol{\gamma}'\left(\mathbf{x}_T - \mathbf{x}_0\right)\right] \\
&= \mathbb{E}\left[\mathrm{tr}\left(\boldsymbol{\Omega}\left[\mathbf{F}_T\mathbf{F}_T' - \mathbf{F}_0\mathbf{F}_0'\right]\right) + \boldsymbol{\gamma}'\left(\mathbf{x}_T - \mathbf{x}_0\right)\right] \\
&= \mathrm{tr}\left(\boldsymbol{\Omega}\left[\boldsymbol{\Sigma}_0 - \mathbf{F}_0\mathbf{F}_0'\right]\right) + \boldsymbol{\gamma}'\left(\mathbf{X}_0 - \mathbf{x}_0\right).
\end{aligned}
$$

□

Theorem 3 characterises the realised profit and loss (P&L) which accrues to the issuer of a DI swap who pays the fixed swap rate $\mathbb{E}\left[\sum_{\boldsymbol{\Pi}_N}\phi\left(\hat{\mathbf{z}}_i\right)\right] = \mathbb{E}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_0\right)\right]$ and receives the floating leg defined by the realised characteristic. The decomposition (2.27) separates the change in the realised characteristic from the change in the implied characteristic. While the value process follows a $\mathbb{Q}$-martingale, the two components are generally not $\mathbb{Q}$-martingales by definition.

Theorem 3 also shows that DI swaps are replicable in discrete time using a static trading strategy in $\boldsymbol{\Sigma} := \{\boldsymbol{\Sigma}_t\}_{t\in\boldsymbol{\Pi}}$ and $\mathbf{X} := \{\mathbf{X}_t\}_{t\in\boldsymbol{\Pi}}$ and a dynamic trading strategy in $\mathbf{F}$. For instance, the realised P&L for a swap on the log characteristic is $\hat{V}_t^{\lambda} = 2\left(\mathrm{e}^{\hat{x}_t} - 1 - \hat{X}_t\right)$ so, for $t \in \boldsymbol{\Pi}_N$, $V_t^{\lambda} = 2\sum_{i=1}^{t} F_{i-1}^{-1}\hat{F}_i - 2\left(X_t - X_0\right)$. Hence this swap can be hedged by buying two log contracts at time zero and shorting $2F_{t-1}^{-1}$ forward contracts from time $t-1$ to t.

We now introduce a canonical choice of implied fundamental contracts for $\mathbf{F}$ that are related to the log-return distribution of a single underlying forward contract with price F, denoting by $X^{(n)} = \left\{X_t^{(n)}\right\}_{t\in\mathbf{\Pi}}$, $X_t^{(n)} := \mathbb{E}_t\left[x_T^n\right]$ the n-th power log contract ($n \geq 2$).[7] According to the replication theorem of Carr and Madan [2001], this expectation can be expressed in terms of vanilla OTM options:

$$X_t^{(n)} = x_t^n + \int_{\mathbb{R}^+} \gamma_n(k) q_t(k) dk, \tag{2.28}$$

where $\gamma_n(k) := n(\ln k)^{n-2} k^{-2}\left[n - 1 - \ln k\right]$ and $q_t(k)$ denotes the time-t forward price of a vanilla OTM option with strike k and maturity T. We may also consider the alternative replication scheme:

$$X_t^{(n)} = x_0^n + n x_0^{n-1}\left(\tfrac{F_t}{F_0} - 1\right) + \int_0^{F_0} \gamma_n(k) P_t(k) dk + \int_{F_0}^{\infty} \gamma_n(k) C_t(k) dk,$$

where $P_t(k)$ and $C_t(k)$ denote the time-t forward prices of vanilla put and call options with strike k and maturity T. The difference between the two replication schemes is that (2.28) is, at any point in time, based only on OTM options because they are more liquidly traded. But due to the stochastic separation strike F_t this portfolio would require continuous rebalancing between puts and calls. The alternative replication scheme involves options that are OTM only at inception but this portfolio describes buy-and-hold strategies that require no dynamic rebalancing. From a theoretical perspective the two representations are interchangeable. However, the OTM scheme may be favorable for computing the fair-value swap

[7] For reasons of space we now focus only those pay-offs that are related to what Neuberger calls the 'geometric world' of log-returns. We have experimented with various DI moment swaps related to the 'arithmetic world' of prices or such contracts that combine prices with log-returns (e.g. the entropy swap). Our theoretical and empirical results, available upon request, suggest that the relevant risk premia are already accessible in the geometric setting.

rate while the alternative scheme may be preferred for hedging. By construction, the price process of any n-th power log contract follows a $\mathbb{Q}$-martingale and can therefore be included in $\mathbf{F}$.

For the next result we suppose that $\mathbf{F}$ contains power log contracts whose corresponding replication portfolios may be derived from (2.28). For instance, for the first four power log contracts we have:

$$\begin{aligned}
\text{log contract:} \quad X_t &= \quad x_t - \int_{\mathbb{R}^+} k^{-2} q_t(k) dk, \\
\text{squared log contract:} \quad X_t^{(2)} &= \quad x_t^2 + 2 \int_{\mathbb{R}^+} (1 - \ln k)\, k^{-2} q_t(k) dk, \\
\text{cubed log contract:} \quad X_t^{(3)} &= \quad x_t^3 + 3 \int_{\mathbb{R}^+} \ln k\, (2 - \ln k)\, k^{-2} q_t(k) dk, \\
\text{quartic log contract:} \quad X_t^{(4)} &= \quad x_t^4 + 4 \int_{\mathbb{R}^+} (\ln k)^2\, (3 - \ln k)\, k^{-2} q_t(k) dk.
\end{aligned}$$

Theorem 4: DI Moment Swaps on the Log Return
Let $\mathbf{F}_t = \left(X_t, X_t^{(2)} \ldots, X_t^{(n-1)} \right)'$ for some $n \geq 2$ and consider the parameters

$$\boldsymbol{\alpha} = \boldsymbol{\beta} = \boldsymbol{\gamma} = \mathbf{0} \quad \text{and} \quad \boldsymbol{\Omega} = \boldsymbol{\Omega}^{(n)} := \begin{bmatrix} \omega_1^{(n)} & \frac{1}{2}\omega_2^{(n)} & \ldots & \frac{1}{2}\omega_{n-1}^{(n)} \\ \frac{1}{2}\omega_2^{(n)} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\omega_{n-1}^{(n)} & 0 & \ldots & 0 \end{bmatrix}$$

with

$$\omega_i^{(n)} := X_0^{n-1-i} \sum_{j=i+1}^{n} \binom{n}{j} (-1)^{n-j} = -X_0^{n-1-i} \sum_{j=0}^{i} \binom{n}{j} (-1)^{n-j} = 0$$

since $\sum_{j=0}^{n}\binom{n}{j}(-1)^{n-j}=0$ for $i\in\{1,\ldots,n-1\}$. Then $v_0^\phi = v_0^{(n)}$ where

$$v_0^{(n)} := \mathbb{E}\left[(x_T - X_0)^n\right] = \sum_{i=1}^{n}\binom{n}{i}\left(-X_0\right)^{n-i}X_0^{(i)} + \left(-X_0\right)^n$$

denotes the n-th (central) moment of the log-return distribution of F.

Proof: Starting with

$$\boldsymbol{\Sigma}_0 - \mathbf{F}_0\mathbf{F}_0' = \begin{bmatrix} X_0^{(2)} - X_0X_0 & \ldots & X_0^{(n)} - X_0X_0^{(n-1)} \\ \vdots & \ddots & \vdots \\ X_0^{(n)} - X_0X_0^{(n-1)} & \ldots & X_0^{(2n-2)} - X_0^{(n-1)}X_0^{(n-1)} \end{bmatrix}$$

for some $n \geq 2$ we apply Theorem 3 as follows:

$$\begin{aligned}
v_0^\phi &= \mathbb{E}\left[\phi\left(\mathbf{z}_T - \mathbf{z}_0\right)\right] = \operatorname{tr}\left(\boldsymbol{\Omega}^{(n)}\left[\boldsymbol{\Sigma}_0 - \mathbf{F}_0\mathbf{F}_0'\right]\right) = \sum_{i=1}^{n-1}\omega_i^{(n)}\left(X_0^{(i+1)} - X_0X_0^{(i)}\right) \\
&= \omega_{n-1}^{(n)}X_0^{(n)} + \sum_{i=2}^{n-1}\left(\omega_{i-1}^{(n)} - \omega_i^{(n)}X_0\right)X_0^{(i)} - \omega_1^{(n)}X_0^2 \\
&= X_0^{(n)} + \sum_{i=2}^{n-1}\binom{n}{i}\left(-X_0\right)^{n-i}X_0^{(i)} + (1-n)\left(-X_0\right)^n \\
&= \sum_{i=1}^{n}\binom{n}{i}\left(-X_0\right)^{n-i}X_0^{(i)} + \left(-X_0\right)^n = \mathbb{E}\left[\sum_{i=0}^{n}\binom{n}{i}\left(-X_0\right)^{n-i}x_T^i\right] \\
&= \mathbb{E}\left[(x_T - X_0)^n\right] = v_0^{(n)},
\end{aligned}$$

where we have used $\omega_{n-1}^{(n)} = 1$ and $\omega_1^{(n)} = \left(-X_0\right)^{n-2}(n-1)$ in the third line. $\square$

2.2.4 Higher-Moment Swaps

We now present specific examples of ϕ-swaps on $\mathbf{z}$ where the characteristic is related to the n-th moment of the log-return distribution. For ease of exposition from henceforth we use the daily partition $\mathbf{\Pi}_{_D}$ in the text while proofs remain for general $\mathbf{\Pi}_{_N}$.

Example 1: Variance Swap. As opposed to squared log-returns, squared price changes in the log contract represent a DI variance characteristic. Let $n = 2$ and consider the characteristic $\hat{X}^2$ which corresponds to $\mathbf{\Omega} = \mathbf{\Omega}^{(2)} = 1$. By construction the fair-value swap rate is $v_{_0}^{(2)} = \mathbb{E}\left[(x_{_T} - X_{_0})^2\right]$ where $X_0 = \mathbb{E}\left[x_{_T}\right]$ and $X_{_T} = x_{_T}$ at maturity. We can write the swap rate in terms of fundamental contracts, i.e. $v_{_0}^{(2)} = X_{_0}^{(2)} - X_{_0}^2$. Now, according to Theorem 3, the P&L on this swap may be written $\hat{V}_t^{(2)} = \hat{X}_t^{(2)} - 2X_{t-1}\hat{X}_t$. Hence, this swap can be hedged by selling a squared log contract and dynamically holding $2X_{t-1}$ log contracts from time $t-1$ to t. We can observe empirically that the risk premium on this variance swap is very highly correlated with that for Neuberger's variance swap.

Example 2: Third-Moment Swap. Let $n = 3$, i.e. $\mathbf{F} = \left(X, X^{(2)}\right)'$, and consider the characteristic $\hat{\mathbf{F}}'\mathbf{\Omega}^{(3)}\hat{\mathbf{F}}$ where

$$\mathbf{\Omega}^{(3)} = \begin{bmatrix} -2X_{_0} & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The fair-value swap rate is $v_{_0}^{(3)} = \mathbb{E}\left[(x_{_T} - X_{_0})^3\right] = X_{_0}^{(3)} - 3X_{_0}^{(2)}X_{_0} + 2X_{_0}^3$. By Theorem 3, $\hat{V}_t^{(3)} = \hat{X}_t^{(3)} - h_{2t}^{(3)}\hat{X}_t^{(2)} - h_{1t}^{(3)}\hat{X}_t$ with $h_{2t}^{(3)} := 2X_{_0} + X_{t-1}$ and $h_{1t}^{(3)} :=$

$X_{t-1}^{(2)} - 4X_0 X_{t-1}$. Hence, the swap can be hedged by selling a cubed log contract and dynamically holding $h_{2t}^{(3)}$ squared log contracts as well as $h_{1t}^{(3)}$ log contracts from time $t-1$ to t.

Example 3: Fourth-Moment Swap. Let $n = 4$, i.e. $\mathbf{F} = \left(X, X^{(2)}, X^{(3)}\right)'$, and consider the characteristic $\hat{\mathbf{F}}'\mathbf{\Omega}^{(4)}\hat{\mathbf{F}}$ where

$$\mathbf{\Omega}^{(4)} = \begin{bmatrix} 3X_0^2 & -\frac{3}{2}X_0 & \frac{1}{2} \\ -\frac{3}{2}X_0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}.$$

Then $v_0^{(4)} = \mathbb{E}\left[(x_T - X_0)^4\right] = X_0^{(4)} - 4X_0^{(3)}X_0 + 6X_0^{(2)}X_0^2 - 3X_0^4$ and $\hat{V}_t^{(4)} = \hat{X}_t^{(4)} - h_{3t}^{(4)}\hat{X}_t^{(3)} - h_{2t}^{(4)}\hat{X}_t^{(2)} - h_{1t}^{(4)}\hat{X}_t$ with $h_{3t}^{(4)} := 3X_0 + X_{t-1}$, $h_{2t}^{(4)} := -3X_0^2 - 3X_0X_{t-1}$ and $h_{1t}^{(4)} := X_{t-1}^{(3)} - 3X_0X_{t-1}^{(2)} + 6X_0^2X_{t-1}$ and the swap can be hedged by selling a quartic log contract and holding $h_{3t}^{(4)}$ cubed log contracts, $h_{2t}^{(4)}$ squared log contracts and $h_{1t}^{(4)}$ log contracts from $t-1$ to t.

Example 4: Alternative Fourth-Moment Swap. Let $\mathbf{F} = \left(X, X^{(2)}\right)'$ and consider the characteristic $\hat{\mathbf{F}}'\tilde{\mathbf{\Omega}}^{(4)}\hat{\mathbf{F}}$ where

$$\tilde{\mathbf{\Omega}}^{(4)} := \begin{bmatrix} X_0^{(2)} + 3X_0^2 & -2X_0 \\ -2X_0 & 1 \end{bmatrix}.$$

It is easy to show that the fair-value swap rate is $v_0^{(4)}$ as in Example 3. Now, by Theorem 3, $\hat{V}_t^{(4)} = \hat{X}_t^{(4)} - 4X_0\hat{X}_t^{(3)} - \tilde{h}_{2t}^{(4)}\hat{X}_t^{(2)} - \tilde{h}_{1t}^{(4)}\hat{X}_t$ with $\tilde{h}_{2t}^{(4)} := -X_0^{(2)} - 3X_0^2 + 2X_{t-1}^{(2)} - 4X_0X_{t-1}$ and $\tilde{h}_{1t}^{(4)} := -4X_0X_{t-1}^{(2)} + 6X_0^2X_{t-1} + 2X_0^{(2)}X_{t-1}$. Hence, the alternative fourth-moment swap can be hedged by selling a quartic log contract, buying $4X_0$ cubed log contracts and dynamically holding $\tilde{h}_{2t}^{(4)}$ squared log contracts as well as $\tilde{h}_{1t}^{(4)}$ log contracts from time $t-1$ to t. This swap has the advantage of not requiring dynamic trading in the cubed log contract.

Later we find empirically that the correlation between risk premia on the swaps defined by Examples 1, 2 are 3 can be quite low but, not surprisingly, the two fourth-moment swaps in Examples 3 and 4 have very highly correlated risk premia. Indeed, there are many other DI moment characteristics which readers can define using different parameterisations and payoff profiles in Theorem 2, but all our DI swaps of the same order moment are essentially capturing the same risks.

Also, similar to the standardisation of the third-moment swap in Kozhan et al. [2013], we standardise an n-th moment swap by dividing the change in both realised and implied by the corresponding power of the implied variance of the log-return, i.e.

$$V_t^{(\bar{n})} = V_t^{(n)} \left(X_0^{(2)} - X_0^2 \right)^{-n/2}. \tag{2.29}$$

In particular we define a skewness and a kurtosis swap on the log-return distribution by setting $V_t^{(\bar{3})} = V_t^{(3)} \left(X_0^{(2)} - X_0^2 \right)^{-3/2}$ and $V_t^{(\bar{4})} = V_t^{(4)} \left(X_0^{(2)} - X_0^2 \right)^{-2}$. The results in our empirical study will shed an interesting new light on the difference between the risk premia associated with standardised and non-standardised moment characteristics.

One of the challenges faced by issuers of standard variance swaps is to hedge the realised variance through dynamic rebalancing of an options portfolio which is tilted towards the low strike options via the weight k^{-2} in the replication formula for the fundamental contracts. These are the illiquid and expensive deep-OTM put options for which demand much exceeds supply during market crashes, because they provide insurance for risk-averse investors. The illiquid market in such options on single-name equities during the financial crisis of 2008-9 is the main reason why equity variance swaps are now focussed mainly on indices, rather than individual stocks. One way to circumvent this problem is to use power contracts on the price rather than the log price, which is consistent with analysing the 'arithmetic' world of price related swaps. The corresponding replication formula lacks the strong tilt present in power log contracts, putting more emphasis on high-strike options, the OTM calls where transactions costs are lower and the market is more liquid.

2.2.5 Strike-Discretisation-Invariant Swaps

All the examples of DI swaps considered so far have fair values which require integration over a continuum of strikes, but in practice options are traded on a relatively small number of discrete strikes. We now introduce strike-discretisation invariant (SDI) swaps that can be priced and replicated exactly based only on the available option prices. Like all other DI swaps they have the same fair value, independent of the partition $\mathbf{\Pi}_{N}$, which is free from both discrete monitoring and model-specific (e.g. jump) error. These swaps can also be hedged exactly without having to replicate the log or any other fundamental contract.

Let $\mathbf{F} = (\mathbf{P}, \mathbf{C})'$ where $\mathbf{P} := \{\mathbf{P}_t\}_{t\in\mathbf{\Pi}}$ and $\mathbf{C} := \{\mathbf{C}_t\}_{t\in\mathbf{\Pi}}$ describe the forward price processes of d vanilla put options and d vanilla call options, with identical, traded strikes $\mathbf{k}$, on an underlying with maturity T, so $\mathbf{P}_t := \mathbb{E}_t\left[(\mathbf{k} - s_{T}\mathbf{1})^+\right]$ and $\mathbf{C}_t := \mathbb{E}_t\left[(s_{T}\mathbf{1} - \mathbf{k})^+\right]$ where $\mathbf{1} := (1, \dots, 1)' \in \mathbb{R}^d$. Assume w.l.o.g. that the traded strikes $\mathbf{k} := (k_1, \dots, k_d)' \in \mathbb{R}^d$ are ordered such that $k_1 < k_2 < \dots < k_d$, and denote by $\hat{\mathbf{P}}$ and $\hat{\mathbf{C}}$ the increments in $\mathbf{P}$ and $\mathbf{C}$, respectively, along some partition of $[0, T]$. Let $\tilde{\mathbf{\Omega}} \in \mathbb{R}^{d\times d}$ be a lower triangular matrix and set

$$\boldsymbol{\alpha} = \boldsymbol{\beta} = \boldsymbol{\gamma} = \mathbf{0}, \quad \mathbf{\Omega} := \begin{bmatrix} \mathbf{0} & \frac{1}{2}\tilde{\mathbf{\Omega}} \\ \frac{1}{2}\tilde{\mathbf{\Omega}}' & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{2d\times 2d}$$

Since strikes are in ascending order either the put or the call has zero pay-off, so

$$\mathbb{E}\left[\mathbf{z}_T'\mathbf{\Omega}\mathbf{z}_T\right] = \mathbb{E}\left[\mathbf{P}_T'\tilde{\mathbf{\Omega}}\mathbf{C}_T\right] = \mathbb{E}\left[(\mathbf{k}' - s_T\mathbf{1}')^+ \tilde{\mathbf{\Omega}}\,(s_T\mathbf{1} - \mathbf{k})^+\right] = 0.$$

Now by Theorem 2: $v_0^\phi = \mathbb{E}\left[(\mathbf{F}_T - \mathbf{F}_0)'\,\mathbf{\Omega}\,(\mathbf{F}_T - \mathbf{F}_0)\right] = -\mathbf{P}_0'\tilde{\mathbf{\Omega}}\mathbf{C}_0$. Therefore an exact swap rate can be derived based only on the current prices $\mathbf{P}_0$ and $\mathbf{C}_0$ of traded vanilla options with strikes $\mathbf{k}$, without using the replication theorem of Carr and Madan [2001]. Next, by Theorem 3, the P&L on this swap is

$$\hat{V}_t^{[\mathbf{k}]} = \left[\boldsymbol{\alpha}_{\mathbf{C}}' - \mathbf{P}_{t-1}'\tilde{\mathbf{\Omega}}\right]\hat{\mathbf{C}}_t + \left[\boldsymbol{\alpha}_{\mathbf{P}}' - \mathbf{C}_{t-1}'\tilde{\mathbf{\Omega}}'\right]\hat{\mathbf{P}}_t$$

where $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_{\mathbf{P}}, \boldsymbol{\alpha}_{\mathbf{C}})'$. Hence, the swap can be hedged exactly by dynamically holding $\left[\mathbf{C}_{t-1}'\tilde{\mathbf{\Omega}}' - \boldsymbol{\alpha}_{\mathbf{P}}'\right]_j$ puts and $\left[\mathbf{P}_{t-1}'\tilde{\mathbf{\Omega}} - \boldsymbol{\alpha}_{\mathbf{C}}'\right]_j$ calls with strike k_j for $j = 1, 2, \dots, d$.

Example 5: Straddle Swap. Let $P := \{P_t\}_{t\in\mathbf{\Pi}}$ and $C := \{C_t\}_{t\in\mathbf{\Pi}}$ describe the forward price processes of a vanilla put and a call option with the same strike k, i.e. $P_t := \mathbb{E}_t\left[(k - s_T)^+\right]$ and $C_t := \mathbb{E}_t\left[(s_T - k)^+\right]$. Then $\mathbf{F} = (P, C)'$ follows a $\mathbb{Q}$-martingale and $\phi\,(\hat{\mathbf{z}}) = \hat{P}\hat{C}$ is a DI characteristic. Furthermore $\mathbb{E}\,[P_T C_T] = 0$

so that the corresponding fair-value swap rate $\mathbb{E}\left[(P_T - P_0)(C_T - C_0)\right] = -P_0 C_0$ can be determined solely based on the price of the put and the call option at inception. By Theorem 3, $\hat{V}_t^{[k]} = -P_{t-1}\hat{C}_t - C_{t-1}\hat{P}_t$ and the swap can be hedged exactly by dynamically holding C_{t-1} puts and P_{t-1} calls from time $t-1$ to t.

Chapter 3

Model-Free Moment Risk Premia

The variance risk premium (VRP) is a measure of how much investors are ready to pay in order to avoid exposure to changes in variance. It is commonly defined as the difference between some realised variance characteristic and its risk-neutral expectation. As we have illustrated in depth in the previous chapter, the standard characteristic – sum of squared log-returns – entails a variety of theoretical and practical problems. We have further shown how a modification of the realised leg can alleviate these problems and allow for a straight-forward generalisation of the swap and risk premium concept to higher-moments of the return distribution. Our higher-moment swaps are model-free and the corresponding fair-value swap rate does not depend on the monitoring scheme. This design makes it possible to analyse unbiased variance- and higher-moment risk premia at any frequency.

Accordingly, to the best of our knowledge, we can provide the first time series analysis based on daily risk premia and detect new empirical effects that are not apparent at lower frequencies. A comprehensive overview of statistical methods applied to financial market data is provided by Alexander [2001].

In the second main part of this thesis we make an important empirical contribution to the literature on variance and higher-moment risk premia in equity markets. We show that, even though fair values are the same whatever the monitoring partition, the salient features and in particular asymmetries and codependencies between risk premia on discretisation-invariant (DI) swaps depend on the frequency at which they are sampled. Our empirical study uses unbiased estimates of Standard & Poor's 500 Stock Market Index (S&P 500) variance and higher-moment risk premia (and risk premia associated with other univariate discretisation-invariant (DI) swaps) analysed at the daily, weekly and monthly frequencies over an 18-year period. We find strong evidence of asymmetric responses to market shocks in variance, skewness and kurtosis risk premia when sampled at the daily frequency. Their correlations also decrease markedly as the sampling frequency increases. These findings are relevant for hedge funds and other diversifiers with short-term investment horizons.

3.1 Literature Review

Most empirical literature on moment risk premia focusses on the variance risk premium (VRP) in the US equity market, where Carr and Wu [2009] provide the benchmark study. They suggest a method for measuring the VRP – based on squared log-returns and a portfolio of vanilla out-of-the-money (OTM) options

– and perform a historical market analysis for five stock indices and 35 individual stocks. According to the authors, this premium is on average negative for stock indexes under both bullish and bearish market conditions. Although mostly negative, the premiums on individual stocks show large cross-sectional variation. On the basis of this observation, Carr and Wu assume that there is a common stochastic variance risk factor in the stock market that causes negative risk premiums. They show that there is indeed a significant negative impact of the so-called 'variance-beta' on the logarithm of the VRP, analogue to the systematic market risk in the capital asset pricing model (CAPM).

3.1.1 Determinants of the Variance Risk Premium

The existence of variance risk premia raises the question whether these can be explained by one or more of the standard equity risk factors which have emerged in the asset pricing literature over the past decades. The most important framework includes the excess return (ER) on the market; the 'small minus big' (*size*) and the 'high minus low' (*growth*) factors introduced by Fama and French [1993]; and the 'up minus down' (*momentum*) factor introduced by Carhart [1997]. The size factor relates to the firm size and represents the historical excess returns of an investment in small firms over the investment in big firms. Historical excess returns of growth stocks over value stocks (as distinguished by the book-to-market ratio) are reflected in the growth factor. According to Fama and French [1993] (p.4), these two factors also cover leverage and earnings-price-ratio effects. Finally, the momentum factor represents a momentum strategy and measures the excess returns of firms that performed well during the last time period over those who performed badly.

It is sensible to assume that the well-studied leverage effect, i.e. the negative correlation between returns and return variance in stock markets, propagates the positive equity risk premium (ERP) to variance risk. For the German stock market, Hafner and Wallmeier [2007] document the presence of a negative VRP as well as a leverage effect. However, the analysis of Carr and Wu [2009] reveals that the equity premium can only account for part of the VRP in the US market. Other common sources of uncertainty such as firm size, book-to-market value or bond market indicators turn out not to have a significant impact on the premium. Kozhan et al. [2013] confirm these results. When analysing the determinants of the VRP for different maturity horizons, Nieto et al. [2014] find that variance risk exposure is not only suitable for portfolio diversification and speculation purposes, but that it can also provide a hedge against economic influence factors.

Carr and Wu [2009] also address the question whether the VRP is constant. By means of hypothesis testing, they provide evidence for a time varying VRP that is correlated with the variance swap rate. They conclude by proving the robustness of their results, stressing various assumptions made throughout the analysis. In particular, they show how the payoff to a continuously monitored variance swap is affected by jumps in the underlying process using a jump diffusion model with stochastic volatility as previously discussed in Bates [1996] and Bakshi et al. [1997]. Their findings further remain the same when they take transaction costs and possible asymmetries of the bid and ask quotes around the mid price into account. Finally, a subsample analysis shows that the VRP is negative under very different market conditions.

3.1.2 Asymmetry and Skewness Risk Premium

In economic terms, a negative VRP can be explained in the presence of risk-averse investors and it is well accepted in literature that representative agents in equity markets are actually risk-averse. Duan and Zhang [2014] estimate the risk-aversion via the generalised method of moments (GMM), using the VRP as well as implied higher-moments of the cumulative return distribution. The authors discuss skewness and kurtosis under the physical distribution as well as the impact of the central limit theorem (CLT). Chabi-Yo [2012] analyses the impact of risk aversion and skewness preference on the VRP. He fits a polynomial function of the market return to the empirical pricing kernel and then extends the function to incorporate skewness and kurtosis as additional stochastic variables. The partial equilibrium model he develops represents an attempt to solve the absolute risk aversion puzzle by taking into account the non-linear nature of empirically observable risk premia. This question leads naturally to the third main chapter of this thesis, where we develop non-linear pricing kernels for stochastic volatility asset pricing models.

In order to tackle the inconsistencies between the traditional CAPM theory and empirical observations Kraus and Litzenberger [1976] propose a three-parameter CAPM which includes the squared excess return as an additional explanatory variable. They find that, in addition to the previously detected risk aversion, investors have a preference for positive skewness, reflecting their fear of negative extreme events. When estimating this model using monthly equity data, Carr and Wu [2009] find no evidence for an asymmetric response of the VRP to market excess returns. However, as will be demonstrated later in this thesis, this finding depends strongly on the frequency of measurement.

