Innovations in Dependence Modelling for Financial Applications

Matthew Ames

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London

Department of Statistical Science
University College London

April 24, 2017

Declaration of Authorship

I, Matthew Ames, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Acknowledgements

First and foremost I would like to thank my supervisor, Gareth Peters, for his much appreciated support and guidance over the past four years. I am extremely grateful for the bountiful research opportunities he has presented me with and for the inspiration that his infectious enthusiasm for the subject brings. I am very much enjoying our interesting research collaborations together and I am looking forward to the fruitful years ahead.

I would also like to thank my second supervisor, Guillaume Bagnarosa, for his expert advice and the ongoing collaboration we enjoy. His insights into the currency carry trade strategy and commodity modelling in practice have no doubt made this thesis much stronger. Moreover, our research benefits tremendously from his positive energy and Gallic flair!

At UCL, I would like to thank my third supervisor, Ioannis Kosmidis, with whom I enjoyed an interesting collaboration on the topic of mixture and rotated copula models.

In addition to thanking my supervisors, I would like to express my gratitude to Pavel Shevchenko for his expert guidance and collaboration, particularly during my six-month visit to CSIRO in Sydney.

Furthermore, I would like to thank Tomoko Matsui at the Institute of Statistical Mathematics for her ongoing collaboration and indeed for her remarkable hospitality during my summer research visit and my current position in Tokyo.

Finally, I would like to thank my family for all their love and support over the years.

Publications

Journal Papers

Published

1. "Upside and Downside Risk Exposures of Currency Carry Trades via Tail Dependence."
   Matthew Ames, Gareth W. Peters, Guillaume Bagnarosa and Ioannis Kosmidis.
   Innovations in Quantitative Risk Management Springer Proceedings in Mathematics & Statistics Volume 99, 2015, pp 163-181
   [http://link.springer.com/chapter/10.1007/978-3-319-09114-3_10]

Accepted

1. "Violations of Uncovered Interest Rate Parity and International Exchange Rate Dependences."
   Matthew Ames, Guillaume Bagnarosa and Gareth W. Peters.
   Journal of International Money and Finance
   [http://www.sciencedirect.com/science/article/pii/S0261560617300025]

2. "Forecasting Covariance for Optimal Carry Trade Portfolio Allocations."
   Matthew Ames, Guillaume Bagnarosa, Gareth W. Peters and Pavel Shevchenko.
   International Conference on Acoustics, Speech, and Signal Processing 2017
   [ssrn.com/abstract=2711586]

Submitted

1. "Understanding the Interplay between Covariance Forecasting Factor Models and Risk Based Portfolio Allocations in Currency Carry Trades."
   Matthew Ames, Guillaume Bagnarosa, Gareth W. Peters and Pavel Shevchenko.

Journal of Forecasting: Forecasting Financial Markets Special Issue
[ssrn.com/abstract=2699020]

In Preparation

1. "Determining the Influence of Macroeconomic Factors on the Oil Price Term Structure in the Short, Medium and Long Term."
Matthew Ames, Gareth W. Peters, Guillaume Bagnarosa, Pavel Shevchenko and Tomoko Matsui.

Research Presentations

1. 2016 International Workshop on Spatial and Temporal Modeling from Statistical, Machine Learning and Engineering perspectives (STM2016),
   20 July 2016 - 23 July 2016
   Institute of Statistical Mathematics, Tokyo, Japan

2. Quantitative Methods in Finance Conference 2015,
   17th December 2015
   Sydney, Australia

3. University of Sydney,
   Business School Seminar,
   4th December 2015
   Sydney, Australia

4. University of New South Wales,
   School of Risk and Actuarial Studies Seminar,
   30th October 2015
   Sydney, Australia

5. University of Technology Sydney,
   School of Mathematical and Physical Sciences Seminar,
   28th October 2015
   Sydney, Australia

6. Macquarie University,
   Department of Statistics Seminar,
   27th October 2015
   Sydney, Australia

7. Commonwealth Scientific and Industrial Research Organisation Seminar,
   14th September 2015
   CSIRO, Sydney, Australia

8. 2015 International Workshop on Spatial and Temporal Modeling from Statistical, Machine Learning and Engineering perspectives (STM2015) and

Workshop on Complex Systems Modeling and Estimation Challenges in Big Data (CSM2015),
13th July 2015 - 17th July 2015
Institute of Statistical Mathematics, Tokyo, Japan

9. 22nd International Forecasting Financial Markets Conference,
20th May - 22nd May 2015
University of Rennes, France

10. 8th International Conference on Computational and Financial Econometrics,
6th - 8th December 2014
University of Pisa, Italy

11. MLSS Machine Learning Summer School,
April 25th - May 4th 2014
Reykjavik University, Iceland

12. 7th International Conference on Computational and Financial Econometrics,
14th - 16th December 2013
Senate House, University of London

13. Risk Management Reloaded Conference,
9th - 13th September 2013
Business Campus München, Technische Universität München

14. IMA Conference on Mathematics in Finance,
8th - 9th April 2013
Edinburgh Conference Centre, Heriot-Watt University

15. Mathematics of Financial Risk Management Workshop,
Poster Presentation,
28th March 2013
Isaac Newton Institute for Mathematical Sciences, Cambridge

International Visits

1. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Sydney, Australia
   July 2015 - January 2016
   Australian Endeavour Research Fellowship Recipient

2. Institute of Statistical Mathematics, Tokyo, Japan
   July 2016 - August 2016

Abstract

The contribution of this thesis is in developing and investigating novel dependence modelling techniques in financial applications. Furthermore, the aim is to understand the key factors driving the dynamic nature of such dependence.

When modelling the multivariate distribution of the returns associated to a portfolio of financial assets one is faced with a multitude of considerations and potential choices. For example, in the currency studies undertaken in this thesis suitably heavy-tailed marginal time series models are developed for the returns of each currency exchange rate, and then the multivariate dependence structure of the returns of multiple-currency baskets at each time instant is considered. These dependence relationships can be studied via numerous concordance measures such as correlation, rank correlations and extremal dependences. Such studies can be undertaken in a static or dynamic setting and either parametrically or non-parametrically.

Another important aspect of financial time series is the enormous amount of financial data available for statistical analysis and financial econometrics that can be used to better understand economic and financial theories. In this thesis, the focus is on the influence of dependence structures in complex financial data in two asset classes: currencies and commodities. These are challenging data structures as they contain temporal serial dependence, cross dependence and term-structural dependences. Each of these forms of dependence are studied in this thesis in both parametric and non-parametric settings.

Statistical Modelling and Estimation Contributions

Three complementary dependence modelling approaches are developed in this thesis. The first two approaches address the challenge of modelling the multivariate distribution of a portfolio of asset returns. The third approach developed concerns commodity price dependence modelling where the link between maturities through the term structure of futures prices is considered. Firstly, a parametric copula modelling approach is considered in order to capture the complex dependence structure present in such data. In particular, flexible mixture copula models, consisting of weighted Archimedean copula members such as Clayton, Frank and Gumbel components, are developed including additional structural flexibility via distortion transforms corresponding to inner and outer-transform variants. These models are estimated via the inference for margins method which consists of a two step fitting procedure for the marginal model and then the dependence structure. In addition, an expectation-maximisation method is considered.

Secondly, a covariance factor regression framework is utilised in order to understand the influence of observed covariates on the covariance of the multivariate distribution of a portfolio of asset returns. This framework provides a number of desirable properties. Crucially, the model is interpretable in a way that GARCH-type models are not and as such, forecasting the covariance matrix is straightforward and transparent. This is achieved by constructing time series models for the observed covariates and calculating forecasts, which are then used as inputs to the covariance matrix forecast. Furthermore, the estimation of the covariance factor model can be performed using a simple and efficient Expectation-Maximization (EM) algorithm. A sensitivity analysis of the covariance matrix to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression.

The resulting forecasts of the covariance matrix of asset returns can then be utilised in portfolio optimisation. In particular, this modelling framework allows one to calculate the sensitivity of the portfolio weights to the observable covariance factors and accordingly helps to devise a global and dynamic hedging strategy for portfolios of assets. Thus, the relationship between interpretable factors and the weightings of assets in a portfolio can be further understood.

Thirdly, a novel Hybrid Multi-Factor (HMF) state-space modelling framework is proposed in order to understand the key factors driving the dependence structure among commodity futures prices along their term structure. A consistent estimation framework is developed, which builds on the familiar two-factor model of Schwartz and Smith (2000), to allow for an investigation of the influence of observable covariates on commodity prices. Using this novel Hybrid Multi-Factor (HMF) model, it is possible to obtain closed form futures prices under standard risk neutral pricing formulations. One can incorporate state-space model estimation techniques to consistently estimate both the structural features related to the convenience yield and spot price dynamics (long and short term stochastic dynamics) and also the structural parameters that relate to the influence on the spot price of the observed exogenous covariates. Such models can then be utilised to gain significant insight into the futures and spot price dynamics in terms of interpretable observed factors that influence speculators and hedgers heterogeneously. This is not attainable with existing modelling approaches.

The proposed HMF modelling framework reconciles two classes of model: the latent multi-factor stochastic differential equation (s.d.e.) models and the alternative class of observable regression econometric factor models, by doing so in a statistically consistent manner from interpretation and estimation perspectives. The novel class of stochastic HMF models developed in this thesis allows for incorporation of exogenous covariate structures in a statistically rigorous manner. Such models are a genuine combination of the two approaches and do not presume any prevalence from one approach or the other. The crux of the matter lies in building a state-space model which allows a one-stage estimation with simultaneous inference of the latent factors dynamic and the covariates coefficients in order to overcome the estimation error associated to the two-stage approach generally proposed in the literature. In such a two-stage model, typically the latent factor estimates are first extracted in order to later regress as a function of a set of covariates. This conditional estimation of the latent factor suffers from several flaws compared to the conditional estimates proposed in this thesis.

The HMF modelling framework also allows one to consider covariate forecasts in order to extrapolate values for the futures prices along the term structure while considering the confidence interval associated to this estimate. This is particularly convenient in risk management and commodity hedging as one needs to consider not only the amount to invest but also the uncertainty associated to this measurement.

Novel Insights into Finance and Econometric Studies

This thesis also contributes to the literature by the application of the dependence structure modelling techniques described above to two challenging financial modelling problems: modelling multiple-currency basket returns and modelling commodity futures price term structure.

In order to perform the empirical analyses considered in this thesis in a robust manner a substantial amount of effort and time was invested into collecting, cleaning and preparing the data.

Multiple Currency Basket Modelling

Firstly, this thesis investigates the well-known financial puzzle of the currency carry trade, which is yet to be satisfactorily explained. It is one of the most robust financial puzzles in international finance and has attracted the attention of academics and practitioners alike for the past 25 years. The currency carry trade is the investment strategy that involves selling low interest rate currencies in order to purchase higher interest rate currencies, thus profiting from the interest rate differentials. Assuming foreign exchange risk is uninhibited and the markets have rational risk-neutral investors, then one would not expect profits from such strategies. That is uncovered interest rate parity (UIP); the parity condition in which exposure to foreign exchange risk, with unanticipated changes in exchange rates, should result in an outcome that changes in the exchange rate should offset the potential to profit from such interest rate differentials.

The two primary assumptions required for interest rate parity are related to capital mobility and perfect substitutability of domestic and foreign assets. Given foreign exchange market equilibrium, the interest rate parity condition implies that the expected return on domestic assets will equal the exchange rate-adjusted expected return on foreign currency assets. However, it has been shown empirically, that investors can actually earn on average arbitrage profits by borrowing in a country with a lower risk free interest rate, exchanging for foreign currency, and investing in a foreign country with a higher risk free interest rate, whilst allowing for any losses (or gains) from exchanging back to their domestic currency at maturity. Therefore trading strategies that aim to exploit the interest rate differentials can be profitable on average.

This research comprises of a comprehensive review of the literature surrounding the forward premium puzzle, a mathematical background to copulas and a review of their various uses in the literature to model dependence, followed by an investigation of the forward premium puzzle via an analysis of the multivariate tail dependence in currency carry trades. A dataset of daily closes on spot and one month forward contracts for 20 currencies from 2000 to 2013 was used to investigate the behaviour of carry portfolios, formed by sorting on the forward premium (a proxy to the interest rate differential to US dollar). A rigorous statistical modelling approach is proposed, which captures the specific statistical features of both the individual currency log-return distributions as well as the joint features, such as the dependence structures prevailing between the exchange rates.

The individual currency returns were transformed to standard uniform margins after fitting appropriately heavy tailed marginal models, namely log-normal and log generalised gamma models. In order to analyse the tail dependence present in the carry portfolios: mixture copula models, consisting of weighted Clayton, Frank and Gumbel components, were fitted on a rolling daily basis to the previous six months of transformed log returns. Extracting and interpreting the multivariate tail dependence present in the rolling daily baskets provided significant evidence that the average excess returns earned from the carry trade strategy can be attributed to compensation for not only individual currency tail risk, but also exposure to significant risk of large portfolio losses due to joint adverse movements.

A key contribution of this thesis is therefore to provide a rationale for the unintuitive excess returns seen empirically in the currency carry trade via the presence of multivariate tail dependence and therefore increased portfolio crash risk. This is a novel and promising approach. A further contribution of this research is the identification of significant periods of carry portfolio construction and unwinding through the analysis of multivariate tail dependence in mixture copula models.

From a fundamental perspective this thesis also explores the impact of speculative trading behaviour on the dependence structure of currency returns. The ratio of speculative open interest (net non-commercial positions) to total open interest, termed the $SPEC$ factor, is shown to provide a good proxy to the behaviour of carry trade investors via a PCA analysis and consequently the resulting complex non linear relation between international exchange rates.

To investigate this phenomena, a covariance regression modelling approach whereby the influence of observed covariates on the covariance of the multivariate returns of a basket of assets is proposed. In particular, the impact of speculative trading behaviour, i.e. the $SPEC$ factors, on the covariance of carry currencies is investigated. These

$SPEC$ factors are shown to hold several orders of magnitude more explanatory power than the price index factors, DOL and HML_{FX}, previously suggested in the literature. Furthermore, it is demonstrated that the time series for the DOL and HML_{FX} factors are very close to white noise and as such are essentially unforecastable. The suggested speculative open interest factors are shown to be amenable to ARIMA model fits and so produce reasonable forecast accuracy.

Thus, time series models for these covariates of interest are built and hence forecasts of the covariance of a basket of currencies can be obtained. Therefore, the inherent heteroskedasticity of the covariance of a basket of currencies can be modelled and forecast whilst maintaining the desirable property of interpretability of the model. This forecasting ability is then useful for risk management, portfolio optimisation and trading strategy development.

A sensitivity analysis of the covariance to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. In addition, a regression of the tail dependence measures, obtained from the mixture copula modelling approach, on the $SPEC$ factors illustrates the influence of carry trade speculative behaviour on the extremal joint currency returns. The DOL and HML_{FX} are shown to hold little explanatory power in the joint tails.

Commodity Price Modelling

In addition, this thesis employs a state-space modelling approach to understand the joint dynamic of the commodity spot price and the related futures prices along the curve. This framework is extended to allow for an investigation of the influence of observed macroeconomical covariates on the commodity term structure and in particular whether these covariates affect the short or long end of the curve. This modelling can be used for risk management, derivatives pricing, real options analysis and (carry) strategy development, e.g. backwardation/contango plays.

In particular, in this thesis the focus is on the behaviour of oil prices. Oil has historically been one of the most closely scrutinized commodities in the market. First and foremost, this is because of the important role this commodity plays in the worldwide economy and international relations, which gives it a prominent role, when compared to other energy, agricultural and metals commodities, in many aspects of the global economy and each country's specific macro, micro and monetary economic policy decisions.

Historically, one has seen the importance that economies have placed on the price variation of oil and understanding the factors that affect such a dynamic in order to better understand the determinants of shocks and volatility regimes in the spot price, demand and supply.

Another determining reason for the continued interest lies in the frequent shocks affecting the supply and demand of the so called "black gold" giving birth to sudden and dramatic price movements, such as during the 1973/74 oil crisis. The price of this exhaustible commodity has indeed been in the past heavily impacted by the discovery of new fields or the conflicts in oil-producing countries. On the other hand, the demand behaviour has generally been more influenced by the business cycles or even the evolution of the extracted oil inventories. That being said, according to the US Department of the Interior (DOI) as well as the US Energy Information Administration (EIA), the technology used for its extraction has recently been the main factor influencing the market supply. Over the last decade, advances in the application of horizontal drilling and hydraulic fracturing in shale have indeed drastically modified the international supply and demand equilibrium as well as the existing international relations by allowing the biggest oil consumer, namely the United States, to become over the same time period less and less dependent on its energy imports. According to the EIA, in 2015, 24% of the petroleum consumed in this country was imported which corresponds to the lowest level since 1970.

From a modelling perspective, such changes in the physical market conditions are significantly impacting the commodity price dynamic and need to be incorporated into any interpretable and realistic commodity futures stochastic model. In addition, if the model is developed, as is the case with the class of Hybrid Multi-Factor (HMF) models introduced in this thesis, to allow for clear closed form representations of structural features such as sensitivity, shock transient response and perturbation influence on the model parameters and the driving exogenous covariates characterizing the features just discussed, then such a class of models has the potential to significantly aid in the study of stochastic variation in oil futures prices and to aid in forecasting and policy decision. The main aim of this research is to provide such a class of models and demonstrate their utility in incorporating a range of exogenous covariates into different structural components that will clearly explain short term and long term speculator and hedger positions in oil futures and their influences.

Finally, the results presented in this thesis shed light upon several topical challenges raised in the literature about the relation between crude oil term structure behaviour and financial or physical information available in the market. One can conclude that the recent increase of the US oil production over the last decade has significantly influenced the behaviour of the crude oil long term equilibrium price and also the dynamics of the futures term structure.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter presents an overview of the thesis. The motivation for researching the currency carry trade puzzle as well as the modelling of commodity prices is presented. Previous related work is then discussed. Subsequently, the research contributions of this thesis are detailed. Finally, the structure of the thesis is outlined.

1.1 Motivation

The main motivation of this thesis is to develop and investigate novel dependence modelling techniques in financial applications. In particular, the aim is to understand the key factors driving the dynamic nature of such dependence. This thesis focuses on two key financial applications: modelling multiple-currency basket returns and modelling commodity prices.

A key motivation of this thesis is to investigate the well-known forward premium puzzle and the associated currency carry trade. The currency carry trade is the investment strategy that involves selling low interest rate currencies in order to purchase higher interest rate currencies, thus profiting from the interest rate differentials. Assuming foreign exchange risk is uninhibited and the markets have rational risk-neutral investors, then one would not expect profits from such strategies. That is uncovered interest rate parity (UIP); the parity condition in which exposure to foreign exchange risk, with unanticipated changes in exchange rates, should result in an outcome that changes in the exchange rate should offset the potential to profit from such interest rate differentials. The two primary assumptions required for interest rate parity are related to capital mobility and perfect substitutability of domestic and foreign assets. Given foreign exchange market equilibrium, the interest rate parity condition implies that the expected return on domestic assets will equal the exchange rate-adjusted expected return on foreign currency assets. However, it has been shown empirically, that investors can actually earn arbitrage profits by borrowing in a country with a lower interest rate, exchanging for foreign currency, and investing in a foreign country with a higher interest rate, whilst allowing for any losses (or gains) from exchanging back to their domestic currency at maturity. Therefore trading strategies that aim to exploit the interest rate differentials can be profitable on average.

The intention of this thesis is therefore to reinterpret the currency carry trade puzzle in light of heavy tailed marginal models coupled with multivariate tail dependence features. To achieve this analysis of the multivariate extreme tail dependence several parametric models are developed and detailed model comparison is performed.

This research thus demonstrates that tail dependences among specific sets of currencies provide other justifications to the carry trade excess return and also allows one to detect construction and unwinding periods of such carry portfolios.

Furthermore, the impact of speculative trader behaviour on currency returns is investigated: in the mean, the covariance and the joint tails. I explore the question of whether this information can be utilised to improve the forecasting of the covariance and hence produce better portfolios.

A second key motivation of this thesis is to investigate and gain significant insight into commodity futures and spot price dynamics in terms of interpretable observed factors that influence speculators and hedgers heterogeneously. This is not attainable with existing modelling approaches. In particular, the HMF modelling framework proposed in this thesis reconciles two classes of model: the latent factor stochastic multi-factor s.d.e. models and the alternative class of observable regression econometric factor models, by doing so in a statistically consistent manner from interpretation and estimation perspectives.

Such models are a genuine combination of the two approaches and do not presume any prevalence from one approach or the other. The crux of the matter lies in building a model which allows a one-stage estimation with simultaneous inference of the latent factors dynamic and the covariates coefficients to overcome the estimation error associated to the two-stage approach generally proposed in the literature. In such a two-stage model, typically the latent factor estimates are first extracted in order to later regress as a function of a set of covariates. This conditional estimation of the latent factor suffers from several flaws compared to the conditional estimates proposed in this thesis.

The HMF modelling framework also allows one to consider covariate forecasts in order to extrapolate values for the futures prices while considering the confidence interval associated to this estimate. This is particularly convenient in risk management and commodity hedging as one needs to consider not only the amount to invest but also the uncertainty associated to this measurement.

1.2 Related Work

The currency carry trade is one of the most robust financial puzzles in international finance and has attracted the attention of academics and practitioners alike for the past 25 years. Numerous empirical studies Hansen and Hodrick [1980]; Fama [1984]; Engel [1996]; Lustig and Verdelhan [2007] have previously demonstrated the excess returns resulting from carry trade strategies.

Such a confounding puzzle has understandably resulted in a vast and varied literature, in which a number of theories have been proposed to justify the phenomenon.

Fama [1984] initially proposed a time varying risk premium within the forward rate relative to the associated spot rate - concluding that, under rational markets, most of the variation in forward rates was due to the variation in risk premium.

Weitzman [2007] demonstrates through a Bayesian approach that the uncertainty about the variance of the future growth rates combined with a thin-tailed prior distribution would generate the fat-tailed distribution required to solve the forward premium puzzle. This could be compared to the argument retained by Menkhoff et al. [2012a] who demonstrate that high interest rate currencies tend to be negatively related to the innovations in global FX volatility, which is considered as a proxy for unexpected changes in the FX market volatility. Menkhoff et al. [2012a] show that sorting currencies by their beta with global FX volatility innovations yields portfolios with large differences in returns, and also similar portfolios to those obtained when sorting by forward discount. Another risk factor shown to be significant, although to a much lesser degree, is liquidity risk.

Burnside et al. [2007] presents an alternative model to a pure risk factor model, in which "adverse selection problems between market makers and traders rationalizes a negative covariance between the forward premium and changes in exchange rates". Here, the authors suggest that the foreign exchange market should not be considered as a Walrasian market and that market makers face a worse adverse selection problem when an agent wants to trade against a public information signal, i.e. to place a contrarian bet as an informed trader.

Another hypothesis, proposed by Farhi and Gabaix [2008], consists of justifying this puzzle through the inclusion of a mean reverting risk premium. According to their model a risky country, which is more sensitive to economic extreme events, represents a high risk of currency depreciation and has thus to propose, in order to compensate this risk, a higher interest rate. Then, when the risk premium reverts to the mean, their exchange rate appreciates while they still have a high interest rate which thus replicates the forward rate premium puzzle.

The causality relation between the interest rate differential and the currency shocks can be presented the other way around as detailed in Brunnermeier and Pedersen [2009]. In this paper, the authors assume that the currency carry trade mechanically attracts investors and more specifically speculators who accordingly increase the probability of a market crash. Tail events among currencies would thus be caused by speculators' need to unwind their positions when they get closer to funding constraints.

This recurrent statement of a relation between tail events and forward rate premium Farhi and Gabaix [2008]; Brunnermeier et al. [2008] has led to the proposal in this thesis of a rigorous measure and estimation of the tail thickness at the level of the marginal distribution associated to each exchange rate. Moreover, the question of the link between the currency's marginal distribution and the associated interest rate differential leads to the consideration more globally of the joint dependence structures between the individual marginal cumulative distribution function (cdf) tails with respect to their respective interest rate differential.

While in the first two applications of this thesis I focus on the currency exchange rates market, in the third application of dependence I investigate another very interesting market which is also crucial for the contemporaneous economies, namely the oil market.

Although one can obtain a coarse picture of the principal fundamental events affecting oil price dynamics throughout history, the modelling and the choice of explanatory variables for oil price dynamics is still fiercely debated in the academic literature. Several reasons for this have been put forward, among which is the microeconomic interactions between very different types of agents who intervene in the market and who are generally classified into two distinct groups labelled respectively hedgers and speculators. The pre-eminent role they can play in the price discovery process of the market has raised unanswered questions about the causality relationship existing between the future prices and the physical or spot price observed in the real economy. As a matter of fact, several papers have demonstrated that not just the speculators but also the commodity-index funds were so influential in the market that the future price was actually leading the spot price and thus disconnecting the oil price from the fundamentals, such as those mentioned earlier (Kaufmann and Ullman [2009]; Silvrio and Szklo [2012]; Kilian and Murphy [2014]). Following this strand of the literature, certain authors (Bessembinder [1992]; Acharya et al. [2013]; Etula [2013]; Adrian et al. [2014]) considered the limits-to-arbitrage as one of the main reasons for the inverted price discovery process. Through such analyses they were able to argue that this demonstrated that any market friction limiting the arbitrage capacity of the financial intermediaries was translating into limits to hedging for the producers and accordingly impacting the real sphere participants' behaviour as well as related variables such as the spot oil prices.

The fact that macro and micro-economic observable variables influence the determination of market price dynamics by directly influencing the decisions and behaviour of speculators and hedgers in the market has naturally led to an alternative proposition from academics consisting of modelling the oil price dynamic through state space models where the log-price can be represented as a combination of several latent processes, which can then be generically interpreted without being necessarily related to any fundamental or microeconomical variables (Gibson and Schwartz [1990]; Schwartz and Smith [2000]; Casassus and Collin-Dufresne [2005]). Among advocates for this approach, authors notably decomposed the future prices as a combination of short term and long term latent components while others have assumed equivalently that the latent process should be associated to the convenience yield and thus determine the basis level or said differently the price difference between the spot and the future contract. Kaldor [1939] explains that the inter-temporal difference between futures and thus between the future price and the spot price are linked to the cost of storage and also the so-called convenience yield which embodies the benefits accrued to the owner of the physical commodity by providing him with a certain flexibility with regards to his reaction in case of market shocks. Schwartz and Smith [2000] demonstrated through a change of variable the linear equivalence between modelling the convenience yield or the dynamic of a long and a short term latent factor in order to model the futures price curve. Another advantage in considering these models resides in the ease of financial change of measure to risk neutral formulations that admit closed form analytical futures prices in terms of stochastic factors assumed to explain the spot price stochastic unobserved dynamics. From this systematic model differentiation between macro, micro and latent factors and given also the fact that the storage cost or the convenience yield are both naturally related to fundamental elements such as the storage capacity in the market, followed several articles dissecting the behaviour of the latent processes relative to a set of fundamental and microeconomic variables (Dempster et al. [2012], Daskalaki et al. [2014]). On the contrary other academics focused on demonstrating that the fundamental factors were not marginally contributing to the explanation provided by the futures prices themselves, and thus the latent processes (Cummins et al. [2016]; Daskalaki et al. [2014]).

1.3 Thesis Contributions

Three complementary dependence modelling approaches are developed in this thesis. The first two approaches address the challenge of modelling the multivariate distribution of a portfolio of asset returns. The third approach developed concerns commodity price dependence modelling where the link between maturities through the term structure of futures prices is considered.

The first approach adopted in this thesis is a statistical framework with a high degree of sophistication, however its fundamental reasoning and justification is indeed analogous in nature to the ideas considered when investigating the "equity risk premium puzzle" coined by Mehra and Prescott [1985] in the late 80's. The equity risk premium puzzle effectively refers to the fact that demand for government bonds which have lower returns than stocks still exists and generally remains high. This poses a puzzle for economists to explain why the magnitude of the disparity between the returns on each of these asset classes, stocks versus bonds, known as the equity risk premium, is so great and therefore implies an implausibly high level of investor risk aversion. In the seminal paper written by Rietz [1988], the author proposes to explain the "equity risk premium puzzle", Mehra and Prescott [1985], by taking into consideration the low but still significant probability of a joint catastrophic event.

Analogously in this thesis, an exploration is presented of the highly leveraged arbitrage opportunities in currency carry trades that arise due to violation of the UIP. However, it is conjectured that if the assessment of the risk associated with such trading strategies was modified to adequately take into account the potential for joint catastrophic risk events accounting for the non-trivial probabilities of joint adverse movements in currency exchange rates, then such strategies may not seem so profitable relative to the risk borne by the investor. A rigorous probabilistic model is proposed in order to quantify this phenomenon and potentially detect when liquidity in FX markets may dry up. This probabilistic measure of dependence can then be very useful for risk management of such portfolios but also for making more tractable the valuation of structured products or other derivatives indexed on this specific strategy. To be more specific, one of the principal contributions of this thesis is indeed to model the dependences between exchange rates using a flexible family of mixture copulae comprised of Archimedean members. This probabilistic approach allows the joint distribution of the vectors of random variables, in this case vectors of exchange rates log-returns in each basket of currencies, to be expressed as functions of each marginal distribution and the copula function itself.

Whereas in the literature mentioned earlier, the tail thickness resulting from the carry trade has been either treated individually for each exchange rate or through the measurement of distribution moments that may not be adapted to a proper estimation of the tail dependences. In this thesis, it is proposed instead to build, on a daily basis, a set of portfolios of currencies with regards to the interest rate differentials of each currency with the US dollar. Using a mixture of copula functions, a measure of the tail dependences within each portfolio is extracted and finally the results are interpreted. Among the outcomes of this study, it is demonstrated that during the crisis periods, the high interest rate currencies tend to display very significant upper tail dependence. Accordingly, it can thus be concluded that the appealing high return profile of a carry portfolio is not only compensating the tail thickness of each individual component probability distribution but also the fact that they tend to occur simultaneously and lead to a portfolio particularly sensitive to the risk of drawdown. Furthermore, it is also shown that high interest rate currency portfolios can display periods during which the tail dependence gets inverted demonstrating when periods of construction of the aforementioned carry positions are being undertaken by investors.

This thesis also explores the impact of speculative trading behaviour on the dependence structure of currency returns. The ratio of speculative open interest (net non-commercial positions) to total open interest, termed the $SPEC$ factor, is shown to provide a good proxy to the behaviour of carry trade investors via a PCA analysis. A covariance regression modelling approach whereby the influence of observed covariates on the covariance of the multivariate returns of a basket of assets is proposed. In particular, the impact of speculative trading behaviour, i.e. the $SPEC$ factors, on the covariance of carry currencies is investigated. These $SPEC$ factors are shown to hold several orders of magnitude more explanatory power than the price index factors, DOL and HML_{FX}, previously suggested in the literature. Furthermore, it is demonstrated that the time series for the DOL and HML_{FX} factors are very close to white noise and as such are essentially unforecastable. The suggested speculative open interest factors are shown to be amenable to ARIMA model fits and so produce reasonable forecast accuracy.

Thus, time series models for these covariates of interest are built and hence forecasts of the covariance of a basket of currencies can be obtained. Therefore, the inherent heteroskedasticity of the covariance of a basket of currencies can be modelled and forecast whilst maintaining the desirable property of interpretability of the model. This forecasting ability is then useful for risk management, portfolio optimisation and trading strategy development.

A sensitivity analysis of the covariance to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. In addition, a regression of the tail dependence measures, obtained from the mixture copula modelling approach, on the $SPEC$ factors illustrates the influence of carry trade speculative behaviour on the extremal joint currency returns. The DOL and HML_{FX} are shown to hold little explanatory power in the joint tails.

In this thesis, I also investigate financial time series dependence structure in commodity markets. The dynamic behaviour of the futures price term structure which combines time series and cross-sectional data has been modelled in this thesis using a so-called Hybrid Multi-Factor (HMF) model. This state-space modelling framework is proposed in order to understand the key factors driving commodity prices. A consistent estimation framework is developed, which builds on the familiar two-factor model of Schwartz and Smith (2000), to allow for an investigation of the influence of observable covariates on commodity prices. Using this novel Hybrid Multi-Factor (HMF) model, it is possible to obtain closed form futures prices under standard risk neutral pricing formulations, and one can incorporate state-space model estimation techniques to consistently estimate both the structural features related to the convenience yield and spot price dynamics (long and short term stochastic dynamics) and also the structural parameters that relate to the influence on the spot price and the futures price term structure of the observed exogenous covariates. Such models can then be utilised to gain significant insight into the futures and spot price dynamics in terms of interpretable observed factors that influence speculators and hedgers differently. This is not attainable with existing modelling approaches.

The proposed HMF modelling framework reconciles two classes of model: the latent multi-factor stochastic differential equation (s.d.e.) models and the alternative class of observable regression econometric factor models, by doing so in a statistically consistent manner from interpretation and estimation perspectives. The novel class of stochastic HMF models developed in this thesis allows for incorporation of exogenous covariate structures in a statistically rigorous manner. Such models are a genuine combination of the two approaches and do not presume any prevalence from one approach or the other. The crux of the matter lies in building a state-space model which allows a one-stage estimation with simultaneous inference of the latent factors dynamic and the covariates coefficients. In order to overcome the estimation error associated to the two-stage approach generally proposed in the literature. In such a two-stage model, typically the latent factor estimates are first extracted in order to later regress as a function of a set of covariates. This conditional estimation of the latent factor suffers from several flaws compared to the conditional estimates proposed in this thesis.

The HMF modelling framework also allows one to consider covariate forecasts in order to extrapolate values for the futures prices along the term structure while considering the confidence interval associated to this estimate. This is particularly convenient in risk management and commodity hedging as one needs to consider not only the amount to invest but also the uncertainty associated to this measurement.

1.4 Thesis Structure

This thesis is structured as follows: Part I introduces the copula modelling framework and its novel application to investigate asymmetric tail dependence in currency carry trade portfolios. Part II introduces the covariance regression framework and its novel application to investigate how observable and interpretable explanatory factors influence the covariance structure of currency returns. Part III introduces portfolio optimisation techniques and then utilises the novel covariance forecasting approach developed in Part II to investigate portfolio optimisation in currency carry portfolios. Finally, Part IV introduces a novel Hybrid Multi-Factor (HMF) stochastic differential equation (s.d.e.) framework to model the term structure dynamic of commodity futures prices.

Part I

Copula Modelling Contributions

Chapter 2

Part I Overview

In the first part of this thesis, the copula modelling framework and its novel application to investigate asymmetric tail dependence in currency carry trade portfolios is introduced.

Chapter 3 reviews the origins and mathematical background of copulae, before discussing the development of copula modelling in the fields of financial mathematics and insurance. Classical measures of dependence are detailed, followed by the concept of tail dependence. Then some key parametric statistical models that directly capture these dependence features are discussed.

The flexible copula modelling framework presented will then be utilised to investigate one of the most robust puzzles in finance, named the forward premium puzzle. This puzzle and the associated currency carry trade have received much attention over recent decades with many theories being proposed to explain the phenomenon. However, a complete and satisfactory explanation has proven illusive.

Chapter 4 presents the forward premium puzzle and then reviews the literature surrounding the puzzle and the associated currency carry trade. The novel approach of analysing both individual tail heaviness and joint tail dependence is proposed.

Chapter 5 presents the investigation of the forward premium puzzle using empirical data. The time-varying dependence structure of currency carry trade baskets is explored. In particular, the multivariate tail dependence characteristics of the baskets are analysed and the results discussed.

Chapter 3

Copula Modelling

In this chapter, the origins and mathematical background of copulae are reviewed, before discussing the development of copula modelling in the fields of financial mathematics and insurance. Classical measures of dependence are detailed, followed by the concept of tail dependence. Then some key parametric statistical models that directly capture these dependence features are discussed.

3.1 Origins

The explosion of interest in copula modelling over the past few decades can largely be attributed to their flexibility and usefulness in a wide range of practical applications, particularly in the world of finance and insurance, see Genest et al. [2009].

The first mathematical use of the word copula can be traced back to Abel Sklar's theorem in 1959, Sklar [1959], in which one-dimensional distribution functions are *joined together* by a copula function to form multivariate distribution functions. However, the roots of copula theory can in fact be traced back further to Hoeffding's work on 'standardised distributions' on the square $[-\frac{1}{2}, \frac{1}{2}]^2$ in the 1940's, Hoeffding [1994b,a]. A more detailed history of the origins and development of copula theory can be found in the introduction of the excellent monograph Nelsen [2006]. Personal recollections by the founders of the field can be found in Schweizer [1991] and Sklar [1996].

So, why are we interested in copulae? As Fisher notes in his article in the Encyclopedia of Statistical Sciences, Fisher [1997], "Copulas [are] of interest to statisticians for two main reasons:

1. as a way of studying scale-free measures of dependence.
2. as a starting point for constructing families of bivariate distributions, sometimes with a view to simulation."

The most natural place to begin this literature review is with Sklar's introduction of the copula function in his famous theorem, Sklar [1959].

A copula is specified according to the following definition.

Definition 1. ***Copula***
A d-dimensional copula is a multivariate cumulative distribution function C with uniform $[0,1]$ margins, i.e.

$$C : [0,1]^d \to [0,1], \tag{3–1}$$

with the following properties:

1. $C(1,\ldots,1,a_i,1,\ldots,1) = a_i$ *for every* $i \leq d$, $\forall a_i$ *in* $[0,1]$;
2. $C(a_1,\ldots,a_d) = 0$ *if* $a_i = 0$ *for* $i \in 1,\ldots,d$
3. C *is d-increasing.*

One of the main attractions for practitioners for the use of copula models is the separation of a multivariate distribution into its marginal distributions and the dependence structure between the margins.

It is also interesting to consider an alternative perspective, i.e. the survival perspective. First the general multivariate survival function is defined, Definition 2, and then the survival copula is defined as the restriction to the unit hypercube, see McNeil and Nešlehová [2009, lemma 1].

Definition 2. ***Multivariate Survival Function***
A survival function $\overline{H}$ *of a probability distribution* H *is a mapping* $\overline{H} : \mathbb{R}^d \mapsto [0,1]$ *if and only if it satisfies*

- $\overline{H}(-\infty, \ldots, -\infty) = 1$ *and* $\overline{H}(\boldsymbol{x}) = 0$ *if* $x_i = \infty$ *for at least one index* $i \in \{1, 2, \ldots, d\}$, *where* $\boldsymbol{x}$ *is a d-dimensional vector with components* x_i*;*
- $\overline{H}$ *is a right continuous function such that for all* $\boldsymbol{x} \in \mathbb{R}^d$ *one has*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall \boldsymbol{y} \geq \boldsymbol{x}\ \ ||\boldsymbol{y} - \boldsymbol{x}||_1 < \delta \Rightarrow |\overline{H}(\boldsymbol{y}) - \overline{H}(\boldsymbol{x})| < \epsilon. \tag{3–2}$$

- $\overline{H}(-\boldsymbol{x})$ *is quasi-monotone on* $\mathbb{R}^d$.

The survival copula is then specified according to the following definition.

Definition 3. ***Survival Copula***
The relationship between a d-dimensional copula $C(u_1, \ldots, u_d)$ *and the survival copula, denoted by* $\check{C}(u_1, \ldots, u_d)$ *is given by*

$$\check{C}(u_1, \ldots, u_d) = \overline{C}(1 - u_1, \ldots, 1 - u_m) \tag{3–3}$$

with

$$\overline{C}(u_1, \ldots, u_d) = \sum_{i=1}^{d} \left[(-1)^i \sum_{v(u_1,\ldots,u_d) \in \mathcal{Z}(M-i,M,0)} C(u_1, \ldots, u_d) \right], \tag{3–4}$$

where $\mathcal{Z}(A, B, \epsilon)$ *denotes the set* $\left\{ v \in [0,1]^B | v_i \in \{u_i, \epsilon\}, \sum_{i=1}^{B} \mathcal{X}_\epsilon(v_i) = A \right\}$.

Sklar's theorem (3–5) provides the foundation to the study of copulae by proving that any multivariate distribution with continuous margins has a unique copula representation.

Theorem 3.1.1. ***Sklar's Theorem (1959)***
Consider a d-dimensional cumulative distribution function H *with marginals* $F_1, \ldots, F_d$. *There exists a copula* C, *s.t.*

$$H(x_1, ..., x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \tag{3–5}$$

for all $x_i \in (-\infty, \infty), i \in 1, \ldots, d$. *Furthermore, if* F_i *is continuous for all* $i = 1, \ldots, d$ *then* C *is unique; otherwise* C *is uniquely determined only on* $RanF_1 \times$

$\cdots \times RanF_d$, *where* $RanF_i$ *denotes the range of the cumulative distribution function (cdf)* F_i.

Sklar's theorem can also be re-expressed in terms of survival functions of a multivariate distribution according to the following result in Theorem 3.1.2; see discussion in McNeil and Nešlehová [2009, theorem 2.1].

Theorem 3.1.2. ***Sklar's Theorem Expressed via Survival Function*** *Considering a d-dimensional survival function* $\overline{H}$ *with marginal distribution survival functions* $\overline{F}_{X_i}$ *for* $i \in \{1, 2, \ldots, d\}$, *then there exists a copula* $\overline{C}$, *called the survival copula of* $\overline{H}$ *such that*

$$\overline{H}(\boldsymbol{x}) = \overline{C}\left(\overline{F}_{X_1}(x_1), \ldots, \overline{F}_{X_d}(x_d)\right), \quad \forall \boldsymbol{x} \in \mathbb{R}^d, \tag{3–6}$$

or conversely one has

$$\overline{C}(\boldsymbol{u}) = \overline{H}\left(\overline{F}_{X_1}^{-1}(u_1), \ldots, \overline{F}_{X_d}^{-1}(u_d)\right), \quad \forall \boldsymbol{u} \in \mathcal{D}, \tag{3–7}$$

with $\mathcal{D} = \left\{\boldsymbol{u} \in [0,1]^d : \boldsymbol{u} \in ran\overline{F}_{X_1} \times \ldots \times ran\overline{F}_{X_d}\right\}$. *The survival copula* $\overline{C}$ *is uniquely determined on the support* $\mathcal{D}$. *Conversely, given a copula* $\overline{C}$ *and marginal survival functions* $\overline{F}_{X_i}$ *for* $i \in \{1, 2, \ldots, d\}$, *then the multivariate survival function* $\overline{H}$ *is uniquely given by Equation 3–6.*

An intuitive pictorial representation of the transformation of marginal distributions to standard uniform margins can be seen in Figure 3.1, as shown in Meucci [2011]. Here, it can be seen that using the individual empirical CDFs, an arbitrary data sample can be transformed to have approximately standard uniform margins.

There already exists an extensive literature on copulae, with publications gathering pace over recent years. Excellent textbooks on the topic include Aglio et al. [1991]; Joe [1997]; Nelsen [2006]. A number of gentle introductions to the world of copulae are available, such as Frees and Valdez [1998]; Bouyé et al. [2000]; Embrechts et al. [2003]; Schmidt [2006]; Genest and Favre [2007]; Meucci [2011].

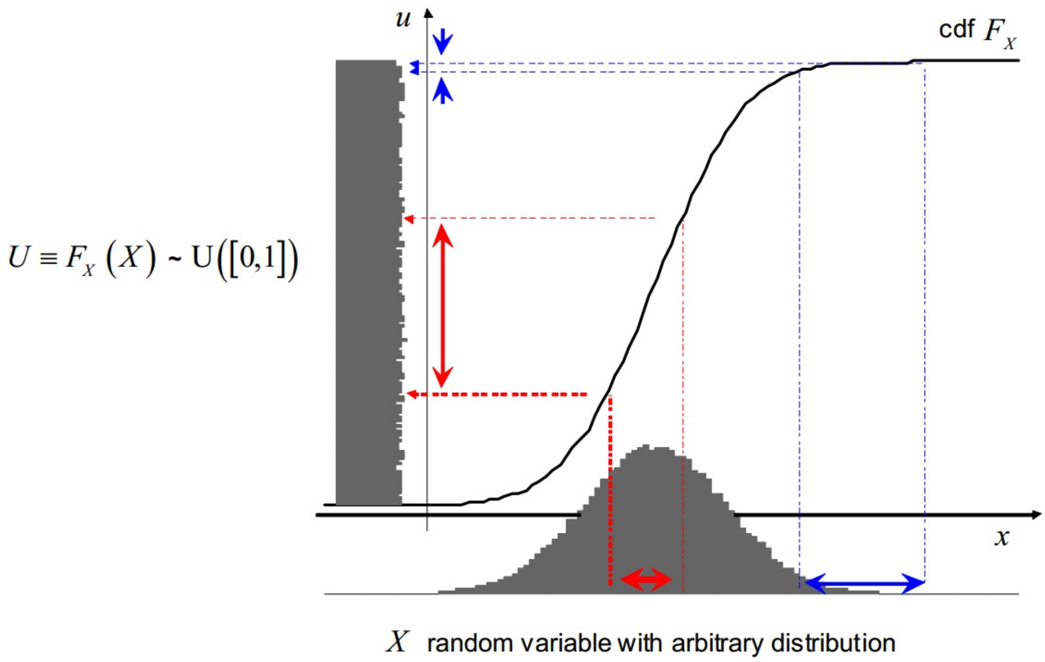


Figure 3.1: Transforming marginal distributions into standard uniform [0,1] margins. (Source: Meucci [2011])

Remark 3.1.3. *In this thesis, copula models are used to study the marginal behaviour of each currency and then separately to focus on developing hypotheses regarding the possible dependence structures between the log returns of the forward exchange rates of the currencies in the portfolios which can be tested through parametrization of a model via a copula and then a process of model selection.*

3.2 Copula Modelling and Its Emergence in Financial Modelling

The explosion of interest in copulae, beginning in the eighties, was in most part due to advances in quantitative risk management methodology in the financial and insurance world. The creation of more complex derivative products and new guidelines on regulation (see Chapter 1 of Embrechts et al. [2005]) contributed heavily to the need for risk management developments.

A notable paper from this era is Embrechts et al. [2002], in which the authors argue for copula approaches over linear correlation for the modelling of dependence for risk management. In particular, the authors point out the pitfalls of using linear correlation in the non-Gaussian world of finance and insurance. Hence, *beyond elliptical multivariate models* we have the following fallacies:

- Fallacy 1 : Marginal distributions and correlation determine the joint distribution.
- Fallacy 2 : Given marginal distributions F_1 and F_2 for X and Y, all linear correlations between -1 and 1 can be attained through suitable specification of the joint distribution.
- Fallacy 3 : The worst case VaR (quantile) for a linear portfolio $X+Y$ occurs when $\rho(X,Y)$ is maximal, i.e. X and Y are comonotonic.

These fallacies are avoided in this thesis through the use of mixture copulae, inner and outer power transforms and appropriately heavy-tailed marginal models to capture the complex non-linear dependence structures inherent in financial data.

In the context of for instance Fallacy 1 - this mistaken understanding typically arises from conceptualization of models constructed with intuition from Gaussian cases, which is the most well known case where this fallacy is correct. In fact, a key example of such a misinterpretation and its potential influence on the economy through misinterpretation of the model features arose in Li [2000]. This practically influential paper utilised a copula modelling framework which was developed on the topic of credit portfolio default modelling. The author proposed the use of copulae to specify the joint distribution of survival times (time until default of a financial instrument) with given marginal distributions (credit curves - giving all the marginal conditional default probabilities over a number of years). However, Li presents the Gaussian copula as the industry standard approach of the time (see Gupton et al. [1997]). It was the use and abuse of this Gaussian copula by the credit rating agencies (Moody's, Standard & Poor's and Finch) and the derivatives departments of investment banks that allowed the CDS (Credit Default Swap) market to balloon out to \$62 trillion in 2007 from \$920 billion in 2001. The CDO (Collateralised Debt Obligation) market saw a similar explosion, from \$275 billion in 2000 to $4.7 trillion by 2006. Li's formula came under much criticism at the time, notably Salmon [2012], for causing the collapse of the global economy. A more detailed analysis of the development and use of the Gaussian copula in this context is given in MacKenzie and Spears [2012], showing the unjustified blame placed on Li. Donnelly and Embrechts [2010] examines the (well-known) shortcomings of the Gaussian copula - explaining the overly simplistic nature of the model for credit derivatives. The authors present a clear analysis of the challenges of applying mathematical models to the constantly changing real world of finance.

The paper of Schönbucher and Schubert [2001] allows for a much more general specification of the dependence between default events than previous works. The modelling framework introduced here is a continuous-time dynamic model, with defaults and default probabilities evolving consistently within the model. The Clayton and Gumbel copulae are proposed to model the default dependence, allowing for more realistic default contagion.

On the topic of portfolio allocation, Patton [2004] explores asymmetries in the dependence structure of stocks across different market conditions. Patton notes that "stock returns appear to be more dependent during market downturns than during market upturns", hence violating the assumption of elliptically distributed asset returns. Dependence models that allow for, but do not impose, greater dependence during bear markets than bull markets are considered. The author finds substantial evidence that skewness and asymmetric dependence are important considerations in portfolio allocation. In particular, the portfolios based on the more flexible copula dependence models outperform both the equally weighted portfolio and the portfolio based on the bivariate normal model.

Hong et al. [2007] introduces a test for asymmetric dependence and then goes on to propose a Bayesian framework for modelling asymmetry via a mixture model of normal and Clayton copulae. The authors conclude that "incorporating assets' asymmetric characteristics can add substantial economic value in portfolio decisions."

The use of copula modelling approaches has started to emerge across various asset classes. For example, Wu et al. [2012] propose dynamic copula-based GARCH models to explore the dependence structure between the oil price and the US dollar exchange rate. In addition, an asset allocation strategy is implemented to evaluate economic value and confirm the efficiency of the copula-based GARCH models. Gronwald et al. [2011] apply various copulae in order to investigate the complex dependence structure between EU emission allowance (EUA) futures returns and those of other commodities, equity and energy indices. The authors consider time-varying copulae, concluding that the estimated copula parameters are not constant over time and that in particular the dependence is stronger during the period of the financial crisis.

Amidst all of this new found excitement for copulae there were some outspoken critics. Most notably was Mikosch [2006a], who cited a concern that copulae were being viewed as *the* solution to all problems in stochastic dependence modelling, whereas in his view "copulas do not contribute to a better understanding of multivariate extremes". There were numerous responses from leaders in the copula field to Mikosch's attack, such as Genest and Rémillard [2006]; Embrechts [2006]; Joe [2006]; de Vries and Zhou [2006]; Lindner [2006]; Peng [2006] and Segers [2006] - leading to a rejoinder by Mikosch, see Mikosch [2006b]. Embrechts [2009] sums up the responses best in his personal review of copulae shortly after:

> "Copulas form a most useful concept for a lot of applied modeling, they do not yield, however, a panacea for the construction of useful and well-understood multivariate dfs, and much less for multivariate stochastic processes. But none of the copula experts makes these claims."

It is useful at this point to discuss the pros and cons of the copula modelling framework.

PROS:

- Separating out the modelling of the marginals and the dependence structure allows for more flexibility in the complete multivariate model.
- The dependence structure as summarized by a copula is invariant under increasing and continuous transformations of the marginals.

- The tail characteristics within the dependence structure can be explicitly modelled using well-known and interpretable parametric models, e.g. Archimedean copulae.

- High dimensional copulae can be reduced to the composition of lower dimensional building block copulae, e.g. pair-copula constructions, to create extremely flexible models of complex dependence structures.

CONS:

- Which copula to choose? Sometimes it is not easy to say which parametric copula fits a dataset best, since some copulae may provide a better fit near the center and others near the tails. However, by focusing on models with suitable characteristics for the application at hand and using goodness-of-fit tests, e.g. AIC, BIC or CIC, one can overcome this issue.

- As with any statistical model, ignorance on the behalf of practitioners can lead to dangerous oversimplification and reliance on inappropriate models.

Thus, when applying these models in practice it is of the utmost importance to carefully consider the assumptions one is making. The key focus in this research is on combining suitable marginal models, i.e. with the capacity to model skewness and tail-heaviness flexibly, with a model of the dependence structure that captures the upper and lower multivariate tail characteristics asymmetrically.

In the context investigated in this thesis, i.e. currency carry trade portfolios, the application of copula models is a novel approach to describe the rationale of the forward premium puzzle. The benefits of such an approach are clear from the pros described above. In order to address the cons of a copula modelling approach, a thorough goodness-of-fit testing procedure was performed in order to select the appropriate marginal (time series or static) and copula model for each data fit. Furthermore, a detailed analysis of the extremal dependence properties of the copulae was carried out.

3.3 Classical Measures of Dependence

Measuring the dependence between random variables has long been of interest to statisticians and practitioners alike. A history of the development of dependence measures can be found in Mari and Kotz [2001]. It is important to note that, in general, the dependence structure between two random variables can *only* be captured in full by their joint probability distribution, and thus any scalar quantity extracted from this structure must be viewed as such. Scarsini [1984] gives the following intuitive definition of dependence:

> "Dependence is a matter of association between X and Y along any measurable function, i.e. the more X and Y tend to cluster around the graph of a function, either $y = f(x)$ or $x = g(y)$, the more they are dependent."

3.3.1 Linear Correlation

The most well-known measure of dependence, Pearson's Product Moment Correlation Coefficient, was developed by Karl Pearson, see Pearson [1896], building on Sir Francis Galton's approach using the median and semi-interquartile range, see Galton [1889].

Pearson's correlation coefficient is a measure of how well the two random variables can be described by a linear function and is defined as follows:

Definition 4. ***Pearson's Correlation Coefficient***

$$\rho := \frac{Cov[X, Y]}{\sqrt{Var[X]Var[Y]}} \tag{3–8}$$

Hence perfect linear dependence gives $\rho = +1$ or $\rho = -1$. The major weakness of linear correlation is its non-invariance under non-linear monotonic transformations of the random variables.

3.3.2 Rank Correlation

Rank correlation measures the relationship between the *rankings* of variables, i.e after assigning the labels "first", "second", "third", etc. to different observations of a particular variable. The coefficient lies in the interval [-1, +1], where +1 indicates the agreement between the two rankings is perfect, i.e. the same; -1 indicates the disagreement between the two rankings is perfect, i.e. one ranking is the reverse of the other; 0 indicates the rankings are completely independent. Due to this scale-invariance, rank correlations thus provide an approach for fitting copulae to data.

The choice of dependence measure is influenced by the type of dependence one seeks to capture, such as lower left quadrant, upper right quadrant etc. However, in non-trivial multivariate distributions it isn't possible to capture all of the possible combinations of dependence patterns within a single dependence measure.

3.3.2.1 Spearman's Rho

Charles Spearman introduced the nonparametric measure of dependence, Spearman's rank correlation coefficient, in Spearman [1904]. This measure assesses how well the dependence between two random variables can be described by a monotonic function. As such it is equivalent to the Pearson's correlation coefficient between the ranked variables, defined as follows:

Definition 5. ***Spearman's Rank Correlation Coefficient***

$$\rho := \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \tag{3–9}$$

where x_i, y_i are the ranks.

Spearman's rank correlation can be directly derived from the copula describing the dependence between random variables X_1 and X_2:

Definition 6. ***Spearman's Rank Correlation via Copula***

$$\rho(X_1, X_2) := 12 \int_0^1 \int_0^1 \left(C\left(u_1, u_2\right) - u_1 u_2\right) du_1 du_2 \tag{3–10}$$

In addition, a general multivariate extension of Spearman's Rank Correlation is developed for d-dimensional loss random vectors and given below, see details in Nelsen [2002].

Definition 7. ***Multivariate Generalized Spearman's Rank Correlation via Copula***
Consider the d-dimensional copula given by C and the permuted copula C^{σ}, then the generalized Spearman's Rho concordance measure of dependence is given according to

$$\rho_d(C) = \alpha_d \left(\int_{[0,1]^d} (C + C^{\sigma})\, d\Pi^d - \frac{1}{2^{d-1}} \right) \tag{3–11}$$

where one has $\alpha_d = \frac{(d+1)2^{d-1}}{2^d-(d+1)}$ and Π^d is the d-dimensional Independence Copula as defined below.

Definition 8. ***Independence Copula***
The d-dimensional independence copula is defined as

$$\Pi^d(u) = \prod_{i=1}^{d} u_i. \tag{3–12}$$

3.3.2.2 Kendall's Tau

Maurice Kendall developed the τ rank correlation coefficient in Kendall [1938], although Gustav Fechner proposed a similar measure in the context of time series in 1897, see Kruskal [1958].

Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables from a joint distribution function F, then Kendall's rank correlation is given by

Definition 9. ***Kendall's Tau***

$$\tau := \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right] \tag{3–13}$$

Similarly, Kendall's rank correlation can be directly derived from the copula describing the dependence between random variables X_1 and X_2:

Definition 10. ***Kendall's Tau via Copula***

$$\tau := 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \tag{3–14}$$

Spearman's ρ and Kendall's τ share a lot of common properties, however "Spearman's ρ is a measure of average quadrant dependence, while Kendall's τ is a measure of average likelihood ratio dependence", see Fredricks and Nelsen [2007]. In layman's terms it can be seen that Kendall's τ penalises rank displacements by the distance of the displacement, whilst Spearman's ρ penalises by the square of the distance. Also, as Newson [2002] notes, "confidence intervals for Spearman's ρ are less reliable and less interpretable than confidence intervals for Kendall's τ-parameters".

3.3.2.3 Blomqvist's Beta

Nils Blomqvist developed a measure of concordance in Blomqvist [1950] known as Blomqvist's β. This is a quadrant measure that is related to medial correlation and is defined as follows.

Definition 11. ***Blomqvist's Beta***

Consider two random variables X_1 and X_2, then Blomqvist's beta is given by

$$\begin{aligned} \rho_\beta \left[X_1, X_2\right] := \mathbb{P}\mathrm{r}\left[\left(X_1 - med\left(X_1\right)\right)\left(X_2 - med\left(X_2\right)\right) > 0\right] \\ - \mathbb{P}\mathrm{r}\left[\left(X_1 - med\left(X_1\right)\right)\left(X_2 - med\left(X_2\right)\right) < 0\right], \end{aligned} \tag{3–15}$$

where $med(X_i)$ is the median of random variable X_i.

For generalizations of Blomqvist's beta to higher dimensions, see discussions in Joe [1990], Nelsen [2002] and Dolati and Úbeda-Flores [2006].

One can also make the following comments regarding the properties of Blomqvist's Beta measure of concordance:

- The empirical version $\hat{\rho}_\beta$ of Blomqvist's beta is a suitably scaled version of the proportion of points whose components are either both smaller, or both larger, than their respective sample medians;

- The computation of $\hat{\rho}_\beta$ involves only $O(n)$ operations, as opposed to $O(n^2)$ for the empirical versions of Kendall's tau and Spearman's rho.

In addition, Blomqvist's Beta can also be specified with regard to a copula as follows.

Definition 12. ***Blomqvist's Beta via Copula***
The bivariate Blomqvist's Beta can be expressed explicitly via the bivariate copula C according to

$$\beta = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1. \tag{3–16}$$

Remark 3.3.1. *Recently in Genest et al. [2013] the authors proposed the inversion of this copula based representation of Blomqvist's Beta to perform explicit parameter estimation for several copula models.*

As with the other popular measures of concordance specified above, there is also a generalization of Blomqvist's Beta to multivariate settings, see discussions in Nelsen [2002].

Definition 13. ***Generalized Blomqvist's Beta via Copula***
Consider a d-dimensional copula C, then the generalized Blomqvist's Beta is given by

$$\beta_d(C) = \alpha_d\left(C(\frac{1}{2}, \ldots, \frac{1}{2}) - \frac{1}{2^d}\right), \tag{3–17}$$

where $\alpha_d = \frac{2^d}{2^{d-1}-1}$

These classical notions of quantifying dependence have been widely considered in some form in financial modelling, especially simple linear dependence models, though they are rarely understood in this literature from the perspective of a copula structure. So it is useful to point out this link when considering ideas of modelling in financial settings via copulae.

The next section outlines less commonly used notions of extremal dependence in financial modelling, namely the strong tail dependence, asymptotic joint extreme dependence measures.

3.4 Tail Dependence

In order to examine the dependence behaviour in the extremes of multivariate distributions we use the concept of tail dependence. The bivariate tail dependence coefficient is defined as the conditional probability that a random variable exceeds a certain threshold given that the other random variable in the joint distribution has exceeded this threshold.

Definition 14. ***Bivariate Tail Dependence***
For random variables X_1 and X_2 with cdfs $F_i, i = 1, 2$ and copula C. We define the coefficient of upper tail dependence by:

$$\lambda_u := \lim_{u \nearrow 1} \mathbb{P}\left(X_2 > F_2^{-1}(u) \,|X_1 > F_1^{-1}(u)\right) = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u} \tag{3–18}$$

and similarly we define the coefficient of lower tail dependence by:

$$\lambda_l := \lim_{u \searrow 0} \mathbb{P}\left(X_2 \leq F_2^{-1}(u) \,|X_1 \leq F_1^{-1}(u)\right) = \lim_{u \searrow 0} \frac{C(u,u)}{u} \tag{3–19}$$

One can consider both closed form expressions for this measure of extremal dependence in terms of copula parameters or non-parametric estimators of these quantities, which don't make explicit dependence on a particular parametric copula family. Both of these approaches are investigated in this thesis.