Based on high-frequency S&P data, Dufour et al. [2012] perform an analysis which distinguishes between the leverage effect and the volatility feedback effect. They confirm the presence of an asymmetric impact of returns on volatility and the VRP and discuss possible causalities for this behaviour. An analysis of lagged variables shows that implied volatility has a considerable feedback effect and therefore market implied expectations are indeed a reasonable forecast of future volatility. More specifically, a positive shock on volatility has about twice the impact of a negative shock on the first day and the effect decays to zero within five days. The analysis is based on a jump-diffusion process as well as squared high-frequency log-returns.

Following the methodology of Carr and Wu [2009], but using the model-free realised characteristics introduced by Neuberger [2012] for the floating leg of a swap, Kozhan et al. [2013] perform a model-free analysis of the variance and skewness risk premia on the S&P 500. They propose a variance swap and a skewness swap that can each be replicated perfectly using hedging strategies in the futures and options markets, thus deriving unbiased estimates for the associated risk premia. Their monthly data leads to the conclusion that the equity skew and VRP are very highly correlated. This study is particularly interesting because it provides the first evidence for a significant skewness risk premium.

The empirically observable long-term skewness in financial market returns can be associated with high default correlations and systemic risk. Engle [2011] attributes this phenomenon to asymmetric volatility in short-period returns, which themselves may even be symmetrically distributed. He concludes that short- and long-term skewness are autonomous indicators for risk and equally important for the purpose of risk management. However, the standard view on short-term skewness as cubed short-period log-returns essentially incorporates the same information as short-term variance or in fact any power of short-term log returns. By taking the autocorrelation of returns and more sophisticated patterns of serial dependence into account, the methodology used by Kozhan et al. [2013] provides a set of trading strategies that are more effective for managing skewness risk.

3.1.3 Trading and Model Specification

When it comes to trading variance swaps, an important practical consideration is the optimal timing of the dynamic replication strategy. Bondarenko [2014] analyses the impact of non-optimal rebalancing times on the VRP on the S&P 500 index and argues that knowledge about the considerable deviations are relevant for exchanges, traders and regulators. In fact, financial derivatives have been developed that exploit risk premia between different monitoring and rebalancing schemes and those who trade, clear or certify such products have to be aware of the risks involved. Bondarenko [2014] compares results for the standard squared log-return characteristic with those for squared simple returns and Neuberger's discretisation-invariant variance characteristic.

While one branch of the literature experiments with different definitions of the realised leg used for defining a swap, more recent studies of variance risk premia such as Egloff et al. [2010] and Konstantinidi and Skiadopoulos [2014] employ market quotes (i.e. CBOE Volatility Index (VIX) futures prices) rather than synthetic variance swap rates for the fixed leg. This is because the latter are subject to a significant bias, as documented by Aït-Sahalia et al. [2015] and many others. The empirical relationship between the realised variance of the S&P index and the VIX index is discussed in Hsu and Murray [2007]. However, no market quotes are yet available for skewness, kurtosis and higher-moment swap rates.

Again following the methodology of Carr and Wu [2009], Ammann and Buesser [2013] analyse the VRP in the foreign exchange market. The authors detect a significant negative premium for intraday realised variance at a low-frequency, however, the picture becomes blurred when they analyse high-frequency data. Both the VIX index and the T-Bills – Eurodollar (TED) spread do have an impact on the VRP. Yet, there are considerable residual premia that are strongly time-varying. This confirms some main results from Guo [1998] who documents a significant, time-varying VRP in the foreign exchange market. Since a rise or drop in the exchange rate can be good news to the one and bad news to the other market participant, unlike with equities, there is no leverage effect in the foreign exchange markets. As a result, the VRP can not be explained by the premium paid for the underlying exchange rate risk. Although Ammann and Buesser [2013] claim that their methodology is model-free, they implicitly assume continuous monitoring and a pure diffusion process for the exchange rate by using squared log-returns for the floating leg of the swap. Also in the foreign exchange market, Bakshi et al. [2008] develop a stochastic discount factor model for the exchange rate triangle spanned by the US Dollar, British Pound and Japanese Yen which takes the variability of return skewness into account. Both the global and the currency-specific risk premia are stochastic and exhibit individual reactions to the economic environment. The authors find that negative, country-specific shocks yield the highest risk premia while global shocks are less priced and upward moves remain widely ignored.

Broadie et al. [2007] discuss model specification issues for the equity market based on a large sample of S&P 500 futures and options prices. In particular, the authors find evidence for jumps in the futures and the volatility process and analyse how these risk factors are priced in the market. They argue that "intuitively, volatility jumps should induce positive skewness and excess kurtosis in volatility increments" (p.1454) and propose a statistic for estimating the phenomenon of jumps in volatility. They conclude that, while introducing price jumps into a stochastic volatility model always yields significantly higher pricing performance, there is an interference between jumps in volatility and the risk premium associated with the volatility of price jump.

When analysing the contribution of jumps to the VRP in the equity market based on high-frequency data, Bollerslev and Todorov [2011] find that more than 50% of the premium can be associated with tail risk. They further report an asymmetry between jumps under the physical and jumps under the risk-neutral measure and attribute the large proportion of downside risk premium to investor fear of extreme negative market events. A new Investor Fear Index, as opposed to the VIX index, distinguishes clearly between common variance uncertainty and investors' fear. Using a new class of discrete-time models, Christoffersen et al. [2012] find that the risk premium associated with uncertainty about the jump intensity has a stronger impact on option prices than the VRP. Their approach allows for time-varying conditional skewness and kurtosis, which both depend on the jump intensity. The affine dynamic the authors use for modelling the pricing kernel is consistent with power utility for a representative investor. The impact of jump fears on the time-varying VRP is also addressed in Todorov [2010], who shows that investors' expectations about jumps change considerably after a market crash.

Although the direct way of getting exposed to variance risk is to trade variance swaps, a delta-hedged options portfolio is an important benchmark strategy. The main difference is the directional risk which is not present in the case of a variance swap investment. Bakshi and Kapadia [2003] compare the VRP with the average returns of such a hedged position and find that excess returns are less negative for OTM than for ATM options and more negative in times of financial distress. Essentially, the gains or losses on the options position depend on the VRP and the (model-dependent) portfolio vega.

3.1.4 Term Structure of the Variance Risk Premium

The study of variance risk premia in different markets is not restricted to its size, variation and determinants. In a recent working paper, Aït-Sahalia et al. [2015] perform a model-dependent analysis of the term-structure of variance risk premia, revealing a downward trend of the premium with increasing time to maturity of a contract. According to their results obtained from a principal component analysis, the two main factors driving the VRP term structure are the level and the slope, accounting for approximately 99.8% of all variance. By making model assumptions, the authors circumvent the lack of complete time series data for deep-OTM put (and call) options with a fixed strike and time to maturity. They also evaluate the effect of a jump risk component on the premium and try to explain how crash scenarios influence investors' behaviour for different investment horizons. All results are based on the sum of squared log-returns.

Egloff et al. [2010] also analyse the term-structure of variance swap rates for the S&P 500 index and deduce profitable trading strategies. They demonstrate that the term-structure of variance swap rates can take a variety of shapes, from contango to backwardating to inverse smile-like structures and argue that the two relevant drivers are the short and the long end of the term structure, which in fact covers the same range of variations as the approach taken by Aït-Sahalia et al. [2015], referring to the short end as the 'instantaneous variance rate variation' and to the long end as the 'central tendency factor'. According to their findings, it is on average more profitable for investors to sell long-term variance swaps than to sell short-term variance swaps (p.12). The inclusion of variance swap investments into the asset allocation improves the investment performance in- and out-of-sample and significantly reduces the necessity of dynamic hedging since variance risk is linearly spanned by the portfolio constituents. It is intuitive to model this linear structure using a general affine jump-diffusion model based on the theory of Duffie et al. [2000]. Like the majority of studies on the subject, this study uses squared log-returns as a measure of realised variance, which – as we discuss in the previous chapter of this thesis – is not consistent with using VIX quotes for the swap rate when the underlying process can jump. Egloff et al. [2010] conclude by remarking that, in order to distinguish between the effects of jumps in the underlying and stochastic volatility, academic literature either assumes constant volatility or pure diffusions and that "integrating these two dimensions can be a challenging but interesting direction for future research" (p.1308).

More recently, Filipović et al. [2016] calibrate a quadratic model for the variance swap rate term structure to S&P market data, proving the compatibility of the model with both upward- and downward-sloping term structures while outperforming some of the standard affine jump-diffusion models. The DI framework introduced in the previous section of this thesis provides academics with a tool that allows us to analyse the effects of stochastic volatility without making any further model assumption, be it the continuity of paths or a specific shape of the VRP term structure.

The term structure of variance swap rates can be used as a predictor variable for the equity premium, the VRP or interest rates, i.e. the bond premium for different maturities. After showing using principal component analysis (PCA) that three main factors, namely level, slope and curvature across maturities, explain 97% of the variation in variance swap rates, Feunou et al. [2014] show that two factors are crucial to explain the interdependence of the three premiums. An extension of the predictor variables to skewness and kurtosis yields no significant increase in explanatory power, which the authors explain via the argument that "the predictive content available from the term structure of different risk measures is broadly overlapping" (p.150). This conclusion may change when looking at higher-moment risk premia rather than swap rates.

3.1.5 Integration of the S&P and VIX Market

Bardgett et al. [2015] use S&P 500 and VIX data to estimate the affine jump-diffusion model previously applied by Egloff et al. [2010] and evaluate the information content of both data sets. They detect complementary information on jumps as well as the mean reversion level of stochastic volatility. Further, in times of market distress, S&P and VIX options contain conflicting information on implied volatilities. They estimate the model parameters for both the physical and the risk-neutral measure and define a set of risk premiums that are based on the differences in $\mathbb{P}$- and $\mathbb{Q}$-parameters. It turns out that the 'central tendency' of the variance term structure improves the fit of the return distribution while jumps in the volatility process allow to explain the upper tail of the variance distribution, i.e. the upward jumps in volatility when prices fall, thus facilitating the joint fit of both sources of information. Accordingly, there is a strong impact of jumps in volatility on the VRP.

The VRP in the VIX market, i.e. the difference between realised and implied VIX variance – the latter extracted from options on the VIX index – can also be understood as a 'variance of variance' risk premium on the underlying S&P 500 index. Using squared simple returns as their measure of realised variance, Barnea and Hogan [2012] determine the sign and size of the VRP in the VIX market. They report a negative VRP that exhibits occasional upward shocks. In particular, it is more negative, on average, than the VRP on the S&P and less time-varying. Trolle and Schwartz [2010] further find significantly negative VRP for energy commodities, crude oil and natural gas in particular. This is intriguing since commodity markets are commonly subject to an inverse leverage effect (that is, volatility increases as prices rise) because market participants are often companies that depend on commodities as an input and therefore high prices are bad news.

3.2 Empirical Results from the S&P Index

This section analyses the risk premia on S&P 500 DI swaps over an 18-year period from January 1996 to December 2013 using term-structure time series of different constant-maturity realised and implied characteristics. Our main purpose is to investigate the common factors influencing the term structure of variance and higher-moment risk premia. In contrast to most previous studies, with the notable exception of Kozhan et al. [2013], we examine the risk premia based on DI realised characteristics. This is because we can derive unbiased estimates of DI risk premia from their fair-value swap rates, i.e. we do not need to rely on market quotes which are anyway not currently available. We find empirical features in these DI risk premia which depend on their monitoring frequency, unlike their fair-value swap rates.

Most previous studies distinguish the sampling frequency of the data from the monitoring frequency of the realised characteristic, typically employing monthly or weekly data on a daily-monitored characteristic.[1] By contrast, we construct our data to match the sampling and monitoring frequencies, using daily data on daily-monitored characteristics, weekly data on weekly-monitored characteristics and monthly data on monthly-monitored characteristics (assuming 5 trading days per week and 20 trading days per month). This way, we can make inference on the properties of risk premia that are relevant for investors who monitor and rebalance positions every few days (e.g. hedge funds) as well as mutual fund and large institutional investors that typically have longer-term investment horizons.

[1] For instance, Kozhan et al. [2013] uses monthly data on daily-monitored skew swaps and Egloff et al. [2010] use weekly data on daily-monitored variance swaps.

Most previous work concerns the second category of investor, but here we are also interested in the potential benefits of short-term diversification and the immediate response of risk premia to market shocks that one can only investigate using daily (or higher frequency) data.

We present results for daily, weekly and monthly monitored characteristics with 30, 90 and 180 days to maturity: by varying the maturity we infer some interesting stylised facts about the term structure of implied moment characteristics; and different monitoring frequencies allow for comparison of daily, weekly and monthly statistical distributions of risk premia.

3.2.1 Data

Following Carr and Wu [2009], Todorov [2010] and others we generate observations on risk premia as the difference between the observed realised characteristic under the physical measure and its synthetic fair value under the risk-neutral measure. As previously mentioned, much previous research on the empirical behaviour of the VRP has used synthetic rates which yield biased estimates. An advantage of our theory is that synthetic swap rates do now yield unbiased estimates of risk premia. However, this typically comes at the cost of including fundamental contacts in our definitions of the realised characteristic and, as a result, the realised moments are not only based on the underlying futures time series but also on option price data.

We obtain daily closing prices P_t and C_t of all traded European put and call options on the S&P 500 between January 1996 and December 2013 and eliminate quotes that fulfil any of the following criteria: less than seven calendar days to maturity, more than 365 calendar days to maturity, zero trading volume, mid-price ≤ 0.5 or an implied Black Scholes volatility $\leq 1\%$ or ≥ 1. For each trading day, we further delete all quotes that refer to the same maturity if less than three different strikes are traded. The forward price is backed out via put-call-parity for each maturity from the pair of quotes whose strike minimises $|P_t - C_t|$. This forward price is also used as the separation strike between OTM put and call options, i.e. we use the put price for $k < F_t$ and the call price for $k \geq F_t$.

In order to preclude static arbitrage between strikes of the same maturity, and between options of different maturities, we apply the cubic spline interpolation algorithm developed by Fengler [2009]. For each day spanned by our sample this interpolation produces an equally distributed grid of OTM option prices with 2000 different strikes for each expiry date.[2] These data are then integrated numerically w.r.t. k to derive time series of daily prices (2.28) for the power log contracts, $n = 1, \ldots 4$. For example, the log contract is approximated by $X_t \approx x_t - \sum_{j=2}^{2000} k_j^{-2} q_t(k_j)(k_j - k_{j-1})$ and similar approximations apply for $X_t^{(n)}$. Next, using the parameterisation of Theorem 4 for DI moment swaps, we apply Theorem 3 (for the special case of power log contracts) to compute trading day, weekly and monthly increments in both the realised and implied characteristic on the r.h.s. of (2.27). Besides the daily partition $\mathbf{\Pi}_D$, we include increments along the partitions $\mathbf{\Pi}_W$ and $\mathbf{\Pi}_M$, reflecting swaps that are monitored on a weekly and monthly basis. This way the time series on risk premia have the same frequency as the monitoring of the swap.

Alternative methodologies for constructing a synthetic time series of risk pre-

[2]The strikes are equally distributed across a six-σ-range around the forward price, σ being the average implied volatility on that day, at a given maturity. Outside the domain of the spline we assume the implied volatility is constant and equal to the implied volatility at the closest strike.

mia over the entire 18-year sample period include: (a) hold a swap until just before maturity and then roll over to another swap with the same initial maturity, tracking observations on its realised characteristic and swap rate; (b) linearly interpolate synthetic constant-maturity swap rates and calculate the corresponding realised characteristic on every monitoring period; or (c) hold a swap for one monitoring period, then roll over to another swap with the same initial maturity.[3] The risk premia obtained using method (a) have a systematically varying maturity. Method (b) is good when the data frequency matches the maturity of the characteristic, but autocorrelation appears as an unwanted artefact when time series of higher frequencies are constructed. We use method (c) because it best facilitates an investigation of the relationship between risk premia, monitoring frequency and maturity. Because linear interpolation between prices produces synthetic constant-maturity contracts which are not truly reflective of investable returns, it is necessary here to apply linear interpolation to the daily, weekly or monthly value increments between the two adjacent traded maturities, as proved by Galai [1979].[4]

---

[3]Kozhan et al. [2013] (p.2184) follow (a), stating that "Our empirical analysis concentrates on trading strategies that run for a month, from the first trading day after one option expires to the next month's expiration date." Carr and Wu [2009] (p.1319) choose the construction method (b): "At each date t, we interpolate the synthetic variance swap rates at the two maturities to obtain the variance swap rate at a fixed 30-day horizon. [...] Corresponding to each 30-day variance swap rate, we also compute the annualised 30-day realised variance [...]."

[4]Thus, the change in price from time $t-1$ to time t of a contract Φ with constant time-to-maturity τ is $\hat{\Phi}_t := (T_u - T_l)^{-1}\left[(T_u - t - \tau)\,\hat{\Phi}_t^l - (T_l - t - \tau)\,\hat{\Phi}_t^u\right]$ where $\hat{\Phi}_t^l$ and $\hat{\Phi}_t^u$ denote the increments in the prices of the contracts with fixed maturity dates T_l and T_u. Note that increments refer now to daily, weekly or monthly increments in the constant-maturity time series, rather than the fixed-maturity series that we have used for developing the theory.

3.2.2 Risk Premia on S&P Moment Swaps

By construction, under the risk-neutral measure

$$\mathbb{E}\left[\hat{F}_t\right] = \mathbb{E}\left[\hat{X}_t\right] = \mathbb{E}\left[\hat{X}_t^{(n)}\right] = \mathbb{E}\left[\hat{V}_t^{\phi}\right] = \mathbb{E}\left[\hat{V}_t^{(n)}\right] = \mathbb{E}\left[\hat{V}_t^{(\bar{n})}\right] = \mathbb{E}\left[\hat{V}_t^{[k]}\right] = 0,$$

$\forall t \in \mathbf{\Pi}_N, n \geq 2$. However, under the physical probability measure the average increment (profit and loss (P&L)) on these contracts and swaps need not be zero, in the presence of a risk premium. Table 3.1 presents annualised estimates of the risk premia on different DI swaps, based on the entire 18-year sample period. The potential variation in P&L decreases as monitoring frequency increases, so to enable comparison between daily, weekly and monthly monitoring each premium is standardised by dividing the average increment ($\hat{F}$, $\hat{X}$, $\hat{V}^{(2)}$, etc.) by its standard deviation.

		F	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
	$\mathbf{\Pi}_D$	0.24	0.36	-0.55	0.32	0.79	-0.13	-0.56	0.34	0.44	0.61
$\tau = 30$	$\mathbf{\Pi}_W$	0.25	0.38	-0.73	0.38	1.12	-0.23	-0.90	0.41	0.60	1.19
	$\mathbf{\Pi}_M$	0.23	0.37	-0.54	0.31	0.17	-0.23	0.01	0.23	0.20	0.37
	$\mathbf{\Pi}_D$	0.23	0.34	-0.33	0.10	0.02	0.04	0.23	0.20	0.30	0.46
$\tau = 90$	$\mathbf{\Pi}_W$	0.25	0.37	-0.50	0.19	0.34	-0.03	-0.07	0.44	0.63	0.88
	$\mathbf{\Pi}_M$	0.22	0.36	-0.43	0.25	0.12	-0.17	0.06	0.26	0.27	0.36
	$\mathbf{\Pi}_D$	0.23	0.34	-0.22	0.06	-0.49	0.01	0.79	0.23	0.29	0.36
$\tau = 180$	$\mathbf{\Pi}_W$	0.24	0.35	-0.33	0.09	0.05	0.02	0.15	0.39	0.44	0.58
	$\mathbf{\Pi}_M$	0.21	0.35	-0.35	0.19	0.14	-0.12	-0.02	0.20	0.25	0.34

Table 3.1: Standardised risk premia between January 1996 and December 2013 on 30-day, 90-day and 180-day constant-maturity contracts based on daily, weekly and monthly monitoring, for: the forward F and log contract X, moment swaps on the log price $V^{(n)}$, the skewness swap $V^{(\bar{3})}$, the kurtosis swap $V^{(\bar{4})}$ as well as straddle swaps with strikes $k_1 = 1000$, $k_2 = 1100$ and $k_3 = 1200$.

The premia for the variance swap are negative while those for the forward, log contract and third-moment swap are positive. The pattern is less clear for skewness, fourth-moment and kurtosis swaps. Straddle swaps exhibit positive risk premia for all maturities and monitoring frequencies.[5] The risk premium associated with a DI moment swap on the S&P 500 tends to decrease in magnitude as the monitoring frequency increases, indicating an upward sloping moment-term-structure of the statistical return distribution. But also, the variance of the associated realised characteristic decreases when the monitoring frequency increases. Hence, the standardised risk premia in Table 3.1 exhibit no systematic pattern with respect to monitoring frequency. However, some of the risk premia do display a systematic pattern with respect to swap maturity. For instance, at each monitoring frequency the standardised risk premium on a 30-day variance swap is greater in magnitude than the corresponding premium on a 90-day variance swap which, in turn, is greater in magnitude than the 180-day swap risk premium. Similar remarks apply to the third-moment and fourth-moment swaps.

Figure 3.1 depicts the cumulative risk premia for 30-day constant-maturity moment swaps over the entire sample period. In each case the total risk premia on the right is disaggregated into realised and implied components, using Theorem 3. These graphs illustrate the dependence of moment risk premia on the monitoring frequency of the realised leg, which is the same as the rebalancing of the implied leg. We use a black line for daily, purple for weekly and red for monthly monitoring. The implied component of the VRP does not depend on the rebalancing frequency.[6] The very small variation evident in the top centre graph is due

[5] We do not list the risk premia on all fundamental contracts here since they are very similar. The main difference is that the price of each power log contract operates on its individual scale, implied by the order of the contract. Further results for power contracts as well as moment swaps on the price distribution are available from the author on request.

[6] That is, when the replication basket of options is rebalanced daily to constant 30-day matu-

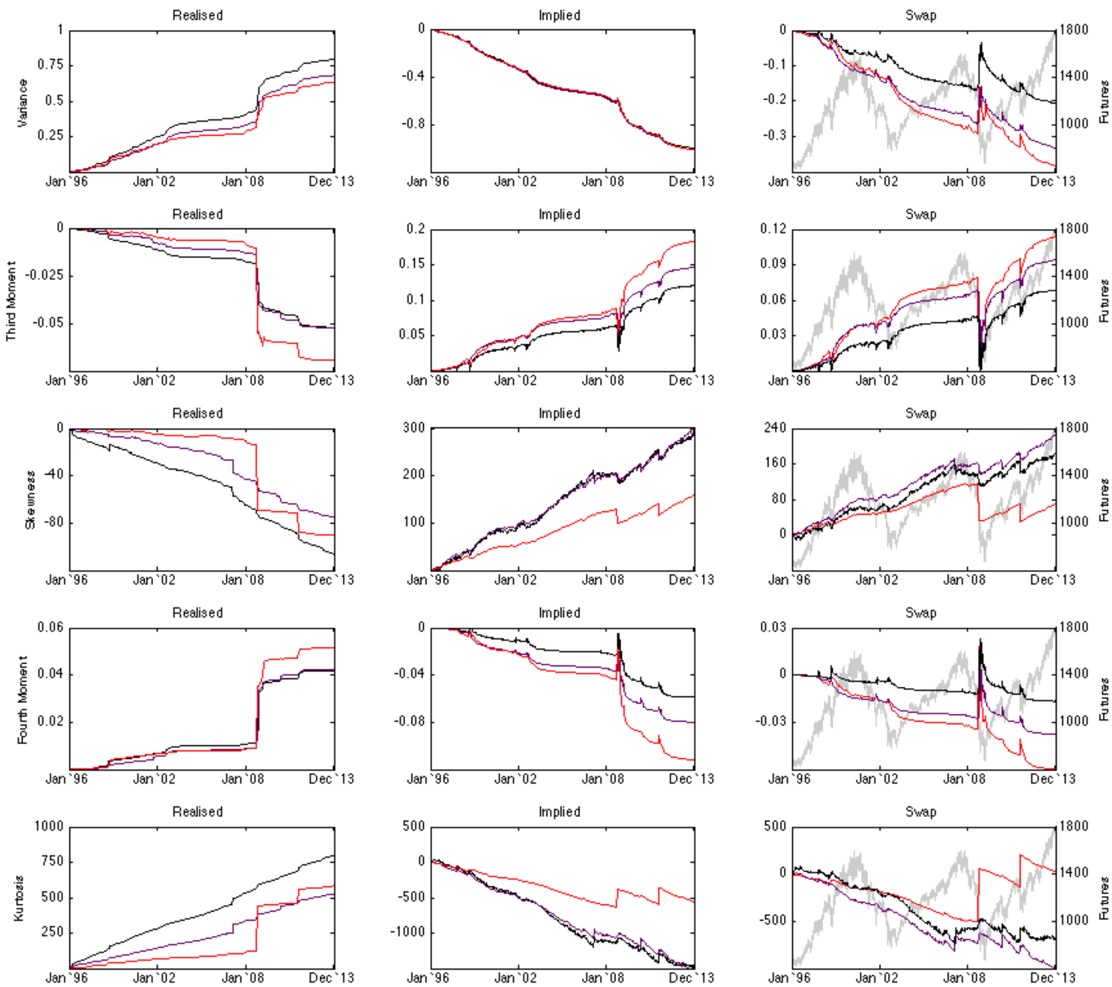

Figure 3.1: Time series of cumulative 30-day variance, third-moment, skewness, fourth-moment and kurtosis risk premia based on daily (black), weekly (purple) and monthly (red) monitoring. The secondary axis on the right refers to the 30-day forward contract plotted in grey. The first and second column of graphs depict the decomposition of the total cumulative risk premia into realised and implied components according to Equation (2.27).

to variation in the separation strike of the replication portfolio. It is the realised variance which drives the dependence of the VRP on the monitoring frequency.

---

rity and valued by marking-to-market (i.e. the black line), the cumulative change in the implied component is approximately the same as if the rebalancing and valuing happens weekly (purple) or monthly (red).

Overall, the VRP becomes smaller and less variable as monitoring frequency increases.[7] It is usually negative but during periods of equity market turmoil (such as the collapse of Lehman Brothers in September 2008 and the breaking news in August 2011 of a European sovereign debt crisis) it is, briefly, highly positive.

By contrast, the third-moment premium is usually positive, but falls sharply during crisis periods when the negative skew in realised returns on equities becomes especially pronounced. This is driven by the large jump down in the realised component during September 2008 (in the left graph in the second row). More generally this premium is dominated by the implied component depicted in the central graph. The effect of rebalancing the separation strike is more evident here than it is in the implied variance. For instance, in the monthly-monitored (red) time series the failure to rebalance the separation strike every day implies using higher-priced in-the-money calls in the replication portfolio during an upwards trending market, or higher-priced in-the-money puts in the replication portfolio during a downward market. A similar but opposite effect is evident in the implied component of the fourth-moment risk premium. As expected, given that the fourth moment captures outliers in a distribution, this premium is dominated by jumps in the index and is strongly positive during crisis periods.

The standardised third and fourth moment swaps have common features with their non-standardised counterparts. In particular, the direction of the realised and implied legs as well as the dependence on monitoring frequency are the same.

[7]This too is clear empirically, from Figure 3.1. Theoretical results to support these observations are model dependent. For instance, when $dF_t = \mu F_t + \sigma F_t dW_t$ where W_t is a Brownian motion it is straightforward to show that the risk premium associated with the conventional realised variance over a regular partition of $[0, T]$ into N elements is $\mu\left(\mu - \sigma^2\right) T^2 N^{-1}$ and the variance of this realised variance is $2\sigma^4 T^2 N^{-1} + 4\mu^2\sigma^2 T^3 N^{-2}$. Further model-dependent results confirm the statement for some other processes and DI variance characteristics.

However, at the daily and weekly monitoring frequency the standardised premia are much less volatile, indicating that changes in the non-standardised higher-moments usually coincide with changes in variance. At the monthly frequency, both the realised and implied leg as well as the risk premium are dominated by the two aforementioned extreme events. The standardised swaps exhibit an inert reaction towards changing market conditions, which is rooted in their construction since the implied volatility used for standardisation is always lagged by one monitoring period.