This concept of bivariate tail dependence has been recently extended to the multivariate setting by De Luca and Rivieccio [2012]. Now one may accurately interpret the tail dependence present between sub-vector partitions of the multivariate random vector with regard to joint tail dependence behaviours. In the context of the applications I consider in this thesis, this allows us to examine the probability that any sub-vector of the log return forward exchange rates for the basket of currencies will exceed a certain threshold given that the log return forward exchange rates for the remaining currencies in the basket have exceeded this threshold, in particular thresholds that are placing an interest in the tails of the multivariate distribution. The interpretation of such results is then directly relevant to assessing the chance of large adverse movements in multiple currencies which could potentially increase the risk associated with currency carry trade strategies significantly, compared to risk measures which only consider the marginal behaviour in each individual currency.

Definition 15. ***Multivariate Tail Dependence***

Let $X = (X_1, ..., X_d)^T$ be a d-dimensional random vector with marginal distribution functions $F_1, ..., F_d$ and copula C. We define the coefficient of multivariate upper tail dependence by:

$$\begin{aligned}\lambda_u^{1,\ldots,h|h+1,\ldots,d} &= \lim_{\nu \nearrow 1} P\left(X_1 > F^{-1}(\nu), \ldots, X_h > F^{-1}(\nu) | X_{h+1} > F^{-1}(\nu), \ldots, X_d > F^{-1}(\nu)\right) \\ &= \lim_{\nu \nearrow 1} \frac{\bar{C}_d(1-\nu, \ldots, 1-\nu)}{\bar{C}_{d-h}(1-\nu, \ldots, 1-\nu)}\end{aligned} \tag{3–20}$$

where $\bar{C}_d$ is the survival copula of a d-dimensional copula C.

Similarly we define the coefficient of multivariate lower tail dependence by:

$$\begin{aligned}\lambda_l^{1,\ldots,h|h+1,\ldots,d} &= \lim_{\nu \searrow 0} P\left(X_1 < F^{-1}(\nu), \ldots, X_h < F^{-1}(\nu) | X_{h+1} < F^{-1}(\nu), \ldots, X_d < F^{-1}(\nu)\right) \\ &= \lim_{\nu \searrow 0} \frac{C_d(\nu, \ldots, \nu)}{C_{d-h}(\nu, \ldots, \nu)}\end{aligned} \tag{3–21}$$

Here, $d - h$ is the number of variables conditioned on (from the d considered).

3.4.1 Non-Parametric Estimators

One may also consider a non-parametric approach to estimating the tail dependence. This can be a useful comparison to the tail dependence coefficients observed using the various parametric copula models. Furthermore, the value of a d-dimensional Archimedean copula is invariant under permutations of its arguments, i.e.

$$C(u_1, \ldots, u_d) = C(u_{\pi(1)}, \ldots, u_{\pi(d)}), \quad u_1, \ldots, u_d \in [0, 1] \tag{3–22}$$

for arbitrary bijections $\pi : \{1, \ldots, d\} \to \{1, \ldots, d\}$.

However, in reality it may not be true, for example that each pair of variables within the multivariate density has the same upper tail dependence. Thus, it is informative to analyse the breakdown of the overall tail dependence within a multivariate density by examining the tail dependence between each of the pairs of variables. In order to estimate the non-parametric tail dependence we need to make use of the empirical copula, which is defined as follows:

Definition 16. *Bivariate Empirical Copula*

$$\hat{C}_n\left(u_1, u_2\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left(\frac{R_{1i}}{n} \leq u_1, \frac{R_{2i}}{n} \leq u_2\right) \tag{3–23}$$

where R_{ji} is the rank of the variable in its marginal dimension that makes up the pseudo data.

As with any estimated quantity, one may estimate it in a number of different ways, each with differing estimator statistical properties. The following are the most widely used non-parametric estimators for extremal tail dependence, see Dobrić and Schmid [2005]; Schmidt and Stadtmüller [2006]; Cruz et al. [2015].

Definition 17. *Non-Parametric Pairwise Estimator of Upper Tail Dependence (Estimator 1)*

$$\hat{\lambda}_u = 2 - \min\left[2\ ,\ \frac{log\,\hat{C}_n\left(\frac{n-k}{n}, \frac{n-k}{n}\right)}{log(\frac{n-k}{n})}\right] \qquad k = 1, 2, \ldots n-1, \tag{3–24}$$

Definition 18. *Non-Parametric Pairwise Estimator of Upper Tail Dependence (Estimator 2)*

$$\hat{\lambda}_u = 2 - \min\left[2\ ,\ \frac{1 - \hat{C}_n\left(\frac{n-k}{n}, \frac{n-k}{n}\right)}{1 - (\frac{n-k}{n})}\right] \qquad k = 1, 2, \ldots n-1, \tag{3–25}$$

Definition 19. *Non-Parametric Pairwise Estimator of Upper Tail Dependence (Estimator 3)*

$$\hat{\lambda}_u = 2 - \min\left[2\ ,\ 2\exp\left\{\frac{1}{n}\sum_{i=1}^{n} ln\left(\frac{\sqrt{ln\frac{1}{U_{1,i}} ln\frac{1}{U_{2,i}}}}{ln\left(\frac{1}{\max\{U_{1,i},U_{2,i}\}^2}\right)}\right)\right\}\right] \qquad k = 1, 2, \ldots n-1, \tag{3–26}$$

Remark 3.4.1. *In order to form a robust estimator of the upper tail dependence a median of the estimates obtained from setting k as the $1^{st}, 2^{nd}, \ldots, 20^{th}$ percentile* *values was used. Similarly, k was set to the $80^{th}, 81^{st}, \ldots, 99^{th}$ percentiles for the lower tail dependence.*

3.4.2 Asymptotic Independence

In the case where the extremes of marginal distributions are asymptotically independent one would find the tail dependence coefficient to be zero. Thus applying extreme value models based on non-zero tail dependence to these cases leads to the over-estimation of probabilities of extreme joint events. In finance this could result in model risk and over confidence in forecast extremal probabilities of events. It is therefore important to carefully check the copula models selected for applications to make sure to avoid as much as possible such situations. One way to help quantify such an event is to examine this class of distributions at finite levels, i.e. non-asymptotic, which allows for a more useful measure of extremal dependence. Coles et al. [1999] defines a new quantity $\bar{\chi}$ as given by equation 3–27.

Definition 20. $\bar{\chi}$ **-** ***Measure of Extremal Dependence***

$$\bar{\chi} := \frac{2\ log\ Pr(U > u)}{log\ Pr(U > u, V > v)} - 1 = \frac{2\ log(1-u)}{log\ \bar{C}(u,\nu)} - 1 \tag{3–27}$$

where $-1 < \bar{\chi}(u) \leq 1$ *for all* $0 \leq u \leq 1$.

Hence, $\bar{\chi}$ increases with dependence strength and equals 1 for asymptotically dependent variables. For Gaussian models of dependence the measure $\bar{\chi}$ is equal to the correlation, providing a benchmark for interpretation in general models of dependence. Coles et al. [1999] thus argues that using this new measure in addition to the tail dependence measure gives a more complete summary of extremal dependence.

3.5 Decomposing Multivariate Distributions

From the previous sections just presented one can start to see that a statistician faced with the task of modelling a multivariate distribution has a multitude of techniques at his disposal. The simplest possible choice one could make is to assume all of the random variables are independent and hence only the marginals need to be modelled and combined to form the multivariate model. Whilst simple, this approach neglects any dependence between the variables and thus is often a very poor model.

A multivariate distribution may be decomposed in all manner of ways, for example via conditional distributions, factor models, tree representations etc. Barber [2012] is a good resource for exploring the possible methods of decomposing multivariate distributions.

The copula modelling framework provides an intuitive method of constructing a multivariate model by carefully considering the marginal models and then the dependence structure between the random variables in two distinct stages.

In the application considered in this research a mixture of d-dimensional copulae has been considered to provide a model with asymmetric tail dependence and the capability of capturing negative dependence between the currencies. Since the carry portfolios only contain four currencies, this mixture of 4-dimensional copulae has sufficient flexibility to accurately model the overall dependence structure, and in particular the upper and lower tails.

In cases of much higher dimensional distributions one may consider the additional flexibility offered by copula models known as vine copula models, since standard multivariate copulae may not always accommodate, with sufficient flexibility and degrees of freedom, dependence structures between pairs of variables. Vine copulae use bivariate copulae (not necessarily from the same parametric family) and a nested set of trees to build up the overall dependence structure more flexibly. Clearly there is a trade-off with the number of parameters here. Kurowicka and Joe [2011] provide an excellent overview of this burgeoning topic. Some key papers include Bedford and Cooke [2002]; Berg and Aas [2009]; Aas et al. [2009].

It is worth noting that one key challenge to be tackled in copula modelling is the construction of dynamic models that capture the time-varying nature of dependence in the real world, such as in finance. For example, Dias and Embrechts [2004] explore the detection of change-points in FX data via a dynamic copula model analysis. More recently, Patton [2012] proposed a new class of dynamic copula models for daily asset returns that exploits information from high frequency (intra-daily) data. The authors augment the generalized autoregressive score (GAS) model of Creal et al. [2013] with high frequency measures such as realized correlation to obtain a "GRAS" model.

3.6 Elliptical Copulae

There is a vast collection of different parametric copulae in the literature, each with associated dependence features. The monograph Nelsen [2006] provides a detailed mathematical background of many important copulae. There are many useful papers reviewing the different families of copulae available to the practitioner, such as Bouyé et al. [2000]; Schmidt [2006]; Trivedi and Zimmer [2007]; Durante and Sempi [2010].

Genest and Neslehova [2007] discusses the issues associated with modelling multivariate distributions with discrete margins, such as in count data. As discussed in Sklar's theorem (3–5), the copula representation of a multivariate distribution is only guaranteed to be unique when the marginal distributions are continuous. This does not present a problem in this thesis as all of the marginals considered for this application are continuous.

Amongst the most popular copulae are elliptical copulae and Archimedean copulae. In general, elliptical copulae arise naturally from their respective elliptical distributions following Sklar's theorem. Although elliptical copulae have no closed form, they have the property that the dependence structure is fully described by the correlation. An elliptical distribution is defined as follows:

Definition 21. ***Elliptical Distribution***
The density function of an elliptical distribution (if it exists) is given by:

$$f(x) = |\Sigma|^{-\frac{1}{2}} g\left[(x-\mu)^T \Sigma^{-1} (x-\mu)\right] \quad , \quad x \in \mathbb{R}^n \tag{3–28}$$

where Σ (dispersion) is a symmetric positive semi-definite matrix, $\mu \in \mathbb{R}^n$ (location) and g (density generator) is a $[0, \infty) \to [0, \infty)$ function.

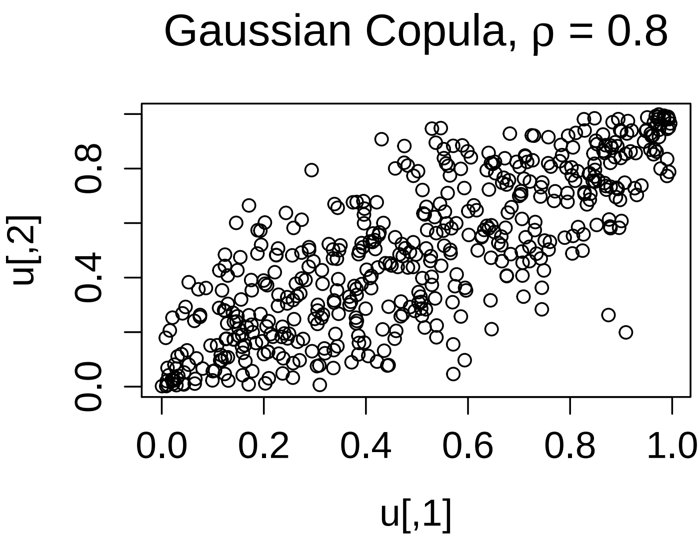


Figure 3.2: Scatterplot of 500 random samples from a Gaussian copula with $\rho = 0.8$.

3.6.1 Gaussian Copula

The Gaussian copula has long been favoured by practitioners due to its simplicity. The bivariate Gaussian copula is defined as follows:

Definition 22. ***Bivariate Gaussian Copula***

$$C^{Gaussian}(u, v) := \Phi_\rho\left(\Phi^{-1}(u), \Phi^{-1}(v)\right) \quad , \tag{3–29}$$

where

$$\Phi_\rho(x, y) := \int_{-\infty}^{x} \int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \frac{2\rho st - s^2 - t^2}{2(1-\rho^2)} dsdt$$

and Φ *denotes the standard normal cdf.*

A random sample from a Gaussian copula with $\rho = 0.8$ can be seen in Figure 3.2. The copula density plot for the Gaussian copula with $\rho = 0.3$ can be seen in Figure 3.3. It is important to note the lack of tail dependence in the Gaussian copula, i.e. in the lower left and upper right corners of the unit square. Hence the Gaussian copula is a very restrictive model of dependence in the real world, since

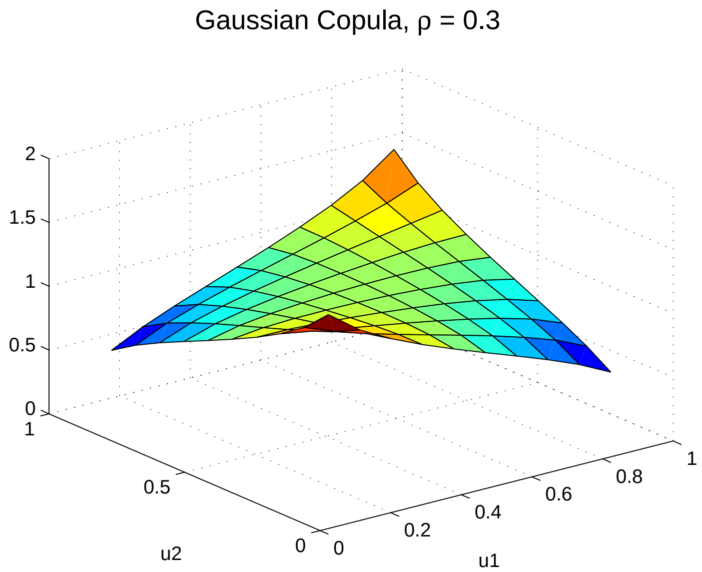


Figure 3.3: Density plot of Gaussian copula with $\rho = 0.3$.

it does not allow for variables to become highly concordant in the extremes, e.g. default contagion.

3.6.2 t-Copula

Student's t-copula retains much of the simplicity of the Gaussian copula, such as in simulation and calibration, but also allows for the modelling of tail dependence between variables. The behaviour of the model at the four corners is quite different from that of the Gaussian copula, while towards the center they are more similar, as can be seen in Figure 3.4 and more clearly in the copula density plot in Figure 3.5 with different parameters. Hence, although having the same correlation as the Gaussian copula, the extreme events are much more likely under the t-copula. This copula has often been referred to as the "desert island copula" by Dr. Paul Embrechts due to its excellent fit to multivariate financial return data. However, it does not allow for asymmetry in the tails, i.e. differing upper and lower tail dependence in a portfolio of currencies. The Student's t-Copula is defines as:

Definition 23. ***Student's t-Copula***

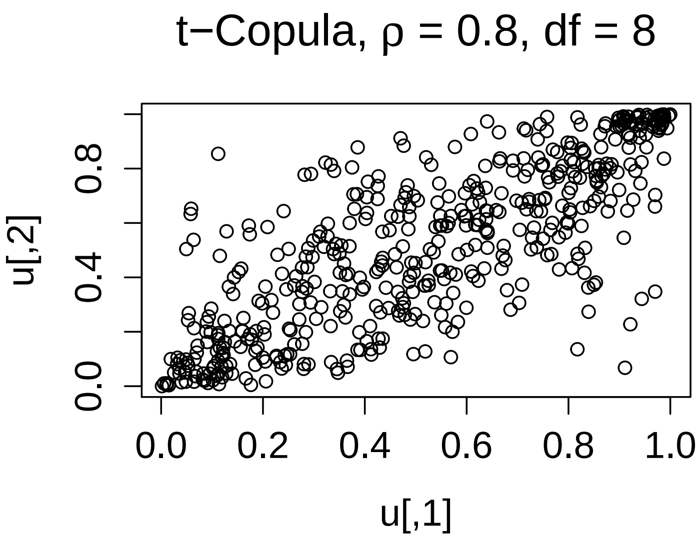


Figure 3.4: Scatterplot of 500 random samples from a t-copula with $\rho = 0.8$, degrees of freedom $= 8$.

$$C^t(u_1, u_2; \nu, \rho) := \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} dsdt \tag{3–30}$$

where $t_\nu^{-1}(u_i)$ denotes the inverse cdf of the standard univariate t-distribution with ν degrees of freedom.

In practice, the use of a standard t-copula comes under fire since it has only a single parameter for the degrees of freedom. This may restrict the flexibility in modelling the tail dependence structure in a multivariate case. The most advanced solution in the literature in this regard is Luo and Shevchenko [2010], in which the authors propose a modified grouped t-copula, "where each group consists of one risk factor only, so that a priori grouping is not required", i.e. each group has only one member and an individual degrees of freedom parameter associated with it.

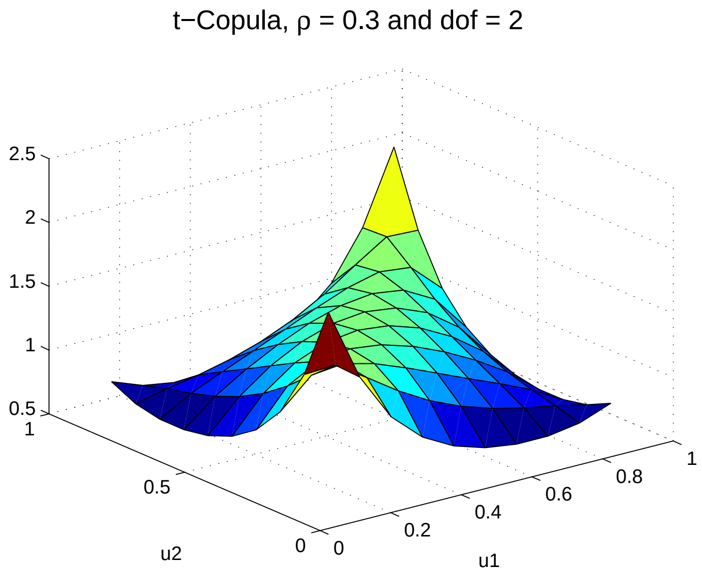


Figure 3.5: Density plot of a t-copula with $\rho = 0.3$, degrees of freedom $= 2$.

3.7 Archimedean Copulae

Archimedean copulae are not derived from multivariate distributions, but can be stated explicitly in a simple form. Many Archimedean copulae have been proposed in the literature, see Nelsen [2006], with many further copulae available as extensions and combinations of these base copulae. Archimedean copulae are attractive to researchers and practitioners due to their directly interpretable tail dependence features and parsimonious representations.

An Archimedean copula is defined as follows:

Definition 24. ***Archimedean Copula***
A d-dimensional copula C is called Archimedean if for some generator ψ it can be represented as:

$$C(\boldsymbol{u}) = \psi\{\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)\} = \psi\{t(\boldsymbol{u})\} \quad \forall \boldsymbol{u} \in [0,1]^d \tag{3–31}$$

where $\psi^{-1} : [0,1] \to [0,\infty]$ is the inverse generator with $\psi^{-1}(0) = inf\{t : \psi(t) = 0\}$.

Note the shorthand notation $t(\mathbf{u}) = \psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)$ that will be used throughout this section.

As we will see later, it is necessary to have formulas for computing the copula densities if one seeks to fit these models using a maximum likelihood approach. equation 3–32 provides such a formula in a generic form for each member of the family of Archimedean copulae.

Definition 25. ***Archimedean Copula Density***
McNeil and Nešlehová [2009] prove that an Archimedean copula C admits a density c if and only if the $(d-1)$th derivative of ψ, i.e. $\psi^{(d-1)}$, exists and is absolutely continuous on $(0,\infty)$. When this condition is satisfied, the copula density c is given by

$$c(\boldsymbol{u}) = \frac{\partial^d C(u_1,\ldots,u_d)}{\partial u_1 \ldots \partial u_d} = \psi^{(d)}\{t(\boldsymbol{u})\}\prod_{j=1}^{d}(\psi^{-1})'(u_j) \ , \quad \boldsymbol{u} \in (0,1)^d \quad (3\text{–}32)$$

Remark 3.7.1. *There are many possible copula models that could be considered in the modelling of the multivariate dependence features of the currency portfolios. The intention of this analysis was to work with well known models which have well understood tail dependence features and are relatively parsimonious with regard to the number of parameters specifying the copula. I obtain flexible dependence relationships by combining such components into mixture models that allow for a range of flexible tail dependence relationships to be studied. In particular, I will focus on the well-known class of Archimedean copulae, as defined in equation 3–31, since they provide a parsimonious approach that allows for the modelling of various tail dependence characteristics.*

Multivariate tail dependence can also be written in terms of the generator derivatives as presented in the next subsection.

3.7.1 Multivariate Archimedean Copula Tail Dependence

As discussed in Section 3.4, it is important to be able to accurately interpret the tail dependence present between sub-vector partitions of the multivariate random vector with regard to joint tail dependence behaviours. Below I give the explicit generalised multivariate expressions for Archimedean copulae, equations 3–33 and 3–34, derived in De Luca and Rivieccio [2012].

Definition 26. ***Generalized Archimedean Upper Tail Dependence***
Let $X = (X_1, ..., X_d)^T$ be a d-dimensional random vector with marginal distribution functions $F_1, ..., F_d$. The coefficient of upper tail dependence is defined as:

$$\begin{aligned}\lambda_u^{1,\dots,h|h+1,\dots,d} &= \lim_{\nu\to 1-} P\left(X_1 > F^{-1}(\nu), ..., X_h > F^{-1}(\nu) | X_{h+1} > F^{-1}(\nu), ..., X_d > F^{-1}(\nu)\right) \\ &= \lim_{t\to 0^+} \frac{\sum_{i=1}^{d}\left(\binom{d}{d-i} i(-1)^i \left[\psi^{-1\prime}(it)\right]\right)}{\sum_{i=1}^{d-h}\left(\binom{d-h}{d-h-i} i(-1)^i \left[\psi^{-1\prime}(it)\right]\right)}\end{aligned} \tag{3–33}$$

where $\psi^{-1\prime}$ is the derivative of the inverse generator. Here, $d-h$ is the number of variables conditioned on (from the d considered).

Definition 27. ***Generalized Archimedean Lower Tail Dependence***
Let $X = (X_1, ..., X_d)^T$ be a d-dimensional random vector with marginal distribution functions $F_1, ..., F_d$. The coefficient of lower tail dependence is defined as:

$$\begin{aligned}\lambda_l^{1,\dots,h|h+1,\dots,d} &= \lim_{\nu\to 0+} P\left(X_1 < F^{-1}(\nu), ..., X_h < F^{-1}(\nu) | X_{h+1} < F^{-1}(\nu), ..., X_d < F^{-1}(\nu)\right) \\ &= \lim_{t\to\infty} \frac{d}{d-h} \frac{\psi^{-1\prime}(dt)}{\psi^{-1\prime}((d-h)t)}\end{aligned} \tag{3–34}$$

where $\psi^{-1\prime}$ is the derivative of the inverse generator. Here, $d-h$ is the number of variables conditioned on (from the d considered).

Remark 3.7.2. *Due to the exchangeability property of Archimedean copulae and the fact that the tail dependence of a mixture copula is equal to the mixture of the component tail dependences, see Nelsen [2006], one does not need to select which currencies to condition on for the proposed Clayton-Frank-Gumbel copula. Therefore the generalised tail dependence (GTD) measures proposed here only require the selection of the number of conditioning variables, i.e. $h = 1$ or $h = 2 \dots$ or $h = d-1$. Furthermore, the choice of the number of conditioning variables h merely scales the resultant tail dependence measure. Therefore the analysis that follows in the thesis is robust to the choice of h. In this light, the choice of $h = 1$ adopted here seems a reasonable approach as it reflects the limiting* *probability of one currency having an extreme move beyond a threshold given that the remaining currencies in the basket have an extreme move beyond this threshold. In Ames et al. [2015c] the authors also investigated robust non-parametric pairwise tail dependence estimators. In this paper, the authors analysed the contribution of each pair of currencies in the basket to the overall model based tail dependence estimate.*

Since the distribution function of an Archimedean copula is specified by a special function $\psi(.)$ called the generator and indeed the multivariate tail dependence can be written in terms of the generator, the next section will first study the mathematical properties that such functions must obey in order to generate a valid copula model.

Then in section Section 3.7.5 I will present examples of models where a particular parametric function is selected for ψ to produce a sub-family of the Archimedean copula family. The choices I consider will also have known properties regarding their linear, rank and extremal tail dependence features in terms of the parametric copula model parameters.

3.7.2 Archimedean Copula Generators

It was shown in the Ph.D. thesis of Ling [1964] that the generator ψ will produce a bivariate copula distribution if and only if it is a convex function. Then in Kimberling [1974] it was shown that in order for the generator ψ to generate any Archimedean copula distribution in any dimension d then it must be a completely monotonic function; see Theorem 3.7.3.

Theorem 3.7.3. ***Completely Monotone Generators and Existence of Archimedean Copulae***
If a generator ψ that is a mapping $\psi : [0, \infty] \mapsto [0, 1]$ is continuous and strictly decreasing such that $\psi(0) = 1$ and $\psi(\infty) = 0$, that is, $\psi \in C^{\infty}(0, \infty)$, i.e. infinitely differentiable, and one has that $(-1)^k \psi^{(k)}(x) \geq 1$ for $k = 1, \ldots$ then this class of generators can create Archimedean copulae models in any dimension. This class of completely monotone generators for Archimedean copula in any dimension are denoted by ψ_{∞}.

It is useful to note the following relevant properties of completely monotone functions in Lemma 3.7.4 see, for instance, discussion in Hofert [2010a].

Lemma 3.7.4. ***Properties of Completely Monotone Functions***
A completely monotone function satisfies the following properties:

- *Closure under multiplication and positive affine transformations (i.e., linear additive combinations with positive coefficients);*
- *If a function f is a Laplace–Stieltjes transform, then the function f^{α} is completely monotone for any power $\alpha \in (0, \infty)$ if and only if the derivative $(-\ln f)'$ is completely monotone;*
- *If a function f is completely monotone and a second function g is non-negative with its first derivative g' completely monotone, then the composite function $f \circ g$ is a completely monotone function;*
- *If a function f is non-negative and its derivative f' is completely monotone, then the reciprocal of the function f given by $1/f$ is a completely monotone function;*
- *If a function f is continuous on $[0, \infty]$, satisfying $\frac{d^k}{dx^k} f(x) \geq 0$ for any integer $k \in \mathbb{J}$ and $x \in (0, \infty)$ and a function g is completely monotone, then the composite function $f \circ g$ is a completely monotone function.*

The requirement for complete monotonicity is only necessary to create a copula of arbitrary dimension, so this was then further relaxed for d-variate Archimedean copula in further studies to include only the positivity of derivatives for $k = 1, 2, \ldots, d$ for a d-variate Archimedean copula; see discussion in McNeil and Nešlehová [2009], where it was shown that one only requires the necessary and sufficient conditions on the generator function to be a d-monotone function as given in Definition 28 in order to create Archimedean copula models up to dimension d.

Definition 28. ***D-Monotone Functions***
A real function $g(\cdot)$ is d-monotone in a range (a, b) for $a, b \in \mathbb{R}$ and $d \geq 2$ if

it is differentiable on this range up to order $d-2$ and the derivatives satisfy the condition that

$$(-1)^k g^{(k)}(x) \geq 0, \quad k = 0, 1, \ldots, d-2 \tag{3–35}$$

for any $x \in (a,b)$ and $(-1)^{d-2} g^{(d-2)}$ is non-increasing and convex in (a,b).

One can then conclude that a function ψ is said to generate an Archimedean copula if it satisfies the following properties.

Definition 29. ***Archimedean Generator***
An Archimedean generator is a continuous, decreasing function $\psi : [0,\infty] \to [0,1]$ that satisfies the following conditions:

1. $\psi : [0,\infty) \mapsto [0,1]$ *with* $\psi(0) = 1$ *and* $\lim_{t\to\infty} \psi(t) = 0$*;*
2. ψ *is a continuous function;*
3. ψ^{-1} *is given by* $\psi^{-1}(t) = \inf\{u : \psi(u) \leq t\}$*;*
4. ψ *is strictly decreasing on* $[0, \inf\{t : \psi(t) = 0\}] = [0, \psi^{-1}(0)]$.

McNeil and Nešlehová [2009] discuss the class of generators, denoted by ψ_∞, which represent all the generators for Archimedean copulae models that produce valid copula distributions in any dimension, that is, those that are completely monotone functions. In this context, they note two representations of such generators: the first based on the Bernstein–Widder theorem and the Laplace transform; and the second based on the Williamson d-transform. I discuss these two representations in the following subsections.

3.7.3 Archimedean Copula Generators and the Laplace Transform of a Non-Negative Random Variable

In understanding the first representation for the completely monotone generator, it will be instructive to first recall the theorem of Bernstein–Widder; see, for instance, a proof in Pollard et al. [1944] or Feller [1971]. This theorem links a completely monotone function to a Laplace transform representation.

Theorem 3.7.5. ***Bernstein–Widder Theorem***
Consider a real function $f(x)$ such that it satisfies

$$f(0) = f(0+), \quad (-1)^k f^{(k)}(x) \geq 0, \quad x \in (0, \infty), \forall k = 0, 1, \ldots. \tag{3–36}$$

Then the function $f(x)$ admits the following representation as a Laplace transform

$$f(x) = \int_0^\infty \exp(-xt) d\alpha(t) \tag{3–37}$$

for $x \geq 0$ and $\alpha(t)$ an increasing and bounded function.

For an Archimedean generator ψ, one can then use this result to link the existence of distributions in all dimensions to the range of complete monotonicity of the generator, see Proposition 3.7.6.

Proposition 3.7.6. ***Complete Monotonicity and Generator Support***
A generator ψ for an Archimedean copula belongs to the class of generators ψ_∞ if and only if it is completely monotone on $[0, \infty)$.