Figure 3.2 provides information on the term-structure of higher-moment risk premia using a black line for the cumulative risk premia on 30-day DI moment swaps, blue for 90-day swaps and green for DI swaps with 180 days to maturity. The implied component (top-centre graph) does not much depend on the time to maturity, so the term structure of implied variance is typically rather flat between 30-days and 180-days. And it is only during excessively volatile periods that the realised variance appears to increase with maturity. If we had used Neuberger's log variance characteristic here, which only depends on the underlying and not on any implied characteristic, then the realised leg would not depend on maturity at all. Note that realised characteristics depend on maturity because they include fundamental contacts, whose values are derived from options of that maturity. The skewness and kurtosis risk premia exhibit similar but opposite effects in both their implied and their realised components, both components become smaller in magnitude as maturity increases, and the implied component dominates the overall risk premium. The 30-day skew premium (black line) tends to be positive, except during turbulent market crises periods. The skew premium at 90 days (blue) is much smaller and close to zero and at 180 days (green) it tends to be

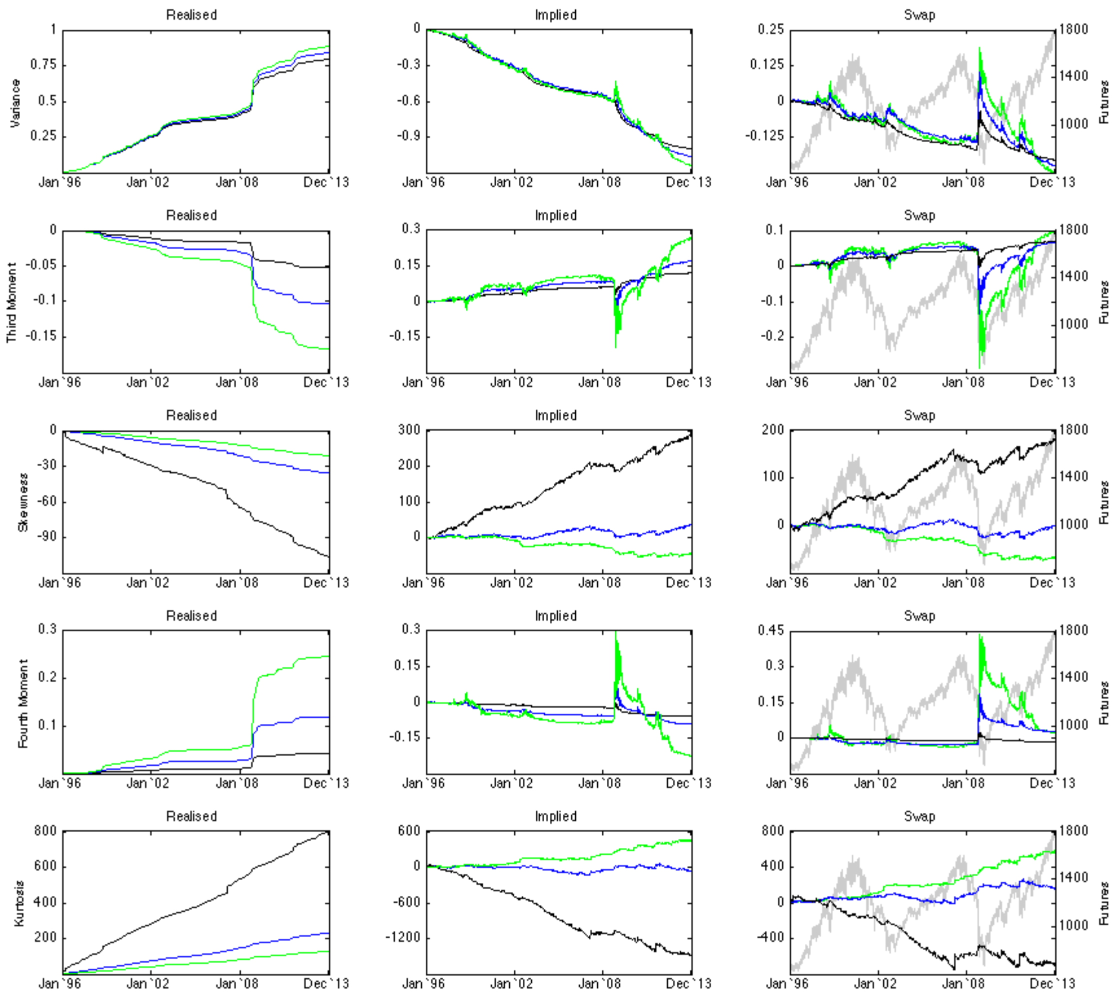

Figure 3.2: Time series for daily-monitored 30-day (black), 90-day (blue) and 180-day (green) variance, third-moment, skewness, fourth-moment and kurtosis cumulative risk premia. The secondary axis on the right refers to the 30-day forward contract plotted in grey. Again, these graphs depict the decomposition of the total cumulative risk premia into realised and implied components according to Equation (2.27).

negative. Similar features are evident in the kurtosis premium but with opposite signs: it is typically negative at 30 days, but sharply increases during periods leading up to a market crisis. Indeed, a fourth-moment swap may be replicated by adding a quartic contract to the portfolio (see Examples 3 or 4) and the quartic contract places even greater weight on the relatively low cost low-strike put options which become attractive to risk-averse investors seeking insurance against a market crash – before it happens. The kurtosis premium is much smaller (near zero) at 90 days and also small but positive at 180 days. Standardising by the implied variance as in (2.29) highlights the concentration of skewness and kurtosis in the very short-term implied distribution. Again it is only when the basket of options are re-balanced back to 30-day maturity on a monthly basis that we observe a difference in behaviour of the implied leg; the daily- and weekly rebalanced portfolios behave very similarly, just as in the skewness case.

3.2.3 Calendar, Frequency and Straddle Swaps

Given that risk premia can exhibit a strong term-structure pattern, as in Figure 3.2, the question arises whether systematic risk premia could be traded by entering a floating-floating 'calendar swap' which exchanges two realised characteristics, monitored at the same frequency, but with different maturities. For example, a 180-for-30-day calendar variance swap pays the forward realised variance, from 30 days after inception of the contract up to 180 days, in exchange for the corresponding fair-value swap rate, which equals the difference between the 180-day and 30-day swap rates.

Table 3.2 summarises the risk premia on some floating-floating swaps. For ease of comparison each premium is standardised by dividing by its standard deviation and annualising (as one does for the Sharpe ratio). The top panel exhibits the standardised risk premia obtained on 180-for-30-day calendar swaps monitored at three different frequencies. The incremental time series are shown in Figure 3.3.

As expected from the very different features of the skewness and kurtosis risk

Calendar		$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
$[\tau = 180]$	$\mathbf{\Pi}_D$	-0.05	0.02	-1.30	0.01	1.12	0.16	0.18	0.20
$-[\tau = 30]$	$\mathbf{\Pi}_W$	-0.03	0.02	-1.54	0.04	1.20	0.25	0.22	0.20
	$\mathbf{\Pi}_M$	-0.02	0.10	-0.18	-0.08	-0.02	0.05	0.18	0.12

Frequency		$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
$\tau = 30$		-0.63	0.37	-0.66	-0.41	0.59	0.27	0.16	0.45
$\tau = 90$	$\mathbf{\Pi}_M - \mathbf{\Pi}_D$	-0.52	0.53	0.31	-0.54	-0.11	0.37	0.30	0.28
$\tau = 180$		-0.46	0.48	1.60	-0.61	-1.77	-0.09	-0.04	0.07

Table 3.2: Standardised risk premia between January 1996 and December 2013 on daily, weekly and monthly monitored 180-for-30-day calendar swaps (above) and 30-day, 90-day and 180-day constant-maturity monthly-daily frequency swaps (below), where the swap rates are exchanged for: moment swaps on the log price $V^{(n)}$, the skewness swap $V^{(\bar{3})}$, the kurtosis swap $V^{(\bar{4})}$ as well as straddle swaps with strikes $k_1 = 1000$, $k_2 = 1100$ and $k_3 = 1200$.

premia displayed in Figure 3.2, the skewness (kurtosis) calendar swaps exhibit large negative (positive) premia at the daily and weekly monitoring frequencies. No other calendar swaps display significant results. The lower panel in Table 3.2 displays standardised risk premia on 'frequency swaps' which exchange two realised legs of the same maturity that are monitored at different frequencies, and the corresponding incremental time series are shown in Figure 3.4.

For instance, a monthly-daily variance frequency swap receives monthly and pays daily realised variance. Conveniently, the aggregation property (AP) implies that the fair-value rate on this type of swap is zero, by definition, but the risk premium may be positive or negative depending on the sample period and underlying characteristic. These frequency swaps tend to give larger risk premia in general and the skewness and kurtosis frequency swaps in particular have large risk premia (1.60 and -1.77 respectively) at the 180-day maturity.

Figure 3.5 depicts the time series of risk premia on straddle swaps with strikes

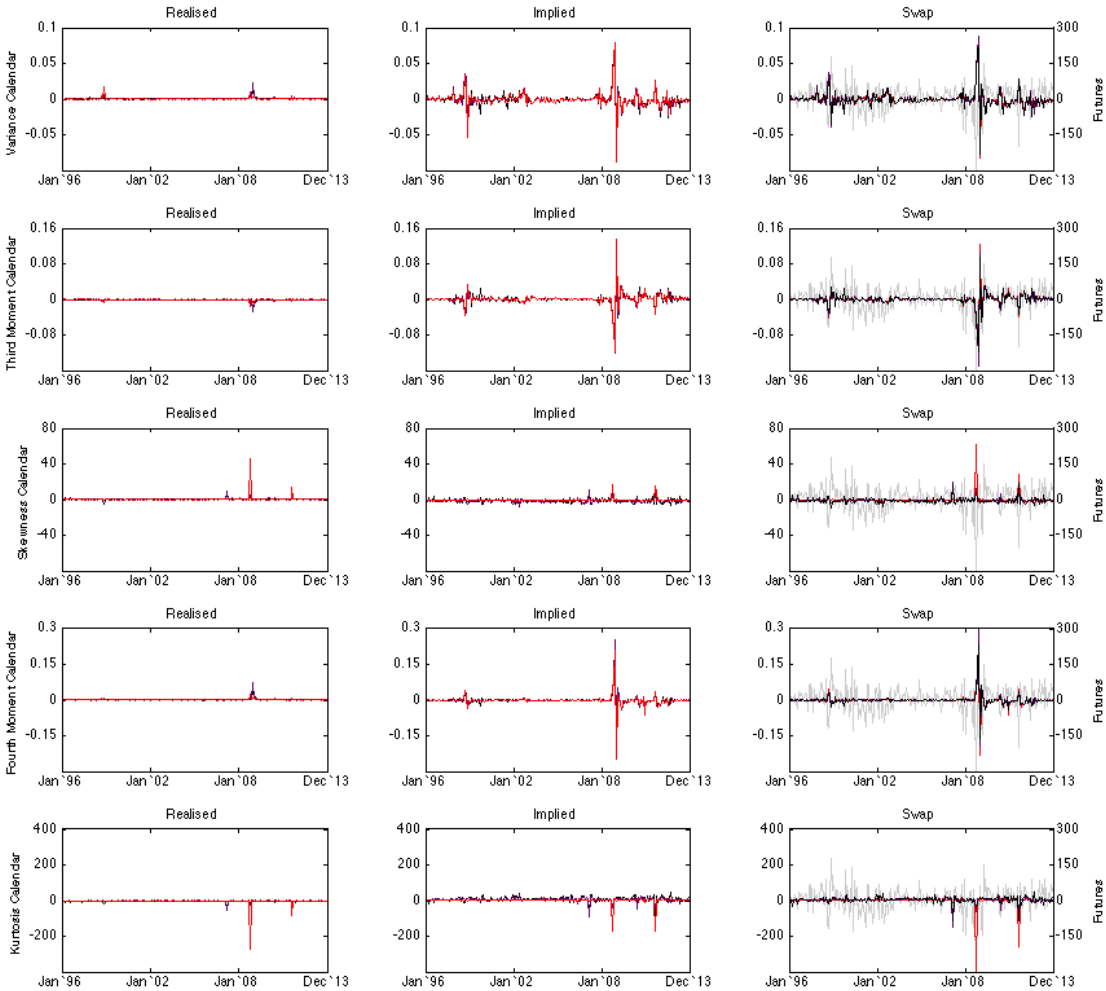

Figure 3.3: Time series of incremental 180-for-30 day constant-maturity variance, third-moment, skewness, fourth-moment and kurtosis calendar risk premia based on daily (black), weekly (purple) and monthly (red) monitoring.

$k_1 = 1000$, $k_2 = 1100$ and $k_3 = 1200$ when monitored at different frequencies.[8] The risk premium on these swaps can be large and negative during a crisis, e.g. in September 2008 and August 2011. Otherwise, the risk premium is small and positive and greater for straddle swaps that are monitored weekly or monthly than

[8]The choice of strike here allows us to investigate the behaviour of the swaps over the 18-year sample period because call and put options at these strikes were traded most of the time. We exclude strangle swaps from this analysis since they are more expensive to trade, due to the concentration of liquidity at the money, but results are available from the authors on request.

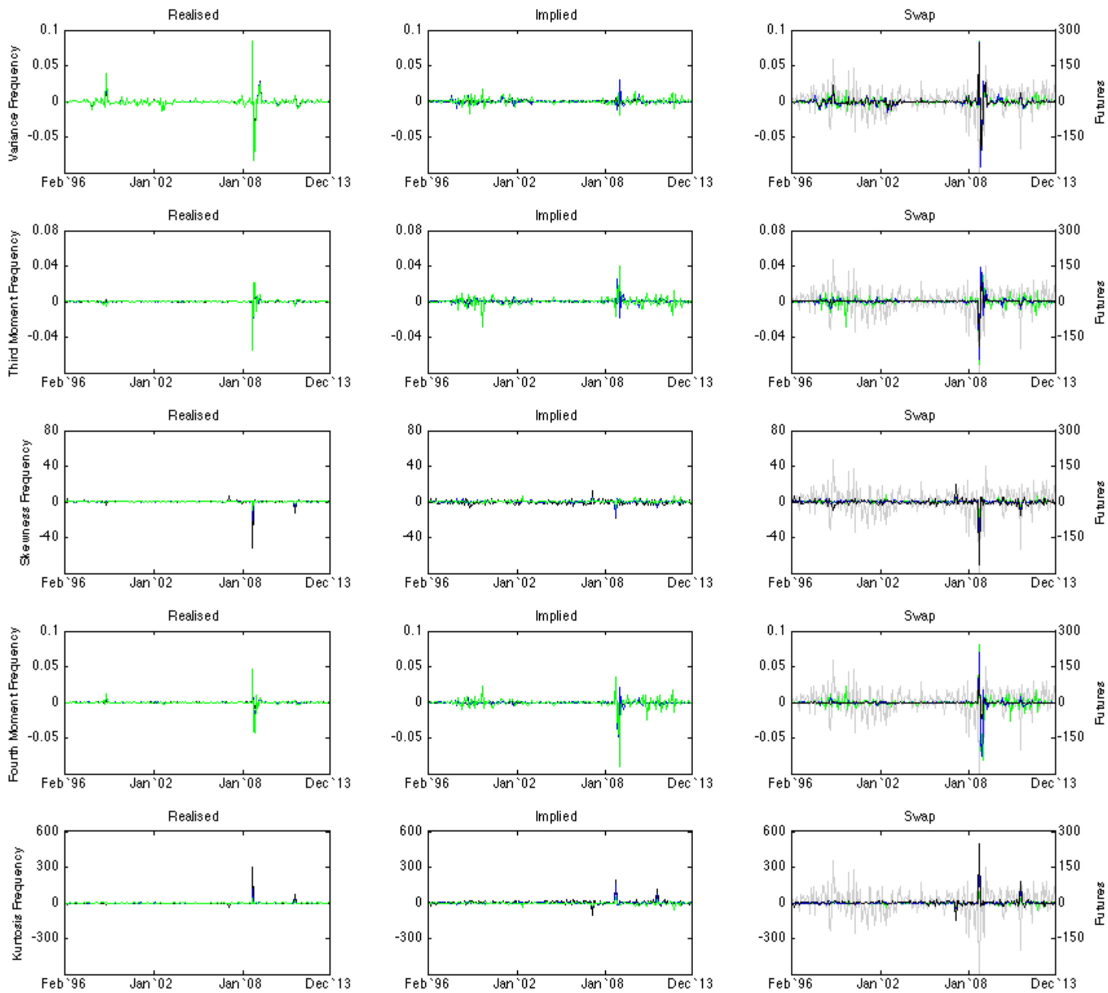

Figure 3.4: Time series of incremental monthly-daily variance, third-moment, skewness, fourth-moment and kurtosis frequency risk premia based on 30 (black), 90 (blue) and 180 (green) days constant maturity.

for straddle swaps that are monitored daily.

3.2.4 Diversification of Risk Premia

How diverse are the risk premia obtainable through trading DI moment characteristics? Tables 3.3, 3.4 and 3.5 present the correlations between daily (top panels), weekly (mid panels) and monthly (bottom panels) monitored risk premia

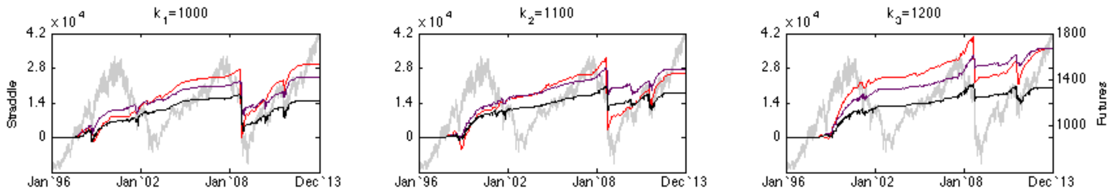

Figure 3.5: Time series for the cumulative risk premia on 30-day constant-maturity straddle swaps with strikes $k_1 = 1000$, $k_2 = 1100$ and $k_3 = 1200$, denoted by $V^{[k_1]}$, $V^{[k_2]}$ and $V^{[k_3]}$ and defined as in Example 5. Black, purple and red lines refer to swaps with realised characteristics that are monitored on a daily, weekly and monthly basis, respectively. Since the implied leg of a straddle swap is always zero, the risk premium is driven entirely by the realised component.

on the DI swaps that we have previously examined for a constant maturity of 30, 90 and 180 days, respectively. In each panel the two rows at the top present the correlations between the S&P 500 forward and log contract with the moment and straddle swaps described earlier; the middle sub-matrix presents cross-correlations between the moment swaps; and the right column presents the correlations with the straddle swaps from Example 5.

As expected from the empirical study of Duffie et al. [2000] and many others since, the correlation between the daily changes in the S&P 500 forward and the variance swap in the top panel of Table 3.3 is around -0.6; the same holds for the correlation between the log contract and the variance swap. Both correlations decrease in magnitude, but only marginally, with the monitoring frequency, reaching the values -0.48 and -0.51 under monthly monitoring in the bottom panel, respectively. At the 180 days maturity horizon, this relationship between correlation and monitoring frequency inverts yet remains at the same overall level, as can be seen in Table 3.5. Thus, as is also evident from Figure 3.1, variance swaps compensate the investor for downward shocks in the forward by a strongly positive realised variance. Further, the VRP is negatively correlated with the

$\mathbf{\Pi}_D$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.61	0.60	0.74	-0.45	-0.52	0.27	0.35	0.41
X	1	-0.66	0.69	0.71	-0.53	-0.49	0.24	0.30	0.33
$V^{(2)}$		1	-0.88	-0.54	0.87	0.46	-0.50	-0.44	-0.35
$V^{(3)}$			1	0.41	-0.96	-0.32	0.19	0.15	0.15
$V^{(\bar{3})}$				1	-0.33	-0.92	0.35	0.40	0.44
$V^{(4)}$					1	0.27	-0.23	-0.12	-0.09
$V^{(\bar{4})}$						1	-0.35	-0.39	-0.40
$V^{[k_1]}$							1	0.80	0.39
$V^{[k_2]}$								1	0.70

$\mathbf{\Pi}_W$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.53	0.57	0.69	-0.45	-0.47	0.38	0.44	0.39
X	1	-0.59	0.66	0.67	-0.53	-0.45	0.39	0.43	0.35
$V^{(2)}$		1	-0.89	-0.53	0.93	0.46	-0.77	-0.71	-0.46
$V^{(3)}$			1	0.45	-0.97	-0.35	0.62	0.58	0.33
$V^{(\bar{3})}$				1	-0.39	-0.95	0.45	0.52	0.47
$V^{(4)}$					1	0.31	-0.66	-0.59	-0.30
$V^{(\bar{4})}$						1	-0.40	-0.47	-0.43
$V^{[k_1]}$							1	0.91	0.46
$V^{[k_2]}$								1	0.69

$\mathbf{\Pi}_M$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.48	0.56	0.57	-0.44	-0.50	0.42	0.44	0.48
X	1	-0.51	0.61	0.58	-0.48	-0.51	0.45	0.45	0.48
$V^{(2)}$		1	-0.90	-0.86	0.94	0.85	-0.94	-0.91	-0.87
$V^{(3)}$			1	0.90	-0.96	-0.87	0.92	0.90	0.85
$V^{(\bar{3})}$				1	-0.84	-0.99	0.91	0.94	0.94
$V^{(4)}$					1	0.82	-0.92	-0.89	-0.81
$V^{(\bar{4})}$						1	-0.89	-0.93	-0.94
$V^{[k_1]}$							1	0.98	0.89
$V^{[k_2]}$								1	0.95

Table 3.3: Correlations between 30-day constant-maturity risk premia ($\tau = 30$) based on daily, weekly and monthly monitoring over the full sample from January 1996 to December 2013.

third-moment, skewness and straddle-swap risk premia and positively correlated with the fourth-moment and kurtosis premia at all monitoring frequencies and maturity horizons. Given its strong positive performance during crisis periods when large losses accrue to short variance swaps positions, the third-moment swap could even be attractive to variance swap issuers as a partial hedge.

$\mathbf{\Pi}_D$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.61	0.53	0.73	-0.35	-0.51	0.35	0.37	0.36
X	1	-0.69	0.62	0.71	-0.44	-0.49	0.29	0.28	0.26
$V^{(2)}$		1	-0.92	-0.63	0.84	0.53	-0.42	-0.34	-0.29
$V^{(3)}$			1	0.50	-0.95	-0.42	0.17	0.10	0.09
$V^{(\bar{3})}$				1	-0.37	-0.93	0.50	0.47	0.43
$V^{(4)}$					1	0.33	-0.11	-0.06	-0.05
$V^{(\bar{4})}$						1	-0.45	-0.43	-0.39
$V^{[k_1]}$							1	0.77	0.49
$V^{[k_2]}$								1	0.77

$\mathbf{\Pi}_W$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.97	-0.61	0.58	0.73	-0.43	-0.47	0.45	0.48	0.41
X	1	-0.69	0.69	0.71	-0.53	-0.45	0.42	0.42	0.34
$V^{(2)}$		1	-0.92	-0.63	0.88	0.50	-0.67	-0.59	-0.42
$V^{(3)}$			1	0.51	-0.96	-0.39	0.41	0.35	0.23
$V^{(\bar{3})}$				1	-0.40	-0.91	0.62	0.63	0.51
$V^{(4)}$					1	0.32	-0.35	-0.27	-0.18
$V^{(\bar{4})}$						1	-0.52	-0.55	-0.43
$V^{[k_1]}$							1	0.87	0.54
$V^{[k_2]}$								1	0.81

$\mathbf{\Pi}_M$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.56	0.63	0.64	-0.47	-0.54	0.49	0.50	0.50
X	1	-0.60	0.70	0.65	-0.53	-0.55	0.50	0.50	0.49
$V^{(2)}$		1	-0.92	-0.86	0.91	0.86	-0.93	-0.90	-0.85
$V^{(3)}$			1	0.85	-0.94	-0.82	0.84	0.82	0.76
$V^{(\bar{3})}$				1	-0.73	-0.99	0.92	0.93	0.89
$V^{(4)}$					1	0.72	-0.80	-0.76	-0.69
$V^{(\bar{4})}$						1	-0.92	-0.93	-0.89
$V^{[k_1]}$							1	0.98	0.91
$V^{[k_2]}$								1	0.96

Table 3.4: Correlations between 90-day constant-maturity risk premia ($\tau = 90$) based on daily, weekly and monthly monitoring over the full sample from January 1996 to December 2013.

Results for the skew and kurtosis risk premia are quite novel.[9] There is a strong positive correlation between the forward and the skew risk premium which increases with monitoring frequency: It is 0.74 at the daily frequency but falls to 0.57 at the monthly frequency, when considering the 30 days maturity horizon. Again, the relationship between monitoring frequency and correlation inverts when looking at the 180 days horizon yet remains at the same overall level. The correlations between the skew and kurtosis premia are strongly negative for all monitoring frequencies and maturities, ranging from -0.88 under daily monitoring and for 180 days to maturity (see Table 3.5) to -0.99 under monthly monitoring for 30 and 90 days to maturity (see Tables 3.3 and 3.4). This indicates that skewness is clearly picking up the asymmetry in the tails of the distribution, rather than asymmetry around the centre. At the monthly frequency, the correlation of -0.86 between the P&L on the variance and skewness swaps is in line with the findings of Neuberger [2012] and Kozhan et al. [2013].[10]

However, our more granular analysis allows for a more discerning conclusion, i.e. that standardised moment risk premia behave quite differently from their non-standardised counterparts when monitored at a higher frequency. The correlation between variance and the third-moment premiums remains almost as high at the daily frequency as it is at the monthly frequency (and similarly for the correlation between variance and the fourth moment). However, the correlation between the

[9] They are similar to the non-standardised third-moment and fourth-moment risk premia, respectively, so we confine our observations to skew and kurtosis.

[10] See Neuberger [2012], p.19: "Both the second and third moments, whether realised or implied, [...] are very highly (negatively) correlated with each other, with correlations in excess of -0.9." See also Kozhan et al. [2013], p.13, Table 2, Panel B: The correlation between excess returns on the variance and cubic swap is 0.874 where the positive sign comes from the fact that, in their setting, a writer of the cubic swap receives fixed an pays floating. The correlation between the variance and skewness swap is even stronger (0.897).

$\mathbf{\Pi}_D$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.97	-0.53	0.40	0.52	-0.23	-0.34	0.26	0.24	0.17
X	1	-0.62	0.50	0.53	-0.32	-0.36	0.22	0.18	0.10
$V^{(2)}$		1	-0.91	-0.64	0.79	0.56	-0.37	-0.26	-0.18
$V^{(3)}$			1	0.52	-0.95	-0.44	0.15	0.06	0.01
$V^{(\bar{3})}$				1	-0.37	-0.88	0.52	0.40	0.26
$V^{(4)}$					1	0.36	-0.07	-0.01	-0.01
$V^{(\bar{4})}$						1	-0.45	-0.36	-0.25
$V^{[k_1]}$							1	0.73	0.36
$V^{[k_2]}$								1	0.66

$\mathbf{\Pi}_W$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.97	-0.60	0.52	0.66	-0.35	-0.50	0.34	0.34	0.30
X	1	-0.69	0.64	0.66	-0.47	-0.50	0.31	0.27	0.22
$V^{(2)}$		1	-0.91	-0.67	0.83	0.61	-0.49	-0.40	-0.31
$V^{(3)}$			1	0.53	-0.96	-0.48	0.21	0.12	0.09
$V^{(\bar{3})}$				1	-0.40	-0.96	0.55	0.49	0.38
$V^{(4)}$					1	0.38	-0.13	-0.05	-0.07
$V^{(\bar{4})}$						1	-0.51	-0.45	-0.37
$V^{[k_1]}$							1	0.84	0.48
$V^{[k_2]}$								1	0.75

$\mathbf{\Pi}_M$	X	$V^{(2)}$	$V^{(3)}$	$V^{(\bar{3})}$	$V^{(4)}$	$V^{(\bar{4})}$	$V^{[k_1]}$	$V^{[k_2]}$	$V^{[k_3]}$
F	0.98	-0.58	0.58	0.73	-0.37	-0.61	0.50	0.51	0.48
X	1	-0.64	0.67	0.73	-0.45	-0.61	0.49	0.49	0.47
$V^{(2)}$		1	-0.91	-0.81	0.83	0.80	-0.85	-0.83	-0.77
$V^{(3)}$			1	0.71	-0.93	-0.68	0.64	0.63	0.60
$V^{(\bar{3})}$				1	-0.51	-0.98	0.88	0.87	0.80
$V^{(4)}$					1	0.50	-0.50	-0.49	-0.47
$V^{(\bar{4})}$						1	-0.88	-0.88	-0.82
$V^{[k_1]}$							1	0.97	0.87
$V^{[k_2]}$								1	0.95

Table 3.5: Correlations between 180-day constant-maturity risk premia ($\tau = 180$) based on daily, weekly and monthly monitoring over the full sample from January 1996 to December 2013.

skew (kurtosis) premium and the VRP decreases in magnitude from -0.86 (0.85) under monthly monitoring, to -0.53 (0.46) with weekly monitoring, and it remains at this level under daily monitoring. Another source of diversification is provided by the straddle swaps. They exhibit relatively low, positive correlations with the forward and the third moment, a strong negative correlation with variance and a relatively small negative correlation with the fourth-moment swaps.