Remark 3.7.7. *One can see from the combination of Theorem 3.7.5 and Proposition 3.7.6 that a generator ψ of an Archimedean copula is completely monotone only when it is formed from the Laplace transform of a non-negative random variable Z. It can then be shown that the resulting Archimedean copula for such a generator $\psi \subset \psi_\infty$ in d-dimensions is given by the survival copula coming from the survival function, which is expressed via the generator of the l_1-norm ($||.||_1$) according to*

$$\begin{aligned}\overline{H}\left(x_1, x_2, \ldots, x_d\right) &= \psi\left(||\max\left(\boldsymbol{x}, \boldsymbol{0}\right)||_1\right) \\ &= \mathbb{E}\left[\exp\left(-||\max\left(\boldsymbol{x}, \boldsymbol{0}\right)||_1 Z\right)\right] \\ &= \mathbb{E}\left[\exp\left(-Z \sum_{i=1}^{d} \max\left(x_i, 0\right)\right)\right],\end{aligned} \tag{3–38}$$

which corresponds to a survival function of a random vector $\boldsymbol{X} = \frac{1}{Z}\boldsymbol{E}$ with $\boldsymbol{E}$ a vector of i.i.d. exponential random variables that are independent of Z; see discussion in McNeil and Nešlehová [2009].

Table 3.1: Generators and inverse Laplace transforms for several copulae from the Archimedean family

Family	ρ Range	Generator $\psi(x;\rho)$	Distribution of $\mathcal{L}^{-1}[\psi]$
Ali–Mikhail–Haq	$[0,1)$	$\frac{1-\rho}{\exp(x)-\rho}$	$(1-\rho)\rho^{k-1}, k \in \mathbb{J}$
Clayton	$(0,\infty)$	$(1+x)^{-1/\rho}$	$\Gamma(1/\rho,1)$
Frank	$(0,\infty)$	$-\frac{1}{\rho}\ln\left(e^{-x}(e^{-\rho}-1)+1\right)$	$\frac{(1-e^{-\rho})^k}{k\rho}, k \in \mathbb{J}$
Gumbel	$[1,\infty)$	$\exp\left(-x^{1/\rho}\right)$	$S_{\frac{1}{\rho}}\left(1,\cos\left(\frac{\pi}{2\rho}\right)^{\rho}, 0; S1\right)$
Joe	$[1,\infty)$	$1-(1-e^{-x})^{\frac{1}{\rho}}$	$(-1)^{k+1}\frac{(1/\rho)!}{k!(1/\rho-k)!}, k \in \mathbb{J}$.

Note: $S_\alpha(\beta,\gamma,\delta;S1)$ is the univariate α-stable distribution with S1 parametrization of Nolan.

One important result of this representation is the ability to simulate exactly Archimedean copula random variates, as discussed in Marshall and Olkin [1988].

Algorithm 1 Simulation from Archimedean Copula via Laplace Transform

1. Sample a random variable $V \sim F$, where the distribution F is given by the inverse Laplace transform of the generator ψ such that $F = \mathcal{L}^{-1}[\psi]$;

2. Sample d i.i.d. draws from a uniform distribution $U_i \sim Uniform(0,1)$ for $i \in \{1,2,\ldots,d\}$;

3. Construct via transformation the d-variate random vector $\boldsymbol{U} = (U_1,\ldots,U_d)$, which is drawn from the Archimedean copula characterized by generator ψ given by

$$X_i = \psi\left(-\frac{1}{V}\ln(U_i)\right), \quad i \in \{1,2,\ldots,d\}. \tag{3–39}$$

The results in Table 3.1 from Hofert [2010b, table 1] demonstrate examples of popular Archimedean copula models for which closed form distributions of such inverse Laplace transforms of the generator are known.

As noted in Hofert [2010b], it is then a trivial consequence to obtain other Archimedean copula model simulation schemes based on, for instance, those presented in Table 3.1 via exponential tilting results presented in Theorem 3.7.8.

Theorem 3.7.8. ***Exponential Tilting of Generator Inverse Laplace Transforms***
Consider an Archimedean copula generator ψ in the family of completely monotone Archimedean generators $\psi \in \psi_\infty$ with a known distribution for the inverse Laplace transform given by $F = \mathcal{L}^{-1}[\psi]$. Then, define a new generator $\widetilde{\psi}(x)$ in terms of $\psi(x)$ according to

$$\widetilde{\psi}(x) = \frac{\psi(x+h;\rho)}{\psi(h;\rho)}, \quad \forall x \in [0,\infty]. \tag{3–40}$$

Then the following holds:

- *$\widetilde{\psi}(x)$ is completely monotone on $x \in [0,\infty]$ and $\widetilde{\psi}(0) = 1$;*

- *The distribution of the inverse Laplace transform for the new generator $\widetilde{F} = \mathcal{L}^{-1}\left[\widetilde{\psi}(x)\right]$ is given in terms of the distribution F by*

$$\widetilde{F}(x) = \frac{1}{\psi(h)}\left(F(0) + \int_0^x \exp(-hu)dF(u)\right), \quad x \in [0,\infty). \tag{3–41}$$

- *If the distribution F admits a density f, the $\widetilde{F}$ admits the exponential tilted density given by*

$$\widetilde{f}(x) = \frac{1}{\psi(h)}\exp(-hx)f(x), \quad x \in [0,\infty). \tag{3–42}$$

3.7.4 Archimedean Copula Generators, l_1-Norm Symmetric Distributions and the Williamson Transform

The second representation developed in McNeil and Nešlehová [2009], which facilitates the simulation exactly of Archimedean copula random variates, utilizes the fact that the random vector discussed in Remark 3.7.7 given by $\boldsymbol{X} = \frac{1}{Z}\boldsymbol{E}$ can be re-represented by utilizing the fact that if one transforms the vector of i.i.d.

exponential random variables according to

$$\boldsymbol{S}_d = \frac{\boldsymbol{E}}{||\boldsymbol{E}||_1}, \tag{3–43}$$

then $\boldsymbol{S}_d$ will be distributed according to a Uniform distribution on the d-dimensional simplex given by the space $\mathcal{S}_d$

$$\mathcal{S}_d = \left\{ \boldsymbol{x} \in \mathbb{R}_+^d : ||\boldsymbol{x}||_1 = 1 \right\}. \tag{3–44}$$

In addition, since $\boldsymbol{S}_d$ and Z are independent, then one can write the random vector $\boldsymbol{X} = R\boldsymbol{S}_d$ with random variable R given by $R = \frac{1}{Z}||\boldsymbol{E}||_1$. The implications of this result for the transformed distribution indicates that the random vector $\boldsymbol{X}$ admits a representation in terms of a mixture of Uniform distributions on simplices.

The significance of this result is that although only completely monotone Archimedean generators will admit representations as survival copulae of random vectors following a particular frailty model, it is clear from the aforementioned result that Archimedean generators which are only d-monotone will produce representations as survival copulae of random vectors with l_1-norm symmetric distributions. As observed in McNeil and Nešlehová [2009], in the case of completely monotone generators of Archimedean copulae one could form a link between the Laplace transform of a particular frailty model and the generator via the Bernstein–Widder theorem. In the case of the d-monotone (not completely monotone) generator functions, one can form an analogous link between d-variate Archimedean copulae and the l_1-norm symmetric distributions via a special class of Mellin–Stieltjes integral transforms known as Williamson transforms; see Definition 30 and Williamson et al. [1956] and McNeil and Nešlehová [2009, proposition 3.1].

Definition 30. ***Williamson d-Transforms***

The Williamson transform of a positive random variable X with distribution F

is a real function on $[0, \infty)$ *given for any integer* $d \geq 2$ *by*

$$f(x) = \mathcal{W}_d\left[F_X(x)\right] = \int\limits_{(x,\infty)} \left(1 - \frac{x}{t}\right)^{d-1} dF(t) = \begin{cases} \mathbb{E}\left[\left(1 - \frac{x}{X}\right)_+^{d-1}\right], & if\ x > 0 \\ 1 - F(0), & if\ x = 0. \end{cases} \tag{3–45}$$

The Williamson d-transform $\mathcal{W}_d$ *will consist of real functions* f *on* $[0, \infty)$ *that are d-monotone on* $[0, \infty)$ *and satisfy boundary conditions that* $\lim_{x\to\infty} f(x) = 0$ *and* $f(0) = p$ *for* $p \in [0, 1]$*. Furthermore, any non-negative random variable's distribution function can be uniquely defined by its Williamson d-transform* $f = \mathcal{W}_d\left[F_X(x)\right]$ *such that* $F_X(x) = \mathcal{W}_d^{-1}\left[f(x)\right]$ *with the inverse given by*

$$F_X(x) = \mathcal{W}_d^{-1}\left[f(x)\right] = 1 - \sum_{k=0}^{d-2} \frac{(-1)^k x^k f^{(k)}(x)}{k!} - \frac{(-1)^{d-1} x^{d-1} f_+^{(d-1)}(x)}{(d-1)!}. \tag{3–46}$$

Remark 3.7.9. *It was therefore observed in McNeil and Nešlehová [2009] that the d-monotone Archimedean copula generators* ψ *will consist of Williamson d-Transforms of distribution functions* F *from non-negative loss random variables that satisfy* $F(0) = 0$*.*

In addition, in Williamson et al. [1956], the result in Proposition 3.7.10 completes the link between l_1-norm symmetric distributions and Archimedean copulae; see McNeil and Nešlehová [2009].

Proposition 3.7.10. ***l_1-Norm Symmetric Distributions and Williamson d-Transforms***

Consider the d-dimensional random vector $\boldsymbol{X}$ *with representation as a* l_1*-norm symmetric distribution* $\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{S}_d$ *with radial distribution* F_R*. Then one has the following relationship between the multivariate survival function of* $\boldsymbol{X}$ *and the Williamson d-transform:*

- $\overline{H}(\boldsymbol{x})$ *is given by*

$$\overline{H}(\boldsymbol{x}) = \mathcal{W}_d\left[F_R\left(||\max\left(\boldsymbol{x}, \boldsymbol{0}\right)||_1\right)\right] + F_R(0)\mathbb{I}\left[\boldsymbol{x} < \boldsymbol{0}\right], \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{3–47}$$

If in addition $F_R(0) = 0$, then $\boldsymbol{X}$ has an Archimedean survival copula with generator given by $\psi = \mathfrak{W}_d\left[F_R(r)\right]$;

- *The density $\boldsymbol{X}$ exists if and only if R has a density, which is given with regard to the density of R denoted $f_R(r)$ by*

$$h\left(||\boldsymbol{x}||_1\right) = \Gamma(d)||\boldsymbol{x}||^{1-d} f_R\left(||\boldsymbol{x}||_1\right). \tag{3–48}$$

- *If $\mathbb{P}\mathrm{r}\left[\boldsymbol{X} = \boldsymbol{0}\right] = 0$, then one has that $R \stackrel{d}{=} ||\boldsymbol{X}||_1$ and $\boldsymbol{S}_d \stackrel{d}{=} \boldsymbol{X}/||\boldsymbol{X}||_1$.*

An important result of this simplectic representation is the ability to simulate exactly Archimedean copula random variates, as discussed in McNeil and Nešlehová [2009] and shown in Section 3.7.4.

Algorithm 2 Simulation from Archimedean Copula via Williamson ***d***-Transform

1. Sample a random variable $R \sim F_R$ where the distribution F_R is given by the inverse Williamson d-transform of the generator ψ such that $F_R = \mathfrak{W}_d^{-1}\left[\psi\right]$, which is given by

$$F_R(x) = \mathfrak{W}_d^{-1}\left[f(x)\right] = 1 - \sum_{k=0}^{d-2} \frac{(-1)^k x^k \psi^{(k)}(x)}{k!} - \frac{(-1)^{d-1} x^{d-1} \psi_+^{(d-1)}(x)}{(d-1)!}. \tag{3–49}$$

2. Sample independently of R the random vector $\boldsymbol{S}_d$ given by transformation of d i.i.d. exponential random variates with $E_i \sim Exp(1)$ such that

$$\boldsymbol{S}_d \stackrel{d}{=} \left(\frac{E_1}{\sum_{i=1}^d E_i}, \ldots, \frac{E_d}{\sum_{i=1}^d E_i}\right). \tag{3–50}$$

3. Construct via transformation the d-variate random vector $\boldsymbol{U} = (U_1, \ldots, U_d)$, which is drawn from the Archimedean copula characterized by generator ψ given by

$$U_i = \psi\left(R\frac{E_i}{\sum_{i=1}^d E_i}\right), \quad i \in \{1, 2, \ldots, d\}. \tag{3–51}$$

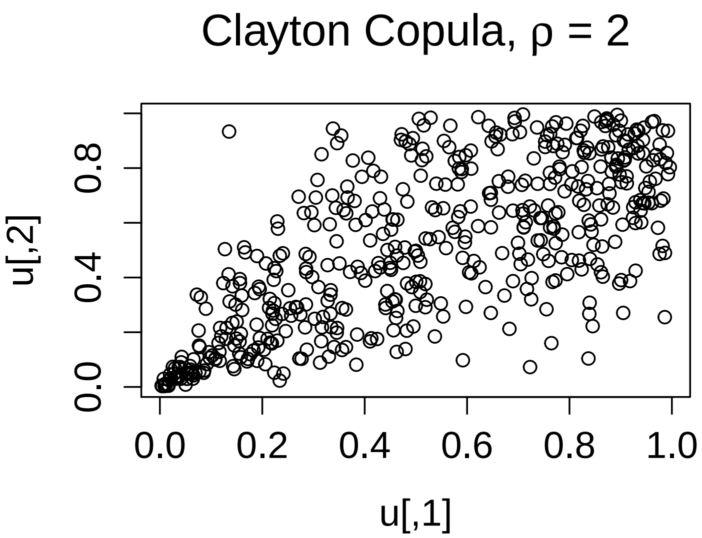


Figure 3.6: Scatterplot of 500 random samples from a Clayton copula with $\rho = 2$.

3.7.5 One-parameter Archimedean Members

In this section, I describe three of the one parameter multivariate Archimedean family copula models which have become popular model choices and are widely used for estimation. This is primarily due to there directly interpretable features. I select these three component members, the Clayton, Frank and Gumbel models, for the mixture models since they each contain differing tail dependence characteristics.

Clayton provides lower tail dependence, as seen in the random sample in Figure 3.6 and the copula density plot in Figure 3.7. The Gumbel copula provides upper tail dependence, as seen in the random sample in Figure 3.10 and the copula density plot in Figure 3.11. The Frank copula also provides dependence in the unit cube with elliptical contours with semi-major axis oriented at either $\pi/4$ or $3\pi/4$ depending on the sign of the copula parameter in the estimation. Therefore the Frank model component will allow one to capture parsimoniously potential negative dependence relationships between the currencies in the portfolio under study, as seen in Figure 3.8 and the copula density plot in Figure 3.9.

Formulas for these copulae, as well as their respective generators, inverse

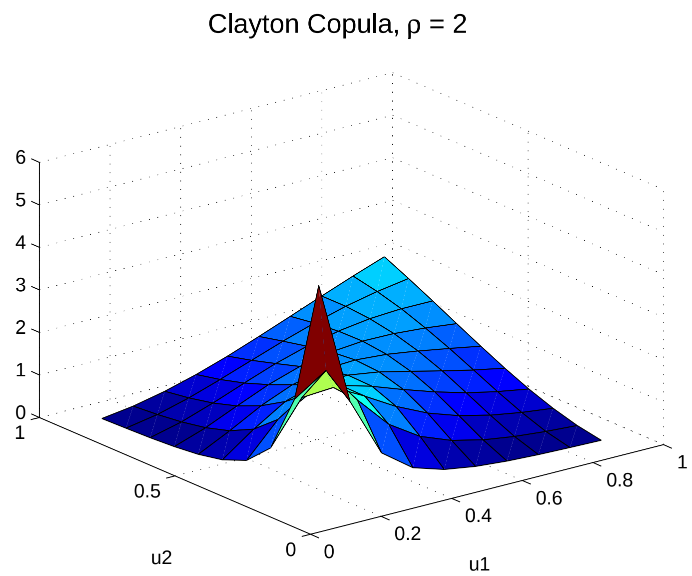


Figure 3.7: Density plot of a Clayton copula with $\rho = 2$.

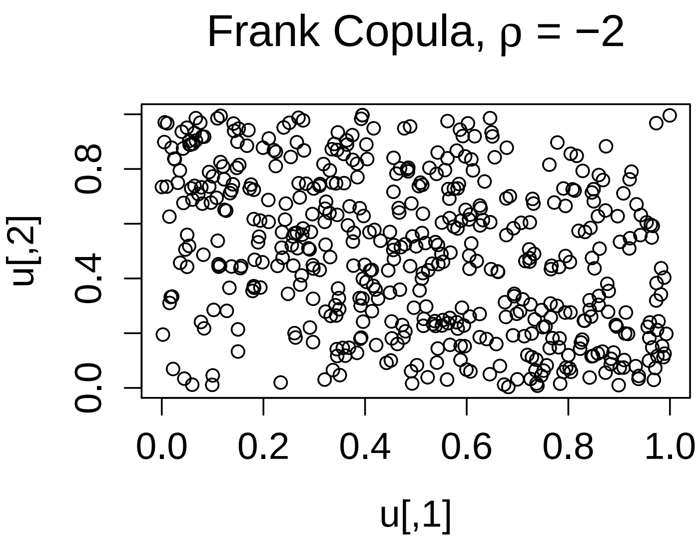


Figure 3.8: Scatterplot of 500 random samples from a Frank copula with $\rho = -2$. The variables show negative dependence here.

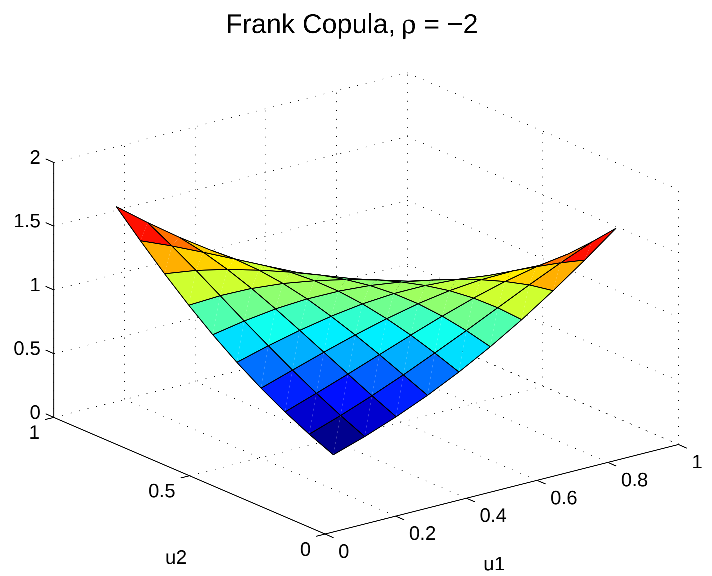


Figure 3.9: Density plot of a Frank copula with $\rho = 2$.

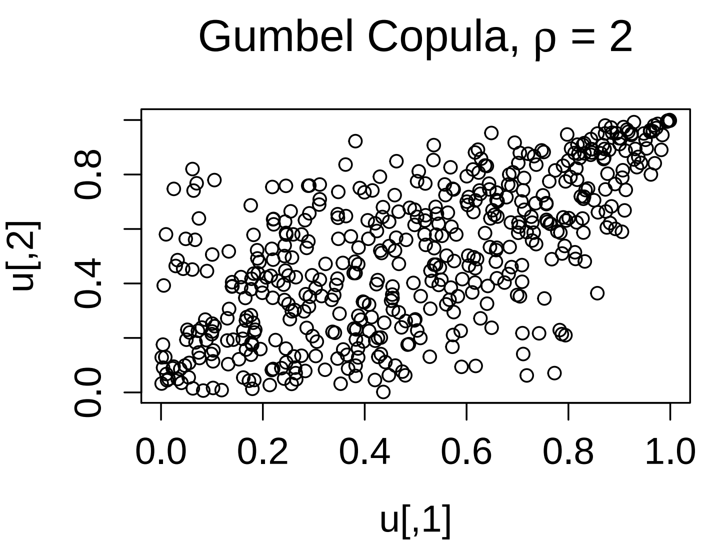


Figure 3.10: Scatterplot of 500 random samples from a Gumbel copula with $\rho = 2$.

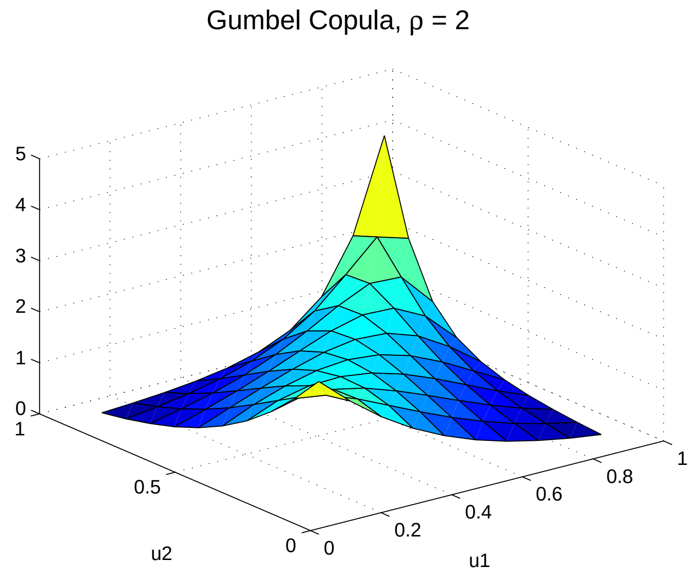


Figure 3.11: Density plot of a Gumbel copula with $\rho = 2$.

generators and the d-th derivatives of their generators (required for the density evaluation) are given in Table 3.3. The explicit formulas for the d-th derivatives for all of the copulae in Table 3.3 were derived in Hofert et al. [2012].

The exact non-linear transformations between the copula parameter ρ and Kendall's rank correlation τ for the Clayton, Frank and Gumbel copulae can be seen in Table 3.2.

Table 3.2: Kendall's tau and tail dependence coefficients.

Family	$\boldsymbol{\tau}$	$\boldsymbol{\lambda_L}$	$\boldsymbol{\lambda_U}$
Clayton	$\frac{\rho}{\rho+2}$	$2^{-\frac{1}{\rho}}$	0
Frank	$1 + \frac{4D_1{}^{1}(\rho)-1}{\rho}$	0	0
Gumbel	$\frac{(\rho-1)}{\rho}$	0	$2 - 2^{\frac{1}{\rho}}$

[1] $D_1 = \int_0^\rho \frac{t}{exp(t)-1} dt/\rho$ is the *Debye function of order one*.

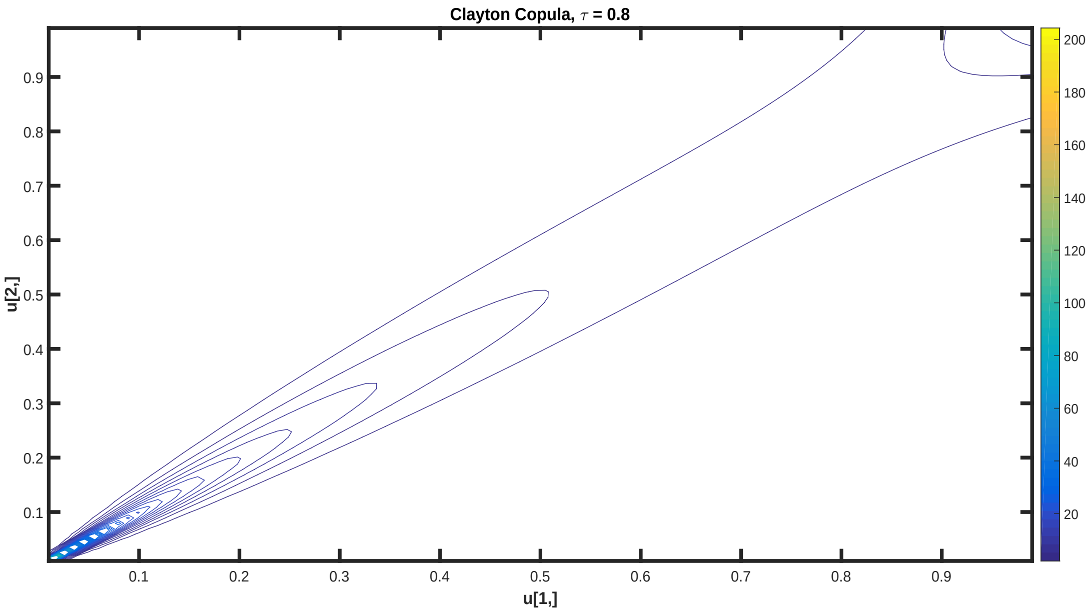


Figure 3.12: Contour plot of Clayton copula with Kendall's $\tau = 0.8$ and copula parameter $\rho = 8$.

Figures 3.12 and 3.13 illustrate the non-linear relationship between the Clayton copula parameter and the Kendall's Tau measure of dependence. Figure 3.12 shows a contour plot for a Clayton copula with $\rho = 8$ and thus $\tau = 0.8$, whereas Figure 3.13 shows a contour plot for a Clayton copula with $\rho = 38$ and thus $\tau = 0.95$. For such a large increase in the copula parameter there is a much smaller increase in Kendall's Tau and also the observable dependence between the variables, as shown by the contour plots, is more similar than perhaps one would expect.

3.7.6 Archimax Copulae

Recently there has been a growing interest in developing archimedean copula models with distortion features, based on the works of Genest and Rivest [2001] and Morillas [2005] which explore ways of distorting a given copula to obtain a new copula with additional features. For instance, they explored the multivariate probability integral transform and its application in distorting existing copula models to obtain new copula models.

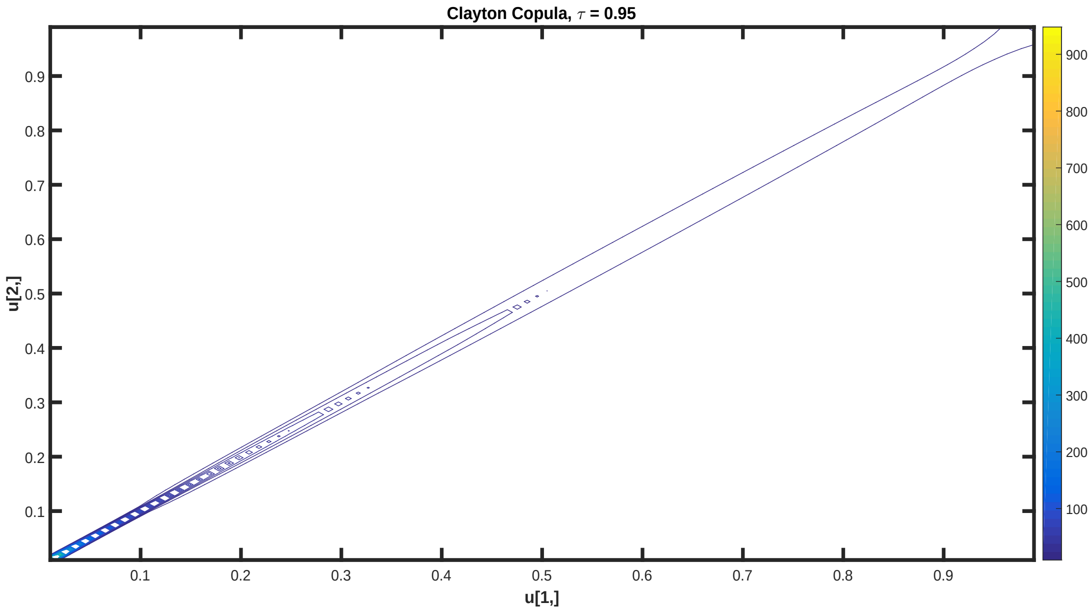


Figure 3.13: Contour plot of Clayton copula with Kendall's $\tau = 0.95$ and copula parameter $\rho = 38$.

For instance in Morillas [2005] they study under what conditions the following distortion copula transform produces a valid copula where $g(\cdot)$ is assumed to be a strictly increasing and continuous function from $[0, 1]$ to $[0, 1]$ such that

$$C_g\left(u_1, \ldots, u_d\right) = g^{-1}\left(C\left(g(u_1), \ldots, g(u_d)\right)\right) \tag{3–52}$$

is a valid distorted copula.

Definition 31. ***Tilted and Distorted Copulae***
Define the function g to be some distortion function, such that $g : [0, 1] \mapsto [0, 1]$ and is defined according to

$$g(t) = \exp[-\varphi(t)], \tag{3–53}$$

where φ is for instance an Archimedean generator function. Now denote C as a base copula that is to be distorted to create a new copula, then

$$C_g(u_1, ..., u_d) = g^{-1}\left(C(g(u_1), , g(u_d))\right) \tag{3–54}$$

is a copula known as the distortion of C.

Several examples of bivariate and multivariate distorted copula models have begun to be studied. Though the current emphasis has focused on their specification and little is know about their properties such as tail dependence features and other concordance measure features they may exhibit.

Here I consider two examples based on ideas developed in Charpentier et al. [2014]. I begin with a bivariate archimax copula given by a parametric model of the distributional form

$$C_{\phi,\mathcal{A}}\left(u_1,u_2\right)=\varphi\left[\left\{\varphi^{-1}\left(u_1\right)+\varphi^{-1}\left(u_2\right)\right\}\mathcal{A}\left\{\frac{\varphi^{-1}\left(u_1\right)}{\varphi^{-1}\left(u_1\right)+\varphi^{-1}\left(u_2\right)}\right\}\right] \tag{3–55}$$

where $\mathcal{A}:[0,1]\rightarrow[1/2,1]$ and $\varphi:[0,\infty)\rightarrow[0,1]$ such that

1. $\mathcal{A}$ is convex and for all $t\in[0,1]$ one has $\max(t,1-t)\leq\mathcal{A}(t)\leq 1$.
2. φ is convex, decreasing and such that $\varphi\left(0\right)=1$ and $\lim_{x\rightarrow\infty}\varphi\left(x\right)=0$ with convention that $\varphi^{-1}\left(0\right)=\inf\left\{x\geq 0:\varphi\left(x\right)=0\right\}$.

Two special cases arise from this model:

- If $\mathcal{A}=1$ one recovers the well known family of Archimedean copula dependence models.
- If $\varphi\left(t\right)=\exp(-t)$ then one recovers the extreme-value copula.

As in Capéraà et al. [2000], I utilise in this study the choice of function

$$\mathcal{A}(t)=\left\{t^{1/\alpha}+(1-t)^{1/\alpha}\right\}^{\alpha}A^{\alpha}\left\{\frac{t^{1/\alpha}}{t^{1/\alpha}+(1-t)^{1/\alpha}}\right\}. \tag{3–56}$$

where $A(\cdot)$ is the Pickands EVT dependence function given by

$$A(t)=1-\min\left\{\beta t,\alpha(1-t)\right\} \tag{3–57}$$

for some parameters $\alpha,\beta\in[0,1]$.

Then I consider a d-variate distortion copula in the Archimax family, examples of such extensions include the works of Bacigál and Mesiar [2012], Mesiar and StupňAnová [2013] and Charpentier et al. [2014]. One possible version of such a d-dimensional Archimax copula is defined through the use of a distortion function based on the stable tail function that must satisfy the properties in Definition 32, see details in Charpentier et al. [2014]. In general a multivariate stable tail function is obtained via the multivariate Generalized Extreme Value distribution G via

$$\log G(x_1, \ldots, x_d) = \mu([\mathbf{0}, \boldsymbol{\infty})[\mathbf{0}, \boldsymbol{x}]), \forall \boldsymbol{x} \in \mathbb{R}_+^d \tag{3–58}$$

such that G is the limiting distribution (max domain of attraction) of the normalized component wise maxima of

$$\boldsymbol{X}_{n:n} = (\max\{X_{1,i}\}, \ldots, \max\{X_{d,i}\}) \tag{3–59}$$

and then the stable tail function is obtained via measure μ or distribution G according to the following

$$l(x_1, \ldots, x_d) = \mu([\mathbf{0}, \boldsymbol{\infty})[\mathbf{0}, \boldsymbol{x}^{-1}]), \forall \boldsymbol{x} \in \mathbb{R}_+^d \tag{3–60}$$

$$-\log G(x_1, \ldots, x_d) = l\left(-\log G_1(x_1), \ldots, -\log G_d(x_d)\right) \tag{3–61}$$

Definition 32. ***Stable Tail Function*** l
A function $l : [0, \infty)^d \mapsto [0, \infty)$ is a d-dimensional stable tail dependence function if and only if it satisfies the following properties:

1. *function l is homogeneous of degree $\lambda = 1$ which means that*

$$l(\lambda x_1, \ldots, \lambda x_d) = \lambda l(x_1, \ldots, x_d), \quad \forall \lambda \in [0, \infty). \tag{3–62}$$

2. *The function l must produce for all $x_1, \ldots, x_d \in [0, \infty)$ that*

$$\overline{G}_l(x_1, \ldots, x_d) = \left[\max\{0, 1 - l(x_1, \ldots, x_d)\}\right]^{d-1} \tag{3–63}$$

defines a d-dimensional survival function with $\mathcal{B}(1, d-1)$ margins.

An example of such a stable tail function involves the transformation of a d-variate extreme value copula C^{EVT} given by

$$l(x_1, x_2, \ldots, x_d) = -\ln\left\{C^{EVT}\left(\exp(-x_1), \exp(-x_2), \ldots, \exp(-x_d)\right)\right\} \quad (3\text{–}64)$$

and we have for instance the Gumbel extreme-value copula (symmetric logistic model) producing for parameter $\theta \geq 1$ the function

$$l_\theta(x_1, x_2, \ldots, x_d) = \left(x_1^\theta + \ldots + x_d^\theta\right)^{1/\theta}. \quad (3\text{–}65)$$

One can then combine this with an Archimedean generator $\varphi(x)$ of an Archimedean copula to produce the resulting family of d-dimensional Archimax copula, given by

$$C_{\varphi,l}(u_1, \ldots, u_d) - \varphi \circ l\left(\varphi^{-1}(u_1), \varphi^{-1}(u_2), \ldots, \varphi^{-1}(u_d)\right). \quad (3\text{–}66)$$

where $\circ$ is the composite function.

One can use this type of representation to develop the d-dimensional extension of the bivariate example above, for instance using the Archimax copula structure of [Charpentier et al., 2014, Corollary 6.3], which is defined as follows:

$$C_{l^*,\mathcal{A}^*}\left(u_1, u_2, \ldots, u_d\right) = \exp\left[\mathcal{A}^*\left(t_1, \ldots, t_{d-1}\right)\sum_{i=1}^{d}\ln(u_i)\right] \quad (3\text{–}67)$$

where $\mathcal{A}^*$ is given by the function

$$\mathcal{A}^* = l^\alpha\left(t_1^{1/\alpha}, \ldots, 1 - \sum_{k=1}^{d-1} t_k^{1/\alpha}\right) \quad (3\text{–}68)$$

where $t_k = |\ln(u_k)| / \left\{\sum_{i=1}^{d}\ln(u_i)\right\}$.

In the next two subsections, I focus on a special class of copula distortion as specified by inner and outer power transforms.

3.7.7 Two-parameter Archimedean Members via Outer Power Transforms

In this section, I consider more flexible generalizations of the single parameter Archimedean members discussed above. To achieve these generalizations I consider the outer-power transforms of the Clayton, Frank and Gumbel members, which is based on a result in Feller [1971].

Definition 33. ***Outer Power Copula***
The copula family generated by $\tilde{\psi}(t) = \psi(t^{\frac{1}{\beta}})$ is called an outer power family, where $\beta \in [1, \infty)$ and $\psi \in \Psi_\infty$ (the class of completely monotone Archimedean generators).

The proof of this follows from Feller [1971], i.e. the composition of a completely monotone function with a non-negative function that has a completely monotone derivative is again completely monotone. Such copula model transforms were also studied in Nelsen [1997], where they are referred to as a beta family associated with the inverse generator ψ^{-1}.

As has been noted above, in performing the estimation of these transformed copula models via likelihood based inference it will be of great benefit to be capable of performing evaluation pointwise of the copula densities. In the case of the outer power transformed models, this will require the utilization of a specific multivariate chain rule result widely known as the Faà di Bruno's Formula, see Faa di Bruno [1857] and discussions in for example Constantine and Savits [1996] and Roman [1980]. To understand how such a result is required consider the following remark.

Remark 3.7.11. *The generator derivatives for the outer power transforms can be calculated using the base generator derivatives and the following multi-dimensional extension to the chain rule for the outer power versions. The densities for the outer power copulae in Table 3.3 can thus be calculated using equation 3–32.*

Before stating Faà di Bruno's Formula for differentiation of multivariate composite functions via a generalized chain rule, it will be convenient notationally to present such results with respect to Bell polynomials. Therefore we recall the definition of such polynomials below, which are widely used in combinatorics analysis, see Mihoubi [2008] for details.

Definition 34. ***Bell Polynomial***
The Bell polynomial with arguments n and k is given by

$$B_{n,k}(x_1, x_2, \dots, x_{n-k+1}) = \sum \frac{n!}{j_1! j_2! \cdots j_{n-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \cdots \left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}} \tag{3–69}$$

where the sum is taken over all sequences j_1, j_2, j_{n-k+1} of non-negative integers such that $j_1 + j_2 + \cdots = k$ and $j_1 + 2j_2 + 3j_3 + \cdots = n$.

These polynomials are then utilised to simplify the expressions for the differentiation of multivariate composite functions in Faà di Bruno's Formula as detailed next.

Faà di Bruno's Formula: Riordan [1946]
If f and g are functions with a sufficient number of derivatives, then

$$\frac{d^n}{dx^n} f(g(x)) = \sum_{k=0}^{n} f^{(k)}(g(x)) \cdot B_{n,k}\left(g'(x), g''(x), ..., g^{n-k+1}(x)\right) \tag{3–70}$$

where $B_{n,k}$ are the Bell polynomials, defined above.

3.7.8 Two-parameter Archimedean Members via Inner Power Transforms

I now consider the inner-power transforms of the Clayton, Frank and Gumbel copulae. The inner power copula is defined as follows.

Definition 35. ***Inner Power Copula***
The copula family generated by $\tilde{\psi}(t) = \psi^{\frac{1}{\alpha}}(t)$ is called an inner power family, where $\alpha \in (0, \infty)$ and $\psi \in \Psi_\infty$ (the class of completely monotone Archimedean generators).

Inner power transforms produce a family of generators associated with the base generator, e.g. the Clayton generator is the inner power transform of the base generator $\psi(t) = (1+t)^{-1}$. The lower tail dependence of the transformed copula is $\lambda_L^{1/\alpha}$, whilst the upper tail dependence remains unchanged.

Inner power copula model transforms were also studied in Nelsen [1997], where they are referred to as an alpha family associated with the inverse generator ψ^{-1}.

3.7.9 Mixtures of Archimedean Copulae

In order to provide flexibility to the possible dependence features available for the currency portfolios, I decided to utilize mixtures of copula models. The advantage of this approach is that I can consider asymmetric dependence relationships in the upper tails and the lower tails in the multivariate model. In addition I can perform a type of model selection purely by incorporating into the estimation the mixture weights associated with each dependence hypothesis. That is the data can be utilised to decide the strength of each dependence feature as interpreted directly through the estimated mixture weight attributed to the feature encoded in the particular mixture component from the Archimedean family.

In particular I have noted that mixture copulae can be used to model asymmetric tail dependence, i.e. by combining the one-parameter or two-parameter families discussed above or indeed by any combination of copulae. This is possible since a linear convex combination of 2 copulae is itself a copula, see discussions on this result in Nelsen [2006].

Definition 36. ***Mixture Copula***
A mixture copula is a linear weighted combination of copulae of the form:

$$C_M(\boldsymbol{u};\Theta) = \sum_{i=1}^{N} \lambda_i C_i(\boldsymbol{u};\theta_i) \tag{3–71}$$

where $0 \leq \lambda_i \leq 1 \;\; \forall i = 1, ..., N$ *and* $\sum_{i=1}^{N} \lambda_i = 1$

Thus we can combine a copula with lower tail dependence, a copula with positive or negative dependence and a copula with upper tail dependence to produce a more flexible copula capable of modelling the multivariate log returns of forward exchange rates of a basket of currencies. For this reason in this analysis I will use the Clayton-Frank-Gumbel (C-F-G) mixture model. In addition to the C-F-G mixture model I will also investigate a mixture of outer power versions of the base copula Clayton, Frank and Gumbel.

Remark 3.7.12. *We note that the tail dependence of a mixture copula can be obtained as the linear weighted combination of the tail dependence of each component in the mixture weighted by the appropriate mixture weight, as discussed in for example Nelsen [2006] and Peters et al. [2014]*

Table 3.3: Archimedean copula generator functions, inverse generator functions and generator function d-th derivatives.

Family	$\boldsymbol{\psi}$	$\boldsymbol{\psi^{-1}}$	$\boldsymbol{(-1)^d\psi^{(d)}}$
Clayton	$(1+t)^{-\frac{1}{\rho}}$	$(s^{-\rho}-1)$	$\frac{\Gamma\left(d+\frac{1}{\rho}\right)}{\Gamma\left(\frac{1}{\rho}\right)}(1+t)^{-\left(d+\frac{1}{\rho}\right)}$
OP-Clayton	$\left(1+t^{\frac{1}{\beta}}\right)^{-\frac{1}{\rho}}$	$(s^{-\rho}-1)^{\beta}$	$\frac{\sum_{k=1}^{d}{}^{1}a_{dk}^{G}(\frac{1}{\beta})\frac{\Gamma\left(k+\frac{1}{\rho}\right)}{\Gamma\left(\frac{1}{\rho}\right)}\left(1+t^{\frac{1}{\beta}}\right)^{-\left(k+\frac{1}{\rho}\right)}\left(t^{\frac{1}{\beta}}\right)^{k}}{t^{d}}$
Frank	$-\frac{1}{\rho}\ln\left[1-e^{-t}(1-e^{-\rho})\right]$	$-\ln\frac{e^{-s\rho}-1}{e^{-\rho}-1}$	$\frac{1}{\rho}{}^{2}Li_{-(d-1)}\{(1-e^{-\rho})e^{-t}\}$
OP-Frank	$-\frac{1}{\rho}\ln\left[1-e^{-t^{\frac{1}{\beta}}}(1-e^{-\rho})\right]$	$\left[-\ln\frac{e^{-s\rho}-1}{e^{-\rho}-1}\right]^{\beta}$	$\frac{\sum_{k=1}^{d}a_{dk}^{G}\left(\frac{1}{\beta}\right)\frac{1}{\rho}Li_{-(k-1)}\left\{(1-e^{-\rho})e^{-t^{\frac{1}{\beta}}}\right\}\left(t^{\frac{1}{\beta}}\right)^{k}}{t^{d}}$
Gumbel	$e^{-t^{\frac{1}{\rho}}}$	$(-\ln s)^{\rho}$	$\frac{\psi_{\rho}(t)}{t^{d}}{}^{3}P_{d,\frac{1}{\rho}}^{G}\left(t^{\frac{1}{\rho}}\right)$
OP-Gumbel	$e^{-t^{\frac{1}{\beta\rho}}}$	$(-\ln s)^{\rho\beta}$	$\frac{\sum_{k=1}^{d}a_{dk}^{G}\left(\frac{1}{\beta}\right)\frac{\psi_{\rho}\left(t^{\frac{1}{\beta}}\right)}{t^{\frac{k}{\beta}}}P_{k,\frac{1}{\rho}}^{G}\left(t^{\frac{1}{\rho\beta}}\right)\left(t^{\frac{1}{\beta}}\right)^{k}}{t^{d}}$

Remark 3.7.13. *The densities for the one-parameter copulae in Table 3.3 can be calculated using equation 3–32. For details of the results contained in this table see Hofert et al. [2012].*

[1] $a_{dk}^{G}(\frac{1}{\rho})=\frac{d!}{k!}\sum_{i=1}^{k}\binom{k}{i}\binom{i/\rho}{d}(-1)^{d-i}\ ,\quad k\in 1,...,d$
[2] $Li_{s}(z)=\sum_{k=1}^{\infty}\frac{z^{k}}{k^{s}}$
[3] $P_{d,\frac{1}{\rho}}^{G}\left(t^{\frac{1}{\rho}}\right)=\sum_{k=1}^{d}a_{dk}^{G}\left(\frac{1}{\rho}\right)(t^{\frac{1}{\rho}})^{k}$

3.8 Estimation Methods for Copulae

Given a family of copula models there are a number of possible approaches available to estimate the parameters. Charpentier et al. [2007] provide a nice overview of the theoretical and practical issues that need to be considered when faced with such a task. Here, I will first present the maximum likelihood estimation (MLE) approach, which is commonly used in the literature. Then the expectation-maximisation (EM) algorithm will be discussed as an alternative approach for the mixture copula.

3.8.1 Maximum Likelihood Estimation

Maximum likelihood estimation is based on the following theory. Given realizations $\boldsymbol{u}_i$, $i \in \{1, \dots, n\}$, of a random sample $\boldsymbol{U}_i$, $i \in \{1, \dots, n\}$, from the copula C, the likelihood is defined as follows:

$$L(\theta; \boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \prod_{i=1}^{n} c_\theta(\boldsymbol{u}_i) \tag{3–72}$$

The log likelihood is thus defined as:

$$l(\theta; \boldsymbol{u}_1, \dots, \boldsymbol{u}_n) = \sum_{i=1}^{n} l(\theta; \boldsymbol{u}_i) = \log c_\theta(\boldsymbol{u}_i), \tag{3–73}$$

which in the case of archimedean copulae is given by

$$\log c_\theta(\boldsymbol{u}_i) = \log\big((-1)^d \psi_\theta(d)(t_\theta(\boldsymbol{u}))\big) + \sum_{j=1}^{d} \log(-(\psi_\theta)'(u_{ij})). \tag{3–74}$$

where $t(\boldsymbol{u}) = \psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)$.

The maximum likelihood estimator $\hat{\theta}_n = \hat{\theta}_n(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ can thus be found by solving the optimization problem

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}}\, l(\theta; \boldsymbol{u}_1, \dots, \boldsymbol{u}_n). \tag{3–75}$$

where the optimization is typically done numerically.

Assuming the derivatives to exist, the score function is defined as

$$s_\theta(\boldsymbol{u}) = \nabla l(\theta; \boldsymbol{u}) = \left(\frac{\partial}{\partial \theta_1} l(\theta; \boldsymbol{u}), \ldots, \frac{\partial}{\partial \theta_p} l(\theta; \boldsymbol{u}) \right)^T \tag{3–76}$$

and the Fisher information is

$$I(\theta) = \mathbb{E}_\theta \left[s_\theta(\boldsymbol{U}) s_\theta(\boldsymbol{U})^T \right] = \mathbb{E}_\theta \left[\left(\frac{\partial}{\partial \theta_i} l(\theta; \boldsymbol{u}) \frac{\partial}{\partial \theta_j} l(\theta; \boldsymbol{u}) \right)_{i,j \in \{1,\ldots,p\}} \right] \tag{3–77}$$

for $\boldsymbol{U} \sim C$.

Under regularity conditions (see for example [Serfling, 2009, pp. 144]), the following results hold.

Theorem 3.8.1. ***(Strong) Consistency of Maximum Likelihood Estimators***

$$\hat{\theta}_n = \hat{\theta}_n(\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n) \xrightarrow[a.s.]{P} \theta_0 \quad as\ n \to \infty. \tag{3–78}$$

Theorem 3.8.2. ***Asymptotic Normality of Maximum Likelihood Estimators***

$$\sqrt{n}\, I(\theta_0)^{1/2} (\hat{\theta}_n - \theta_0) \xrightarrow{d} N(\mathbf{0}, I_p), \tag{3–79}$$

where I_p denotes the identity matrix in $\mathbb{R}^{p \times p}$.

3.8.2 Expectation-Maximisation

In order to estimate the parameters in a mixture copula model it is interesting to also consider the expectation algorithm, as introduced in Dempster et al. [1977] and correctly proven to converge in Wu [1983]. In the case of a mixture copula model the mixture weights λ_j are assumed to be latent unobserved variables. Thus a two stage iteration procedure can be applied as follows:

The framework of copula-based finite mixture models utilising the expectation-maximisation algorithm is explored in Kosmidis and Karlis [2014]. The authors show that the use of copulae in model-based clustering offers two direct advantages over current methods: i) the appropriate choice of copulae provides the ability

Algorithm 3 EM Algorithm for Mixture Copulae

1. Initialise the copula parameters $\boldsymbol{\theta}^{(0)}$ and the copula weights $\lambda_j^{(0)}$ $(j = 1, \ldots, k)$.

2. Iterate until some convergence criterion is satisfied:

   (a) *E-step:* Calculate

$$w_{ij}^{(l+1)} = \frac{\lambda_j^{(l)} c_j(\boldsymbol{u}_i; \boldsymbol{\theta}_j^{(l)})}{\sum_{j=1}^{k} \lambda_j^{(l)} c_j(\boldsymbol{u}_i; \boldsymbol{\theta}_j^{(l)})} \quad (i = 1, \ldots, n; \quad j = 1, \ldots, k)\,.$$

   (b) *M-step 1:* Set $\lambda_j^{(l+1)} = \sum_{i=1}^{n} w_{ij}^{(l+1)}/n$ $(j = 1, \ldots, k)$.

   (c) *M-step 2:* Maximise w.r.t $\boldsymbol{\theta}$

$$\boldsymbol{\theta}^{(l+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{n} \log \sum_{j=1}^{k} w_{ij}^{(l+1)} \left\{c_j(\boldsymbol{u}_i; \boldsymbol{\theta}_j)\right\}\,,$$

to obtain a range of exotic shapes for the clusters, and ii) the explicit choice of marginal distributions for the clusters allows the modelling of multivariate data of various modes (either discrete or continuous) in a natural way.

Chapter 4

Currency Carry Trade Literature Review

In this chapter, the forward premium puzzle is presented and the literature surrounding the puzzle and the associated currency carry trade is reviewed. The novel approach of analysing both individual tail thickness and joint tail dependence, as proposed in this thesis, is then discussed.

4.1 The Forward Premium Puzzle

This phenomenon introduced initially by Hansen and Hodrick [1980, 1983]; Fama [1984]; Engel [1984] is directly linked to the arbitrage relation existing between the spot and the forward prices of a given currency, namely the Covered Interest Parity. This relation states that the price of a forward rate can be expressed according to the relationship:

$$F_t^T = e^{(r_t - r_t^f)(T-t)} S_t \tag{4–1}$$

where F_t^T and S_t denote respectively the forward and the spot prices at time t. While r_t and r_t^f represent respectively the local risk free rate[1] and the foreign risk free rate. I denote by T the maturity of the forward contract considered.

[1] I mean by local risk free rate the interest rate prevailing in the reference country which would be for instance the dollar for an American investor.

It is worth emphasizing that under the absence of an arbitrage hypothesis, this relation is directly resulting from the replication of the forward contract payoff using a self financed strategy. Moreover, it has been demonstrated empirically the validity of this arbitrage relation in the currency market Juhl et al. [2006]; Akram et al. [2008] when we consider daily data. The highly unusual period following the onset of the financial crisis in August 2007 saw large deviations from CIP due to the funding constraints of arbitrageurs and uncertainty about counterparty risk, see Coffey et al. [2009] for a thorough analysis of this period, though this was an exceptional case and typically CIP holds. In this thesis the associated concept of the Uncovered Interest Rate Parity (UIP) condition is investigated.

Definition 4.1.1. *Uncovered Interest Parity (UIP)*

The uncovered interest parity states that under the historical probability distribution the expected change in the currency spot rates equals the differential of interest rates such that:

$$E\left[\frac{S_T}{S_t}\middle|\mathcal{F}_t\right] = e^{\left(r_{t,T}-r^f_{t,T}\right)(T-t)}$$

where $\mathcal{F}_t$ is the filtration associated to the stochastic process S_t. Furthermore, if one assumes that the covered interest parity described in (equation 4–1) is holding, which is commonly admitted in the literature for daily data (Juhl et al. [2006]; Akram et al. [2008]), one can then rewrite the previous relation accordingly:

$$E\left[\frac{S_T}{S_t}\middle|\mathcal{F}_t\right] = \frac{F_t^T}{S_t}$$

which means that according to the UIP, and admitting that the CIP holds, if until the forward contract's maturity date the associated spot rate varies on average more or less than its initial difference with the forward contract's price, an abnormal profit can be captured and the UIP condition is violated.

4.2 Currency Carry Trade

Numerous empirical studies (Hansen and Hodrick [1980]; Fama [1984]; Engel [1996]; Backus et al. [2001]; Lustig and Verdelhan [2007]; Brunnermeier et al. [2008];

Burnside et al. [2011]; Christiansen et al. [2011]; Lustig et al. [2011]; Menkhoff et al. [2012a]; Ames et al. [2015c]) have previously demonstrated that investors can actually earn arbitrage profits by borrowing in a country with a lower interest rate, exchanging for foreign currency, and investing in a foreign country with a higher interest rate, whilst allowing for any losses (or gains) from exchanging back to their domestic currency at maturity. Therefore, trading strategies that aim to exploit the interest rate differentials can be profitable on average. This is notably the case for the currency carry trade which is thus the simple investment strategy of selling a low interest rate currency forward and then buying a high interest rate currency forward. The idea is that the interest rate returns will outweigh any potential adverse moves in the exchange rate. Historically the Japanese yen and Swiss franc have been used as "funding currencies", since they have maintained very low interest rates for a long period. The currencies of developing nations, such as the South African rand and Brazilian real have been typically used as "investment currencies". Whilst this sounds like an easy money making strategy there is of course a downside risk. This risk comes in the form of currency crashes in periods of high global FX volatility and liquidity shortages. A prime example of this is the sharp yen carry trade reversal in 2007.

Remark 4.2.1. *In addition to the currency carry trade studied in this thesis, numerous other high volume trading strategies are performed by speculative investors in practice. In particular, time series momentum trading strategies, involving buying assets with high recent returns and selling assets with low recent returns, have been shown to be very profitable investment strategies, see Jegadeesh and Titman [1993, 2001]; Moskowitz et al. [2012]; Menkhoff et al. [2012b]; Baltas and Kosowski [2013, 2015]; Baltas [2015]. Understanding the dependence structure of asset returns is of key importance in optimising the asset weights and portfolio rebalancing. Furthermore, in Baltas [2016] a seasonality-adjusted trend-following strategy, which actively incorporates seasonality signals by switching off long and short positions, is shown to constitute a significant improvement over the basic strategy.*

4.3 A Review of the Literature

If the UIP relationship held, then there should indeed not be on average any yield difference between a risk-free investment in a reference currency and a risk-free investment in another currency after converting it back to the reference currency. Accordingly, the depreciation of a currency relative to another should be equal to the risk free interest rates differential between them. However, Hansen and Hodrick [1980, 1983]; Fama [1984] among other recent articles Lustig and Verdelhan [2007]; Lustig et al. [2011]; Menkhoff et al. [2012a], demonstrate that this relation is not observed empirically in markets data and that the "currency carry trade" strategy discussed above takes advantage of this market irregularity.

Over the last few decades there have been many theories proposed for the justification of this phenomenon. Fama [1984] initially proposed a time varying risk premium within the forward rate relative to the associated spot rate - concluding that, under rational markets, most of the variation in forward rates was due to the variation in risk premium.

Weitzman [2007] demonstrates through a Bayesian approach that the uncertainty about the variance of the future growth rates combined with a thin-tailed prior distribution would generate the fat-tailed distribution required to solve the forward premium puzzle. This could be compared to the argument retained by Menkhoff et al. [2012a] who demonstrate that high interest rate currencies tend to be negatively related to the innovations in global FX volatility, which is considered as a proxy for unexpected changes in the FX market volatility. Menkhoff et al. [2012a] show that sorting currencies by respect to their beta with (or sensitivity to) global FX volatility innovations yields portfolios with large differences in returns, and also similar portfolios to those obtained when sorting by forward discount, i.e. forward price minus spot price. Another risk factor shown to be significant, although to a much lesser degree, is liquidity risk. These findings are supported by Shehadeh et al. [2016], who analyse the relationship between currency carry return and volatility and liquidity risk factors. Namely, the global FX volatility, VIX, the global FX bid-ask spread and TED spread. Furthermore, Orlov [2016] empirically examines the effect of equity market illiquidity on the excess returns of currency carry and momentum trading strategies.

Burnside et al. [2007] present an alternative model to a pure risk factor model, in which "adverse selection problems between market makers and traders rationalizes a negative covariance between the forward premium and changes in exchange rates". Here, the authors suggest that the foreign exchange market should not be considered as a Walrasian market and that market makers face a worse adverse selection problem when an agent wants to trade against a public information signal, i.e. to place a contrarian bet as an informed trader.

Another hypothesis, proposed by Farhi and Gabaix [2008], consists of justifying this puzzle through the inclusion of a mean reverting risk premium. According to their model a risky country, which is more sensitive to economic extreme events, represents a high risk of currency depreciation and has thus to propose, in order to compensate this risk, a higher interest rate. Then, when the risk premium reverts to the mean, their exchange rate appreciates while they still have a high interest rate which thus replicates the forward rate premium puzzle.

The causality relation between the interest rate differential and the currency shocks can be presented the other way around as detailed in Brunnermeier and Pedersen [2009]. In this paper, the authors indeed assume that the currency carry trade mechanically attracts investors and more specifically speculators who accordingly increase the probability of a market crash. Tail events among currencies would thus be caused by speculators' need to unwind their positions when they get closer to funding constraints.

This recurrent statement of a relation between tail events and forward rate premium (Farhi and Gabaix [2008]; Brunnermeier et al. [2008]) has led to the proposal in this thesis of a rigorous measure and estimation of the tail thickness at the level of the marginal distribution associated to each exchange rate. Moreover, the question of the link between the currency's marginal distribution and the associated interest rates differential leads to the consideration more globally of the joint dependence structures between the individual marginal cumulative distribution function (cdf) tails with respect to their respective interest rate differential.

Ready et al. [2017] introduce a simple two-country model that captures the asymmetry between commodity exporting (higher interest rate) countries and commodity consuming (lower interest rate) countries. In the model, persistent changes in trade costs arise from low frequency movements in shipping capacity, leading to endogenous time-varying dynamics in global market segmentation.

Burnside et al. [2011] try to explain the positive average returns of the carry trade via a "peso problem", i.e. the effects on inference caused by low-probability events that do not occur in the sample. The authors find that a peso event is characterized by modest negative payoffs to an unhedged carry trade and a large value of the stochastic discount factor, when compared to a strategy that employs currency options to protect an investor from the downside risk associated with large, adverse movements in exchange rates.

Lempérière et al. [2017] introduce a new measure of skewness, which overcomes the issue of the classically defined skewness, i.e. the third cumulant of the distribution of returns, which is that a few extreme events can completely dominate the empirical determination. The authors conclude that for a wide spectrum of "risk premia" strategies, skewness rather than volatility is a determinant of returns.

The hypothesis that skewness and crash risk explains carry returns is challenged by Bekaert and Panayotov [2016], who introduce the concept of "good" and "bad" carry trades constructed from G-10 currencies. The good trades exhibit higher Sharpe ratios and sometimes positive return skewness, in contrast to the bad trades that have both substantially lower Sharpe ratios and highly negative return skewness.

Daniel et al. [2014] provide an analysis on the risks of currency carry trades that differs from the conventional wisdom in the literature. The authors find that the three Fama-French equity market risk factors do significantly explain the returns to an equally weighted carry trade that has no direct exposure to the dollar. Also they find that carry trade strategies with alternative weighting schemes are not fully priced by the HML_{FX} risk factor proposed by Lustig et al. [2011]. In addition, the authors argue that the time varying dollar exposure of the carry trade is at the core of the carry trade puzzle. Finally, they find that the exposure of carry trades to downside market risk is not statistically significantly different from the unconditional exposure.

Chapter 5

Investigating Multivariate Tail Dependence in Currency Carry Trade Portfolios via Copula Models

In this chapter, the forward premium puzzle is investigated using empirical data. The time-varying dependence structure of currency carry trade baskets is explored. In particular, the multivariate tail dependence characteristics of the baskets are analysed and the results discussed.

5.1 Research Contribution: Tail Dependence and Forward Premium Puzzle

The approach adopted in this chapter is a statistical framework with a high degree of sophistication that I developed in order to accommodate the Forward premium puzzle, which is indeed analogous in nature to the ideas considered when investigating the "equity risk premium puzzle" coined by Mehra and Prescott [1985] in the late 80's. The equity risk premium puzzle effectively refers to the fact that demand for government bonds which have lower returns than stocks still exists and generally remains high. This poses a puzzle for economists to explain why the magnitude of the disparity between the returns on each of these asset classes, stocks versus bonds, known as the equity risk premium, is so great and therefore implies an implausibly high level of investor risk aversion. In the seminal paper written by Rietz [1988], the author proposes to explain the "equity risk premium puzzle" Mehra and Prescott [1985] by taking into consideration the low but still significant probability of a joint catastrophic event.