3.2.5 Determinants of Moment Risk Premia

Following the study by Carr and Wu [2009] on the determinants of variance risk premia, we now question whether significant common factors influencing our moment risk premia can be found among standard equity risk factors, namely: the excess return on the market (ER); the 'small minus big' (*size*) and the 'high minus low' (*growth*) factors introduced by Fama and French [1993]; as well as the 'up minus down' (*momentum*) factor introduced by Carhart [1997].

Using monthly data on the VRP in the S&P 500 market from January 1996 through February 2003, Carr and Wu [2009] find no significant effect for anything other than the market excess return as a driver of the VRP. They also add a squared market factor as explanatory variable, as in the three-moment CAPM of Kraus and Litzenberger [1976], but find no evidence of an asymmetric response to market shocks. Our data construction methodology allows us to investigate the same phenomenon using higher frequency data. Given that Engle [2011] and others document the importance of an asymmetric response in volatility to market shocks at the daily frequency, it seems likely that daily or even weekly data would be sufficient to detect this effect. Using monthly data over the same period as Carr and Wu [2009], we also find no empirical evidence for an asymmetric response in the VRP. However, using daily data over the same period the regression coefficient on the squared market factor is significantly different from zero at 0.1%. This finding leads us to question whether similar asymmetric responses are evident in third and fourth moment, and skewness and kurtosis risk premia, when measured at the daily frequency.

Figure 3.6 presents time series on daily changes in the S&P 500 30-day constant-maturity synthetic futures price (in grey, measured on the right-hand scale) with a black line (measured on the left-hand scale) depicting daily changes in the 30-day, daily-monitored, DI VRP (above), skewness risk premium (middle) and kurtosis risk premium (below). The DI VRP displays the well-known features common to the standard VRP: it is typically small and negative but occasionally large and positive, in particular during the onset of a period of market turbulence. Notably, it has returned to very small levels ever since the Eurozone crisis in August 2011 – not dissimilar to its behaviour during the credit boom period of mid-2003 to mid-2007. By contrast, the skew risk premium is typically small and positive, but occasionally takes large negative values. For instance, on 27 February 2007 it reached -22.26. On that day the S&P 500 index fell by 3.5%, its biggest one-day fall since March 2003. The same day also marked a significant jump in the kurtosis risk premium, when it exceeded 160. Otherwise, like the VRP, the kurtosis risk premium is usually small and negative. However, unlike the VRP, the kurtosis risk premium has clearly increased in variability during the latter part of the sample.

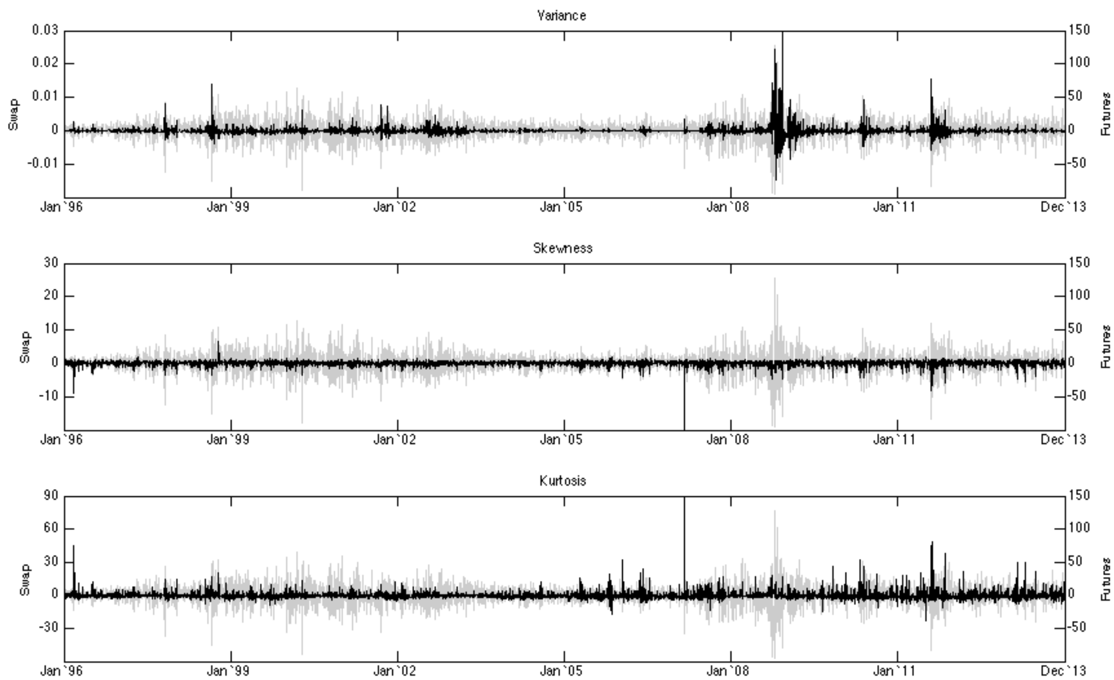

Figure 3.6: Time series for the daily incremental risk premia on 30-day constant-maturity swaps. The black bars refer to variance, skewness and kurtosis, respectively, while the grey bars represent the 30-day constant-maturity futures.

Following Carr and Wu [2009] we now specify the regression model:[11]

$$\hat{V} = \alpha + \beta_{ER} ER + \beta_{ER^2} ER^2 + \beta_s \text{size} + \beta_g \text{growth} + \beta_m \text{momentum}, \qquad (3.1)$$

where $\hat{V}$ denotes the daily change in the 30-day risk premium under consideration. We estimate this model using daily data on risk factors from Kenneth French's website but also present results for a restricted model where $\beta_s = \beta_g = \beta_m = 0$. We perform the analysis for the entire sample and separately for the financial crisis

[11] No significant autocorrelation is observed in the dependent and independent variables. Neumann and Skiadopoulos [2013] do observe autocorrelation in daily data – on risk-neutral higher-moments, as opposed to higher-moment risk premia – but it is not sufficiently significant to be exploitable after transactions costs.

period between July 2008 and June 2009. We standardise all time series to make coefficients commensurate in size. As a result the intercept cannot be interpreted as an expected risk premium, but the beta coefficients can be interpreted as the number of standard deviations a risk premium is expected to change per standard deviation change in the corresponding factor.

96-13	Variance		3rd Moment		Skewness		4th Moment		Kurtosis	
α	-0.14	-0.14	0.04	0.04	0.06	0.06	-0.08	-0.08	-0.07	-0.07
	(-13.45)	(-13.57)	(3.73)	(3.74)	(5.76)	(5.71)	(-6.10)	(-6.10)	(-5.40)	(-5.17)
β_{ER}	-0.61	-0.63	0.65	0.65	0.70	0.74	-0.49	-0.50	-0.48	-0.52
	(-62.73)	(-61.06)	(57.90)	(56.20)	(68.00)	(68.91)	(-39.97)	(-38.28)	(-37.71)	(-39.39)
β_{ER^2}	0.14	0.14	-0.04	-0.04	-0.06	-0.06	0.08	0.08	0.07	0.07
	(42.33)	(42.38)	(-11.74)	(-11.67)	(-18.13)	(-17.84)	(19.21)	(19.07)	(17.01)	(16.16)
β_s		0.06		-0.07		-0.02		0.05		-0.03
		(6.43)		(-6.56)		(-2.14)		(4.10)		(-2.21)
β_g		-0.09		0.12		0.03		-0.11		-0.07
		(-8.89)		(10.68)		(2.43)		(-8.51)		(-4.86)
β_m		0.00		-0.05		0.14		0.05		-0.14
		(0.39)		(-4.31)		(12.79)		(3.93)		(-10.52)
R^2	0.567	0.581	0.439	0.469	0.528	0.544	0.309	0.331	0.280	0.298
F	(50.1)		(83.8)		(56.0)		(50.9)		(40.0)	

08-09	Variance		3rd Moment		Skewness		4th Moment		Kurtosis	
α	-0.65	-0.76	0.30	0.44	0.10	0.11	-0.52	-0.65	-0.10	-0.11
	(-4.61)	(-5.66)	(2.01)	(3.22)	(2.88)	(2.91)	(-2.99)	(-3.95)	(-2.97)	(-3.09)
β_{ER}	-1.04	-1.24	1.40	1.69	0.37	0.36	-1.23	-1.53	-0.23	-0.23
	(-18.64)	(-16.19)	(23.63)	(21.62)	(25.72)	(17.01)	(-17.60)	(-16.13)	(-16.94)	(-11.68)
β_{ER^2}	0.15	0.17	-0.05	-0.07	-0.03	-0.04	0.10	0.12	0.03	0.03
	(11.84)	(13.76)	(-3.49)	(-5.35)	(-10.28)	(-10.35)	(6.16)	(7.89)	(9.66)	(9.87)
β_s		0.26		-0.22		-0.05		0.22		0.04
		(3.68)		(-3.10)		(-2.43)		(2.46)		(2.05)
β_g		0.30		-0.38		-0.00		0.47		0.02
		(3.16)		(-3.94)		(-0.04)		(3.97)		(0.73)
β_m		-0.07		0.13		-0.01		-0.07		0.01
		(-0.72)		(1.38)		(-0.42)		(-0.66)		(0.22)
R^2	0.658	0.704	0.694	0.750	0.753	0.756	0.579	0.639	0.600	0.603
F	(13.7)		(19.6)		(2.0)		(14.8)		(1.6)	

Table 3.6: Estimates and t-statistics (in brackets) as well as adjusted R^2 and F-test (in brackets) on joint significance for the restricted and unrestricted regression of the constant 30-days-to-maturity moment risk premia from January 1996 to December 2013 as well as for the crisis period from July 2008 to June 2009.

Being based on more than 4500 observations, our analysis over the entire period provides some highly significant results. The first blocks of both panels in Table 3.6 report our results for the VRP. The linear and quadratic excess return factors have highly significant loadings, the negative $\hat{\beta}_{ER}$ being compensated by a positive $\hat{\beta}_{ER^2}$. Thus the VRP increases more when there is a negative market return than it decreases when there is a positive return of the same size. This asymmetric response is particularly pronounced during the financial crisis period (bottom panel). Over the whole 18-year period (top panel) the coefficients on the size and growth factors are small but significant, indicating that firm size has a positive impact and firm growth a negative impact on the VRP, respectively. The addition of the Fama-French factors only marginally increases the adjusted R^2 from 0.567 to 0.581 but the F-statistic for addition of these factors is significant.

During the financial crisis the R^2 increases considerably relative to its value over the full sample as the VRP becomes more sensitive to market shocks. The Fama-French factors, however, remain only marginally significant. Of these only the size factor has a significant coefficient of the same sign as for the full sample estimate. The change in sign of the coefficient on growth underlines the fact that July 2008 – June 2009 represents a very particular market regime. The momentum factor appears to be irrelevant for both periods considered.

The second column block of Table 3.6 displays estimates for the third-moment risk premium. Here, the directional effects of the linear and quadratic factors are opposite to those observed in the variance premium regression: a market shock now has a greater impact on the third-moment premium when positive than when negative. The contribution of the size and momentum factors is relatively small but the growth factor has a significant positive effect which again changes sign during the financial crisis. Conclusions regarding the skewness premium (in the third column block of the table) are similar, except that it is momentum rather than growth that has a positive effect on the skewness premium. It is remarkable that the explanatory power during the crisis period for the third moment is as high as 0.694 (0.750 for the unrestricted model) and even higher (0.753) for the skewness risk premium. In fact, during the financial crisis the Fama-French and Carhart factors have almost no impact on the skewness premium: the F-statistic for joint significance is only 2.0. The fourth and fifth column blocks of Table 3.6 analyse the determinants of the fourth-moment and kurtosis risk premia. The much lower R^2 here indicates that these premia may be driven, to a large extent, by so far unknown risk factors. Otherwise the conclusions drawn are similar to – yet weaker than – those drawn about the variance premium. Apart from the excess market return the only consistently significant effect is exhibited by the growth factor for the fourth-moment risk premium and by the momentum factor for the kurtosis risk premium, where the signs of the corresponding coefficients are opposite to those for the regression on the third-moment and skewness risk premia.

Chapter 4

Variance Swaps in Affine Models

The purpose of the third main part of this thesis is to derive the dynamics of variance swaps in affine stochastic volatility (SV) models as well as variance risk premia that result from different $\mathbb{Q}$- and $\mathbb{P}$-parameterisations. We compare the risk premia for standard variance swaps with those based on Neuberger's alternative definition of realised variance as well as the squared changes in the price of the log contract we use in the empirical section.

After giving an overview of the existing literature on such model-dependent evaluations, we derive explicit formulae for the Heston model as well as a stochastic volatility model with contemporaneous jumps in the underlying and variance process. We discuss the impact of the diverse model parameters of both the risk-neutral and the physical price process on the dynamics of a variance swap as well as the variance risk premium (VRP). Particular attention is given to the effect of jumps in the underlying process, which have a different impact on the discretisation-invariant characteristic we use and the standard characteristic, squared log returns, which does not satisfy the aggregation property.

We further provide an explicit derivation for the joint characteristic function of the price and volatility processes in the Heston model, which makes it possible to analytically evaluate the prices of power log contracts other than the log contract. These prices can be used to derive similar analytic expressions for higher-moment swap dynamics and risk premia. However, we spare the reader from displaying these rather bulky formulae. The main references for all technical derivations are Cont and Tankov [2004], Hull [2009], Jacod and Shiryaev [2003], Oksendal [2003] and Shreve [2004].

4.1 Literature Review

Besides the evaluation of empirically observable risk premia, a considerable body of literature is concerned with the implications of model assumptions for the underlying process on the pricing and hedging of variance swaps. The objects of interest are the risk-neutral and physical dynamics of both the realised leg and the premium paid for a swap contract as well as the swap rate and implied higher-moments.[1] The purpose of this literature review is to summarise existing results

[1]The traditional derivatives pricing literature develops models under a risk-neutral measure for the purpose of pay-off valuation, based on no-arbitrage considerations, and calibrates them to a snapshot of options data. By contrast, mainstream asset pricing literature looks at (possibly cross-sectional) time-series data and tries to identify common behaviour and determinants of the price development under the physical measure. Although these two are fundamentally different concepts, a more recent branch of literature tries to incorporate both views in unified 'equilibrium' models. In these theories the change from the risk-neutral to the physical measure

on these $\mathbb{Q}$ and $\mathbb{P}$-dynamics for a variety of stochastic volatility models and draw a connection to the market price of risk.

When pricing a swap in a model, the first question to ask is about the finiteness of the fair-value swap rate. For a variety of stochastic volatility models Andersen and Piterbarg [2007] assess whether these are well-posed, which is the case if a sufficiently large number of risk-neutral moments of the price distribution are finite for finite time horizons and given model parameters. They argue that popular models have been abused in order for their dynamics to incorporate a variety of features, leading to a lack of price bounds for otherwise common derivative securities. In the following we will focus on analytically tractable models where the variance swap rate exists for finite horizons.

In the classical Black and Scholes [1973] model volatility is assumed to be constant and hence the only risk premium which can be captured is the equity risk premium (ERP). In order to explain the presence of a VRP, one must assume variance to be stochastic. One of the simplest and most popular stochastic volatility model that preserves analytical tractability is given by Heston [1993]. However, fair-value variance swap rates can also be derived for other models including those introduced by Cox and Ross [1976], Emanuel and MacBeth [1982], Bates [1996], Scott [1997], Bates [2000], Barndorff-Nielsen and Shephard [2001], Carr and Schoutens [2008], Christoffersen et al. [2009] and Goard [2011]. The following paragraphs provide a review of closed-form as well as approximate pricing formulae that are available for both discretely and continuously monitored

and vice-versa, which in the derivatives pricing literature is referred to as the Radon-Nikodým derivative and in the asset pricing literature as the stochastic discount factor, is called the pricing kernel. All these concepts are essentially equivalent to the market price of risk.

variance swaps.

4.1.1 Results for the Heston Model

The Heston [1993] model allows one to reproduce stylised facts from empirical observations such as the volatility smile implicit in equity, commodity and other option prices. It is built on a mean-reverting variance process, which is introduced in Cox et al. [1985] and has a quasi-analytical representation (i.e. an explicit solution up to an integral along the path of a Brownian Motion). The fair strike price for a variance swap in the Heston [1993] model is derived in Broadie and Jain [2008a]; so also is an upper bound for the fair strike of a volatility swap, which pays the square root of realised variance in exchange for a fixed swap rate. Broadie and Jain [2008b] generalise the results from Broadie and Jain [2008a] for volatility and variance swaps in the models by Merton [1973], Bates [1996] and Scott [1997]. They state that the convexity correction for volatility swaps is not adequate in the presence of jumps in the underlying, and that jumps have a stronger impact on swap rates than discrete monitoring.

Rujivan and Zhu [2014] derive a closed-form solution for the price of a discretely-monitored variance swap, exploiting the tower rule for conditional expectations, and comment on the transferability of this solution to other affine models. Guillaume and Schoutens [2014] show how calibrating the Heston [1993] model to variance swap rates can yield stable parameter estimates over time. Detlefsen and Haerdle [2013] address problems that arise when a static model is used for explaining the term structure of variance. In particular, the authors comment on insufficient out-of-sample performance and a lack of capturing observable dynamics. Zhu and Lian [2015] derive pricing formulae for discretely-monitored variance swaps in the Heston model based on both squared log returns and squared simple returns. Using Monte-Carlo simulation and an Euler-style discretisation of the model, they analyse the effects of sampling frequency, forward start, mean reversion and realised characteristic on the fair-value swap rate.

Also in the Heston [1993] model, yet in an incomplete market with regime switching, Elliott et al. [2007] derive a partial differential equation that can be solved in order to approximate the prices of continuously monitored variance and volatility swaps. The results are extended to the discrete monitoring case by Elliott and Lian [2013], who develop analytic pricing formulae for variance and volatility swaps. They discuss the impact of the monitoring frequency and relate the results to the continuous limit. Zhu and Lian [2012] derive a computationally efficient pricing formula for discretely monitored variance swaps and find that the discrete monitoring error is exponentially increasing as the frequency decreases. They generalise their results to forward starting variance swaps in Zhu and Lian [2015]. Carr and Schoutens [2008] consider a modification of the Heston [1993] model which incorporates a jump-to-default feature and demonstrate how power payoffs as well as European options can be hedged using variance swaps and credit default swaps. The authors apply the theory of orthogonal polynomials to relate Gamma payoffs, Dirac payoffs and European options to the purpose of deriving an approximation hedge between the corresponding derivative instruments.

4.1.2 Affine Stochastic Volatility Models

Many analytical results from the Heston [1993] model can be transfered to more general affine models, and the fundamental properties of this class of models are presented in Duffie et al. [2000]. The authors provide a method for finding analytical solutions to a wide range of asset and derivatives pricing problems. For a very general class of affine jump-diffusions, they derive a closed-form representation for what they call the 'extended transform'. Evaluating the characteristic function of the log price (and all other state variables) reduces to solving a set of ordinary differential equations, which is computationally easy. The pricing of options on quadratic variation in affine models is discussed in Kallsen et al. [2011], with particular emphasis on analytical tractability. Egloff et al. [2010] show that the variance term structure is driven by two main factors, namely the short end and the long end, and can be modelled using an affine model. According to the authors, a portfolio consisting of a long position in long-term variance swaps and short positions in short-term variance swaps as well as the underlying index is optimal for investors. They follow the affine model specification of Duffie et al. [2000] and compare the Heston [1993] model with a two-factor variance rate model that incorporates a stochastic mean level.

In affine models, a number of structure-preserving specifications of the pricing kernel have been identified. The market price of risk in a Heston-style and other affine asset pricing models depends on the difference between the mean-reversion and long-term mean parameter under the physical and the risk-neutral measure. Based on the definition of a completely affine specification of the market price of risk by Dai and Singleton [2000] and the later generalisation to essentially affine specifications by Duffee [2002], which allows the market price of risk to vary independently from the current level of volatility, Cheridito et al. [2007] introduce an extended affine specification. This way of modelling the market price of risk incorporates the previous two cases and is always more general than the completely affine specification. Chernov and Ghysels [2000] develop an approach to jointly estimate the risk-neutral and physical parameters by tackling the challenge of latent variables involved in the model selection process. The estimation approach is illustrated using the Heston model as well as the standard change of measure where the mean-reversion level and speed can differ between the two measures but the correlation of the driving Brownian motions is the same. Their results indicate the relative importance of the filtered volatility estimate over the choice of a particular option pricing formula.

4.1.3 Other Diffusion Processes

Beyond affine models, the 3/2 stochastic volatility model has attracted the attention of researchers in recent years. As opposed to the Heston [1993] model, the drift and diffusion terms of the volatility process in the 3/2 model are quadratic in and proportional to the 3/2 power of the current volatility level, respectively. A range of publications address the pricing of variance swaps under these model assumptions. Jarrow et al. [2013] analyse the convergence behaviour of the price of a discretely monitored variance and volatility swap towards the continuously monitored limit. Goard [2011] considers an extension of the model where the mean-reversion level can be time-dependent, improving calibration properties. Chan and Platen [2015] discuss the pricing and hedging features of long-dated contracts, based on a specification of the model under both the risk-neutral and the physical measure, and derive an analytical pricing formula. Again under the 3/2 volatility model, Drimus [2012] derives a pricing formula for options on realised variance using Laplace transform techniques as well as hedging ratios and draws a comparison to the Heston model benchmark.

Alternatively, Jordan and Tier [2009] derive pricing formulas for continuously monitored variance swaps under the constant elasticity of volatility (CEV) model. In particular, they present a closed-form solution for the price of a log contract as well as an approximate price for a shifted CEV process, where a lower threshold triggers default. In order to circumvent problems with certain parameterisations where the underlying price can reach zero and therefore the log contract is not well defined, Wang et al. [2015] develop an alternative numerical approach for calculating the fair-value swap rate. Javaheri et al. [2004] use a generalised autoregressive conditional heteroscedasticity (GARCH) model as well as its continuous-time limit to price and hedge volatility swaps, providing an approximation for the convexity correction. Also starting from a continuous-time GARCH model, Swishchuk and Xu [2011] further include a jump component into the underlying price process and provide approximate pricing formulae.

Benth et al. [2007] derive swap price dynamics for powers of realised volatility under the Non-Gaussian Ornstein-Uhlenbeck-type model proposed by Barndorff-Nielsen and Shephard [2001]. The authors show that the prices which they derive for continuously monitored variance swaps are very close to those for discretely monitored contracts. They also derive option pricing formulae as well as approximations for volatility swaps. Again for a Non-Gaussian Ornstein-

Uhlenbeck process, Barndorff-Nielsen and Veraart [2013] study the impact of stochastic volatility-of-volatility on the leverage effect as well as the VRP. They also study structure-preserving change of measures in an incomplete market, where the stochastic differential equations (SDEs) explaining the $\mathbb{P}$ and $\mathbb{Q}$-dynamics correspond to different parameterisations of the same model.

4.1.4 The Effect of Jumps

Jumps are particularly important when it comes to pricing continuously monitored swap derivatives. Carr et al. [2005] study the pricing of options on realised variance for underlying processes that are pure sequences of jumps and argue that in a risk-neutral setting quadratic variation, which corresponds to the limit of realised variance as the monitoring frequency becomes infinite, is crucial for the pricing of financial derivatives. The authors compare Lévy and Sato processes with both finite and infinite jump activity as well as a special Carr-Geman-Madan-Yor (CGMY) Lévy process and derive an option pricing formula using the Laplace transform of realised variance. In a subsequent publication, Carr et al. [2011] discuss term-structure monotonicity of call option prices and relate empirical observations to stochastic volatility models and Lévy processes. Related results based on stochastic time change can be found in Itkin and Carr [2010].

Bates [2000] finds that including jumps leads to more plausible parameters for the stochastic volatility and therefore jumps are necessary to explain skewness consistently with the time-series characteristics of futures prices. However, he argues that "since unconstrained risk premia can potentially explain any deviations between actual and risk-neutral distributions, it is important to have some idea of plausible values for signs and magnitudes". Bakshi et al. [1997] perform a similar analysis for option pricing models with stochastic volatility, jumps and interest rates and assess whether the risk-neutral parameterisations are compatible with the time-series characteristics. They find that jumps are important for improving the pricing performance but not for hedging, while introducing correlation between interest rates and stock prices does not yield any improvement in performance. Also for stochastic volatility models with jumps Bregantini [2013] develops a moment-based estimation method which relies on lagged realised variance.

4.2 Heston Model

The Heston [1993] model is defined by the forward dynamics of the underlying as well as the stochastic variance process under the risk-neutral measure:

$$
\begin{aligned}
dF_t &:= F_t \sqrt{v_t} dW_t^{F(\mathbb{Q})}, \\
dv_t &:= \kappa^{(\mathbb{Q})} \left(\theta^{(\mathbb{Q})} - v_t\right) dt + \sqrt{v_t} \sigma dW_t^{v(\mathbb{Q})},
\end{aligned}
$$

where the condition $2\kappa^{(\mathbb{Q})}\theta^{(\mathbb{Q})} > \sigma^2$ guarantees positivity of the variance and $dW_t^{F(\mathbb{Q})} dW_t^{v(\mathbb{Q})} =: \rho dt$ denotes the correlation of the Wiener processes. The log forward price $x_t := \ln F_t$ follows the dynamics

$$
dx_t = \sqrt{v_t} dW_t^{F(\mathbb{Q})} - \tfrac{v_t}{2} dt,
$$

with quadratic variation $\langle x \rangle_t = \int_0^t v_u du$. Whenever we omit the measure superscript for a model parameter, it corresponds to that of the corresponding Wiener process. If we omit the measure superscript for a Wiener process, the corresponding equation holds under all measures under consideration.

4.2.1 Variance Process and Log Contract

First we derive the explicit solution for v by applying Itô's formula to $\mathrm{e}^{\kappa t}v_t$:

$$d\left(\mathrm{e}^{\kappa t}v_t\right) = \mathrm{e}^{\kappa t}\kappa v_t dt + \mathrm{e}^{\kappa t}dv_t = \mathrm{e}^{\kappa t}\kappa\theta dt + \mathrm{e}^{\kappa t}\sqrt{v_t}\sigma dW_t^v.$$

Integrating both sides from t to $u \geq t$ yields

$$\mathrm{e}^{\kappa u}v_u - \mathrm{e}^{\kappa t}v_t = \theta\left(\mathrm{e}^{\kappa u} - \mathrm{e}^{\kappa t}\right) + \sigma\int_t^u \mathrm{e}^{\kappa s}\sqrt{v_s}dW_s^v,$$

and we can solve for v_u:

$$\begin{aligned} v_u &= \mathrm{e}^{-\kappa(u-t)}v_t + \theta\left(1 - \mathrm{e}^{-\kappa(u-t)}\right) + \mathrm{e}^{-\kappa u}\sigma\int_t^u \mathrm{e}^{\kappa s}\sqrt{v_s}dW_s^v \\ &= \theta + \mathrm{e}^{-\kappa(u-t)}\left(v_t - \theta\right) + \sigma\int_t^u \mathrm{e}^{-\kappa(u-s)}\sqrt{v_s}dW_s^v. \end{aligned}$$

Then $\mathbb{E}\left[v_u\middle|\mathcal{F}_t\right] = \theta + \mathrm{e}^{-\kappa(u-t)}\left(v_t - \theta\right)$. We shall further use the notation $d\left\langle x\right\rangle_t = v_t dt$. The quadratic variation of the log price, which corresponds to integrated variance, yields

$$\begin{aligned} \left\langle x\right\rangle_t &= \int_0^t v_u du = \int_0^t\left(\theta + \mathrm{e}^{-\kappa u}\left(v_0 - \theta\right) + \sigma\int_0^u \mathrm{e}^{-\kappa(u-s)}\sqrt{v_s}dW_s^v\right)du \\ &= \left[\theta u - \tfrac{1}{\kappa}\mathrm{e}^{-\kappa u}\left(v_0 - \theta\right)\right]_0^t + \sigma\int_0^t\int_s^t \mathrm{e}^{-\kappa(u-s)}du\sqrt{v_s}dW_s^v \\ &= \theta t - \tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa t} - 1\right)\left(v_0 - \theta\right) - \tfrac{\sigma}{\kappa}\int_0^t\left(\mathrm{e}^{-\kappa(t-s)} - 1\right)\sqrt{v_s}dW_s^v, \end{aligned}$$

where we have changed the order of integration in the second line.