Analogously in this thesis, an exploration is presented of the highly leveraged arbitrage opportunities in currency carry trades that arise due to violation of the UIP. However, it is conjectured that if the assessment of the risk associated with such trading strategies was modified to adequately take into account the potential for joint catastrophic risk events accounting for the non-trivial probabilities of joint adverse movements in currency exchange rates, then such strategies may not seem so profitable relative to the risk borne by the investor. A rigorous probabilistic model is proposed in order to quantify this phenomenon and potentially detect when liquidity in FX markets may dry up and thus simultaneously impact a whole set of currencies. This probabilistic measure of dependence can then be very useful for risk management of such portfolios but also for making more tractable the valuation of structured products or other derivatives indexed on this specific strategy. To be more specific, one of the contributions of this thesis is indeed to model the dependences between exchange rates using a flexible family of mixture copulae comprised of Archimedean members. This probabilistic approach allows the joint distribution of the vectors of random variables, in this case vectors of exchange rates log-returns in each basket of currencies, to be expressed as functions of each marginal distribution and the copula function itself.

In the literature mentioned earlier, the tail thickness resulting from the carry trade has been either treated individually for each exchange rate or through the measurement of distribution moments that may not be adapted to a proper estimation of the tail dependences. In this thesis, it is proposed instead to build, on a daily basis, a set of portfolios of currencies with regards to the interest rate differentials of each currency with the US dollar. Using a mixture of copula functions, a measure of the tail dependences within the high interest rate and low interest rate baskets is extracted and finally the results are interpreted. Among the outcomes of this study, it is demonstrated that during the crisis periods, the high interest rate currencies tend to display very significant upper tail dependence. Accordingly, it can thus be concluded that the appealing high return profile of a carry portfolio is not only compensating the tail thickness of each individual component probability distribution but also the fact that they tend to occur simultaneously and lead to a portfolio particularly sensitive to the risk of drawdown. Furthermore, it is also shown that high interest rate currency portfolios can display periods during which the tail dependence gets inverted demonstrating when periods of construction of the aforementioned carry positions are being undertaken by investors.

5.2 Data Description and Portfolio Construction

In this section, I describe the set of data used for this empirical study and describe the macro-economic specificities associated to some of the currencies I considered. Furthermore, I present the method I retained in this thesis to build the portfolios that are combined to build a carry trade position.

5.2.1 Data Description

I consider for this empirical analysis a set of 20 currency exchange rates relative to the USD. I indeed considered the point of view of an American investor as this is generally the hypothesis retained in the literature Brunnermeier et al. [2008]; Menkhoff et al. [2012a]. However the same analysis could be carried out from any other investor standpoint as the phenomenon I will describe does not only depend on a specific currency but more on two sets of currencies. These sets of currencies correspond to the high interest rate currencies which are used to obtain the highest return (named the "investment currencies") and the low interest rate currencies which allow for borrowing at a low cost the amount of money necessary for this investment (named the "financing currencies").

The currency data for this analysis was obtained from Bloomberg. The time series analysed ranges from 04/01/2000 to 02/01/2013 and comprises the following currencies: Euro (EUR), Turkish lira (TRY), Japanese yen (JPY), British pound sterling (GBP), Australian dollar (AUD), Canadian dollar (CAD), Norwegian krone (NOK), Swiss franc (CHF), Swedish krona (SEK), Mexican peso (MXN), Polish zloty (PLN), Malaysian ringgit (MYR), Singaporean dollar (SGD), Indian rupee (INR), South African rand (ZAR), New Zealand dollar (NZD), Thai baht (THB), South Korean won (KRW), Taiwanese dollar (TWD), Brazilian real (BRL). I have been provided, on a daily basis, with the settlement prices for each currency exchange rate as well as the simultaneous price for the associated 1 month forward contract. Due to differing market closing days, e.g. national holidays, there was missing data for a couple of currencies and for a small number of days. For missing prices, the previous day's closing prices were retained.

The reason why I based this analysis upon a constant maturity 1 month forward is twofold. Firstly, I do not try in this investigation to replicate as realistically as possible a currency carry trade portfolio to see if there is a recurrently high average return. The main inconvenience of such analysis comes from the loss of data points. As a matter of fact, to build a carry portfolio, the position has to be held until the maturity of the forward contract which leads in this case to retain only one point for each month. However, in this case I have at my disposal one point per day which makes this analysis of individual tails and their interdependences more robust. Secondly, tail behaviour of monthly data is naturally different from the tail behaviour of daily data, one reason for this difference is that individual currencies can display a mean reversion in the mid-term and thus reduce the amplitude of the movement.

Among the currencies under scrutiny, some of them have displayed very large variations in the last decade mainly for macro-economic reasons. Therefore, I considered it insightful to mention some of the most meaningful. The Brazilian real displays in its time series two important periods of shocks, the first in 2001 and the second in 2002. Naturally the first of them was due to the terrorist attacks against the world trade center in September. However the Brazilian real has been also impacted by the market's concerns of a contagion after the rumours of default of the Argentinian government. The second shock on the Brazilian real in 2002 was related to the potential election of the Workers' Party leader Luiz Inacio Lula da Silva which prompted concern he might spark a default by overspending to meet promises of spurring growth and employment. In 2001, the South African rand slumped 29% after the events of September 11 2001 and the market's concern of a global recession and a slump in commodity prices to which the South African economy is particularly exposed to. As a third example of a shock in an instrumental currency in a carry trade strategy one can note the 30% daily loss of the Turkish lira on the 22nd of February 2001. This was due to Turkey's decision to abandon the defence of their currency in order to reduce the cost of financing lira-denominated debt. It is worth mentioning that I did not remove these data points from the time series given that different events may have impacted the other exchange rates at a different time but this analysis does not focus only on the tail events associated to a particular currency but more on the events impacting simultaneously a set of currencies.

5.2.2 Data Preparation

In order to perform the empirical analyses considered in this chapter a substantial amount of effort and time was invested into collecting, cleaning and preparing the data. In particular, the following key steps were performed:

1. Collect daily currency spot price data: closing price, bid and ask price.
2. Collect daily currency forward price data - at maturities of one week, two weeks, three weeks and 1 month: closing price, bid and ask price.
3. Pre-process the price data to deal with missing data, i.e. if data is missing copy previous day's price.
4. Match one month forward contracts with closing spot price on the correct date of delivery for the contract.
5. Calculate the forward premium (interest rate proxy) as the difference between the forward price and the spot price.

5.2.3 Currency Portfolio Construction

As described earlier, the currency carry trade results from the differential of interest rates prevailing in different countries. By borrowing a certain amount of money in low interest rate countries and investing it in high interest rate countries, a recurrent profit can be generated given that the UIP condition is on average not satisfied. In order to differentiate the “financing currencies” from the “investment currencies”, I start by classifying each currency relative to its differential of risk free rate with the US dollar. The following basic explanation of the high interest rates and low interest rates can be noted. In general countries that are considered ‘safe’ can borrow at a lower interest rate, which may explain why historically the US dollar or Swiss franc interest rates were low (Gourinchas and Rey [2007]) while the Turkish lira rates were historically high as this country is not considered as financially secure. Dimic et al. [2016] explores the risk profile of individual currency carry trades, finding that carry trade profitability depends on a country’s political risk, supporting the risk-based view on forward bias.

Moreover I demonstrated in expression (4.1.1) that the differential of interest rates between two countries can be estimated through the ratio of the forward contract price and the spot price. It is worth mentioning that Juhl et al. [2006] demonstrate this relationship to hold empirically on a daily basis but not necessarily on an intraday basis. Accordingly, instead of considering the differential of risk free rates between the reference and the foreign countries, I build the respective baskets of currencies with respect to the ratio of the forward and the spot prices for each currency. On a daily basis I compute this ratio for each currency and then build five portfolios of four currencies each. The first portfolio gathers the four currencies with the highest positive differential of interest rate with the US dollar. The selected currencies over the period 04/01/2000 to 02/01/2013 for the high interest rate basket are displayed in Figure 5.1. These currencies are thus representing the “investment” currencies, through which one invests the money to benefit from the currency carry trade. The last portfolio will gather the four currencies with the highest negative differential (or at least the lowest differential) of interest rate. As with the high interest rate basket, I also display the low interest interest rate currency selections in Figure 5.2. These currencies are thus representing the “financing” currencies, through which one borrows the money to build the currency carry trade.

It can be noted here that during the period investigated in this analysis there is a strong presence of emerging markets currencies, e.g. Brazilian real and Turkish lira, in the long basket whereas the short basket is dominated by developed market currencies. A discussion of the effects of this asymmetry can be seen in Bekaert and Panayotov [2016].

Conditionally to this classification I investigate then the joint distribution of each group of currencies to understand the impact of the currency carry trade, embodied by the differential of interest rates, on currencies returns. In this analysis I concentrate on the high interest rate basket (investment currencies) and the low interest rate basket (funding currencies), since typically when implementing a carry trade strategy one would go short the low basket and go long the high basket.

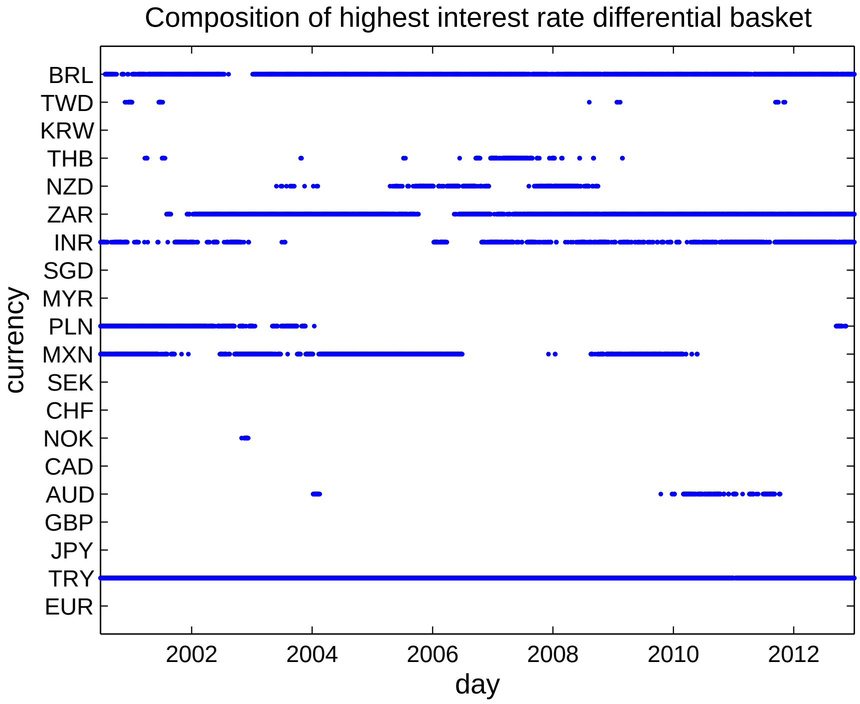


Figure 5.1: Basket 5 (highest IR) composition.

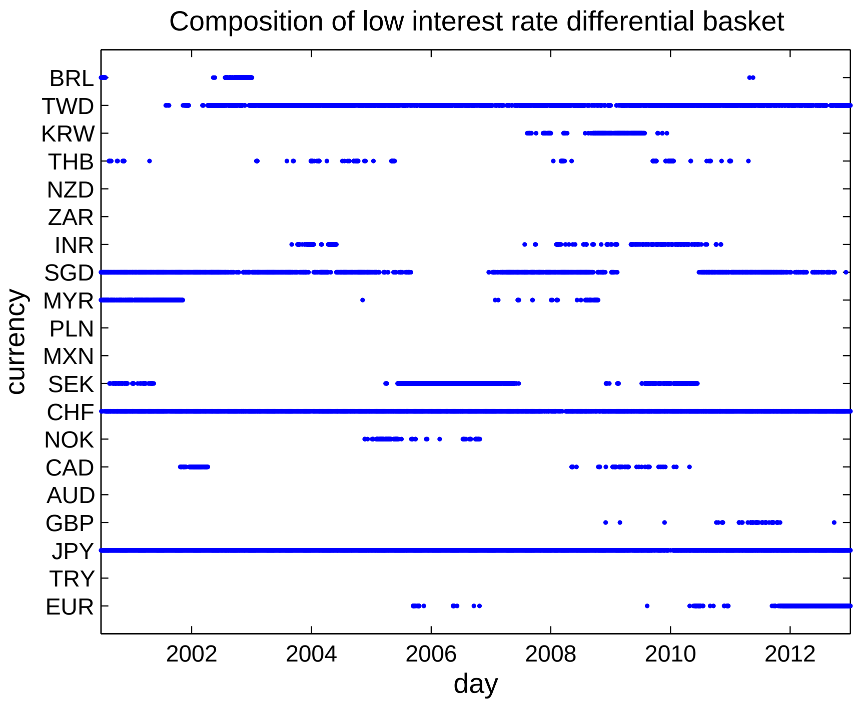


Figure 5.2: Basket 1 (lowest IR) composition.

5.3 Interpreting Tail Dependence as Financial Risk Exposure in Carry Trade Portfolios

In order to fully understand the tail risks of joint exchange rate movements present when one invests in a carry trade strategy it is important to look at both the downside extremal tail exposure and the upside extremal tail exposure within the funding and investment baskets that comprise the carry portfolio. The downside tail exposure can be seen as the crash risk of the basket, i.e. the risk that one will suffer large joint losses from each of the currencies in the basket. These losses would be the result of joint appreciations of the currencies one is short (meaning the currencies that one has sold) in the low interest rate basket and/or joint depreciations of the currencies one is long (meaning the currencies that one has bought) in the high interest rate basket.

Definition 37. ***Downside Tail Risk Exposure in Carry Trade Portfolios***

Consider the funding currency (short) basket with n-exchange rates relative to base currency, on day t, with currency log-returns $(X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(n)})$. Then the downside tail exposure risk for the carry trade will be defined as the conditional probability of adverse currency movements in the short basket, corresponding to its upper tail dependence, given by

$$\lambda_u^{(i)}(u) := \mathbb{P}r\big(X_t^{(i)} > F_i^{-1}(u) | X_t^{(1)} > F_1^{-1}(u), \ldots, X_t^{(i-1)} > F_{i-1}^{-1}(u), \\ X_t^{(i+1)} > F_{i+1}^{-1}(u), \ldots, X_t^{(n)} > F_n^{-1}(u)\big) \tag{5–1}$$

for a currency of interest $i \in \{1, 2, \ldots, n\}$.

The downside tail exposure for the investment (long) basket with n currencies will be defined as the conditional probability of adverse currency movement in the long basket, given by

$$\lambda_l^{(i)}(u) := \mathbb{P}r\big(X_t^{(i)} < F_i^{-1}(u) | X_t^{(1)} < F_1^{-1}(u), \ldots, X_t^{(i-1)} < F_{i-1}^{-1}(u), \\ X_t^{(i+1)} < F_{i+1}^{-1}(u), \ldots, X_t^{(n)} < F_n^{-1}(u)\big). \tag{5–2}$$

In general then a basket's upside or downside risk exposure would be quantified by the probability of a loss (or gain) arising from an appreciation or depreciation jointly of magnitude u and the dollar cost associated to a given loss/gain of this magnitude. The standard approach in economics would be to associate say a linear cost function in u to such a probability of loss to get say the downside risk exposure in dollars according to $E(u) = C_u(F_{X_t^{(i)}}(u)) \times \lambda_u(u)$, which will be a function of the level u. As λ_u becomes independent of the marginals, i.e. as $u \to 0$ or $u \to 1$, C also becomes independent of the marginals.

Conversely, the upside tail exposure contributes to profitable returns in the carry trade strategy when extreme movements are in favour of the carry position held. These would correspond to precisely the probabilities discussed above applied in the opposite direction. That is the upside risk exposure in the funding (short) basket is given by Equation (5–2) and the upside risk exposure in the investment (long) basket is given by Equation (5–1). That is the upside tail exposure of the carry trade strategy is defined to be the risk that one will earn large joint profits from each of the currencies in the basket. These profits would be the result of joint depreciations of the currencies one is short in the low interest rate basket and/or joint appreciations of the currencies one is long in the high interest rate basket.

Remark 5.3.1. *In a basket with n currencies, $n \geq 2$, if one considers capturing the upside and downside financial risk exposures from a model based calculation of these extreme probabilities then if the parametric model is exchangeable, such as an Archimedean copula, then swapping currency i in Equation (5–1) and Equation (5–2) with another currency from the basket, say j will not alter the downside or upside risk exposures. This exchangeability property can be visualised intuitively by considering the symmetry with respect to the diagonal in the two-dimensional case, and hence the plot is invariant under a switch of axes. If they are not exchangeable then one can consider upside and downside risks for each individual currency in the carry trade portfolio.*

These tail upside and downside exposures of the carry trade strategy can be considered as features that show that even though average profits may be made from the violation of UIP, it comes at significant tail exposure.

The notion of the dependence behaviour in the extremes of the multivariate distribution can be formalised through the concept of tail dependence, i.e. the limiting behaviour of Equations (5–1) and (5–2), as $u \uparrow 1$ and $u \downarrow 0$ asymptotically. The interpretation of such quantities is then directly relevant to assessing the chance of large adverse movements in multiple currencies which could potentially increase the risk associated with currency carry trade strategies significantly, compared to risk measures which only consider the marginal behaviour in each individual currency. Under certain statistical dependence models these extreme upside and downside tail exposures can be obtained analytically. Here, I develop flexible copula mixture models that have such properties.

5.4 Likelihood Based Estimation of the Mixture Copula Models

Let me begin this section with a discussion on the choices I make for the marginal distributions for each of the currencies specified in the baskets constructed for the high interest rate differentials and also the baskets for the low interest rate differentials.

In modelling parametrically the marginal features of the log return forward exchange rates, I wanted flexibility to capture a broad range of skew-kurtosis relationships as well as potential for sub-exponential heavy tailed features. In addition, I wished to keep the models to a selection which is efficient to perform inference and easily interpretable. I therefore considered a first analysis utilizing log-normal distributions for the monthly forward exchange rate returns, which would be equivalent to specification of a Normality assumption on the distribution for the log return forward exchange rates. This model is given by the following parametric density, for a random variable $X \sim F(x; \mu, \sigma)$, in Equation 5–3 below.

$$f_X(x; \mu, \sigma) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \tag{5–3}$$

with the shape parameter $\sigma^2 > 0$ and the log-scale parameter $\mu \in \mathbb{R}$ and the support $x \in (0, \infty)$.

I found when analysing the goodness-of-fit for this log-normal model on each of the assets in the 20 currencies considered, over both 6 month and 1 year sliding windows, that the fit of the log-normal model would be systematically rejected as a suitable model for a couple of currencies. In the majority of cases over these sliding windows (locally stationary time series) the log-normal model was more than adequate. However, since some of the currencies that were rejecting this fit were appearing regularly in the high interest rate baskets I also decided to consider a more flexible three parameter model for the marginal distributions given by the Log-Generalized-Gamma distribution (l.g.g.d.), see details in Consul and Jain [1971] and Lawless [1980].

The l.g.g.d. is a parametric model based on the generalized gamma distribution which is highly utilized in lifetime modelling and survival analysis. The density for the generalized gamma distribution and the l.g.g.d are given respectively by Equations 5–4 and 5–5.

$$f_X(x;k,\alpha,\beta) = \frac{\beta}{\Gamma(k)} \frac{x^{\beta k-1}}{\alpha^{\beta k}} \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) \tag{5–4}$$

with parameter ranges $k > 0$, $\alpha > 0$ and $\beta > 0$ and a support of $x \in (0, \infty)$. Then the log transformed g.g.d. random variable $Y = \ln X$ is given by the density of the l.g.g.d. as follows.

$$f_Y(y;k,u,b) = \frac{1}{b\Gamma(k)} \exp\left[k\left(\frac{y-u}{b}\right) - \exp\left(\frac{y-u}{b}\right)\right] \tag{5–5}$$

with $u = \log(\alpha)$, $b = \beta^{-1}$ and the support of the l.g.g.d. distribution is $y \in \mathbb{R}$.

This more flexible three parameter model is particularly interesting in the context of the marginal modelling considered here since the log-normal model is nested within the g.g.d. family as a limiting case. In addition the g.g.d. also includes the exponential model ($\beta = k = 1$), the Weibul distribution with ($k = 1$) and the Gamma distribution with ($\beta = 1$). Next I discuss how one can perform inference for the multivariate currency basket models using these marginal models and the mixture copula discussed previously.

5.4.1 Two Stages: Inference For the Margins

The inference function for margins (IFM) technique introduced in Joe [2005] provides a computationally faster method for estimating parameters than Full Maximum Likelihood, i.e. simultaneously maximising all model parameters and produces in many cases a more stable likelihood estimation procedure. An alternative approach to copula model parameter estimation that is popular in the literature is known as the Maximum Partial Likelihood Estimator (MPLE) detailed in Genest et al. [1995].

The procedure I adopt for likelihood based estimation is the two stage estimation known as Inference on the Margins which is studied with regard to the asymptotic relative efficiency of the two-stage estimation procedure compared with maximum likelihood estimation in Joe [2005] and in Hafner and Manner [2010]. It can be shown that the IFM estimator is consistent under weak regularity conditions. However, it is not fully efficient for the copula parameters. Nevertheless, it is widely used for its ease of implementation and efficiency in large data settings such as the models I consider in this study.

To complete this discussion on general IFM, before providing the MLE estimation expressions, it can be first noted that in this study copula models are fit to the high interest rate (IR) basket and the low IR basket updated for each day in the period 04/01/2000 to 02/01/2013 using log return forward exchange rates at one month maturities for data covering both the previous 6 months and previous year as a sliding window analysis on each trading day in this period. Next I discuss briefly the marginal MLE estimations for the log-normal and the l.g.g.d. models.

5.4.1.1 Stage 1: Fitting the Marginal Distributions via MLE

In the first step I fit the marginal distributions to either the log-normal model or the l.g.g.d model. In the case of the log-normal model this is achieved effortlessly since one may utilise the well-known analytic expressions for the MLE estimates:

$$
\begin{aligned}
\hat{\mu}_j &= \frac{1}{n}\sum_j \log\left(x_j\right) \\
\hat{\sigma}_j &= \sqrt{\frac{1}{n}\sum_j \log\left(x_j\right)^2 - \hat{\mu}_j^2}
\end{aligned}
\tag{5–6}
$$

In the case of the l.g.g.d. distribution the estimation for the three model parameters can be significantly more challenging due to the fact that a wide range of model parameters, especially for k can produce similar resulting density shapes, see discussions in Lawless [1980]. To overcome this complication and to make the estimation efficient it is proposed to utilise a combination of profile likelihood methods over a grid of values for k and perform profile likelihood based MLE estimation for each value of k, then for the other two parameters b and u. The differentiation of the profile likelihood for a given value of k produces the system of two equations given by

$$
\begin{aligned}
\exp(\tilde{\mu}) &= \left[\frac{1}{n}\sum_{i=1}^{n}\exp\left(\frac{y_i}{\tilde{\sigma}\sqrt{k}}\right)\right]^{\tilde{\sigma}\sqrt{k}} \\
\frac{\sum_{i=1}^{n} y_i \exp\left(\frac{y_i}{\tilde{\sigma}\sqrt{k}}\right)}{\sum_{i=1}^{n}\exp\left(\frac{y_i}{\tilde{\sigma}\sqrt{k}}\right)} &- \overline{y} - \frac{\tilde{\sigma}}{\sqrt{k}} = 0
\end{aligned}
\tag{5–7}
$$

with n the number of observations, $y_i = \log x_i$ and the parameter transformations $\tilde{\sigma} = \frac{b}{\sqrt{k}}$ and $\tilde{\mu} = u + b\ln k$. The second equation is solved directly via a simple root search for the estimation of $\tilde{\sigma}$ and then substitution into the first equation provides the estimation of $\tilde{\mu}$. Note, for each value of k selected in the grid, one gets the pair of parameter estimates $\tilde{\mu}$ and $\tilde{\sigma}$, which can then be plugged back into the profile likelihood to make it purely a function of k, with the estimator for k then selected as the one with the maximum likelihood score.

5.4.1.2 Stage 2: Fitting the Mixture Copula via MLE

In order to fit the Clayton-Frank-Gumbel (C-F-G) model the copulae parameters $(\rho_{Clayton},\ \rho_{Frank}, \rho_{Gumbel})$ and the copulae mixture parameters $(\lambda_{Clayton}, \lambda_{Frank},$

λ_{Gumbel}) are estimated using maximum likelihood on the data after conditioning on the selected marginal distribution models and their corresponding estimated parameters obtained in Stage 1. These models are utilised to transform the data using the cdf function with the mle parameters ($\hat{\mu}$ and $\hat{\sigma}$) if the log-normal model is used or ($\hat{k}$, $\hat{u}$ and $\hat{b}$) if the l.g.g.d is considered.

Therefore, in this second stage of MLE estimation one aims to estimate either the one parameter mixture of C-F-G components with parameters $\underline{\theta} = (\rho_{Clayton}, \rho_{Frank}, \rho_{Gumbel}, \lambda_{Clayton}, \lambda_{Frank}, \lambda_{Gumbel})$ or the two parameter mixture of outer power transformed mixture components, i.e. Outer power Clayton - Outer power Frank - Gumbel (OC-OF-G) components with parameters $\underline{\theta} = (\rho_{Clayton}, \rho_{Frank}, \rho_{Gumbel}, \lambda_{Clayton}, \lambda_{Frank}, \lambda_{Gumbel}, \beta_{Clayton}, \beta_{Frank})$. It can be noted that in fact an Outer power Gumbel copula is equivalent to a standard Gumbel copula but with a superfluous additional parameter. Therefore, I use a standard Gumbel parameter here. This is achieved in each case by the conditional maximum likelihood. To achieve this one needs to maximise the log likelihood expressions for the mixture copula models, which in this framework are given generically by the following function for which one needs to find the mode,

$$l(\underline{\theta}) = \sum_{i=1}^{n} log\ c^{C-F-G}(F_1(X_{i1}; \hat{\mu}_1, \hat{\sigma}_1), \ldots, F_d(X_{id}; \hat{\mu}_d, \hat{\sigma}_d)) + \sum_{i=1}^{n}\sum_{j=1}^{d} log\ f_j(X_{ij}; \hat{\mu}_j, \hat{\sigma}_j) \quad (5\text{–}8)$$

with respect to the parameter vector $\underline{\theta}$.

For example in the case of the Clayton-Frank-Gumbel mixture copula one needs to maximise on the log-scale the following expression.

$$\begin{aligned} l(\underline{\theta}) = \sum_{i=1}^{n} log\ \Big[& \lambda_C * \left(c^{C}_{\rho_C}\left(F_1\left(X_{i1}; \hat{\mu}_1, \hat{\sigma}_1\right) \ldots, F_d\left(X_{id}; \hat{\mu}_d, \hat{\sigma}_d\right)\right)\right) \\ & + \lambda_F * \left(c^{F}_{\rho_F}\left(F_1\left(X_{i1}; \hat{\mu}_1, \hat{\sigma}_1\right) \ldots, F_d\left(X_{id}; \hat{\mu}_d, \hat{\sigma}_d\right)\right)\right) \\ & + \lambda_G * \left(c^{G}_{\rho_G}\left(F_1\left(X_{i1}; \hat{\mu}_1, \hat{\sigma}_1\right) \ldots, F_d\left(X_{id}; \hat{\mu}_d, \hat{\sigma}_d\right)\right)\right) \Big] \end{aligned} \quad (5\text{–}9)$$

This optimization is achieved via a gradient descent iterative algorithm which was found to be quite robust given the likelihood surfaces considered in these models with the real data. To illustrate this point, at this stage it is instructive to present some examples of the shapes of the profile likelihoods that are being optimized over for some of the important copula model parameters in the C-F-G mixture example for a 6 month window of data randomly selected from the data set for both the high interest rate basket and the low interest rate basket. Two example plots of the profile likelihood for the 6-dimensional optimisation space for two different example days can be seen in Figures 5.3 and 5.4. From the contour plots here it can be seen that gradient descent could be expected to perform well and hence converge to a global optimal solution. Thorough testing of the gradient descent algorithm with multiple starting parameter values showed that this was indeed the case for the data considered in this chapter.

5.4.2 Goodness-of-Fit Tests

In this section I briefly comment on the model selection aspects of the analysis I undertook. As mentioned I first undertook a process of fitting the marginal log-normal model to all of the 20 currencies considered in the analysis over a sliding window of 6 months and 1 year. For each of these fits I then performed a formal hypothesis test in which I postulated that the null distribution is the log-normal model and then look for evidence in the data to reject this hypothesis at a level of significance of 5%. To undertake this test I considered the standard Kolmogorov-Smirnov (KS) test. As I will present in the results I found strong evidence to reject the null systematically for a few important developing countries' marginal models, hence I also undertook estimation of the l.g.g.d. models for all of the 20 currencies. I am particularly interested in this case in the optimal choice of the model parameter k which as it asymptotically gets large $k \to \infty$ will produce a log-normal model. I found as expected the estimated model fits were significantly improved when fitting the l.g.g.d. models for the cases in which the log-normal was rejected by the K-S test. In addition the estimated k parameter in the periods of rejection of the log-normal hypothesis were estimated at values significantly lower than the upper bound in the search space. I assessed the optimal choice of marginal model between the log-normal and the l.g.g.d. models then via a standard information criterion based on the Akaike Information Criterion (AIC).

In terms of the selection of the copula mixture models, between the mixture

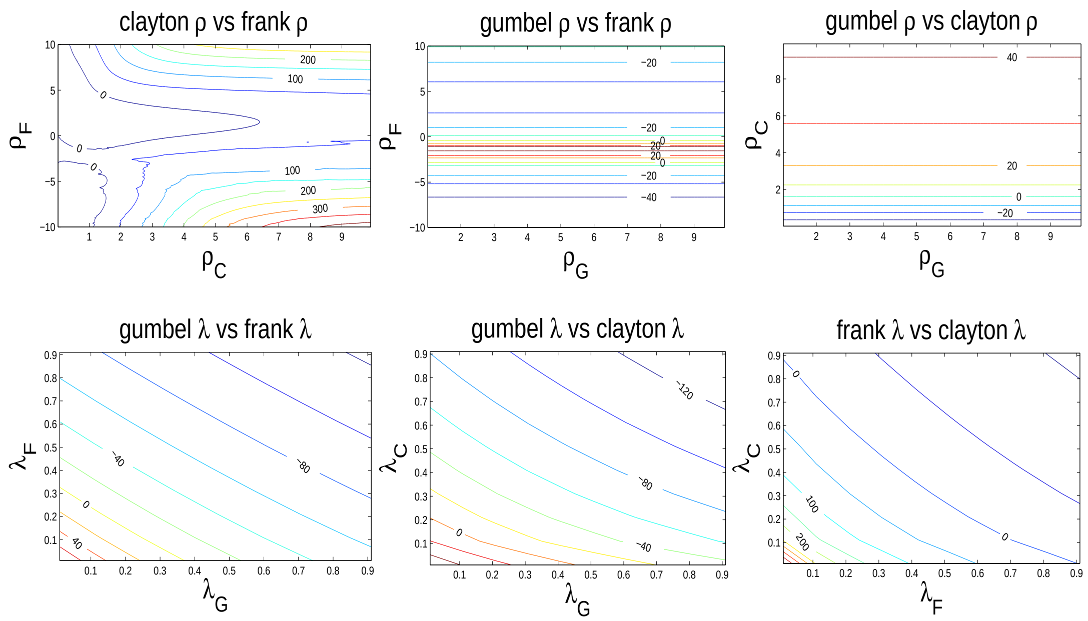


Figure 5.3: Example 1: Profile likelihood plots for C-F-G mixture model.

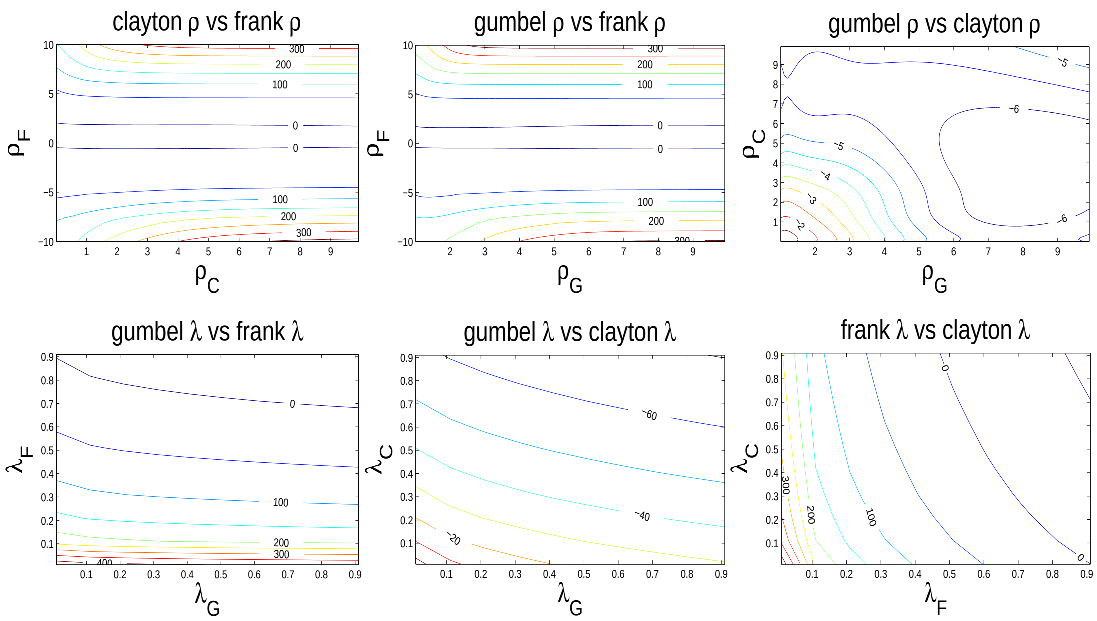


Figure 5.4: Example 2: Profile likelihood plots for C-F-G mixture model.

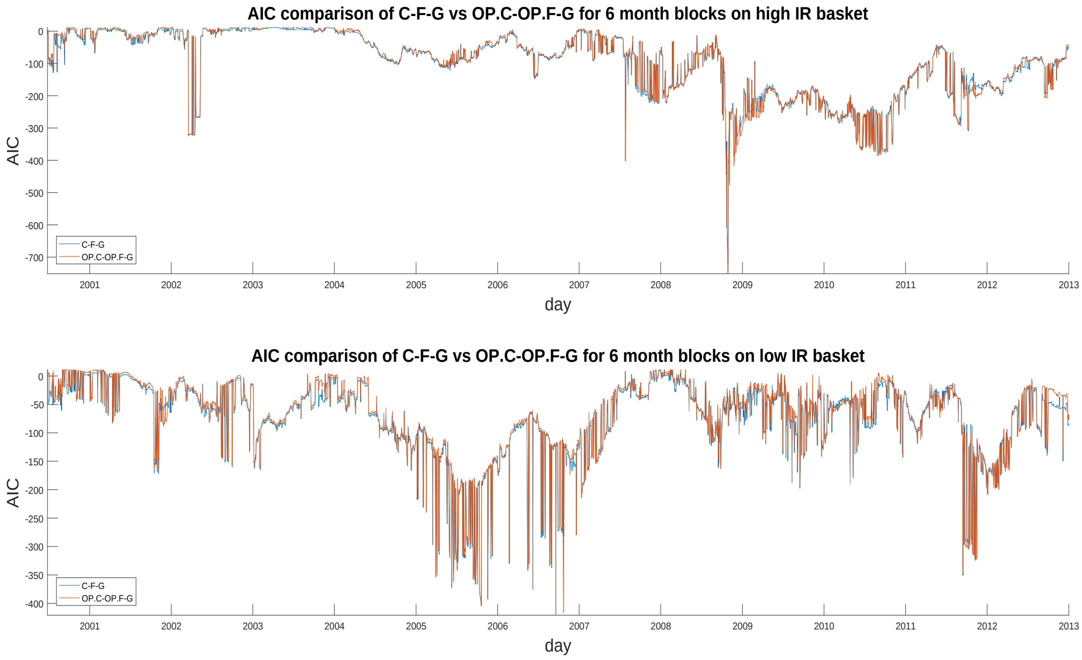


Figure 5.5: AIC comparison of C-F-G vs OP.C-OP.F-G for 6 month blocks on high and low IR baskets.

of one parameter C-F-G model versus the two parameter mixtures of OC-OF-G models, I again used a scoring via the AIC. It can be noted that there are also alternative information criterion developed for copula models to assess the joint suitability of the copula model incorporating both the marginal and the joint copula structure which are modifications of the AIC, adjusting the penalty term for the approach adopted in the estimation, see for example the Copula-Information-Criterion (CIC) in Grønneberg [2010] for details. The results are presented for this comparison in Figure 5.5 in the top panel for the high interest rate basket and in Figure 5.5 in the lower panel for the low interest rate basket, over time based on the 6 month sliding window.

To further analyse this comparison of optimal copula mixtures I plot the AIC differentials for each of the currency baskets in Figure 5.6. Figures 5.5 and 5.6 show it is not unreasonable to consider the C-F-G model for this analysis, since the mean difference between the two AIC scores for the models is 2.05 in favour of the C-F-G. However, it can be noted that the OP.C-OP.F-G model seems to fit

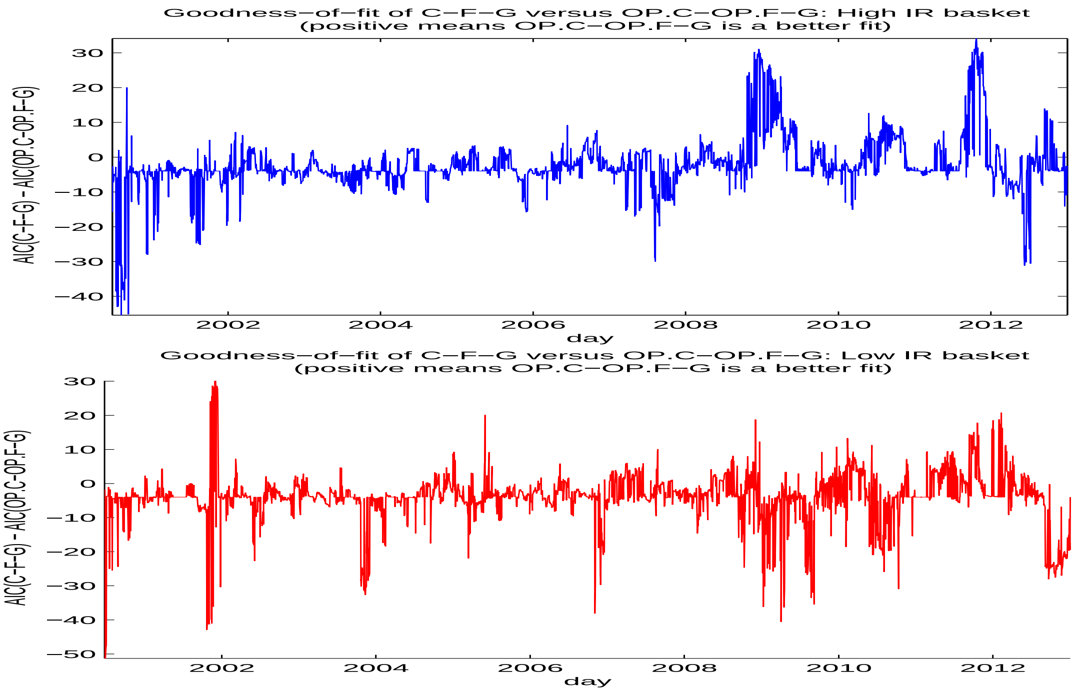


Figure 5.6: AIC differences: C-F-G vs OP.C-OP.F-G for 6 month blocks on high and low IR baskets.

better during crisis periods.

5.5 Results and Analysis

In this section, I present a detailed analysis of the estimation of the marginal distributional models and the mixture copula models for both the high interest rate basket and the low interest rate basket. Firstly, I investigate the properties of the marginal distributions of the exchange rate log-returns for the 20 currencies. I then interpret the time-varying dependence characteristics of the fitted copula models to the high interest rate basket and the low interest rate basket across the period 04/01/2000 to 02/01/2013. Note, all results presented below are for the case in which I considered a 6 month sliding window, since results for the 1 year sliding window were similar in nature and so are omitted.

5.5.1 Modelling the Marginal Exchange Rate Log-Returns

In order to model the marginal exchange rate log-returns I first fit log-normal models to each of the 20 currencies considered in the analysis, updating the fits for every trading day in the period 04/01/2000 to 02/01/2013 based on the 6 months sliding window. The log-normal model was selected due to the fact it has a positive support, represents a range of skew-kurtosis characteristics and can display sub-exponential tail features (i.e. heavy tailed features) should such attributes be present in the data. I assessed the quality of the fits for each currency using a standard Kolmogorov-Smirnov goodness-of-fit test, at the 5% significance level. A summary of the results of this analysis are presented in Table 5.1 which shows the proportion of rejections of the null hypothesis, that the marginal distribution is log-normal for each of the currencies on a given 6 months block of trading days.

One learns from this analysis that the majority of the currencies demonstrate reasonable marginal distribution fits under a log-normal family, however there are a few notable exceptions. Specifically the Turkish lira, Brazilian real, Malaysian ringgit, Indian rupee, Thai baht, South Korean won and Taiwanese dollar demonstrated sustained periods in the analysis in which the log-normal model would be unsuitable to capture the features of the time series adequately. This is significant in this analysis since these currencies actually correspond to the currencies that have a strong presence in the high interest rate baskets, as seen in Figure 5.1. Therefore, they will play an important role in the multivariate analysis of the currency carry trade. As such, it is important to accurately model the features of each of these particular currencies' marginal distributions, before undertaking the multivariate mixture copula analysis, I proposed to generalize the marginal model analysis to a more flexible three parameter family of models given by the log generalized gamma distribution, as discussed in Section 5.4.

The log-generalised gamma distribution (l.g.g.d.) should improve the fit for all currencies since it allows for more flexibility in the tails of the distribution and a wider range of skew-kurtosis relationships when compared to the log-normal model family. In addition, as noted in Section 5.4, for those currencies in which the log-normal model was a suitable fit, then they will still obtain such distributional characteristics since the log-normal model is a limiting case of the l.g.g.d. as k tends to infinity. Hence, the log-normal model for the currencies that were a good fit can still be incorporated.

The maximum likelihood parameters $(\hat{\mu}, \hat{\sigma}, \hat{k})$ of the fitted l.g.g.d. margins for each of the currencies can be seen in Figures 5.7 to 5.11. These plots demonstrate the time varying attributes of the marginal distributions for each currency, illustrating interesting changes in tail behaviour and skewness-kurtosis characteristics over time, especially in heightened periods of volatility in some of these currencies. In particular, there are three standout periods (2003, 2009 and 2012) of heightened μ and σ parameter values across most of the currencies. Hence, during these periods the exchange rate log-returns may demonstrate heavier tails, and increased volatility in the parameter estimates. In addition, it is observed that a few important currencies for the currency carry trade analysis demonstrate sustained differences in their marginal distribution attributes relative to the other currencies. An important example of this is the μ estimates in Figures 5.7 and 5.8 for the TRY, the NZD and the BRL. Similar significant differences between these particular currencies and the rest of the currencies are observed in the estimates of σ in Figures 5.9 and 5.10.

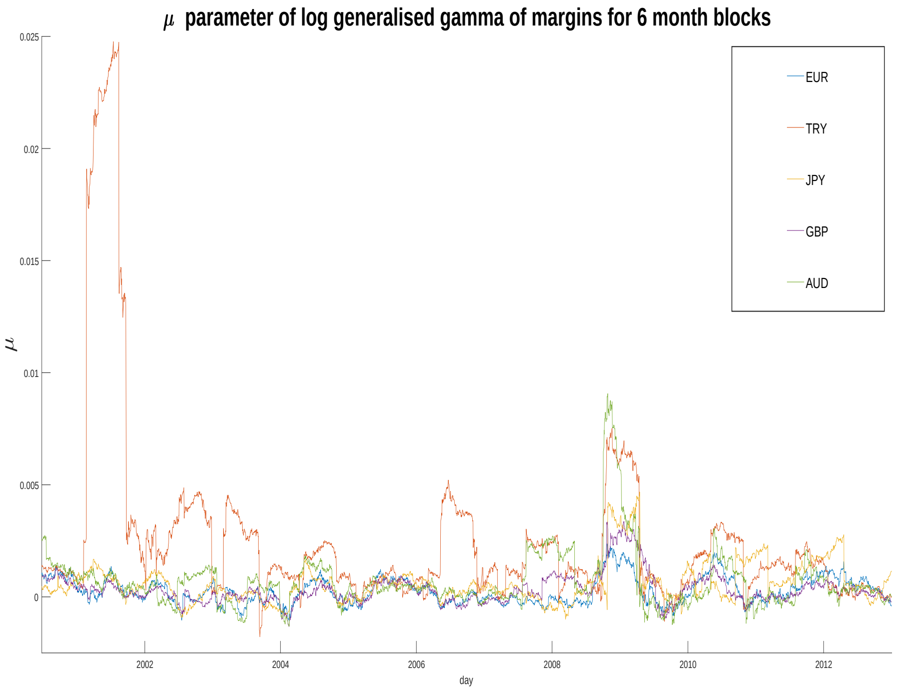
As the value of the parameter k in the l.g.g.d. gets large I expect the log-normal fit to be a suitable model structure for the marginal distributions. As illustrated in the K-S test results certain currencies systematically did not have a suitable fit with the log-normal model. Examples of this are clear when one considers the estimates of k in Figure 5.11 and Table 5.2, which contains the median and interquartile ranges of the estimated k parameter. Again systematically smaller values for the estimate of k in the TRY and the BRL are observed. The periods of time during which the currencies display non log-normal behaviour can clearly be seen in Figure 5.11. The most prominent example being the Turkish lira (orange), which shows consistently low values of k. As noted in Section 5.4, for small values of $k \approx 1$ one obtains Weibull like tail behaviour and in addition, in the cases when $\sigma \approx 1$ jointly with small values of k, I expect the light tailed exponential models to be suitable. As a consequence of this analysis and comparison of AIC results I proceeded with the joint estimation utilising the l.g.g.d. marginal models for every currency.

A noticeable period for the Turkish lira is early in 2001 during which low values of the parameter k clearly provides evidence of heavy tail log-returns distribution for this specific currency. As mentioned earlier in this investigation the Turkish government's decision in February 2001 to stop draining reserves to bolster its currency led the same day to a 30% devaluation of the Turkish lira relative to the dollar.

Table 5.1: Proportion of rejections of the null hypothesis that the sample is from a log-normal distribution, measured using a k-s test at the 5% level.

Block length	**EUR**	**TRY**	**JPY**	**GBP**	**AUD**	**CAD**	**NOK**	**CHF**	**SEK**	**MXN**
6 month	0.001	0.198	0.043	0.000	0.023	0.000	0.000	0.031	0.012	0.032
Year	0.000	**0.553**	0.107	0.007	0.120	0.018	0.006	0.084	0.018	0.128

Block length	**PLN**	**MYR**	**SGD**	**INR**	**ZAR**	**NZD**	**THB**	**KRW**	**TWD**	**BRL**
6 month	0.018	0.494	0.000	0.234	0.025	0.012	0.221	0.130	0.192	0.086
Year	0.094	**0.651**	0.071	**0.549**	0.124	0.113	**0.504**	0.350	0.381	0.403



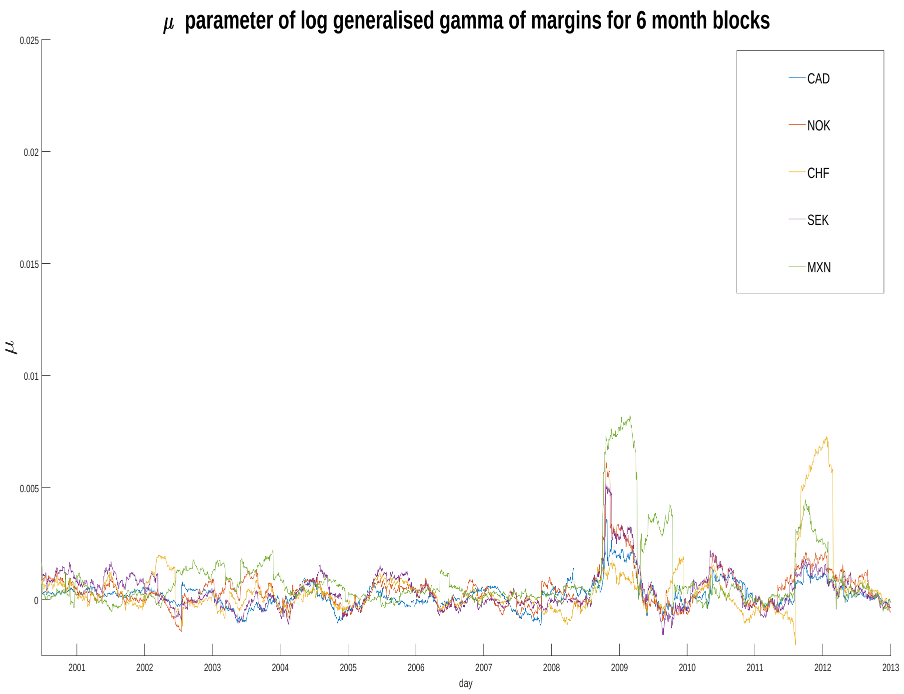
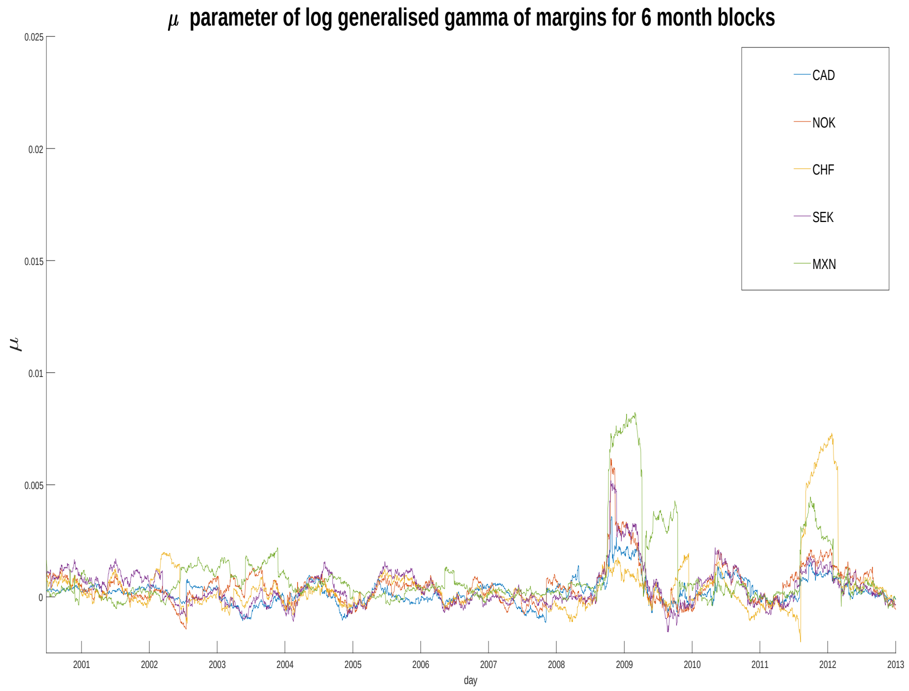


Figure 5.7: μ parameter of log generalised gamma margins using 6 month blocks



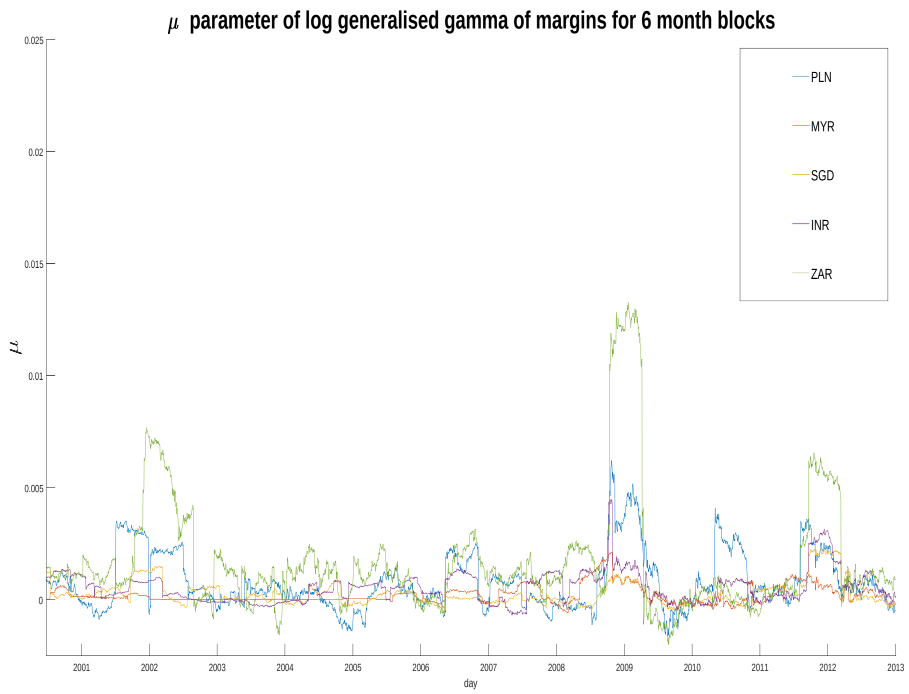


Figure 5.8: μ parameter of log generalised gamma margins using 6 month blocks

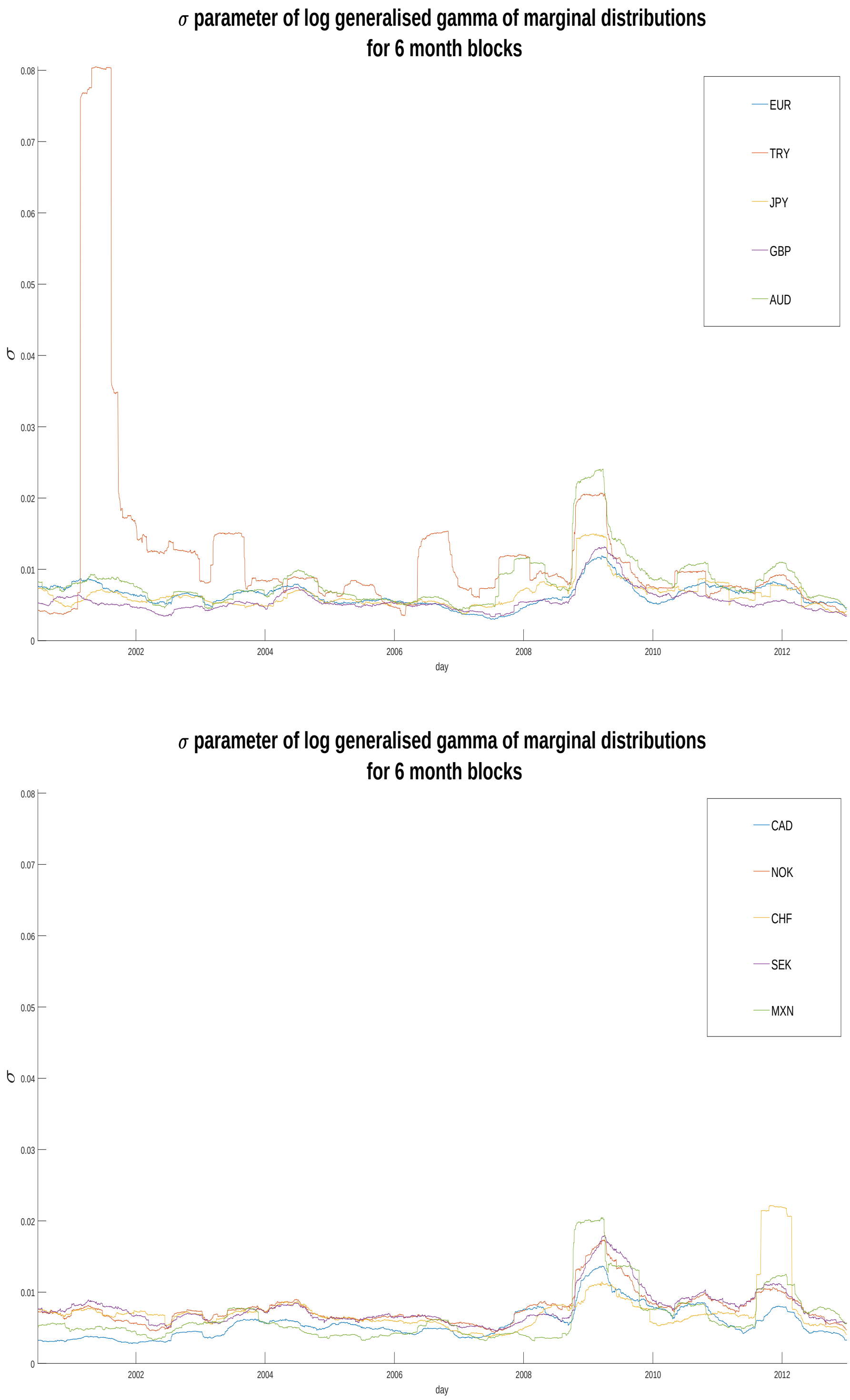
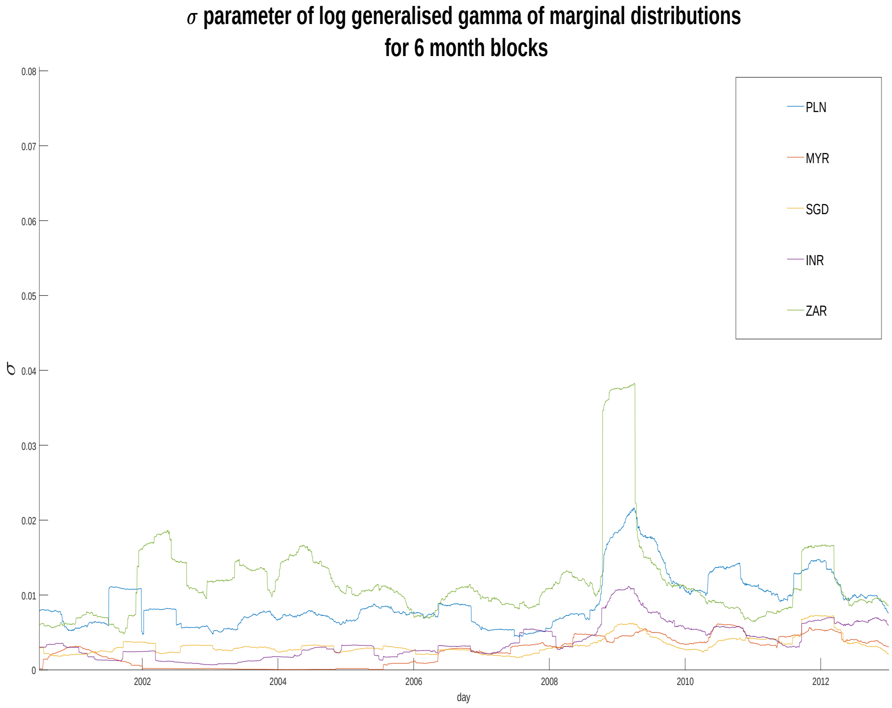


Figure 5.9: σ parameter of log generalised gamma margins using 6 month blocks



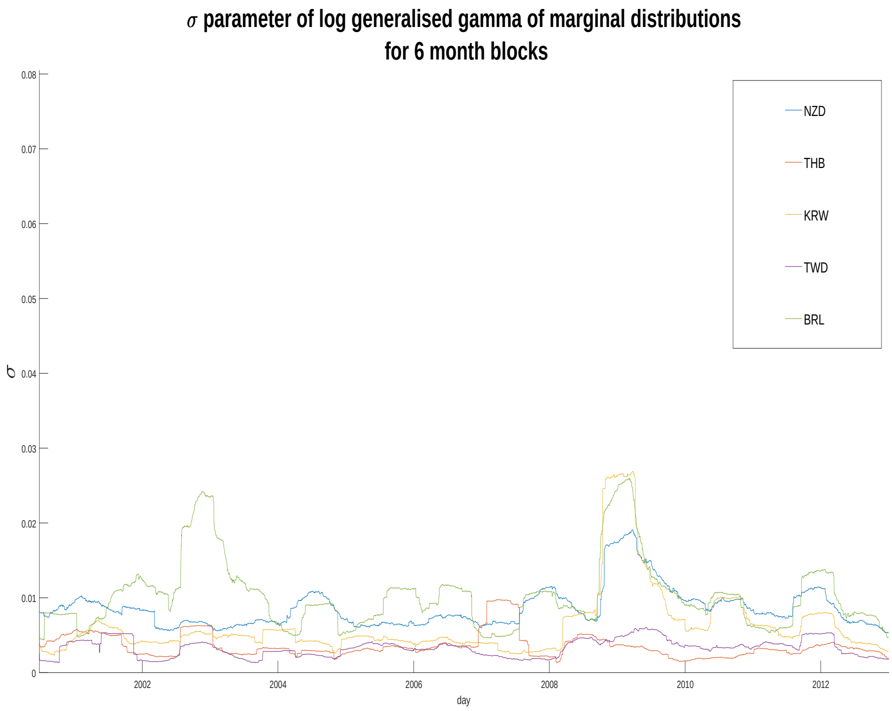


Figure 5.10: σ parameter of log generalised gamma margins using 6 month blocks

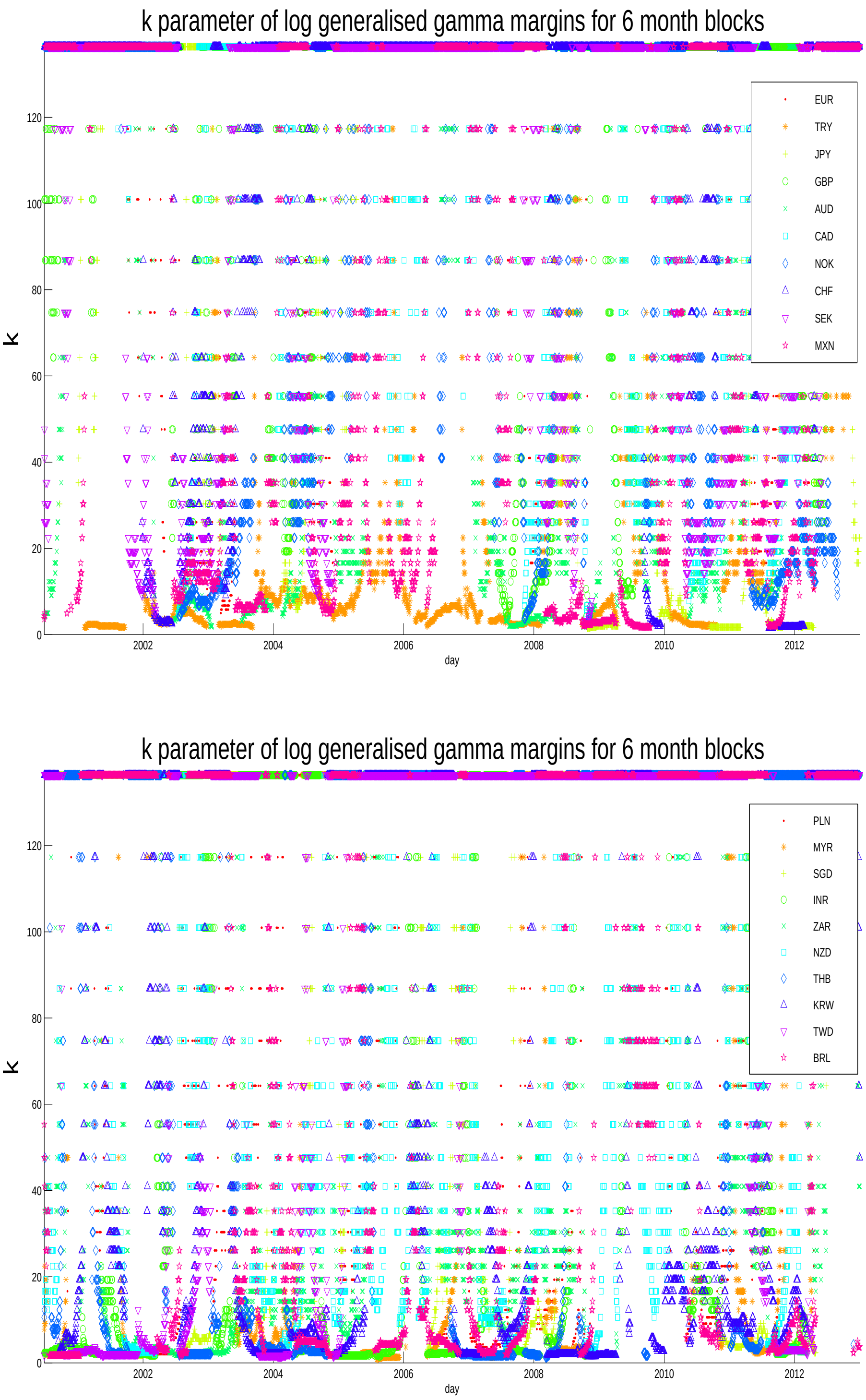


Figure 5.11: K parameter of log generalised gamma margins using 6 month blocks

Table 5.2: Median and interquartile ranges of the estimated k parameter.