The price process of the log contract $X_t := \mathbb{E}^{\mathbb{Q}}\left[x_T \middle| \mathcal{F}_t \right]$ with maturity $T \geq t$ is

$$\begin{aligned} X_t &= x_t + \mathbb{E}^{\mathbb{Q}}\left[\int_t^T dx_u \middle| \mathcal{F}_t \right] = x_t - \tfrac{1}{2}\int_t^T \mathbb{E}^{\mathbb{Q}}\left[v_u \middle| \mathcal{F}_t \right] du \\ &= x_t - \tfrac{1}{2}\int_t^T \left[\theta^{(\mathbb{Q})} + \mathrm{e}^{-\kappa^{(\mathbb{Q})}(u-t)}\left(v_t - \theta^{(\mathbb{Q})} \right) \right] du \\ &= x_t - \tfrac{\theta^{(\mathbb{Q})}}{2}(T-t) + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1 \right)\left(v_t - \theta^{(\mathbb{Q})} \right), \end{aligned}$$

with dynamics

$$\begin{aligned} dX_t &= dx_t + \tfrac{\theta^{(\mathbb{Q})}}{2}dt + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\Big(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}dv_t + \mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}v_t\kappa^{(\mathbb{Q})}dt \\ &\quad -dv_t - \theta^{(\mathbb{Q})}\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}\kappa^{(\mathbb{Q})}dt \Big) \\ &= \sqrt{v_t}\left[dW_t^{F(\mathbb{Q})} + \tfrac{\sigma}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1 \right) dW_t^{v(\mathbb{Q})} \right], \end{aligned}$$

and quadratic variation

$$\begin{aligned} \langle X \rangle_t &= \int_0^t v_u \left[1 + \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1 \right) + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1 \right)^2 \right] du \\ &= \left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2} \right)\int_0^t v_u du + \left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} - \tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2} \right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\int_0^t \mathrm{e}^{\kappa^{(\mathbb{Q})}u}v_u du \\ &\quad + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\int_0^t \mathrm{e}^{2\kappa^{(\mathbb{Q})}u}v_u du. \end{aligned}$$

Using that $v_u = \theta + \mathrm{e}^{-\kappa(u-t)}(v_t - \theta) + \sigma \int_t^u \mathrm{e}^{-\kappa(u-s)} \sqrt{v_s} dW_s^v$ we can further simplify

$$\begin{aligned} \int_t^T \mathrm{e}^{\kappa^{(\mathbb{Q})}u} v_u du &= \int_t^T \left[\mathrm{e}^{\kappa^{(\mathbb{Q})}u} \theta^{(\mathbb{Q})} + \mathrm{e}^{\kappa^{(\mathbb{Q})}t} \left(v_t - \theta^{(\mathbb{Q})} \right) + \sigma \int_t^u \mathrm{e}^{\kappa^{(\mathbb{Q})}s} \sqrt{v_s} dW_s^v \right] du \\ &= \tfrac{\theta^{(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}} \left(\mathrm{e}^{\kappa^{(\mathbb{Q})}T} - \mathrm{e}^{\kappa^{(\mathbb{Q})}t} \right) + \mathrm{e}^{\kappa^{(\mathbb{Q})}t} \left(v_t - \theta^{(\mathbb{Q})} \right) (T-t) \\ &\quad + \sigma \int_t^T (T-s) \mathrm{e}^{\kappa^{(\mathbb{Q})}s} \sqrt{v_s} dW_s^v, \end{aligned}$$

as well as

$$\begin{aligned} \int_t^T \mathrm{e}^{2\kappa^{(\mathbb{Q})}u} v_u du &= \int_t^T \left[\mathrm{e}^{2\kappa^{(\mathbb{Q})}u} \theta^{(\mathbb{Q})} + \mathrm{e}^{\kappa^{(\mathbb{Q})}(u+t)} \left(v_t - \theta^{(\mathbb{Q})} \right) \right. \\ &\quad \left. + \sigma \mathrm{e}^{\kappa^{(\mathbb{Q})}u} \int_t^u \mathrm{e}^{\kappa^{(\mathbb{Q})}s} \sqrt{v_s} dW_s^v \right] du \\ &= \tfrac{\theta^{(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}} \left(\mathrm{e}^{2\kappa^{(\mathbb{Q})}T} - \mathrm{e}^{2\kappa^{(\mathbb{Q})}t} \right) + \tfrac{v_t - \theta^{(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}} \left(\mathrm{e}^{\kappa^{(\mathbb{Q})}(T+t)} - \mathrm{e}^{2\kappa^{(\mathbb{Q})}t} \right) \\ &\quad + \tfrac{\sigma}{\kappa^{(\mathbb{Q})}} \int_t^T \left(\mathrm{e}^{\kappa^{(\mathbb{Q})}(T+s)} - \mathrm{e}^{2\kappa^{(\mathbb{Q})}s} \right) \sqrt{v_s} dW_s^v. \end{aligned}$$

4.2.2 Squared Log Contract

The price process of the squared log contract $X_t^{(2)} := \mathbb{E}^{\mathbb{Q}}\left[x_T^2 \middle| \mathcal{F}_t \right]$, $T \geq t$ is

$$\begin{aligned} X_t^{(2)} &= \mathbb{E}^{\mathbb{Q}}\left[X_T^2 \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}}\left[(X_T - X_t + X_t)^2 \middle| \mathcal{F}_t \right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[(X_T - X_t)^2 + X_t^2 \middle| \mathcal{F}_t \right] = X_t^2 + \mathbb{E}^{\mathbb{Q}}\left[\left(\int_t^T dX_u \right)^2 \middle| \mathcal{F}_t \right] \\ &= X_t^2 + \mathbb{E}^{\mathbb{Q}}\left[\int_t^T (dX_u)^2 \middle| \mathcal{F}_t \right] = X_t^2 + \mathbb{E}^{\mathbb{Q}}\left[\langle X \rangle_T - \langle X \rangle_t \middle| \mathcal{F}_t \right], \end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\langle X\rangle_T - \langle X\rangle_t \middle| \mathcal{F}_t\right] \;=\; & \left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\theta^{(\mathbb{Q})}(T-t)\\
& -\left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\tfrac{1}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}-1\right)\left(v_t-\theta^{(\mathbb{Q})}\right)\\
& +\left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}-\tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\tfrac{\theta^{(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{\kappa^{(\mathbb{Q})}T}-\mathrm{e}^{\kappa^{(\mathbb{Q})}t}\right)\\
& +\left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}-\tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}\left(v_t-\theta^{(\mathbb{Q})}\right)(T-t)\\
& +\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\tfrac{\theta^{(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{2\kappa^{(\mathbb{Q})}T}-\mathrm{e}^{2\kappa^{(\mathbb{Q})}t}\right)\\
& +\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\tfrac{v_t-\theta^{(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{\kappa^{(\mathbb{Q})}(T+t)}-\mathrm{e}^{2\kappa^{(\mathbb{Q})}t}\right).
\end{aligned}$$

4.2.3 VIX Volatility Index

The CBOE Volatility Index (VIX) volatility index is defined in a way such that $vix_t^2 := \frac{2}{T-t}\left(x_t - X_t\right)$ and accordingly

$$vix_t = \sqrt{\theta^{(\mathbb{Q})} - \frac{\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}-1}{\kappa^{(\mathbb{Q})}(T-t)}\left(v_t-\theta^{(\mathbb{Q})}\right)}.$$

4.2.4 Variance Swaps

The value process of an idealised standard variance swap, as e.g. considered by Carr and Wu [2009], that pays the quadratic variation of the log-price for a fixed swap rate yields $V_{tT}^{(S)} := \mathbb{E}^{\mathbb{Q}}\left[\langle x\rangle_T \middle| \mathcal{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[\langle x\rangle_T \middle| \mathcal{F}_0\right]$, which implies $V_{0T}^{(S)} = 0$

at inception. In the Heston model we have

$$\begin{aligned}\mathbb{E}\left[\left.\langle x\rangle_T\right|\mathcal{F}_t\right] &= \mathbb{E}\Big[\theta T-\tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa T}-1\right)\left(v_0-\theta\right) \\ &\quad -\tfrac{\sigma}{\kappa}\int_0^T\left(\mathrm{e}^{-\kappa(T-s)}-1\right)\sqrt{v_s}dW_s^v\Big|\,\mathcal{F}_t\Big] \\ &= \theta T-\tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa T}-1\right)\left(v_0-\theta\right)-\tfrac{\sigma}{\kappa}\int_0^t\left(\mathrm{e}^{-\kappa(T-s)}-1\right)\sqrt{v_s}dW_s^v,\end{aligned}$$

and therefore

$$V_{tT}^{(S)}=-\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\int_0^t\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-s)}-1\right)\sqrt{v_s}dW_s^{v(\mathbb{Q})},$$

with dynamics $dV_{tT}^{(S)}=-\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}-1\right)\sqrt{v_t}dW_t^{v(\mathbb{Q})}$. A perpetual idealised standard variance swap that pays the quadratic variation of the log-price for a fixed swap rate up to infinity ($T\to\infty$) follows the price process

$$V_{t\infty}^{(S)}=\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\int_0^t\sqrt{v_s}dW_s^{v(\mathbb{Q})},$$

with dynamics $dV_{t\infty}^{(S)}=\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\sqrt{v_t}dW_t^{v(\mathbb{Q})}$. In a pure diffusion model, an idealised (i.e. continuously monitored) variance swap based on the realised leg as defined by Neuberger [2012] yields the same payoff as the idealised standard variance swap since

$$\int_0^T 2\left(\mathrm{e}^{dx_t}-dx_t-1\right)=\int_0^T 2\left(1+dx+\tfrac{1}{2}\left(dx_t\right)^2-dx_t-1\right)=\int_0^T\left(dx_t\right)^2=\langle x\rangle_t\,,$$

where all higher powers of dx_t vanish for a continuous process.

By contrast, the value process of our continuously monitored discretisation-invariant variance swap yields $V_{tT}^{(DI)} := \mathbb{E}^{\mathbb{Q}}\left[\langle X\rangle_T \middle| \mathcal{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[\langle X\rangle_T \middle| \mathcal{F}_0\right]$, which again implies $V_{0T}^{(DI)} = 0$. Using the dynamics and quadratic variation of the log contract in the Heston model, we have

$$\begin{aligned}
\mathbb{E}\left[\langle X\rangle_T \middle| \mathcal{F}_t\right] &= \left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathbb{E}\left[\int_0^T v_u du \middle| \mathcal{F}_t\right] \\
&\quad + \left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} - \tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\mathbb{E}\left[\int_0^T \mathrm{e}^{\kappa^{(\mathbb{Q})}u} v_u du \middle| \mathcal{F}_t\right] \\
&\quad + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\mathbb{E}\left[\int_0^T \mathrm{e}^{2\kappa^{(\mathbb{Q})}u} v_u du \middle| \mathcal{F}_t\right],
\end{aligned}$$

and therefore

$$\begin{aligned}
V_{tT}^{(DI)} &= -\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\int_0^t \left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-s)} - 1\right)\sqrt{v_s}dW_s^{v(\mathbb{Q})} \\
&\quad + \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(\rho\sigma - \tfrac{\sigma^2}{2\kappa^{(\mathbb{Q})}}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\int_0^t (t-s)\mathrm{e}^{\kappa^{(\mathbb{Q})}s}\sqrt{v_s}dW_s^{v(\mathbb{Q})} \\
&\quad + \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\int_0^t \left(\mathrm{e}^{\kappa^{(\mathbb{Q})}(t+s)} - \mathrm{e}^{2\kappa^{(\mathbb{Q})}s}\right)\sqrt{v_s}dW_s^{v(\mathbb{Q})}.
\end{aligned}$$

The perpetual variance swap that pays the quadratic variation of the log-contract for a fixed swap rate up to infinity ($T \to \infty$) follows the price process

$$V_{t\infty}^{(DI)} = \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\int_0^t \sqrt{v_s}dW_s^{v(\mathbb{Q})},$$

with dynamics $dV_{t\infty}^{(DI)} = \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\sqrt{v_t}dW_t^{v(\mathbb{Q})}$. Thus, the dynamics of the perpetual idealised discretisation-invariant variance swap corresponds to the dynamics of the perpetual idealised standard variance swap multiplied with an adjustment factor that depends on the mean reversion speed, the variance of variance as well as the correlation between the changes in price and variance. In particular, $\left(1 - \frac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \frac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)$ is positive as long as $|\rho| < 1$, which covers all relevant cases.

4.2.5 Change of Measure and Pricing Kernel

We now assume that the price process also follows a Heston model under the physical probability measure, albeit with a different set of parameters. Then

$$\begin{aligned} dF_t &:= F_t\left(\mu_t - r_t\right) dt + F_t\sqrt{v_t}dW_t^{F(\mathbb{P})}, \\ dv_t &:= \kappa^{(\mathbb{P})}\left(\theta^{(\mathbb{P})} - v_t\right) dt + \sqrt{v_t}\sigma dW_t^{v(\mathbb{P})}, \end{aligned}$$

where the condition $2\kappa^{(\mathbb{P})}\theta^{(\mathbb{P})} > \sigma^2$ guarantees positivity of v and the change of measure is given by $dW_t^{F(\mathbb{P})} = dW_t^{F(\mathbb{Q})} - \lambda_t^F dt$ as well as $dW_t^{v(\mathbb{P})} = dW_t^{v(\mathbb{Q})} - \lambda_t^v dt$, while the correlation ρ between the Wiener processes remains unchanged.[2]

Consequently, and in accordance with Chernov and Ghysels [2000], the market price of risk in the underlying yields $\lambda_t^F = \frac{\mu_t - r_t}{\sqrt{v_t}}$ and the market price of variance risk yields $\lambda_t^v = \frac{\sqrt{v_t}}{\sigma}\left(\kappa^{(\mathbb{Q})} - \kappa^{(\mathbb{P})}\right) - \frac{\kappa^{(\mathbb{Q})}\theta^{(\mathbb{Q})} - \kappa^{(\mathbb{P})}\theta^{(\mathbb{P})}}{\sqrt{v_t}\sigma}$, with the pricing kernel being

$$m_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{-\int_0^t \lambda_s^F dW_s^{F(\mathbb{P})} - \int_0^t \lambda_s^v dW_s^{v(\mathbb{P})} - \tfrac{1}{2}\int_0^t \left[\left(\lambda_s^F\right)^2 + \left(\lambda_s^v\right)^2\right] ds\right\},$$

or equivalently the inverse pricing kernel being

$$m_t^{-1} = \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \exp\left\{\int_0^t \lambda_s^F dW_s^{F(\mathbb{Q})} + \int_0^t \lambda_s^v dW_s^{v(\mathbb{Q})} - \tfrac{1}{2}\int_0^t \left[\left(\lambda_s^F\right)^2 + \left(\lambda_s^v\right)^2\right] ds\right\}.$$

[2]No arbitrage implies that ρ be the same under both measures, see Broadie et al. [2007] and Bardgett et al. [2015]. Details regarding change of measure in a multivariate setting are provided in the Appendix.

4.2.6 Equity and Variance Risk Premia

The ERP can be defined as $\mathbb{E}^{\mathbb{P}}\left[\left(1-m_t\right)F_{T}\middle|\,\mathcal{F}_t\right]$, and the instantaneous ERP in the Heston model yields $F_t\left(\mu_t-r_t\right)$. This corresponds to the annualised amount which an investor expects to receive for holding one share of the underlying over the infinitesimally small time period $[t,t+dt]$. Alternative definitions can be found in Bollerslev and Todorov [2011], where the ERP is $\mathbb{E}^{\mathbb{P}}\left[\left(1-m_t\right)F_{T}/F_t\middle|\,\mathcal{F}_t\right]$ and the instantaneous ERP yields μ_t-r_t, as well as in Bardgett et al. [2015], where the ERP yields $\mathbb{E}^{\mathbb{P}}\left[\left(1-m_t\right)x_{T}\middle|\,\mathcal{F}_t\right]$, with the instantaneous ERP being $\mu_t-r_t-\frac{v_t}{2}$.

The VRP is commonly defined as $\mathbb{E}^{\mathbb{P}}\left[\left(1-m_t\right)\int_t^T v_s ds\middle|\,\mathcal{F}_t\right]$, and the instantaneous VRP in the Heston model yields

$$v_t-\theta^{(\mathbb{Q})}-\tfrac{\kappa^{(\mathbb{P})}}{\kappa^{(\mathbb{Q})}}\left[v_t-\theta^{(\mathbb{P})}\right].$$

This corresponds to the annualised amount which an investor expects to receive for shorting a delta-hedged log contract (with infinite maturity), i.e. investing in a perpetual standard variance swap, over the infinitesimally small time period $[t,t+dt]$. We suggest the alternative definition $\mathbb{E}^{\mathbb{P}}\left[\left(1-m_t\right)x_{T}^2\middle|\,\mathcal{F}_t\right]$, which is consistent with the definition of the ERP by Bardgett et al. [2015]. The instantaneous VRP in the Heston model yields

$$\left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\left(v_t-\theta^{(\mathbb{Q})}-\tfrac{\kappa^{(\mathbb{P})}}{\kappa^{(\mathbb{Q})}}\left[v_t-\theta^{(\mathbb{P})}\right]\right),$$

and corresponds to the annualised amount which an investor expects to receive for investing in a perpetual idealised discretisation-invariant variance swap over the infinitesimally small time period $[t,t+dt]$. Assuming that the physical mean reversion level is considerably lower than the risk-neutral mean reversion level, and that the mean reversion speed is roughly the same under the physical and risk-neutral measure, both variance risk premia are negative.

4.2.7 Joint Characteristic Function

We now turn to the joint characteristic function of the log price and the variance:

$$\Psi\left(t, x_t, v_t, \xi, \eta, T\right) := \mathbb{E}\left[\mathrm{e}^{i\xi x_T + i\eta v_T} \,\middle|\, \mathcal{F}_t\right].$$

Using Itô's formula for complex functions we derive the dynamics of Ψ as

$$\begin{aligned} d\Psi &= \Psi_t dt + \Psi_x dx_t + \tfrac{1}{2}\Psi_{xx} d\left\langle x\right\rangle_t + \Psi_v dv_t + \tfrac{1}{2}\Psi_{vv} d\left\langle v\right\rangle_t + \Psi_{xv} dx_t dv_t \\ &= \left[\Psi_t - \tfrac{v_t}{2}\Psi_x + \tfrac{v_t}{2}\Psi_{xx} + \kappa\left(\theta - v_t\right)\Psi_v + \tfrac{v_t}{2}\sigma^2\Psi_{vv} + \rho\sigma v_t \Psi_{xv}\right] dt \\ &\quad + \sqrt{v_t}\Psi_x dW_t^F + \sigma\sqrt{v_t}\Psi_v dW_t^v, \end{aligned}$$

where the subscripts denote the respective derivatives of Ψ. Now Ψ follows a martingale by construction and therefore the drift must be zero, i.e.

$$\Psi_t - \tfrac{v_t}{2}\Psi_x + \tfrac{v_t}{2}\Psi_{xx} + \kappa\left(\theta - v_t\right)\Psi_v + \tfrac{v_t}{2}\sigma^2\Psi_{vv} + \rho\sigma v_t \Psi_{xv} = 0.$$

Since the Heston model is affine in all state variables we know from Duffie et al. [2000] that the characteristic function must have a representation of the form

$$\Psi\left(t, x_t, v_t, \xi, \eta, T\right) = \mathrm{e}^{\mathcal{A}(t) + \mathcal{B}(t) x_t + \mathcal{C}(t) v_t},$$

with the relevant derivatives being

$$\begin{aligned}
\Psi_t\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta\right)\left[\mathcal{A}'(t)+\mathcal{B}'(t)x_t+\mathcal{C}'(t)v_t\right],\\
\Psi_x\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta,T\right)\mathcal{B}(t),\\
\Psi_{xx}\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta,T\right)\mathcal{B}(t)^2,\\
\Psi_v\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta,T\right)\mathcal{C}(t),\\
\Psi_{vv}\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta,T\right)\mathcal{C}(t)^2,\\
\Psi_{xv}\left(t,x_t,v_t,\xi,\eta,T\right) &= \Psi\left(t,x_t,v_t,\xi,\eta,T\right)\mathcal{B}(t)\mathcal{C}(t).
\end{aligned}$$

Inserting these derivatives into the drift condition and dividing by Ψ yields

$$\begin{aligned}
&\mathcal{A}'(t)+\mathcal{B}'(t)x_t+\mathcal{C}'(t)v_t-\tfrac{v_t}{2}\mathcal{B}(t)+\tfrac{v_t}{2}\mathcal{B}(t)^2\\
&+\kappa\left(\theta-v_t\right)\mathcal{C}(t)+\tfrac{v_t}{2}\sigma^2\mathcal{C}(t)^2+\rho\sigma v_t\mathcal{B}(t)\mathcal{C}(t) \quad = \quad 0,
\end{aligned}$$

and since this must hold for all t, x_t and v_t we have three equations:

$$\begin{aligned}
\mathcal{A}'(t)+\kappa\theta\mathcal{C}(t) &= 0,\\
\mathcal{B}'(t) &= 0,\\
\mathcal{C}'(t)-\tfrac{1}{2}\mathcal{B}(t)+\tfrac{1}{2}\mathcal{B}(t)^2-\kappa\mathcal{C}(t)+\tfrac{\sigma^2}{2}\mathcal{C}(t)^2+\rho\sigma\mathcal{B}(t)\mathcal{C}(t) &= 0.
\end{aligned}$$

Further, for $t=T$ we have $\Psi\left(T,x_T,v_T,\xi,\eta,T\right)=\mathrm{e}^{\mathcal{A}(T)+\mathcal{B}(T)x_T+\mathcal{C}(T)v_T}=\mathrm{e}^{i\xi x_T+i\eta v_T}$ which implies the boundary conditions $\mathcal{A}(T)=0$, $\mathcal{B}(T)=i\xi$ and $\mathcal{C}(T)=i\eta$. From the second equation it follows immediately that $\mathcal{B}(t)=i\xi$. Then the third equation yields

$$\mathcal{C}'(t)=\tfrac{\alpha}{2}+\beta\mathcal{C}(t)-\tfrac{\sigma^2}{2}\mathcal{C}(t)^2,$$

with $\beta = \kappa - \rho\sigma i\xi$ and $\alpha = i\xi(1 - i\xi)$, which can be reduced to a linear ordinary differential equation by performing the substitution $\mathcal{C}(t) = \frac{2}{\sigma^2}\frac{\tilde{\mathcal{C}}'(t)}{\tilde{\mathcal{C}}(t)}$, so that $\mathcal{C}'(t) = \frac{2}{\sigma^2}\frac{\tilde{\mathcal{C}}(t)\tilde{\mathcal{C}}''(t) - \tilde{\mathcal{C}}'(t)^2}{\tilde{\mathcal{C}}(t)^2}$ and $\tilde{\mathcal{C}}'(T) = \frac{1}{2}i\eta\sigma^2\tilde{\mathcal{C}}(T)$, and multiplying both sides of the equation with $\frac{\sigma^2}{2}\tilde{\mathcal{C}}(t)$ yields

$$\tilde{\mathcal{C}}''(t) - \beta\tilde{\mathcal{C}}'(t) - \tfrac{\alpha\sigma^2}{4}\tilde{\mathcal{C}}(t) = 0.$$

The roots of the characteristic polynomial $z^2 - \beta z - \frac{\alpha\sigma^2}{4}$ are $z_\pm = \frac{1}{2}(\beta \pm \gamma)$, with $\gamma = \sqrt{\beta^2 + \alpha\sigma^2}$, and are distinct and non-zero for $0 < \xi < 1$. Therefore the closed-form solution is given by $\tilde{\mathcal{C}}(t) = \phi_+ \mathrm{e}^{z_+ t} + \phi_- \mathrm{e}^{z_- t}$ with derivative $\tilde{\mathcal{C}}'(t) = \phi_+ z_+ \mathrm{e}^{z_+ t} + \phi_- z_- \mathrm{e}^{z_- t}$ and together with the boundary condition $\frac{\phi_-}{\phi_+} = -\frac{2z_+ - i\eta\sigma^2}{2z_- - i\eta\sigma^2}\mathrm{e}^{(z_+ - z_-)T}$ we have

$$\begin{aligned}
\tilde{\mathcal{C}}(t) &= \phi_+ \mathrm{e}^{z_+ t}\left(1 - \tfrac{2z_+ - i\eta\sigma^2}{2z_- - i\eta\sigma^2}\mathrm{e}^{(z_+ - z_-)(T-t)}\right), \\
\tilde{\mathcal{C}}'(t) &= \phi_+ \mathrm{e}^{z_+ t}\left(z_+ - z_-\tfrac{2z_+ - i\eta\sigma^2}{2z_- - i\eta\sigma^2}\mathrm{e}^{(z_+ - z_-)(T-t)}\right).
\end{aligned}$$

Inserting this into the original function yields

$$\begin{aligned}
\mathcal{C}(t) &= \frac{2}{\sigma^2}\frac{z_+ - z_-\frac{2z_+ - i\eta\sigma^2}{2z_- - i\eta\sigma^2}\mathrm{e}^{(z_+ - z_-)(T-t)}}{1 - \frac{2z_+ - i\eta\sigma^2}{2z_- - i\eta\sigma^2}\mathrm{e}^{(z_+ - z_-)(T-t)}} \\
&= \frac{1}{\sigma^2}\frac{(\beta+\gamma)(\beta-\gamma-i\eta\sigma^2) - (\beta-\gamma)(\beta+\gamma-i\eta\sigma^2)\,\mathrm{e}^{\gamma(T-t)}}{\beta-\gamma-i\eta\sigma^2 - (\beta+\gamma-i\eta\sigma^2)\,\mathrm{e}^{\gamma(T-t)}} \\
&= \frac{(\alpha+\beta i\eta)\left(1-\mathrm{e}^{\gamma(T-t)}\right) + \gamma i\eta\left(1+\mathrm{e}^{\gamma(T-t)}\right)}{2\gamma - (\beta+\gamma-i\eta\sigma^2)\left(1-\mathrm{e}^{\gamma(T-t)}\right)}.
\end{aligned}$$

We obtain $\mathcal{A}(t)$ by integrating w.r.t. t and using the boundary condition:

$$\begin{aligned}
\mathcal{A}(t) &= \mathcal{A}(T) + \kappa\theta \int_t^T \mathcal{C}(u)du = \frac{2\kappa\theta}{\sigma^2}\int_t^T \frac{\tilde{\mathcal{C}}'(u)}{\tilde{\mathcal{C}}(u)}du = \frac{2\kappa\theta}{\sigma^2}\left[\ln \tilde{\mathcal{C}}(u)\right]_t^T \\
&= \tfrac{\kappa\theta}{\sigma^2}\left[(\beta+\gamma)\,u + 2\ln\left(1 - \tfrac{\beta+\gamma-i\eta\sigma^2}{\beta-\gamma-i\eta\sigma^2}\mathrm{e}^{\gamma(T-u)}\right)\right]_t^T \\
&= \tfrac{\kappa\theta}{\sigma^2}\left[(\beta+\gamma)\,(T-t) - 2\ln\left(1 - \tfrac{\beta+\gamma-i\eta\sigma^2}{2\gamma}\left(1-\mathrm{e}^{\gamma(T-t)}\right)\right)\right].
\end{aligned}$$

On the one hand, when we are interested in the characteristic function of the log-price only, i.e. set $\eta = 0$, the coefficients simplify to:

$$\begin{aligned}
\mathcal{A}(t)|_{\eta=0} &= \tfrac{\kappa\theta}{\sigma^2}\left[(\beta+\gamma)\,(T-t) - 2\ln\left(1 - \tfrac{\beta+\gamma}{2\gamma}\left(1-\mathrm{e}^{\gamma(T-t)}\right)\right)\right], \\
\mathcal{C}(t)|_{\eta=0} &= \frac{\alpha\left(1-\mathrm{e}^{\gamma(T-t)}\right)}{2\gamma - (\beta+\gamma)\left(1-\mathrm{e}^{\gamma(T-t)}\right)}.
\end{aligned}$$

On the other hand, when we want to look at the characteristic function of the variance only, i.e. set $\xi = 0$ so that $\alpha = 0$ and $\beta = \gamma = \kappa$, the coefficients simplify to

$$\begin{aligned}
\mathcal{A}(t)|_{\xi=0} &= \tfrac{2\kappa\theta}{\sigma^2}\left[\kappa(T-t) - \ln\left(\tfrac{i\eta\sigma^2}{2\kappa} + \left(1-\tfrac{i\eta\sigma^2}{2\kappa}\right)\mathrm{e}^{\kappa(T-t)}\right)\right], \\
\mathcal{C}(t)|_{\xi=0} &= \frac{i\eta}{\left(2\kappa - i\eta\sigma^2\right)\left(1-\mathrm{e}^{\gamma(T-t)}\right)}.
\end{aligned}$$

4.2.8 Power Log Contracts

Having found this representation of Ψ, we can derive the price of the n-th power log contract $X_t^{(n)} := \mathbb{E}^{\mathbb{Q}}\left[x_T^n \middle| \mathcal{F}_t\right]$ by making use of the relationship

$$\begin{aligned} X_t^{(n)} &= \mathbb{E}^{\mathbb{Q}}\left[x_T^n \middle| \mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[i^{-n}\left(\tfrac{\partial}{\partial\xi}\right)^n \mathrm{e}^{i\xi x_T + i\eta v_T}\Big|_{\xi,\eta=0} \middle| \mathcal{F}_t\right] \\ &= i^{-n}\left(\tfrac{\partial}{\partial\xi}\right)^n \Psi\left(t, x_t, v_t, \xi, \eta, T\right)\Big|_{\xi,\eta=0}. \end{aligned}$$

For the log contract, recalling that $\Psi\left(t, x_t, v_t, \xi, \eta, T\right) = \mathrm{e}^{\mathcal{A}(t) + i\xi x_t + \mathcal{C}(t) v_t}$, we have

$$\begin{aligned} X_t &= -i\ \Psi\left(t, x_t, v_t, \xi, \eta, T\right) \tfrac{\partial}{\partial\xi}\left[\mathcal{A}(t) + i\xi x_t + \mathcal{C}(t) v_t\right]\Big|_{\xi,\eta=0} \\ &= -i\ \tfrac{\partial}{\partial\xi}\mathcal{A}(t)\Big|_{\xi,\eta=0} + x_t - i\ \tfrac{\partial}{\partial\xi}\mathcal{C}(t)\Big|_{\xi,\eta=0} v_t, \end{aligned}$$

since $\mathcal{A}(t)|_{\xi,\eta=0} = \mathcal{C}(t)|_{\xi,\eta=0} = 0$. Further

$$\begin{aligned} \tfrac{\partial}{\partial\xi}\mathcal{A}(t)\Big|_{\eta=0} &= \tfrac{\kappa\theta}{\sigma^2}\Bigg[\left(\beta' + \gamma'\right)(T-t) \\ &\quad -2\frac{\frac{\beta+\gamma}{2\gamma}\mathrm{e}^{\gamma(T-t)}\gamma'(T-t) - \frac{\gamma\beta' - \beta\gamma'}{2\gamma^2}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)}{1 - \frac{\beta+\gamma}{2\gamma}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)}\Bigg], \\ \tfrac{\partial}{\partial\xi}\mathcal{C}(t)\Big|_{\eta=0} &= \frac{\alpha'\left(1 - \mathrm{e}^{\gamma(T-t)}\right) - \alpha\mathrm{e}^{\gamma(T-t)}\gamma'(T-t)}{2\gamma - (\beta+\gamma)\left(1 - \mathrm{e}^{\gamma(T-t)}\right)} \\ &\quad - \frac{\alpha\left(2\gamma' - \left(\beta' + \gamma'\right)\left(1 - \mathrm{e}^{\gamma(T-t)}\right) + (\beta+\gamma)\,\mathrm{e}^{\gamma(T-t)}\gamma'(T-t)\right)}{\left(1 - \mathrm{e}^{\gamma(T-t)}\right)^{-1}\left(2\gamma - (\beta+\gamma)\left(1 - \mathrm{e}^{\gamma(T-t)}\right)\right)^2}, \end{aligned}$$

with $\alpha' := \frac{\partial}{\partial\xi}\alpha = i + 2\xi$, $\beta' := \frac{\partial}{\partial\xi}\beta = -i\rho\sigma$ and $\gamma' := \frac{\partial}{\partial\xi}\gamma = \frac{2\beta\beta' + \alpha'\sigma^2}{2\sqrt{\beta^2 + \alpha\sigma^2}}$. Now we have $\alpha|_{\xi=0} = 0$ and $\beta|_{\xi=0} = \gamma|_{\xi=0} = \kappa$ as well as $\alpha'|_{\xi=0} = i$, $\beta'|_{\xi=0} = -i\rho\sigma$ and

$\gamma'|_{\xi=0} = \frac{i\sigma^2}{2\kappa} - i\rho\sigma$ and therefore evaluation at ξ yields

$$\begin{aligned} \frac{\partial}{\partial\xi}\mathcal{A}(t)\Big|_{\xi,\eta=0} &= -\frac{i\theta}{2}(T-t) - \frac{i\theta}{2\kappa}\left(\mathrm{e}^{-\kappa(T-t)} - 1\right), \\ \frac{\partial}{\partial\xi}\mathcal{C}(t)\Big|_{\xi,\eta=0} &= \frac{i}{2\kappa}\left(\mathrm{e}^{-\kappa(T-t)} - 1\right). \end{aligned}$$

Together we have

$$\begin{aligned} X_t &= -\frac{\theta}{2}(T-t) - \frac{\theta}{2\kappa}\left(\mathrm{e}^{-\kappa(T-t)} - 1\right) + x_t - \frac{1}{2\kappa}\left(\mathrm{e}^{-\kappa(T-t)} - 1\right) v_t \\ &= x_t - \frac{\theta}{2}(T-t) + \frac{1}{2\kappa}\left(\mathrm{e}^{-\kappa(T-t)} - 1\right)(v_t - \theta), \end{aligned}$$

in accordance with the direct solution found in the log contract section. For the squared log contract we have

$$\begin{aligned} X_t^{(2)} &= i^{-2}\,\Psi\left[\left(\frac{\partial}{\partial\xi}\right)^2 [\mathcal{A}(t) + i\xi x_t + \mathcal{C}(t)v_t] + \left(\frac{\partial}{\partial\xi}[\mathcal{A}(t) + i\xi x_t + \mathcal{C}(t)v_t]\right)^2\right]\Bigg|_{\xi,\eta=0} \\ &= X_t^2 - \left(\frac{\partial}{\partial\xi}\right)^2 \mathcal{A}(t)\Big|_{\xi,\eta=0} - \left(\frac{\partial}{\partial\xi}\right)^2 \mathcal{C}(t)\Big|_{\xi,\eta=0} v_t. \end{aligned}$$

Further

$$\begin{aligned} \left(\frac{\partial}{\partial\xi}\right)^2 \mathcal{A}(t)\Big|_{\eta=0} = \frac{\kappa\theta}{\sigma^2}\Bigg[&(\beta'' + \gamma'')(T-t) - 4\frac{\frac{\gamma\beta' - \beta\gamma'}{2\gamma^2}\mathrm{e}^{\gamma(T-t)}\gamma'(T-t)}{1 - \frac{\beta+\gamma}{2\gamma}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)} \\ &-2\frac{\frac{\beta+\gamma}{2\gamma}\left(\mathrm{e}^{\gamma(T-t)}\gamma''(T-t) + \mathrm{e}^{\gamma(T-t)}(\gamma')^2(T-t)^2\right)}{1 - \frac{\beta+\gamma}{2\gamma}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)} \\ &+2\frac{\left(\frac{\gamma\beta'' - \beta\gamma''}{2\gamma^2} - \frac{\gamma'(\gamma\beta' - \beta\gamma')}{\gamma^3}\right)\left(1 - \mathrm{e}^{\gamma(T-t)}\right)}{1 - \frac{\beta+\gamma}{2\gamma}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)} \\ &+2\left(\frac{\frac{\beta+\gamma}{2\gamma}\mathrm{e}^{\gamma(T-t)}\gamma'(T-t) - \frac{\gamma\beta' - \beta\gamma'}{2\gamma^2}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)}{1 - \frac{\beta+\gamma}{2\gamma}\left(1 - \mathrm{e}^{\gamma(T-t)}\right)}\right)^2\Bigg], \end{aligned}$$

as well as (recalling that $\alpha|_{\xi=0} = 0$ and hence we can disregard the α term)

$$\left(\tfrac{\partial}{\partial\xi}\right)^2 \mathcal{C}(t)\Big|_{\eta=0} = \alpha\cdot\ldots + \frac{\alpha''\left(1-\mathrm{e}^{\gamma(T-t)}\right) - 2\alpha'\mathrm{e}^{\gamma(T-t)}\gamma'(T-t)}{2\gamma-(\beta+\gamma)\left(1-\mathrm{e}^{\gamma(T-t)}\right)} - \frac{2\alpha'\left(2\gamma'-(\beta'+\gamma')\left(1-\mathrm{e}^{\gamma(T-t)}\right)+(\beta+\gamma)\,\mathrm{e}^{\gamma(T-t)}\gamma'(T-t)\right)}{\left(1-\mathrm{e}^{\gamma(T-t)}\right)^{-1}\left(2\gamma-(\beta+\gamma)\left(1-\mathrm{e}^{\gamma(T-t)}\right)\right)^2},$$

where $\alpha'' := \frac{\partial}{\partial\xi}\alpha' = 2$, $\beta'' := \frac{\partial}{\partial\xi}\beta' = 0$ and

$$\gamma'' := \tfrac{\partial}{\partial\xi}\gamma' = \frac{2\sqrt{\beta^2+\alpha\sigma^2}\left(2(\beta')^2+2\beta\beta''+\alpha''\sigma^2\right)-\left(2\beta\beta'+\alpha'\sigma^2\right)^2\sqrt{\beta^2+\alpha\sigma^2}^{\,-1}}{4(\beta^2+\alpha\sigma^2)}.$$

Now we have $\alpha''|_{\xi=0} = 2$, $\beta''|_{\xi=0} = 0$ and $\gamma''|_{\xi=0} = \frac{\sigma^2}{\kappa}\left(1-\frac{\rho\sigma}{\kappa}+\frac{\sigma^2}{4\kappa^2}\right)$ and therefore evaluating at $\xi = 0$ yields

$$\left(\tfrac{\partial}{\partial\xi}\right)^2 \mathcal{A}(t)\Big|_{\xi,\eta=0} = \theta\left[-\left(1-\tfrac{2\rho\sigma}{\kappa}+\tfrac{3\sigma^2}{4\kappa^2}\right)\left(\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}+T-t\right) - \left(\tfrac{\sigma^2}{2\kappa}-\rho\sigma\right)(T-t)\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa} - \tfrac{\sigma^2}{8\kappa}\left(\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}\right)^2\right],$$

and

$$\left(\tfrac{\partial}{\partial\xi}\right)^2 \mathcal{C}(t)\Big|_{\xi,\eta=0} = \tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa} + \left(\tfrac{\sigma^2}{2\kappa^2}-\tfrac{\rho\sigma}{\kappa}\right)(T-t) + \tfrac{\sigma^2}{4\kappa^2}\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}\mathrm{e}^{-\kappa(T-t)} + \left(\tfrac{\sigma^2}{4\kappa^2}-\tfrac{\rho\sigma}{\kappa}\right)\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa} + \left(\tfrac{\sigma^2}{2\kappa}-\rho\sigma\right)(T-t)\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}.$$

Together we have

$$
\begin{aligned}
X_t^{(2)} &= X_t^2 + \left(1 - \tfrac{\rho\sigma}{\kappa} + \tfrac{\sigma^2}{4\kappa^2}\right)(T-t)\,\theta \\
&\quad - \left(1 - \tfrac{\rho\sigma}{\kappa} + \tfrac{\sigma^2}{4\kappa^2}\right)\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}\,(v_t - \theta) \\
&\quad - \left(\tfrac{\rho\sigma}{\kappa} - \tfrac{\sigma^2}{2\kappa^2}\right)\tfrac{\mathrm{e}^{-\kappa(T-t)}-1}{\kappa}\theta \\
&\quad + \left(\tfrac{\rho\sigma}{\kappa} - \tfrac{\sigma^2}{2\kappa^2}\right)(T-t)\mathrm{e}^{-\kappa(T-t)}\,(v_t - \theta) \\
&\quad + \tfrac{\sigma^2}{8\kappa^3}\left(1 - \mathrm{e}^{-2\kappa(T-t)}\right)\theta \\
&\quad - \tfrac{\sigma^2}{4\kappa^3}\left(\mathrm{e}^{-2\kappa(T-t)} - \mathrm{e}^{-\kappa(T-t)}\right)(v_t - \theta)\,,
\end{aligned}
$$

in accordance with the direct solution found in the squared log contract section.

4.3 SV with Contemporaneous Jumps

The stochastic volatility with contemporaneous jumps (SVCJ) model by Duffie et al. [2000] is an extension of the Heston model that allows for discontinuities in both the underlying and the variance process. It is more general than the Bates [1996] model, which only allows for jumps in the underlying. Formally,

$$
\begin{aligned}
\tfrac{dF_t}{F_{t-}} &:= -\psi dt + \sqrt{v_{t-}}\,dW_t^{F(\mathbb{Q})} + \left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right)dN_t \\
dv_t &:= \kappa^{(\mathbb{Q})}\left(\theta^{(\mathbb{Q})} - v_{t-}\right)dt + \sqrt{v_{t-}}\,\sigma dW_t^{v(\mathbb{Q})} + Z_t^{v(\mathbb{Q})}dN_t
\end{aligned}
$$

where $dW_t^{F(\mathbb{Q})}dW_t^{v(\mathbb{Q})} = \rho dt$, $Z^{F(\mathbb{Q})} \sim \mathcal{N}\left(\mu^{F(\mathbb{Q})}, \sigma^{F(\mathbb{Q})}\right)$, $Z^{v(\mathbb{Q})} \sim \mathcal{E}\left(\mu^{v(\mathbb{Q})}\right)$, N_t is a Poisson process with constant intensity λ and F_{t-} denotes the value of F prior to any jump at time t (and analogously for v). For the underlying to follow a martingale under the risk-neutral measure, we must define the jump compensator as $\psi := \lambda \left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - 1 \right)$. The log price $x_t = \ln F_t$ follows the dynamics

$$\begin{aligned} dx_t &= -\psi dt + \sqrt{v_{t^-}} dW_{t^-}^F - \tfrac{1}{2} v_{t^-} dt + \Delta x_t \\ &= \sqrt{v_{t^-}} dW_t^F - \left(\psi + \tfrac{1}{2} v_{t^-}\right) dt + Z_t^F dN_t, \end{aligned}$$

where $\Delta x_t := \ln F_t - \ln F_{t^-} = Z_t^F dN_t$, having used that $dW_{t^-} = dW_t$ by predictability in the second line.

4.3.1 Variance Process and Log Contract

First we derive the explicit solution of the variance process by applying Itô's formula for jump diffusion processes (see e.g. Cont and Tankov [2004] and Appendix) to $\mathrm{e}^{\kappa t} v_t$:

$$\begin{aligned} d\left(\mathrm{e}^{\kappa t} v_t\right) &= \mathrm{e}^{\kappa t} \kappa v_{t^-} dt + \mathrm{e}^{\kappa t} dv_{t^-} + \mathrm{e}^{\kappa t} \Delta v_t \\ &= \mathrm{e}^{\kappa t} \kappa \theta dt + \mathrm{e}^{\kappa t} \sqrt{v_{t^-}} \sigma dW_t^v + \mathrm{e}^{\kappa t} Z_t^v dN_t, \end{aligned}$$

where $\Delta v_t := v_t - v_{t^-} = Z_t^v dN_t$. Integrating both sides from t to $u > t$ yields

$$\mathrm{e}^{\kappa u} v_u - \mathrm{e}^{\kappa t} v_t = \theta \left(\mathrm{e}^{\kappa u} - \mathrm{e}^{\kappa t}\right) + \sigma \int_t^u \mathrm{e}^{\kappa s} \sqrt{v_{s^-}} dW_s^v + \int_t^u \mathrm{e}^{\kappa s} Z_s^v dN_s$$

and we can solve for v_u:

$$\begin{aligned} v_u &= \mathrm{e}^{-\kappa(u-t)} v_t + \theta \left(1 - \mathrm{e}^{-\kappa(u-t)}\right) + \mathrm{e}^{-\kappa u} \sigma \int_t^u \mathrm{e}^{\kappa s} \sqrt{v_{s^-}} dW_s^v + \mathrm{e}^{-\kappa u} \int_t^u \mathrm{e}^{\kappa s} Z_s^v dN_s \\ &= \theta + \mathrm{e}^{-\kappa(u-t)} \left(v_t - \theta\right) + \sigma \int_t^u \mathrm{e}^{-\kappa(u-s)} \sqrt{v_{s^-}} dW_s^v + \sum_{t < s \leq u}^{\Delta v_s \neq 0} \mathrm{e}^{-\kappa(u-s)} \Delta v_s. \end{aligned}$$

Also, since

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[\left.\sum_{t<s\leq u}^{\Delta v_s\neq 0} \mathrm{e}^{-\kappa(u-s)}\Delta v_s\right| \mathcal{F}_t\right] &= \int_t^u \mathrm{e}^{-\kappa(u-s)}\mathbb{E}^{\mathbb{Q}}\left[\left. Z_s^v dN_s\right| \mathcal{F}_t\right] \\ &= \lambda\mu^{v(\mathbb{Q})}\int_t^u \mathrm{e}^{-\kappa^{(\mathbb{Q})}(u-s)}ds, \end{aligned}$$

we have $\mathbb{E}^{\mathbb{Q}}\left[\left. v_u\right| \mathcal{F}_t\right] = \theta^{(\mathbb{Q})} + \mathrm{e}^{-\kappa^{(\mathbb{Q})}(u-t)}\left(v_t - \theta^{(\mathbb{Q})}\right) + \frac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\left(1 - \mathrm{e}^{-\kappa^{(\mathbb{Q})}(u-t)}\right)$. By extension of the corresponding calculations for the Heston model, the quadratic variation of the log price yields

$$\begin{aligned} \langle x\rangle_t &= \int_0^t v_{u^-}du + \int_0^t \left(Z_u^F\right)^2 dN_u \\ &= \theta t - \tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa t} - 1\right)\left(v_{_0} - \theta\right) - \tfrac{\sigma}{\kappa}\int_0^t \left(\mathrm{e}^{-\kappa(t-s)} - 1\right)\sqrt{v_{s^-}}dW_s^v \\ &\quad + \int_0^t \sum_{0<s<u}^{\Delta v_s\neq 0} \mathrm{e}^{-\kappa(u-s)}\Delta v_s du + \sum_{0<u\leq t}^{\Delta x_t\neq 0}\left(\Delta x_u\right)^2, \end{aligned}$$

since $dW_t^F dN_t = dW_t^v dN_t = 0$ because diffusions and jumps are orthogonal.

The price process of the log contract is

$$\begin{aligned} X_t &= x_t + \mathbb{E}^{\mathbb{Q}}\left[\left.\int_t^T dx_u\right| \mathcal{F}_t\right] \\ &= x_t - \int_t^T \mathbb{E}^{\mathbb{Q}}\left[\left.\psi + \tfrac{1}{2}v_{u^-}\right| \mathcal{F}_t\right]du + \mathbb{E}^{\mathbb{Q}}\left[\left.\sum_{t<u\leq T}^{\Delta x_u\neq 0}\Delta x_u\right| \mathcal{F}_t\right] \\ &= x_t - \left(\psi + \tfrac{\theta^{(\mathbb{Q})}}{2} + \tfrac{\lambda\mu^{v(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}}\right)(T-t) - \tfrac{1}{2}\left(v_t - \theta^{(\mathbb{Q})} - \tfrac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)\int_t^T \mathrm{e}^{-\kappa^{(\mathbb{Q})}(u-t)}du \\ &\quad + \lambda\mu^{F(\mathbb{Q})}(T-t) \\ &= x_t - \left[\tfrac{\theta^{(\mathbb{Q})}}{2} + \tfrac{\lambda\mu^{v(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}} + \psi - \lambda\mu^{F(\mathbb{Q})}\right](T-t) \\ &\quad + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(v_t - \theta^{(\mathbb{Q})} - \tfrac{\lambda\mu^{v(\mathbb{Q})}}{\kappa}\right)\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1\right), \end{aligned}$$

using that $\mathbb{E}^{\mathbb{Q}}\left[v_{u^-}\right|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}\left[v_u\right|\mathcal{F}_t]$ in the second line, with dynamics

$$\begin{aligned}
dX_t &= dx_t + \left[\tfrac{1}{2}\left(\theta^{(\mathbb{Q})} + \tfrac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right) + \psi - \lambda\mu^{F(\mathbb{Q})}\right] dt \\
&\quad + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left[\left(v_{t^-} - \theta^{(\mathbb{Q})} - \tfrac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}\kappa^{(\mathbb{Q})}dt\right. \\
&\quad \left. + \left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1\right)\left(dv_{t^-} + \Delta v_t\right)\right] \\
&= \sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} - \left(\psi + \tfrac{1}{2}v_{t^-}\right)dt + Z_t^{F(\mathbb{Q})}dN_t \\
&\quad + \left[\tfrac{\theta^{(\mathbb{Q})}}{2} + \tfrac{\lambda\mu^{v(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}} + \psi - \lambda\mu^{F(\mathbb{Q})} + \tfrac{1}{2}\left(v_{t^-} - \theta^{(\mathbb{Q})} - \tfrac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}\right]dt \\
&\quad + \tfrac{\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}-1}{2\kappa^{(\mathbb{Q})}}\left[\kappa^{(\mathbb{Q})}\left(\theta^{(\mathbb{Q})} - v_{t^-}\right)dt + \sqrt{v_{t^-}}\sigma dW_t^{v(\mathbb{Q})} + Z_t^{v(\mathbb{Q})}dN_t\right] \\
&= \sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} + Z_t^{F(\mathbb{Q})}dN_t - \lambda\mu^{F(\mathbb{Q})}dt \\
&\quad + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1\right)\left(\sqrt{v_{t^-}}\sigma dW_t^{v(\mathbb{Q})} + Z_t^{v(\mathbb{Q})}dN_t - \lambda\mu^{v(\mathbb{Q})}dt\right).
\end{aligned}$$

The quadratic variation of the log contract yields

$$\begin{aligned}
\langle X\rangle_t &= \int_0^t v_{u^-}\left[1 + \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1\right) + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1\right)^2\right]du \\
&\quad + \int_0^t \left(Z_u^{F(\mathbb{Q})} + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1\right)Z_u^{v(\mathbb{Q})}\right)^2 dN_u \\
&= \left(1 - \tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\int_0^t v_u du + \left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}} - \tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\int_0^t \mathrm{e}^{\kappa^{(\mathbb{Q})}u}v_u du \\
&\quad + \tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\int_0^t \mathrm{e}^{2\kappa^{(\mathbb{Q})}u}v_u du \\
&\quad + \int_0^t \left[Z_u^{F(\mathbb{Q})} + \tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)} - 1\right)Z_u^{v(\mathbb{Q})}\right]^2 dN_u,
\end{aligned}$$

where again we have used that $dW_t^F dN_t = dW_t^v dN_t = 0$.

4.3.2 VIX Volatility Index

In the SVCJ model the VIX volatility index is given by

$$vix_t = \sqrt{\theta^{(\mathbb{Q})} + \frac{\lambda\mu^{v(\mathbb{Q})}}{\kappa} + \frac{\mathrm{e}^{\mu^{F(\mathbb{Q})}+\frac{1}{2}\sigma_F^2}-1-\mu^{F(\mathbb{Q})}}{(2\lambda)^{-1}} - \frac{\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)}-1}{\kappa^{(\mathbb{Q})}(T-t)}\left(v_t - \theta^{(\mathbb{Q})} - \frac{\lambda\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)}.$$

4.3.3 Variance Swaps

Note that, in the SVCJ model, changing the order of integration for the first jump term of the quadratic variation of the log price yields

$$\begin{aligned} \langle x\rangle_t &= \theta t - \tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa t}-1\right)(v_0-\theta) - \tfrac{\sigma}{\kappa}\int_0^t \left(\mathrm{e}^{-\kappa(t-s)}-1\right)\sqrt{v_{s-}}dW_s^v \\ &\quad -\tfrac{1}{\kappa}\int_0^t \left(\mathrm{e}^{-\kappa(t-s)}-1\right)Z_s^v dN_s + \int_0^t \left(Z_u^F\right)^2 dN_u. \end{aligned}$$

Then, provided that the jump distribution is unconditional for all t, we have

$$\begin{aligned} \mathbb{E}\left[\left.\langle x\rangle_T\right|\mathcal{F}_t\right] &= \theta T - \tfrac{1}{\kappa}\left(\mathrm{e}^{-\kappa T}-1\right)(v_0-\theta) - \tfrac{\sigma}{\kappa}\int_0^t \left(\mathrm{e}^{-\kappa(T-s)}-1\right)\sqrt{v_{s-}}dW_s^v \\ &\quad -\tfrac{1}{\kappa}\int_0^t \left(\mathrm{e}^{-\kappa(t-s)}-1\right)Z_s^v dN_s + \int_0^t \left(Z_u^F\right)^2 dN_u \\ &\quad -\tfrac{\lambda\mu_v}{\kappa^2}\left(\mathrm{e}^{\kappa(T-t)}-1\right) + \lambda\left(\mu_F^2+\sigma_F^2+\tfrac{\mu_v}{\kappa}\right)(T-t), \end{aligned}$$

and therefore the value process $V_{tT}^{(S)} := \mathbb{E}^{\mathbb{Q}}\left[\langle x\rangle_{T}\middle|\mathcal{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[\langle x\rangle_{T}\middle|\mathcal{F}_0\right]$ of an idealised standard variance swap yields

$$\begin{aligned}
V_{tT}^{(S)} &= -\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\int_0^t\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-s)}-1\right)\sqrt{v_{s-}}dW_s^{v(\mathbb{Q})}\\
&\quad+\tfrac{1}{\kappa^{(\mathbb{Q})}}\int_0^t\left(1-\mathrm{e}^{-\kappa^{(\mathbb{Q})}(t-s)}\right)Z_s^{v(\mathbb{Q})}dN_s+\int_0^t\left(Z_u^{F(\mathbb{Q})}\right)^2dN_u\\
&\quad-\tfrac{\lambda\mu^{v(\mathbb{Q})}}{\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{\kappa^{(\mathbb{Q})}T}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}t}-1\right)-\lambda\left[\left(\mu^{F(\mathbb{Q})}\right)^2+\left(\sigma^{F(\mathbb{Q})}\right)^2+\tfrac{\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right]t.
\end{aligned}$$

Note that the first jump compensator term in the third line is unbounded for $T\to\infty$ and therefore the price process of a perpetual variance swap is not well defined.

In the presence of jumps an idealised variance swap based on the realised leg as defined by Neuberger [2012] differs from the idealised standard variance swap. We have the payoff

$$\begin{aligned}
\int_0^T 2\left(\mathrm{e}^{dx_t}-dx_t-1\right) &= \int_0^T 2\mathrm{e}^{dx_{t^-}}\mathrm{e}^{\Delta x_t}-2\left(x_{T}-x_{0}\right)-2T\\
&= \int_0^T 2\mathrm{e}^{dx_{t^-}}+\int_0^T 2\left(\mathrm{e}^{Z_t^F}-1\right)dN_t-2\left(x_{T}-x_{0}\right)-2T\\
&= \int_0^T 2\left(1+\sqrt{v_{t-}}dW_t^F-\psi dt\right)+\int_0^T 2\left(\mathrm{e}^{Z_t^F}-1\right)dN_t\\
&\quad-2\left(x_{T}-x_{0}\right)-2T\\
&= \int_0^T 2\sqrt{v_{t-}}dW_t^F-2\psi T+\int_0^T 2\left(\mathrm{e}^{Z_t^F}-1\right)dN_t\\
&\quad-2\left(x_{T}-x_{0}\right),
\end{aligned}$$

where $dx_{t^-} = \sqrt{v_{t^-}}dW_t^F - \left(\psi + \frac{1}{2}v_{t^-}\right)dt$ and $\mathrm{e}^{dx_{t^-}} = 1 + \sqrt{v_{t^-}}dW_t^F - \psi dt$. Then

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\int_0^T 2\left(\mathrm{e}^{dx_t} - dx_t - 1\right)\middle|\mathcal{F}_t\right] &= \int_0^t 2\sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} - 2\psi T \\
&\quad + \int_0^t 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right)dN_t - 2\left(X_t - x_0\right) \\
&\quad + 2\lambda\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right](T-t) \\
&= \int_0^t 2\sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} - 2\psi T \\
&\quad + \int_0^t 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right)dN_t - 2\left(X_t - x_0\right) \\
&\quad + 2\lambda\left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - 1\right)(T-t),
\end{aligned}
$$

and therefore the value process

$$
V_{tT}^{(N)} := \mathbb{E}^{\mathbb{Q}}\left[\int_0^T 2\left(\mathrm{e}^{dx_t} - dx_t - 1\right)\middle|\mathcal{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[\int_0^T 2\left(\mathrm{e}^{dx_t} - dx_t - 1\right)\middle|\mathcal{F}_0\right]
$$

of Neuberger's idealised variance swap yields

$$
\begin{aligned}
V_{tT}^{(N)} &= \int_0^t 2\sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} - 2\left(X_t - X_0\right) \\
&\quad + \int_0^t 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right)dN_t - 2\lambda\left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - 1\right)t,
\end{aligned}
$$

with dynamics

$$
\begin{aligned}
dV_{tT}^{(N)} &= 2\sqrt{v_{t^-}}dW_t^{F(\mathbb{Q})} + 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - 1\right)dN_t - 2\lambda\left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - 1\right)dt - 2dX_t \\
&= 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - Z_t^{F(\mathbb{Q})} - 1\right)dN_t - 2\lambda\left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - \mu^{F(\mathbb{Q})} - 1\right)dt \\
&\quad - \tfrac{1}{\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-t)} - 1\right)\left(\sqrt{v_{t^-}}\sigma dW_t^{v(\mathbb{Q})} + Z_t^{v(\mathbb{Q})}dN_t - \lambda\mu^{v(\mathbb{Q})}dt\right).
\end{aligned}
$$

Like the standard variance swap, the dynamics of Neuberger's variance swap also depends on jumps in the underlying price process. However, only the part of the dynamics that relates to the variance process depends on the contract maturity. Taking the limit as $T \to \infty$ yields the dynamics of Neuberger's perpetual idealised variance swap, i.e.

$$\begin{aligned} dV_{t\infty}^{(N)} &= 2\left(\mathrm{e}^{Z_t^{F(\mathbb{Q})}} - Z_t^{F(\mathbb{Q})} - 1\right) dN_t - 2\lambda \left(\mathrm{e}^{\mu^{F(\mathbb{Q})} + \frac{1}{2}\left(\sigma^{F(\mathbb{Q})}\right)^2} - \mu^{F(\mathbb{Q})} - 1\right) dt \\ &\quad + \tfrac{1}{\kappa^{(\mathbb{Q})}} \left(\sqrt{v_{t-}}\sigma dW_t^{v(\mathbb{Q})} + Z_t^{v(\mathbb{Q})} dN_t - \lambda \mu^{v(\mathbb{Q})} dt\right), \end{aligned}$$

and in contrast with the equivalent calculation for the standard variance swap this limit is well defined. Two new features are evident compared to the pure diffusion case: in addition to diffusive changes in variance, Neuberger's perpetual idealised variance swap reacts to jumps in the price (first line) as well as to jumps in the variance process (last two terms of the second line).

By contrast, the value process of the continuously monitored discretisation-invariant variance swap yields $V_{tT}^{(DI)} := \mathbb{E}^{\mathbb{Q}}\left[\left.\langle X\rangle_T \right| \mathcal{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[\left.\langle X\rangle_T \right| \mathcal{F}_0\right]$, which again implies $V_{0T}^{(DI)} = 0$. Using the dynamics and quadratic variation of the log contract in the SVCJ model, and by extending the corresponding calculations for the Heston model, we have

$$
\begin{aligned}
\mathbb{E}\left[\left.\langle X\rangle_{T}\right|\mathcal{F}_t\right] &= \left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathbb{E}\left[\left.\int_0^T v_u du\right|\mathcal{F}_t\right] \\
&\quad+\left(\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}-\tfrac{\sigma^2}{2\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\mathbb{E}\left[\left.\int_0^T \mathrm{e}^{\kappa^{(\mathbb{Q})}u}v_u du\right|\mathcal{F}_t\right] \\
&\quad+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\mathbb{E}\left[\left.\int_0^T \mathrm{e}^{2\kappa^{(\mathbb{Q})}u}v_u du\right|\mathcal{F}_t\right] \\
&\quad+\mathbb{E}\left[\left.\int_0^T\left(Z_u^{F(\mathbb{Q})}+\tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)}-1\right)Z_u^{v(\mathbb{Q})}\right)^2 dN_u\right|\mathcal{F}_t\right],
\end{aligned}
$$

and therefore

$$
\begin{aligned}
V_{tT}^{(DI)} &= -\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\int_0^t\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-s)}-1\right)\sqrt{v_s}dW_s^{v(\mathbb{Q})} \\
&\quad+\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(\rho\sigma-\tfrac{\sigma^2}{2\kappa^{(\mathbb{Q})}}\right)\mathrm{e}^{-\kappa^{(\mathbb{Q})}T}\int_0^t(t-s)\mathrm{e}^{\kappa^{(\mathbb{Q})}s}\sqrt{v_s}dW_s^{v(\mathbb{Q})} \\
&\quad+\tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\mathrm{e}^{-2\kappa^{(\mathbb{Q})}T}\int_0^t\left(\mathrm{e}^{\kappa^{(\mathbb{Q})}(t+s)}-\mathrm{e}^{2\kappa^{(\mathbb{Q})}s}\right)\sqrt{v_s}dW_s^{v(\mathbb{Q})} \\
&\quad+\int_0^t\left(Z_u^{F(\mathbb{Q})}+\tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)}-1\right)Z_u^{v(\mathbb{Q})}\right)^2 dN_u \\
&\quad-\lambda\mathbb{E}\left[\left.\int_0^t\left(Z_u^{F(\mathbb{Q})}+\tfrac{1}{2\kappa^{(\mathbb{Q})}}\left(\mathrm{e}^{-\kappa^{(\mathbb{Q})}(T-u)}-1\right)Z_u^{v(\mathbb{Q})}\right)^2 du\right|\mathcal{F}_t\right].
\end{aligned}
$$

The perpetual variance swap that pays the quadratic variation of the log-contract for a fixed swap rate up to infinity ($T\to\infty$) follows the price process

$$
\begin{aligned}
V_{t\infty}^{(DI)} &= \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\int_0^t\sqrt{v_s}dW_s^{v(\mathbb{Q})}+\int_0^t\left(Z_u^{F(\mathbb{Q})}-\tfrac{Z_u^{v(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}}\right)^2 dN_u \\
&\quad-\lambda\left(\left[\left(\sigma^{F(\mathbb{Q})}\right)^2+\left(\mu^{F(\mathbb{Q})}\right)^2\right]-\tfrac{\mu^{F(\mathbb{Q})}\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}+\tfrac{1}{2}\left(\tfrac{\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)^2\right)t,
\end{aligned}
$$

with dynamics

$$
\begin{aligned}
dV_{t\infty}^{(DI)} &= \tfrac{\sigma}{\kappa^{(\mathbb{Q})}}\left(1-\tfrac{\rho\sigma}{\kappa^{(\mathbb{Q})}}+\tfrac{\sigma^2}{4\left(\kappa^{(\mathbb{Q})}\right)^2}\right)\sqrt{v_t}dW_t^{v(\mathbb{Q})}+\left(Z_t^{F(\mathbb{Q})}-\tfrac{Z_t^{v(\mathbb{Q})}}{2\kappa^{(\mathbb{Q})}}\right)^2 dN_t \\
&\quad -\lambda\left(\left[\left(\sigma^{F(\mathbb{Q})}\right)^2+\left(\mu^{F(\mathbb{Q})}\right)^2\right]-\tfrac{\mu^{F(\mathbb{Q})}\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}+\tfrac{1}{2}\left(\tfrac{\mu^{v(\mathbb{Q})}}{\kappa^{(\mathbb{Q})}}\right)^2\right)dt.
\end{aligned}
$$

Again this limit is well defined, as in the case of Neuberger's variance swap, and again the swap reacts to both jumps in the underlying and in the variance process. However, the impact of jumps is different due to the alternative definition of realised variance.

The explicit representations for variance swap dynamics derived above may be useful for model calibration. In particular, it may be possible to estimate the parameters of a Heston or SVCJ style asset pricing model from empirically observable risk premia on discretisation-invariant (DI) moment swap contracts.

Chapter 5

Conclusions and Outlook

Fair-value rates for conventional variance swaps are biased due to discretisation, jump and truncation errors. As a result market rates can deviate substantially from their fair values. The possibility for arbitrage opportunities and the concomitant market uncertainties have been a catalyst for considerable research on finding arbitrage bounds for these errors. A more recent strand of research concerns different definitions for the realised variance for which more precise fair values may be obtained; our research develops this second strand to derive a general theory for variance, higher-moment and other so-called 'discretisation-invariant' (DI) characteristics for which exact fair values are derived in a model-free setting.

Assuming only that the forward price follows a martingale we have followed the lead set out in the concluding remarks of Neuberger [2012] to define a whole vector space of DI characteristics. Theorem 1 allows us to find all characteristics which have this property, by solving a second order system of partial differential equations, for any set of deterministic functions of a multivariate martingale process. Theorem 2 focusses on a particular sub-class of these swaps, i.e. those for which the characteristic depends only on a multivariate martingale itself, and its logarithm. In this case we have found analytic solutions that can be used to define a rich variety of DI characteristics. Theorem 3 shows how the value of these swaps can be replicated by dynamically rebalancing portfolios of the underlying and certain fundamental contracts and Theorem 4 considers some special DI swaps which correspond to second, third and higher-order moments of a single log-return distribution.

Model-free DI variance swaps have several advantages over conventional variance swaps: (i) there is no jump or other model dependence error in their theoretical fair-value swap rate; consequently (ii) issuers would face smaller residual hedging risks; and (iii) the absence of arbitrage should yield market prices that are within the bid-ask spread of the fair-value; (iv) unbiased estimates for the variance risk premium (VRP) can be derived from fair values rather than market quotes; and (v) issuers would have greater flexibility to choose the monitoring frequency of the realised leg because the fair-value swap rate is the same for all frequencies – the monitoring does not even need to be regular. All these advantages apply to higher-order moment risk premia also.

The calculation of the fair-value for a model-free DI variance or higher-moment swap is still subject to a computation error because the replication theorem requires numerical integration over option prices at traded strikes to approximate an integral formula. However, a sub-space of DI swaps can be defined for which even this error is zero. These 'strike-discretisation-invariant' (SDI) swaps have characteristics defined by bi-linear forms of traded call and put prices. Again, an infinite variety of such SDI swaps exists and we have only investigated so-called 'straddle swaps' empirically. Their fair-value rates are simply (minus) the product of the prices of one put and one call of the same strike.

Our empirical analysis, spanning an 18-year sample period, demonstrates that a diverse variety of risk premia are available to trade via these swaps. By contrast with the realised skewness swap introduced by Neuberger [2012], and later analysed empirically by Kozhan et al. [2013], we find higher-moment risk premia that are not necessarily highly correlated with the VRP, in particular when they are monitored and sampled at the weekly or daily frequencies. However, the correlation between the skewness and kurtosis risk premia remains very large and negative, even when monitored and sampled daily. We conclude that the skew risk premium reflects asymmetry in the tails of the S&P 500 distribution, rather than asymmetry around the centre. The empirical dependence of risk premia on monitoring frequency, as well as on the maturity of the swap, motivates monthly-for-daily 'frequency swaps' and 180-for-30 day calendar swaps. Being based on realised and implied term structures respectively, these swaps can yield large pay-offs when based on skewness or kurtosis.

We have extended the results of Carr and Wu [2009] on the determinants of the VRP in three ways: (1) we replicate their main finding for a longer and more recent time period, namely that the excess return on the market is the only really significant equity-factor determinant of the S&P 500 variance premium; (2) when monitored on a daily basis we show that the VRP exhibits a highly significant asymmetric response to the market factor, especially during the year surrounding the financial crisis (July 2008 – June 2009); (3) we find that the market (and the squared market) factor is also the major driver of the skewness and kurtosis premia. However, despite their very high correlation, the market factor has much lower explanatory power for kurtosis than for the skewness. The market-only asymmetric factor model for the skew premium has an R^2 of almost 70% during the financial crisis period, which is higher than for the VRP. Over the entire period the factor-model R^2 remain high except for the kurtosis premium, which is even less than 30%. We conclude that largely unexplained factors are driving this. During 2012 and 2013 the kurtosis premium was exceptionally variable, yet the variance premium remained small and almost as stable as it was during the credit-boom years in the mid 2000's.

Some novel sources of risk become tradable via the creative use of these new swaps and they should be attractive to investors seeking new sources of diversification. Furthermore, the lack of error in the pricing formulas for DI swaps, plus the exact dynamic hedging portfolios that can be used to replicate them, considerably reduce the uncertainties faced by their issuers.

We hope that the general concepts and specific results presented in this paper will lay the foundations for a profitable agenda of research on new and diverse sources of risk which become tradable via the creative use of DI characteristics. Theoretical and empirical examples for interesting bivariate swaps, such as swaps on realised joint characteristics of S&P 500 and VIX futures, could open up a new strand of research on correlation and covariance swaps. More generally, we could investigate moments of univariate and multivariate distributions based on factors such as equity, bond and commodity index futures, and the addition of foreign exchange rates might lay the ground for new types of currency-protected products. We could extend results on the VRP by Ammann and Buesser [2013], Bakshi et al. [2008], Tian [2011] and Trolle and Schwartz [2010] to higher-moments and joint moments, e.g. using a covariance swap. The construction of multi-asset swaps as well as their replication using single-asset and spread options are discussed in Carr and Corso [2001]. Based on the explicit representations of dynamics and prices for discretisation-invariant (DI) swap contracts, it may be possible to develop new calibration procedures for asset pricing models that yield comparably stable parameter estimates.

Further empirical work would be interesting on straddle and other SDI swaps, and on the frequency and calendar swaps which trade on the term structures of the realised and implied characteristics, respectively. Oomen [2006] analyses the optimal monitoring frequency given market microstructure noise. The results of this paper could be improved by using a model-free DI variance characteristic. Empirical work on swaps that are monitored at irregular frequencies might include deriving a VRP from a realised characteristic that is monitored in transaction time. Such a swap could be monitored whenever cumulative trading in the underlying reaches a pre-defined level. If the S&P 500 'transaction time' VRP is small and negative, but less prone to brief periods of extremely high values at the onset of a crisis, then banks would take much less risk in paying these rather than standard realised variance. Investors would still have the incentive to receive that premium as a source of diversification, assuming it has a high negative correlation with returns on the S&P 500. In fact, Ané and Geman [2000] discuss swaps where the rebalancing frequency is driven by a 'transaction clock', showing that under an adequate change of time asset returns can be assumed to be normal. One challenge for our framework is that we assume a deterministic partition for monitoring. Under the assumption that the partition and the underlying asset are independent, a generalisation to a stochastic partition is feasible. However, transaction volume and asset price are negatively correlated in practice.

Finally, it would be interesting to construct optimal portfolios which diversify variance risk through skew or kurtosis swaps. In the S&P 500 index we know that both skew and kurtosis premia have quite low correlation with the VRP, but only when monitored at relatively high frequency (daily or weekly). So this research would be interesting for hedge funds and other investors with relatively short-term horizons.

Chapter 6

Appendix

6.1 Itô Formula for Jump Diffusions

The classical Itô formula for diffusions can be generalised to jump diffusions. Assume that the process X has a finite number of jumps ΔX on a finite interval and that the process behaves like a pure diffusion between jumps. In integral form we have

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t f_t(s, X_{s^-})\, ds + \int_0^t f_x(s, X_{s^-})\, dX_{s^-} \\ &\quad + \frac{1}{2}\int_0^t f_{xx}(s, X_{s^-})\, d\langle X\rangle_{s^-} + \sum_{0<s\le t}^{\Delta f\neq 0} \Delta f(s, X_s) \\ &= f(0, X_0) + \int_0^t f_t(s, X_s)\, ds + \int_0^t f_x(s, X_{s^-})\, dX_s \\ &\quad + \frac{1}{2}\int_0^t f_{xx}(s, X_{s^-})\, d\langle X\rangle_{s^-} + \sum_{0<s\le t}^{\Delta f\neq 0} \left[\Delta f(s, X_s) - f_x(s, X_{s^-})\, \Delta X_s\right] \end{aligned}$$

where X_{t^-} denotes the value of X prior to any jump at time t, dX_{t^-} and $d\langle X\rangle_{t^-}$ denote the continuous parts of the dynamics and instantaneous quadratic variation, respectively, and Δf denotes the jump in f that follows on a jump in X. In differential form we can write

$$\begin{aligned} df(t, X_t) &= f_t(t, X_{t^-})\, dt + f_x(t, X_{t^-})\, dX_{t^-} + \tfrac{1}{2} f_{xx}(t, X_{t^-})\, d\langle X\rangle_{t^-} + \Delta f(t, X_t) \\ &= f_t(t, X_t)\, dt + f_x(t, X_{t^-})\, dX_t + \tfrac{1}{2} f_{xx}(t, X_{t^-})\, d\langle X\rangle_{t^-} \\ &\quad + \left[\Delta f(t, X_t) - f_x(t, X_{t^-})\, \Delta X_t\right]. \end{aligned}$$

We shall further use the following shorthand notation:

$\cdot$	dt	dW_t	dN_t
dt	0	0	0
dW_t	0	dt	0
dN_t	0	0	dN_t

While the quadratic variation of the Wiener process is deterministic, the quadratic variation of the Poisson process is unpredictable. Diffusions and jumps are orthogonal.

6.2 Girsanov Change of Measure

Let $W^{(\mathbb{Q})} := \left\{W_t^{(\mathbb{Q})}\right\}_{t\in\mathbf{\Pi}}$ with $\mathbf{\Pi} := [0, T > 0]$ be a $\mathbb{Q}$-Brownian Motion and consider the pricing kernel $m := \{m_t\}_{t\in\mathbf{\Pi}}$ defined by

$$m_t^{-1} := \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} := \exp\left\{\int_0^t \lambda_s dW_s^{(\mathbb{Q})} - \tfrac{1}{2}\int_0^t \lambda_s^2 ds\right\},$$

where $\lambda := \{\lambda_t\}_{t\in\mathbf{\Pi}}$ satisfies the Novikov condition $\mathbb{E}^{\mathbb{Q}}\left[\exp\left\{\frac{1}{2}\int_0^t \lambda_s^2 ds\right\}\right] < \infty$ and therefore $\mathbb{E}^{\mathbb{Q}}\left[m_t^{-1}\right] = 1$ for all $t \in \mathbf{\Pi}$. Then the stochastic process $W^{(\mathbb{P})} := \left\{W_t^{(\mathbb{P})}\right\}_{t\in\mathbf{\Pi}}$ defined by $W_t^{(\mathbb{P})} := W_t^{(\mathbb{Q})} - \int_0^t \lambda_s ds$ with dynamics $dW_t^{(\mathbb{P})} = dW_t^{(\mathbb{Q})} - \lambda_t dt$ is a $\mathbb{P}$-Brownian Motion. Further

$$m_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{-\int_0^t \lambda_s dW_s^{(\mathbb{P})} - \tfrac{1}{2}\int_0^t \lambda_s^2 ds\right\},$$

where $\mathbb{E}^{\mathbb{P}}[m_t] = 1$ provided that $\mathbb{E}^{\mathbb{P}}\left[\exp\left\{\frac{1}{2}\int_0^t \lambda_s^2 ds\right\}\right] < \infty$, as e.g. in Carr and Wu [2009]. For technical details regarding Novikov's condition see e.g. Ruf [2013].

Proof: $W^{(\mathbb{P})}$ is a $\mathbb{P}$-Brownian Motion since the paths are continuous by construction and

1. $W_0^{(\mathbb{P})} = W_0^{(\mathbb{Q})} = 0$
2. $dW_t^{(\mathbb{P})} dW_t^{(\mathbb{P})} = dW_t^{(\mathbb{Q})} dW_t^{(\mathbb{Q})} = dt$
3. $dW_s^{(\mathbb{P})} dW_t^{(\mathbb{P})} = dW_s^{(\mathbb{Q})} dW_t^{(\mathbb{Q})} = 0$ for $s \neq t$
4. $W_t^{(\mathbb{P})} m_t^{-1}$ follows a $\mathbb{Q}$-martingale.

Property (4) holds since $d\left(m_t^{-1}\right) = m_t^{-1}\left(\lambda_t dW_t^{(\mathbb{Q})} - \frac{1}{2}\lambda_t^2 dt + \frac{1}{2}\lambda_t^2 dt\right) = m_t^{-1}\lambda_t dW_t^{(\mathbb{Q})}$ and

$$\begin{aligned} d\left(W_t^{(\mathbb{P})} m_t^{-1}\right) &= W_t^{(\mathbb{P})} d\left(m_t^{-1}\right) + m_t^{-1} dW_t^{(\mathbb{P})} + dW_t^{(\mathbb{P})} d\left(m_t^{-1}\right) \\ &= W_t^{(\mathbb{P})} m_t^{-1}\lambda_t dW_t^{(\mathbb{Q})} + m_t^{-1}\left(dW_t^{(\mathbb{Q})} - \lambda_t dt\right) + m_t^{-1}\lambda_t dt \\ &= m_t^{-1}\left(W_t^{(\mathbb{P})}\lambda_t + 1\right) dW_t^{(\mathbb{Q})}. \end{aligned}$$

Further $dm_t = m_t\left(-\lambda_t dW_t^{(\mathbb{P})} - \frac{1}{2}\lambda_t^2 dt + \frac{1}{2}\lambda_t^2 dt\right) = -m_t\lambda_t dW_t^{(\mathbb{P})}$ s.t. $\mathbb{E}^{\mathbb{P}}\left[m_t\right] = 1$.

6.3 Multivariate Change of Measure

Let $\mathbf{w}^{(\mathbb{Q})} := \left\{\mathbf{w}_t^{(\mathbb{Q})}\right\}_{t\in\mathbf{\Pi}}$ with $\mathbf{\Pi} := [0, T > 0]$ be an n-dimensional $\mathbb{Q}$-Brownian Motion with the invertible (instantaneous) correlation matrix $\mathbf{\Sigma}_t \in \mathbb{R}^{n\times n}$, i.e. $d\left\langle \mathbf{w}_t^{(\mathbb{Q})}\right\rangle =: \mathbf{\Sigma}_t dt$, and consider the pricing kernel $m := \{m_t\}_{t\in\mathbf{\Pi}}$ defined by

$$m_t^{-1} := \left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} := \exp\left\{\int_0^t \boldsymbol{\lambda}_s' d\mathbf{w}_s^{(\mathbb{Q})} - \tfrac{1}{2}\int_0^t \boldsymbol{\lambda}_s'\mathbf{\Sigma}_s\boldsymbol{\lambda}_s ds\right\},$$

where $\boldsymbol{\lambda} := \{\boldsymbol{\lambda}_t\}_{t\in\mathbf{\Pi}} \in \mathbb{R}^n$ satisfies the generalised Novikov condition

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left\{\tfrac{1}{2}\int_0^t \boldsymbol{\lambda}_s'\mathbf{\Sigma}_s\boldsymbol{\lambda}_s ds\right\}\right] < \infty$$

s.t. $\mathbb{E}^{\mathbb{Q}}\left[m_t^{-1}\right] = 1$ for all $t \in \mathbf{\Pi}$. Then the stochastic process $\mathbf{w}^{(\mathbb{P})} := \left\{\mathbf{w}_t^{(\mathbb{P})}\right\}_{t\in\mathbf{\Pi}}$ defined by

$$\mathbf{w}_t^{(\mathbb{P})} := \mathbf{w}_t^{(\mathbb{Q})} - \int_0^t \mathbf{\Sigma}_s\boldsymbol{\lambda}_s ds,$$

follows the dynamics

$$d\mathbf{w}_t^{(\mathbb{P})} := d\mathbf{w}_t^{(\mathbb{Q})} - \mathbf{\Sigma}_t \boldsymbol{\lambda}_t dt$$

and is a multivariate $\mathbb{P}$-Brownian Motion with the same instantaneous correlation $\mathbf{\Sigma}_t$, i.e. $d\left\langle \mathbf{w}_t^{(\mathbb{P})} \right\rangle = \mathbf{\Sigma}_t dt$. Further the pricing kernel yields

$$m_t = \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \exp\left\{ -\int_0^t \boldsymbol{\lambda}_s' d\mathbf{w}_s^{(\mathbb{P})} - \tfrac{1}{2} \int_0^t \boldsymbol{\lambda}_s' \mathbf{\Sigma}_s \boldsymbol{\lambda}_s ds \right\},$$

and $\mathbb{E}^{\mathbb{P}}[m_t] = 1$ provided that

$$\mathbb{E}^{\mathbb{P}}\left[\exp\left\{ \tfrac{1}{2} \int_0^t \boldsymbol{\lambda}_s' \mathbf{\Sigma}_s \boldsymbol{\lambda}_s ds \right\} \right] < \infty.$$

Proof: The stochastic process $\mathbf{w}^{(\mathbb{P})}$ is a multivariate $\mathbb{P}$-Brownian Motion with correlation matrix $\mathbf{\Sigma}_t$ since the paths are continuous by construction and

1. $\mathbf{w}_0^{(\mathbb{P})} = \mathbf{w}_0^{(\mathbb{Q})} = 0$
2. $d\mathbf{w}_t^{(\mathbb{P})} d\mathbf{w}_t^{(\mathbb{P})\prime} = d\mathbf{w}_t^{(\mathbb{Q})} d\mathbf{w}_t^{(\mathbb{Q})\prime} = \mathbf{\Sigma}_t dt$
3. $d\mathbf{w}_s^{(\mathbb{P})} d\mathbf{w}_t^{(\mathbb{P})\prime} = d\mathbf{w}_s^{(\mathbb{Q})} d\mathbf{w}_t^{(\mathbb{Q})\prime} = \mathbf{0}$ for $s \neq t$
4. $\mathbf{w}_t^{(\mathbb{P})} m_t^{-1}$ follows a multivariate $\mathbb{Q}$-martingale.

Property (4) holds since

$$d\left(m_t^{-1}\right) = m_t^{-1}\left(\boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{Q})} - \tfrac{1}{2} \boldsymbol{\lambda}_t' \mathbf{\Sigma}_t \boldsymbol{\lambda}_t dt + \tfrac{1}{2} \boldsymbol{\lambda}_t' \mathbf{\Sigma}_t \boldsymbol{\lambda}_t dt \right) = m_t^{-1} \boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{Q})},$$

which is uniformly integrable and therefore defines an exponential martingale, and

$$\begin{aligned}
d\left(\mathbf{w}_t^{(\mathbb{P})} m_t^{-1}\right) &= \mathbf{w}_t^{(\mathbb{P})} d\left(m_t^{-1}\right) + m_t^{-1} d\mathbf{w}_t^{(\mathbb{P})} + d\mathbf{w}_t^{(\mathbb{P})} d\left(m_t^{-1}\right) \\
&= \mathbf{w}_t^{(\mathbb{P})} m_t^{-1} \boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{Q})} + m_t^{-1}\left(d\mathbf{w}_t^{(\mathbb{Q})} - \boldsymbol{\Sigma}_t \boldsymbol{\lambda}_t dt\right) \\
&\quad + \left(d\mathbf{w}_t^{(\mathbb{Q})} - \boldsymbol{\Sigma}_t \boldsymbol{\lambda}_t dt\right) m_t^{-1} \boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{Q})} \\
&= m_t^{-1}\left(\mathbf{w}_t^{(\mathbb{P})} \boldsymbol{\lambda}_t' + \mathbf{I}\right) d\mathbf{w}_t^{(\mathbb{Q})} - m_t^{-1} \boldsymbol{\Sigma}_t \boldsymbol{\lambda}_t dt \\
&\quad + m_t^{-1} d\mathbf{w}_t^{(\mathbb{Q})} d\mathbf{w}_t^{(\mathbb{Q})\prime} \boldsymbol{\lambda}_t dt \\
&= m_t^{-1}\left(\mathbf{w}_t^{(\mathbb{P})} \boldsymbol{\lambda}_t' + \mathbf{I}\right) d\mathbf{w}_t^{(\mathbb{Q})}.
\end{aligned}$$

Further

$$dm_t = m_t\left(-\boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{P})} - \tfrac{1}{2}\boldsymbol{\lambda}_t' \boldsymbol{\Sigma}_t \boldsymbol{\lambda}_t dt + \tfrac{1}{2}\boldsymbol{\lambda}_t' \boldsymbol{\Sigma}_t \boldsymbol{\lambda}_t dt\right) = -m_t \boldsymbol{\lambda}_t' d\mathbf{w}_t^{(\mathbb{P})},$$

and thus $\mathbb{E}^{\mathbb{P}}\left[m_t\right] = 1$. For $n = 1$ we have $\boldsymbol{\Sigma} = 1$ as well as $\boldsymbol{\lambda} = \lambda$ and this theorem corresponds to the standard Girsanov change of measure in one dimension.

6.4 Replication Theorem

Carr and Madan [2001] show that any twice differentiable function $\phi : \mathbb{R}^+ \to \mathbb{R}$ can be expressed as

$$\phi(z) = \phi(k^*) + \phi'(k^*)(z - k^*) + \int_0^{k^*} \phi''(k)(k - z)^+ dk + \int_{k^*}^{\infty} \phi''(k)(z - k)^+ dk.$$

Setting $k^* = F_t$, which corresponds to the standard forward-at-the-money separation strike, as well as $z = F_T$ yields

$$\phi(F_T) = \phi(F_t) + \phi'(F_t)(F_T - F_t) + \int_0^{F_t} \phi''(k)(k - F_T)^+ dk + \int_{F_t}^{\infty} \phi''(k)(F_T - k)^+ dk,$$

and taking the conditional expectation at time t implies

$$\mathbb{E}_t^{\mathbb{Q}}[\phi(F_T)] = \phi(F_t) + \int_{\mathbb{R}^+} \phi''(k) q_t(k) dk.$$

6.5 Aggregation of Moments

According to the central limit theorem (CLT), both the sum and average of a large number of i.i.d. random variables are approximately normally distributed. This behaviour holds irrespective of the distributional properties of a single random variable. In the following we analyse the convergence rates of the variance, third and fourth moment as well as of the skewness and excess kurtosis of a single random variable to the corresponding parameters of the normal distribution for the sum and average as the number of random variables increases.

Let F_t be the forward price of a financial asset and denote by $X_t := \mathbb{E}_t[x_T]$ the price of the log contract, which follows a martingale by construction, with value increments $\hat{X}_i := X_{t_i} - X_{t_{i-1}}$ along a partition $\mathbf{\Pi}_N := \{t_0 := 0, \ldots, t_N := T\}$. The study of the distribution of value increments in the log contract is interesting since this derivative merges the crucial features of forward prices and log returns: its price process follows a martingale under the risk-neutral measure and is approximately normally distributed. Under the assumption that a continuum of option strikes is tradable the log contract is tradable.

Assume that $\hat{X}_i$ are i.i.d. with moment generating function $\hat{\chi}(\xi) := \mathbb{E}\left[e^{\xi \hat{X}}\right]$ and consider the four central moments

$$
\begin{aligned}
\hat{\mu} &:= \left.\tfrac{d}{d\xi}\hat{\chi}(\xi)\right|_{\xi=0} := \left.\hat{\chi}'(\xi)\right|_{\xi=0} = 0,\\
\hat{\sigma}^2 &:= \left.\left(\tfrac{d}{d\xi}\right)^2 \hat{\chi}(\xi)\right|_{\xi=0} := \left.\hat{\chi}^{(2)}(\xi)\right|_{\xi=0},\\
\hat{\mu}_3 &:= \left.\left(\tfrac{d}{d\xi}\right)^3 \hat{\chi}(\xi)\right|_{\xi=0} := \left.\hat{\chi}^{(3)}(\xi)\right|_{\xi=0},\\
\hat{\mu}_4 &:= \left.\left(\tfrac{d}{d\xi}\right)^4 \hat{\chi}(\xi)\right|_{\xi=0} := \left.\hat{\chi}^{(4)}(\xi)\right|_{\xi=0},
\end{aligned}
$$

as well as $\hat{\tau} := \hat{\mu}_3/\hat{\sigma}^3$ and $\hat{\kappa} := \hat{\mu}_4/\hat{\sigma}^4 - 3$. We now consider the total value increment $X_T - X_0 = \sum_{i=1}^N \hat{X}_i$ with moment generating function $\chi(\xi) := \mathbb{E}\left[e^{\xi(X_T - X_0)}\right]$. The i.i.d. assumption implies

$$
\chi(\xi) = \mathbb{E}\left[e^{\xi(X_T - X_0)}\right] = \mathbb{E}\left[e^{\xi \sum_{i=1}^N \hat{X}_i}\right] = \prod_{i=1}^N \mathbb{E}\left[e^{\xi \hat{X}_i}\right] = \prod_{i=1}^N \hat{\chi}(\xi) = \hat{\chi}(\xi)^N,
$$

and therefore, using that $\hat{\chi}(0) = 1$, we have

$$
\begin{aligned}
\mu &:= \left.\tfrac{d}{d\xi}\chi(\xi)\right|_{\xi=0} = \left.\tfrac{d}{d\xi}\hat{\chi}(\xi)^N\right|_{\xi=0} = \left. N\hat{\chi}(\xi)^{N-1}\hat{\chi}'(\xi)\right|_{\xi=0} = 0,\\
\sigma^2 &:= \left.\left(\tfrac{d}{d\xi}\right)^2 \chi(\xi)\right|_{\xi=0} = \left.\left(\tfrac{d}{d\xi}\right)^2 \hat{\chi}(\xi)^N\right|_{\xi=0} = \left.\tfrac{d}{d\xi}\left(N\hat{\chi}(\xi)^{N-1}\hat{\chi}'(\xi)\right)\right|_{\xi=0}\\
&= \left[N\hat{\chi}(\xi)^{N-1}\hat{\chi}^{(2)}(\xi) + N(N-1)\hat{\chi}(\xi)^{N-2}\hat{\chi}'(\xi)^2\right]_{\xi=0}\\
&= N\hat{\sigma}^2 + N(N-1)\hat{\mu}^2 = N\hat{\sigma}^2,
\end{aligned}
$$

$$
\begin{aligned}
\mu_3 &:= \left(\tfrac{d}{d\xi}\right)^3 \chi(\xi)\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^3 \hat{\chi}(\xi)^N\Big|_{\xi=0} = \tfrac{d}{d\xi}\left(\tfrac{d}{d\xi}\right)^2 \hat{\chi}(\xi)^N\Big|_{\xi=0} \\
&= \tfrac{d}{d\xi}\left(N\hat{\chi}(\xi)^{N-1}\hat{\chi}^{(2)}(\xi) + N(N-1)\hat{\chi}(\xi)^{N-2}\hat{\chi}'(\xi)^2\right)\Big|_{\xi=0} \\
&= \Big[N\hat{\chi}(\xi)^{N-1}\hat{\chi}^{(3)}(\xi) + N(N-1)\hat{\chi}(\xi)^{N-2}\hat{\chi}^{(2)}(\xi)\hat{\chi}'(\xi) \\
&\quad +2N(N-1)\hat{\chi}(\xi)^{N-2}\hat{\chi}'(\xi)\hat{\chi}^{(2)}(\xi) \\
&\quad +N(N-1)(N-2)\hat{\chi}(\xi)^{N-3}\hat{\chi}'(\xi)^3\Big]_{\xi=0} \\
&= N\hat{\mu}_3 + 3N(N-1)\hat{\sigma}^2\hat{\mu} + N(N-1)(N-2)\hat{\mu}^3 = N\hat{\mu}_3, \\
\mu_4 &:= \left(\tfrac{d}{d\xi}\right)^4 \chi(\xi)\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^4 \hat{\chi}(\xi)^N\Big|_{\xi=0} = \tfrac{d}{d\xi}\left(\tfrac{d}{d\xi}\right)^3 \hat{\chi}(\xi)^N\Big|_{\xi=0} \\
&= \Big[N\hat{\chi}(\xi)^{N-1}\hat{\chi}^{(4)}(\xi) + N(N-1)\hat{\chi}(\xi)^{N-2}\hat{\chi}^{(3)}(\xi)\hat{\chi}'(\xi) \\
&\quad +3N(N-1)\Big(\hat{\chi}(\xi)^{N-2}\left\{\hat{\chi}^{(2)}(\xi)^2 + \hat{\chi}^{(3)}(\xi)\hat{\chi}'(\xi)\right\} \\
&\quad +(N-2)\hat{\chi}(\xi)^{N-3}\hat{\chi}^{(2)}(\xi)\hat{\chi}'(\xi)^2\Big) \\
&\quad +3N(N-1)(N-2)\hat{\chi}(\xi)^{N-3}\hat{\chi}'(\xi)^2\hat{\chi}^{(2)}(\xi) \\
&\quad +N(N-1)(N-2)\hat{\chi}(\xi)^{N-4}\hat{\chi}'(\xi)^4\Big]_{\xi=0} \\
&= N\hat{\mu}_4 + N(N-1)\hat{\mu}_3\hat{\mu} + 3N(N-1)\left\{\hat{\sigma}^4 + \hat{\mu}_3\hat{\mu} + (N-2)\hat{\sigma}^2\hat{\mu}^2\right\} \\
&\quad +N(N-1)(N-2)\left(3\hat{\sigma}^2\hat{\mu}^2 + \hat{\mu}^4\right) \\
&= N\hat{\mu}_4 + 3N(N-1)\hat{\sigma}^4.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\tau &:= \mu_3/\sigma^3 = N\hat{\mu}_3\sqrt{N\hat{\sigma}^2}^{-3} = \hat{\mu}_3/\hat{\sigma}^3\sqrt{N}^{-1} = \hat{\tau}\sqrt{N}^{-1}, \\
\kappa &:= \mu_4/\sigma^4 - 3 = \left(N\hat{\mu}_4 + 3N(N-1)\hat{\sigma}^4\right)/\left(N\hat{\sigma}^2\right)^2 - 3 = \hat{\kappa}N^{-1}.
\end{aligned}
$$

When short period price changes in the log contract are independent and non-normally distributed with skewness $\bar{\tau}$ and excess kurtosis $\bar{\kappa}$, the skewness and excess kurtosis of long period price changes converge to zero as $\sqrt{N}^{-1}$ and N^{-1}, respectively.

We further consider the average increment $\left(X_T - X_0\right)/N = N^{-1}\sum_{i=1}^{N}\hat{X}_i$ with moment generating function $\bar{\chi}\left(\xi\right) := \mathbb{E}\left[\mathrm{e}^{\xi\left(X_T - X_0\right)/N}\right]$. The i.i.d. assumption implies

$$\bar{\chi}\left(\xi\right) = \mathbb{E}\left[\mathrm{e}^{\xi N^{-1}\sum_{i=1}^{N}\hat{X}_i}\right] = \prod_{i=1}^{N}\mathbb{E}\left[\mathrm{e}^{\xi N^{-1}\hat{X}_i}\right] = \prod_{i=1}^{N}\hat{\chi}\left(\xi N^{-1}\right) = \hat{\chi}\left(\xi N^{-1}\right)^N$$

and therefore, using that $d/d\left(\xi N^{-1}\right) = Nd/d\xi$, we have

$$\begin{aligned}
\bar{\mu} &:= \tfrac{d}{d\xi}\bar{\chi}\left(\xi\right)\Big|_{\xi=0} = \tfrac{d}{d\xi}\hat{\chi}\left(\xi N^{-1}\right)^N\Big|_{\xi=0} = N^{-1}\,\tfrac{d}{d\xi}\hat{\chi}\left(\xi\right)^N\Big|_{\xi=0} = N^{-1}\mu = 0,\\
\bar{\sigma}^2 &:= \left(\tfrac{d}{d\xi}\right)^2\bar{\chi}\left(\xi\right)\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^2\hat{\chi}\left(\xi N^{-1}\right)^N\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^2\hat{\chi}\left(\xi\right)^N\Big|_{\xi=0}/N^2 = \hat{\sigma}^2/N,\\
\bar{\mu}_3 &:= \left(\tfrac{d}{d\xi}\right)^3\bar{\chi}\left(\xi\right)\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^3\hat{\chi}\left(\xi N^{-1}\right)^N\Big|_{\xi=0} = \left(\tfrac{d}{d\xi}\right)^3\hat{\chi}\left(\xi\right)^N\Big|_{\xi=0}/N^3 = \hat{\mu}_3/N^2,\\
\bar{\mu}_4 &:= \left(\tfrac{d}{d\xi}\right)^4\bar{\chi}\left(\xi\right)\Big|_{\xi=0} = N^{-4}\left(\tfrac{d}{d\xi}\right)^4\hat{\chi}\left(\xi\right)^N\Big|_{\xi=0} = N^{-3}\hat{\mu}_4 + 3N^{-3}(N-1)\hat{\sigma}^4.
\end{aligned}$$

Hence,

$$\begin{aligned}
\bar{\tau} &:= \bar{\mu}_3/\bar{\sigma}^3 = N^{-2}\hat{\mu}_3/\sqrt{N^{-1}\hat{\sigma}^2}^3 = \hat{\mu}_3/\hat{\sigma}^3\sqrt{N}^{-1} = \hat{\tau}\sqrt{N}^{-1} = \tau,\\
\bar{\kappa} &:= \bar{\mu}_4/\bar{\sigma}^4 - 3 = \left(N^{-3}\hat{\mu}_4 + 3N^{-3}(N-1)\hat{\sigma}^4\right)/\left(N^{-1}\hat{\sigma}^2\right)^2 - 3\\
&= \hat{\mu_4}/\hat{\sigma}^4N^{-1} - 3N^{-1} = \hat{\kappa}N^{-1} = \kappa.
\end{aligned}$$

That is, skewness and kurtosis are the same for the total and average increment while the non-standardised moments grow with N for the total increment and decrease with N^{-m+1} for the average increment, m being the order of the moment.

Bibliography

Y. Aït-Sahalia, M. Karaman, and L. Mancini. The term structure of variance swaps and risk premia. *Working Paper*, 2015.

C. Alexander. *Market Models: A Guide to Financial Data Analysis*. Wiley, 2001.

C. Alexander and S. Leontsinis. Model risk in variance swap rates. *Working Paper*, 2011.

M. Ammann and R. Buesser. Variance risk premiums in foreign exchange markets. *Journal of Empirical Finance*, 23:16–32, 2013.

L. B. G. Andersen and V. V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

T. Ané and H. Geman. Order flow, transaction clock, and normality of asset returns. *Journal of Finance*, 55(5):2259–2284, 2000.

G. Bakshi and N. Kapadia. Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies*, 16(2):527–566, 2003.

G Bakshi and D. Madan. Spanning and derivative-security valuation. *Journal of Financial Economics*, 55:205–238, 2000.

G. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52(5):2003–2049, 1997.

G. Bakshi, N. Kapadia, and D. Madan. Stock return characteristics, skew laws and the differential pricing of individual equity options. *Review of Financial Studies*, 16(1):101–143, 2003.

G. Bakshi, P. Carr, and L. Wu. Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. *Journal of Financial Economics*, 87(1):132–156, 2008.

C. Bardgett, E. Gourier, and M. Leippold. Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Working Paper*, 2015.

O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society*, B(63):167–241, 2001.

O. E. Barndorff-Nielsen and N. Shephard. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17:457–477, 2002.

O. E. Barndorff-Nielsen and A. E. D. Veraart. Stochastic volatility of volatility and variance risk premia. *Journal of Financial Econometrics*, 11(1):1–46, 2013.

A. Barnea and R. Hogan. Quantifying the variance risk premium in VIX options. *Journal of Portfolio Management*, 38(3):143–148, 2012.

D. S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, 9(1):69–107, 1996.

D. S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94:181–238, 2000.

F. E. Benth, M. Groth, and R. Kufakunesu. Valuing volatility and variance swaps for a non-Gaussian Ornstein-Uhlenbeck stochastic volatility model. *Applied Mathematical Finance*, 14(4):347–363, 2007.

C. Bernard and Z. Cui. Prices and asymptotics for discrete variance swaps. *Applied Mathematical Finance*, 21:140–173, 2014.

C. Bernard, Z. Cui, and D. Mcleish. Convergence of the discrete variance swap in time-homogeneous diffusion models. *Quantitative Finance Letters*, 2(1):1–6, 2014.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

T. Bollerslev and V. Todorov. Tails, fears, and risk premia. *Journal of Finance*, 66(6):2165–2211, 2011.

O. Bondarenko. Variance trading and market price of variance risk. *Journal of Econometrics*, 180:81–97, 2014.

D. T. Breeden and R. H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978.

D. Bregantini. Moment-based estimation of stochastic volatility. *Journal of Banking & Finance*, 37:4755–4764, 2013.

L.M. Bregman. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200–217, 1967.

M. Britten-Jones and A. Neuberger. Option prices, implied price processes, and stochastic volatility. *Journal of Finance*, 55(2):839–866, 2000.

M. Broadie and A. Jain. Pricing and hedging volatility derivatives. *Journal of Derivatives*, pages 7–24, 2008a.

M. Broadie and A. Jain. The effect of jumps and discrete sampling on volatility and variance swaps. *International Journal of Theoretical and Applied Finance*, 11(8):761–979, 2008b.

M. Broadie, M. Chernov, and M. Johannes. Model specification and risk premia: Evidence from futures options. *Journal of Finance*, 62(3):7–24, 2007.

M. M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997.

P. Carr and A. Corso. Covariance contracting for commodities. *EPRM*, pages 42–45, 2001.

P. Carr and R. Lee. Volatility derivatives. *Annual Review of Financial Economics*, 1:1–21, 2009.

P. Carr and R. Lee. Variation and share-weighted variation swaps on time-changed Lévy processes. *Finance and Stochastics*, 17(4):685–716, 2013.

P. Carr and D. B. Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 1(1):19–37, 2001.

P. Carr and W. Schoutens. Hedging under the Heston model with jump-to-default. *International Journal of Theoretical and Applied Finance*, 11(4):403–414, 2008.

P. Carr and L. Wu. The finite moment log stable process and option pricing. *Journal of Finance*, 58(2):753–777, 2003.

P. Carr and L. Wu. Variance risk premiums. *Review of Financial Studies*, 22(3): 1311–1341, 2009.

P. Carr, H. Geman, D. B. Madan, and M. Yor. Pricing options on realized variance. *Finance and Stochastics*, 9(4):453–475, 2005.

P. Carr, H. Geman, D. B. Madan, and M. Yor. Options on realized variance and convex orders. *Quantitative Finance*, 11(11):1685–1694, 2011.

F. Chabi-Yo. Pricing kernels with stochastic skewness and volatility risk. *Management Science*, 58(3):624–640, 2012.

L. Chan and E. Platen. Pricing and hedging of long dated variance swaps under a 3/2 volatility model. *Journal of Computational and Applied Mathematics*, 278: 181–196, 2015.

P. Cheridito, D. Filipović, and R. L. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83: 123–170, 2007.

M. Chernov and E. Ghysels. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics*, 56(3):407–458, 2000.

Chicago Board Options Exchange. The CBOE volatility index - VIX. *White Paper*, 2009.

G. H. Choe and K. Lee. High moment variations and their application. *Journal of Futures Markets*, 34(11):1040–1061, 2014.

P. Christoffersen, S. Heston, and K. Jacobs. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, 55(12):1914–1932, 2009.

P. Christoffersen, K. Jacobs, and C. Ornthanalai. Dynamic jump intensities and risk premiums: Evidence from S&P 500 returns and options. *Journal of Financial Economics*, 106:447–472, 2012.

R. Cont and P. Tankov. *Financial Modeling with Jump Processes*. Chapman & Hall, first edition, 2004.

J. C. Cox and S. A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.

J. C. Cox, J. E. Ingersoll Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978, 2000.

M. Davis, J. Obłój, and V. Raval. Arbitrage bounds for weighted variance swap prices. *Mathematical Finance*, 24(4):821–854, 2014.

K. Demeterfi, E. Derman, M. Kamal, and J. Zou. A guide to volatility and variance swaps. *Journal of Derivatives*, 6(4):9–32, 1999.

P. Dennis and S. Mayhew. Risk-neutral skewness: Evidence from stock options. *Journal of Financial and Quantitative Analysis*, 37(3):471–493, 2002.

E. Derman and I. Kani. Riding on a smile. *RISK Magazine*, 7(2):32–39, 1994.

K. Detlefsen and W. K. Haerdle. Variance swap dynamics. *Quantitative Finance*, 13(5):675–685, 2013.

G. Drimus. Options on realized variance by transform methods: A non-affine stochastic volatility model. *Quantitative Finance*, 12(11):1679–1694, 2012.

J.-C. Duan and W. Zhang. Forward-looking market risk premium. *Management Science*, 60(2):521–538, 2014.

G. R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 52(1):405–443, 2002.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

J.-M. Dufour, R. Garcia, and A. Taamouti. Measuring high-frequency causality between returns, realized volatility, and implied volatility. *Journal of Financial Econometrics*, 10(1):124–163, 2012.

B. Dupire. Pricing with a smile. *RISK Magazine*, 7, 1994.

D. Egloff, M. Leippold, and L. Wu. The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis*, 45(5):1279–1310, 2010.

R. J. Elliott and G.-H. Lian. Pricing variance and volatility swaps in a stochastic volatility model with regime switching: Discrete observations case. *Quantitative Finance*, 13(5):687–698, 2013.

R. J. Elliott, T. K. Siu, and L. Chan. Pricing volatility swaps under Heston's stochastic volatility model with regime switching. *Applied Mathematical Finance*, 14(1):41–62, 2007.

D. C. Emanuel and J. D. MacBeth. Further results on the constant elasticity of variance call option pricing model. *Journal of Financial and Quantitative Analysis*, 17(4):533–554, 1982.

R. F. Engle. Long-term skewness and systemic risk. *Journal of Financial Economics*, 9(3):437–468, 2011.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

M. R. Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.

B. Feunou, J.-S. Fontaine, A. Taamouti, and R. Tédongap. Risk premium, variance premium, and the maturity structure of uncertainty. *Review of Finance*, 18(1):219–269, 2014.

D. Filipović, E. Gourier, and L. Mancini. Quadratic variance swap models. *Journal of Financial Economics*, 119(1):44–68, 2016.

D. Galai. A proposal for indexes for traded call options. *Journal of Finance*, 34 (5):1157–1172, 1979.

J. Gatheral. *The Volatility Surface*. Wiley Finance, 2006.

J. Goard. A time-dependent variance model for pricing variance and volatility swaps. *Applied Mathematical Finance*, 18(1):51–70, 2011.

F. Guillaume and W. Schoutens. Heston model: The variance swap calibration. *Journal of Optimization Theory and Applications*, 161(1):76–89, 2014.

D. Guo. The risk premium of volatility implicit in currency options. *Journal of Business and Economic Statistics*, 16(4):498–507, 1998.

R. Hafner and M. Wallmeier. Volatility as an asset class: European evidence. *European Journal of Finance*, 13(7):621–644, 2007.

J. M. Harrison and D. M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2): 327–343, 1993.

D. Hobson and M. Klimmek. Model independent hedging strategies for variance swaps. *Finance and Stochastics*, 16:611–649, 2012.

S. D. H. Hsu and B. M. Murray. On the volatility of volatility. *Physica A: Statistical Mechanics and its Applications*, 380(1-2):366–376, 2007.

J. C. Hull. *Options, Futures and Other Derivatives*. Prentice Hall, seventh edition, 2009.

A. Itkin and P. Carr. Pricing swaps and options on quadratic variation under stochastic time change models: Discrete observations case. *Review of Derivatives Research*, 13(2):141–176, 2010.

J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2003.

R. Jarrow, Y. Kchia, M. Larsson, and P. Protter. Discretely sampled variance and volatility swaps versus their continuous approximations. *Finance and Stochastics*, 17:305–324, 2013.

A. Javaheri, P. Wilmott, and E. G. Haug. GARCH and volatility swaps. *Quantitative Finance*, 4(5):589–595, 2004.

G. J. Jiang and Y. S. Tian. The model-free implied volatility and its information content. *Review of Financial Studies*, 18(4):1305–1342, 2005.

R. Jordan and C. Tier. The variance swap contract under the CEV process. *International Journal of Theoretical and Applied Finance*, 12(5):709–743, 2009.

J. Kallsen, J. Muhle-Karbe, and M. Voss. Pricing options on variance in affine stochastic volatility models. *Mathematical Finance*, 21(4):627–641, 2011.

M. Keller-Ressel and J. Muhle-Karbe. Asymptotic and exact pricing of options on variance. *Finance and Stochastics*, 17:107–133, 2013.

E. Konstantinidi and G. Skiadopoulos. How does the market variance risk premium vary over time? Evidence from S&P 500 variance swap investment returns. *Working Paper*, 2014.

R. Kozhan, A. Neuberger, and P. Schneider. The skew risk premium in the equity index market. *Review of Financial Studies*, 26(9):2174–2203, 2013.

A. Kraus and R. H. Litzenberger. Skewness preferences and the valuation of risk assets. *Journal of Finance*, 31(4):1085–1100, 1976.

G. Le and L. Yang. Impact of truncation on model-free implied moment estimator. *Working Paper*, 2015.

R. Lee. Gamma swap. In R. Cont, editor, *Encyclopedia of Quantitative Finance, IV Volume Set*. Wiley, 2010a.

R. Lee. Weighted variance swap. In R. Cont, editor, *Encyclopedia of Quantitative Finance, IV Volume Set*. Wiley, 2010b.

I. Martin. Simple variance swaps. *Working Paper*, 2013.

R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

A. Neuberger. The log contract. A new instrument to hedge volatility. *Journal of Portfolio Management*, 20(2):74–80, 1994.

A. Neuberger. Realized skewness. *Review of Financial Studies*, 25(11):3423–3455, 2012.

M. Neumann and G. Skiadopoulos. Predictable dynamics in higher-order risk-neutral moments: Evidence from the S&P 500 options. *Journal of Financial and Quantitative Analysis*, 48(3):947–977, 2013.

B. Nieto, A. Novales, and G. Rubio. Variance swaps, non-normality and macroeconomic and financial risks. *Quarterly Review of Economics and Finance*, 54: 257–270, 2014.

B. K. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2003.

R. C. A. Oomen. Properties of realized variance under alternative sampling schemes. *Journal of Business & Economic Statistics*, 24(2):219–237, 2006.

L. S. Rompolis and E. Tzavalis. Retrieving risk neutral moments and expected quadratic variation from option prices. *Working Paper*, 2013.

J. Ruf. A new proof for the conditions of Novikov and Kazamaki. *Stochastic Processes and their Applications*, 123(2):404–421, 2013.

S. Rujivan and S.-P. Zhu. A simple closed-form formula for pricing discretely-sampled variance swaps under the Heston model. *ANZIAM Journal*, 56(1): 1–27, 2014.

P. Schneider and F. Trojani. Divergence and the price of uncertainty. *Swiss Finance Institute Research Paper*, 15(60), 2015a.

P. Schneider and F. Trojani. Fear trading. *Swiss Finance Institute Research Paper*, 15(3), 2015b.

W. Schoutens. Moment swaps. *Quantitative Finance*, 5(6):525–530, 2005.

L. O. Scott. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates. Applications of Fourier inversion methods. *Mathematical Finance*, 7(4):413–424, 1997.

S. E. Shreve. *Stochastic Calculus for Finance II: Continuous Time Models*. Springer Finance, 2004.

A. Swishchuk and L. Xu. Pricing variance swaps for stochastic volatilities with delay and jumps. *International Journal of Stochastic Analysis*, 2011.

Y. S. Tian. Extracting risk-neutral density and its moments from American option prices. *Journal of Derivatives*, 18(3):17–34, 2011.

V. Todorov. Variance risk premium dynamics. *Review of Financial Studies*, 23 (1):345–383, 2010.

L. Torricelli. Pricing joint claims on an asset and its realized variance in stochastic volatility models. *International Journal of Theoretical and Applied Finance*, 16 (1), 2013.

A. B. Trolle and E. S. Schwartz. Variance risk premia in energy commodities. *Journal of Derivatives*, 17(3):15–32, 2010.

H. Wang, J. G. O'Hara, and N. Constantinou. A path-independent approach to integrated variance under the CEV model. *Mathematics and Computers in Simulation*, 109:130–152, 2015.

L. Wu. Dampened power law: Reconciling the tail behavior of financial security returns. *Journal of Business*, 79(3):1445–1473, 2006.

S.-P. Zhu and G.-H. Lian. On the valuation of variance swaps with stochastic volatility. *Applied Mathematics and Computation*, 219(4):1654–1669, 2012.

S.-P. Zhu and G.-H. Lian. Pricing forward-start variance swaps with stochastic volatility. *Applied Mathematics and Computation*, 250:920–933, 2015.