	EUR	TRY	JPY	GBP	AUD	CAD	NOK	CHF	SEK	MXN
Median	136.4	9.1	136.4	136.4	55.3	136.4	136.4	136.4	136.4	47.6
IQR	31.3	43.9	101.1	19.0	124.1	72.1	106.0	0.0	95.4	124.1

	PLN	MYR	SGD	INR	ZAR	NZD	THB	KRW	TWD	BRL
Median	74.7	136.4	136.4	40.9	30.3	55.3	136.4	22.4	136.4	64.3
IQR	119.8	132.1	127.3	133.2	125.8	117.1	132.7	130.6	122.1	130.6

5.5.2 Copula Modelling Results

I now utilised each of the l.g.g.d. marginal distribution fits for a given day's set of currencies in the high interest rate and low interest rate baskets to analyse the joint multivariate features. To achieve this for each of the currencies, the exchange rate log-return data was transformed via the l.g.g.d. marginal model's distribution function to uniform [0, 1] margins. Then the mixture Clayton-Frank-Gumbel copula (denoted C-F-G) and the outer-power versions were fitted each day to a sliding window of 6 months and one year log-returns data for both the high interest rate and low interest rate baskets. Below the time-varying parameters of the maximum likelihood fits of this mixture C-F-G copula model will be examined. Furthermore, the results for the outer-power transform cases did not demonstrate discernible differences from the base C-F-G model and so were excluded. This can be seen from the figures displaying the AIC for each of these models (Figures 5.5 and 5.6).

In this analysis there are several attributes to be considered for the mixture copula model, such as the relevant copula structures for the high and low interest rate baskets and how these copula dependence structures may change over time. In addition, there is the strength of the tail dependence in each currency basket and how this changes over time, especially in periods of heightened market volatility. The first of these attributes I will consider to be a structure analysis studying the relevant forms of dependence in the currency baskets and the second of these attributes that I shall study will be the strength of dependence present in the currency baskets, given the particular copula structures in the mixture.

Therefore I first consider the structural components of the multivariate copula model. To achieve this, I begin with a form of model selection in a mixture context, in which I consider the estimated relative contributions of each of the copula components (and their associated dependence features) to the joint relationship in the high and low interest rate currency baskets over time. This is reflected in the estimated mixture component weights, which can be seen in Figures 5.12 and 5.13 for the high interest rate basket and low interest rate basket respectively. The λ values show the relevance of each of the component copulae to the data. Thus a small λ value indicates the lack of a need for that particular copula component in order to model the associated 6 months or one year block of data. In contrast, for example a λ value for the Gumbel component very close to 1 indicates the block of data could be well modelled by a Gumbel copula alone. Hence, these plots convey the time varying significance of hypotheses about the presence of upper and lower tail dependence in each of the baskets over time. Examining these plots shows that in general the Clayton mixture weight tends to be lower when the Gumbel mixture weight is higher. It can also be seen that the Frank copula is systematically present in the mixture. In addition, there are periods which are dominated by one of the structural component copulae. That is, there is an asymmetric tendency for the presence of particular copula components over time when comparing the high and low interest rate baskets. The implications of this will be discussed in further detail in the discussions.

In terms of the second attribute, the strength of the copula dependence, I analyse this in several ways. Firstly through an analysis of the estimation copula parameter components over time, then through an analysis of the transformation of these copula parameters to rank correlations and finally through an analysis of the multivariate strength of the mixture copula tail dependence over time.

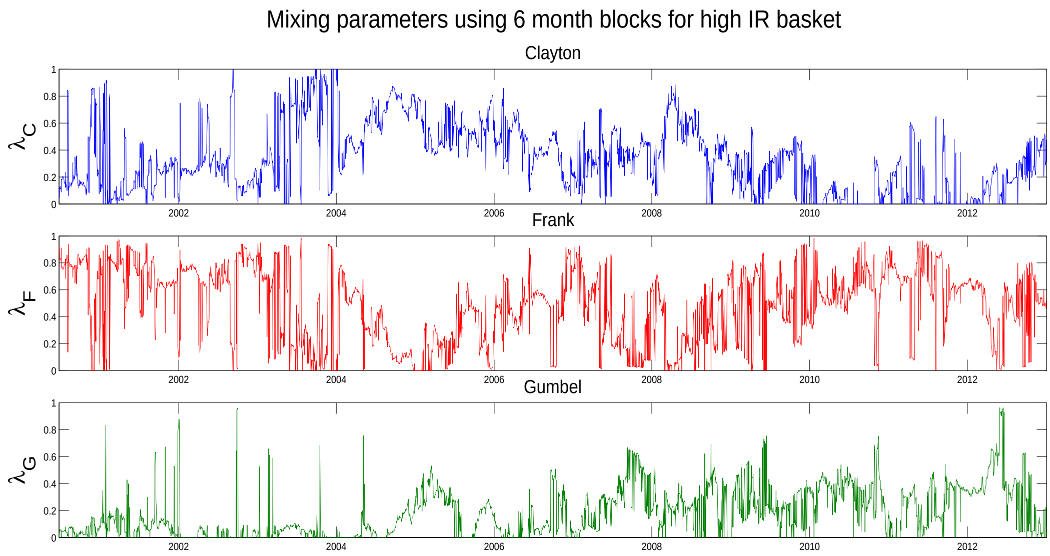


Figure 5.12: λ Mixing proportions of the respective Clayton, Frank and Gumbel copulae on the high interest rate basket, using 6 month blocks.

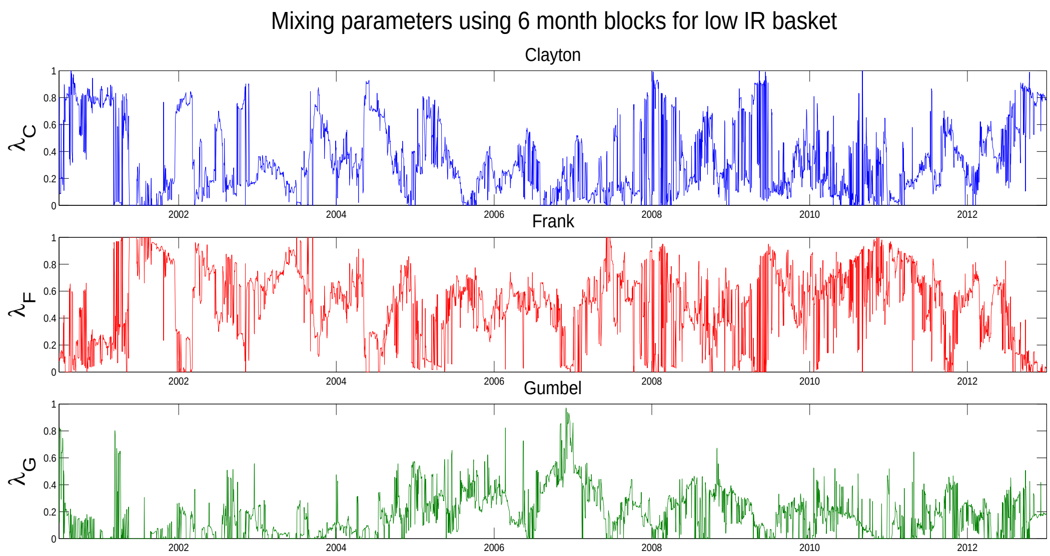


Figure 5.13: λ Mixing proportions of the respective Clayton, Frank and Gumbel copulae on the low interest rate basket, using 6 month blocks.

The individual component copula parameters can be seen in Figures 5.14 and 5.15 for the high interest rate basket and low interest rate basket respectively. The strength of the copula parameters in the baskets shows a large degree of variance during the period 04/01/2000 to 02/01/2013. One interesting observation is the very large spikes in the Gumbel copula parameter observed for the high interest rate basket from 2006 to 2007 and again in 2009. This was significant as it also corresponds to periods in which the Gumbel copula mixture weight was non-trivial.

The measure of concordance as captured by Kendall's tau is decomposed in this analysis according to each of the mixture components, scaled by the mixture weights λ, and can be seen in Figure 5.16 for the high interest rate basket and Figure 5.17 for the low interest rate basket. These plots provide a more intuitive picture of the time-varying contributions of the individual copulae to the dependence structure present in each of the baskets. Interestingly, one sees that the rank correlation contribution from the Frank copula indicates the presence of negative as well as positive rank correlations. In addition, as discussed with the mixture weights, there is perhaps some asymmetry present between the high and low interest rate baskets over time.

Perhaps the most interesting and revealing representation of the tail dependence characteristics of the currency baskets can be seen in Figures 5.18 to 5.23. Here it can be seen that there are indeed periods of heightened upper and lower tail dependence in the high interest rate basket. There is a noticeable increase in upper tail dependence at times of global FX volatility. Specifically, during late 2007, i.e. the global financial crisis, there is a sharp peak in upper tail dependence. Preceding this, there is an extended period of heightened lower tail dependence from 2004 to 2007, which could tie in with the building of the leveraged carry trade portfolio positions.

In understanding this analysis it is important to note that Figures 5.18 and 5.19 show the probability that one currency in the high interest rate basket or low interest rate basket respectively (which contain four or five currencies depending on the data availability during that time period) will have a move above/below a certain extreme threshold given that the other remaining currencies in the basket have had a move beyond this threshold. Then in Figures 5.20 and 5.21 I show the

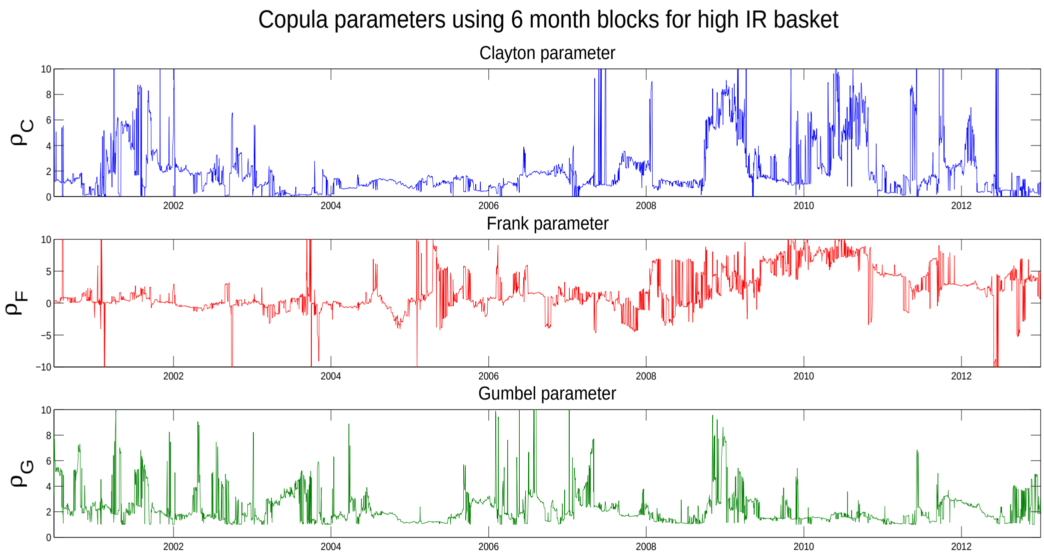


Figure 5.14: ρ Copula parameters for the Clayton, Frank and Gumbel copulae on the high interest rate basket, using 6 month blocks.

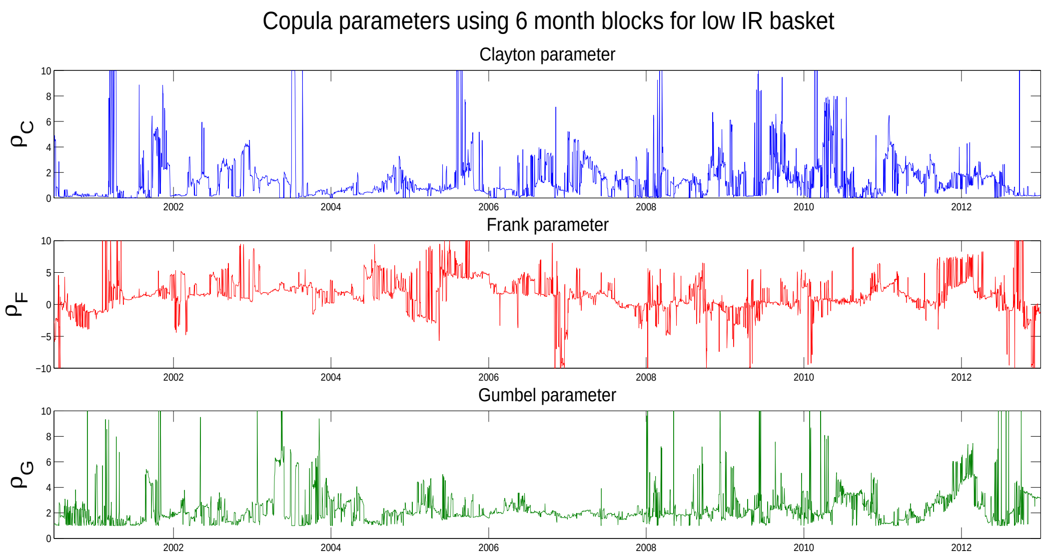


Figure 5.15: ρ Copula parameters for the Clayton, Frank and Gumbel copulae on the low interest rate basket, using 6 month blocks.

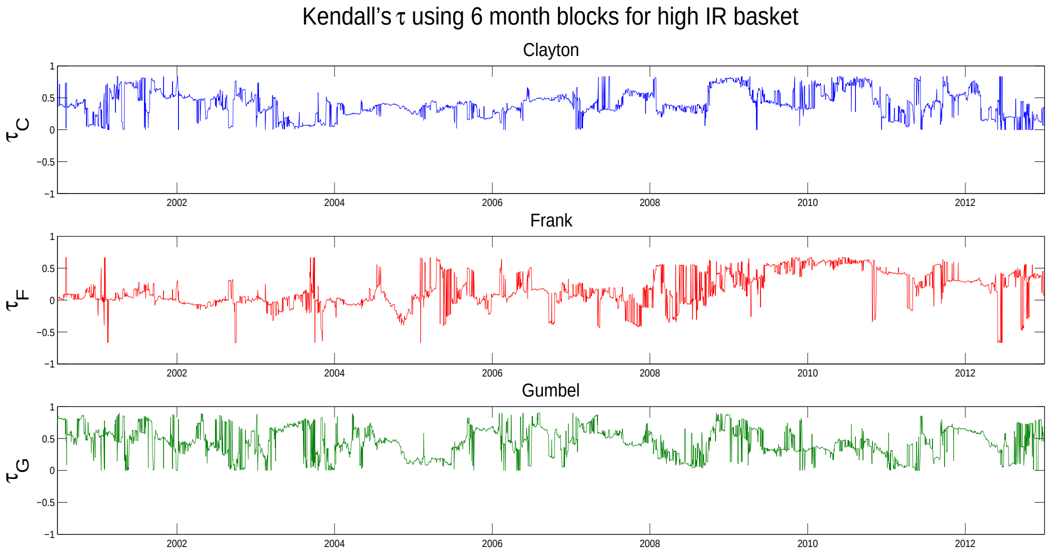


Figure 5.16: Kendall's τ for the Clayton, Frank and Gumbel copulae on the high interest rate basket, using 6 month blocks.

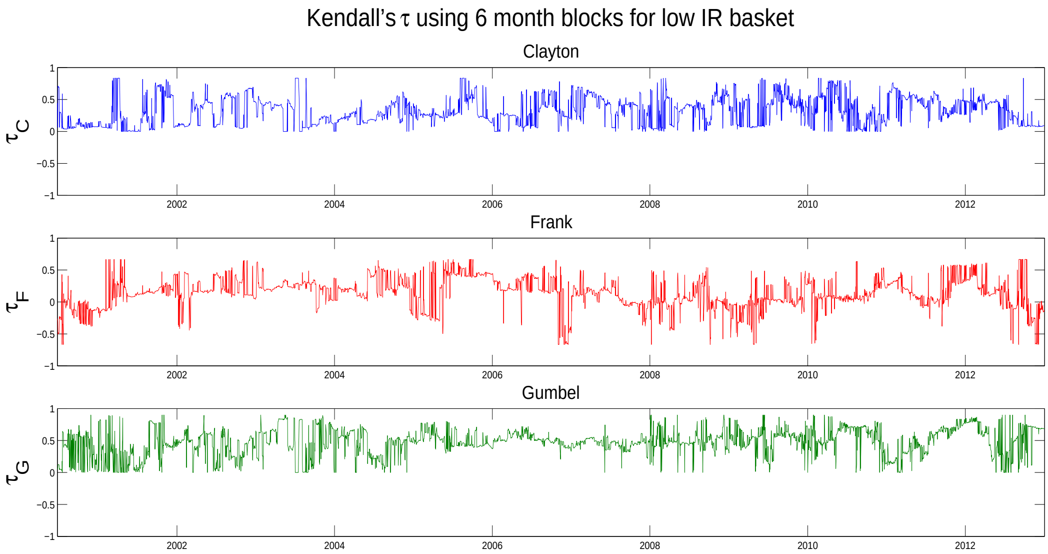


Figure 5.17: Kendall's τ for the Clayton, Frank and Gumbel copulae on the low interest rate basket, using 6 month blocks.

probability that two currencies in the basket will have a move above/below such an extreme threshold given that the other two currencies have had a move beyond this threshold. Finally, in Figures 5.22 and 5.23 I show the probability that three currencies in the basket will have a move above/below a certain threshold given that the remaining currency has had a move beyond this threshold.

To illustrate the relationship between heightened periods of significant upper and lower tail dependence features over time and to motivate the clear asymmetry present in the upper and lower tail dependence features between the high and low interest rate baskets over time I consider a further analysis. In particular, I compare in Figures 5.24 and 5.25 the tail dependence plotted against the daily average FX volatility (as calculated in Menkhoff et al. [2012a] on a monthly basis) and given by:

$$\sigma_t^{FX} = \sum_{k \in K_t} \left(\frac{|r_t^k|}{K_t} \right) \tag{5–10}$$

where $|r_t^k|$ is the absolute daily log return of currency k on day t, and K_t denotes the number of available currencies on day t.

In addition, in these figures I plot the VIX volatility index for the high interest rate basket and the low interest rate basket respectively for the period under investigation. The VIX is a popular measure of the implied volatility of S&P 500 index options - often referred to as the *fear index*. As such it is one measure of the market's expectations of stock market volatility over the next 30 days. It can clearly be seen here that in the high interest rate basket there are upper tail dependence peaks at times when there is increased stock market volatility, particularly post-crisis. However, I would not expect the two to match exactly since the VIX is not a direct measure of global FX volatility, but US equities volatility. Thus it can be concluded that investors' risk aversion clearly plays an important role in the tail behaviour of high interest rate currencies and more importantly in their dependence structure. This statement can also be associated to the globalization of financial markets and the resulting increase of the contagion risk between countries. This conclusion corroborates some of the recent results in the literature with regards to the skewness and the kurtosis features characterizing the currency carry trade portfolios, see Farhi and Gabaix [2008]; Brunnermeier

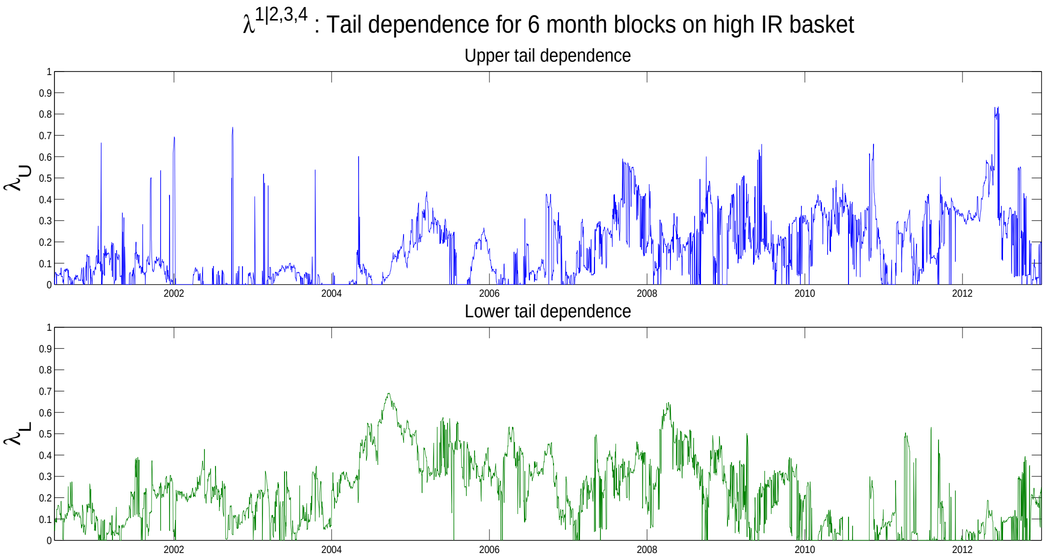


Figure 5.18: $\lambda^{1|234}$: 6 month blocks on high interest rate basket.

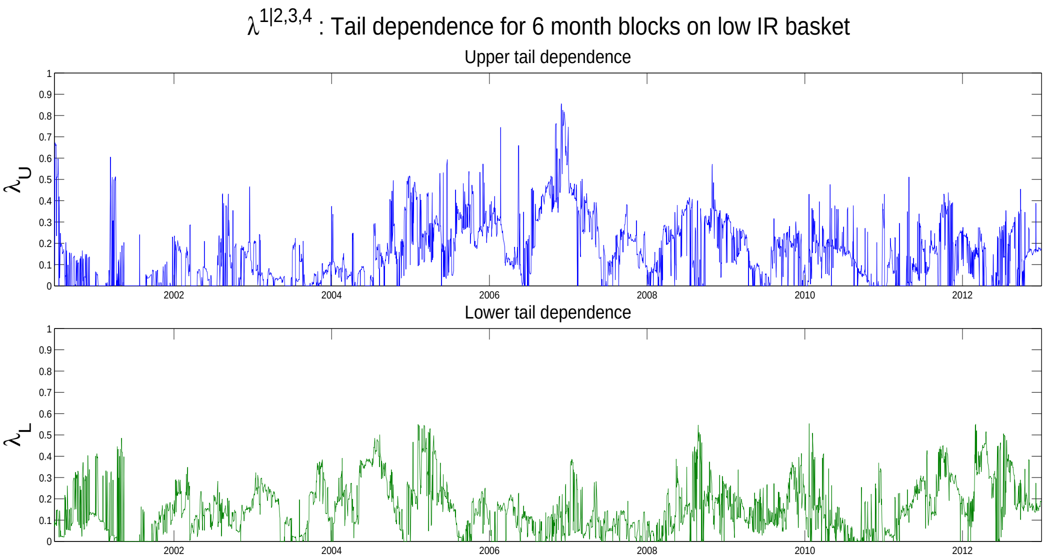


Figure 5.19: $\lambda^{1|234}$: 6 month blocks on low interest rate basket.

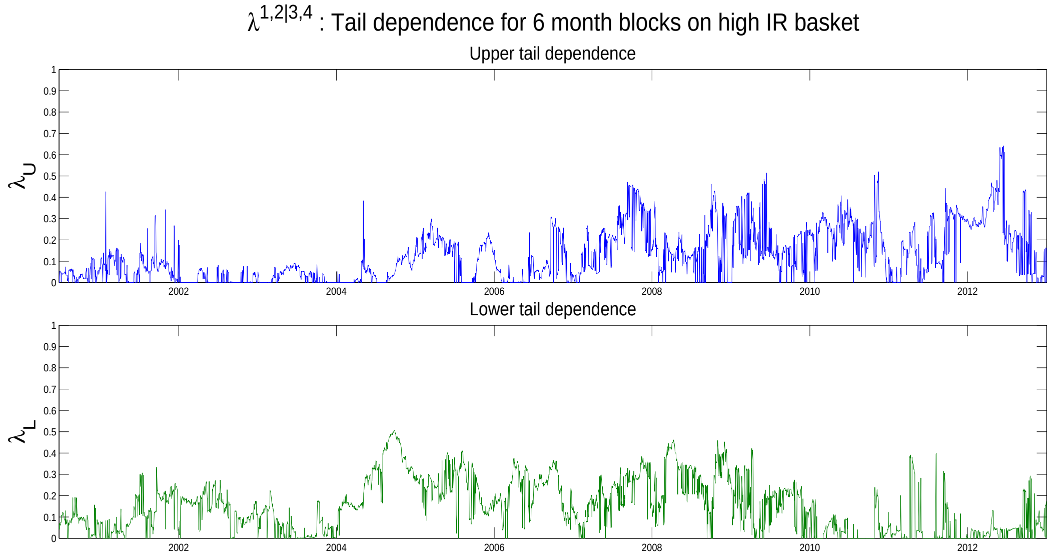


Figure 5.20: $\lambda^{12|34}$: 6 month blocks on high interest rate basket.

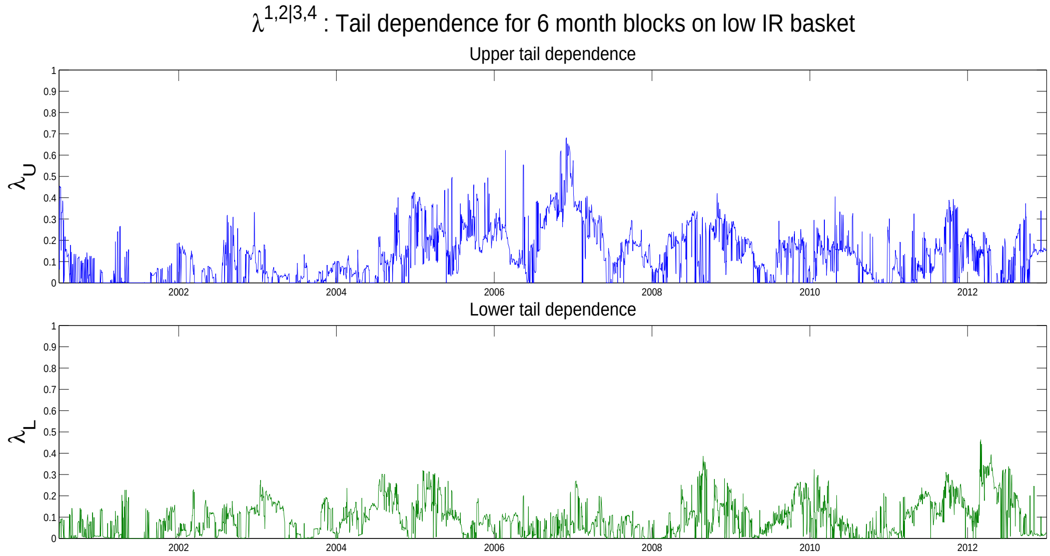


Figure 5.21: $\lambda^{12|34}$: 6 month blocks on low interest rate basket.

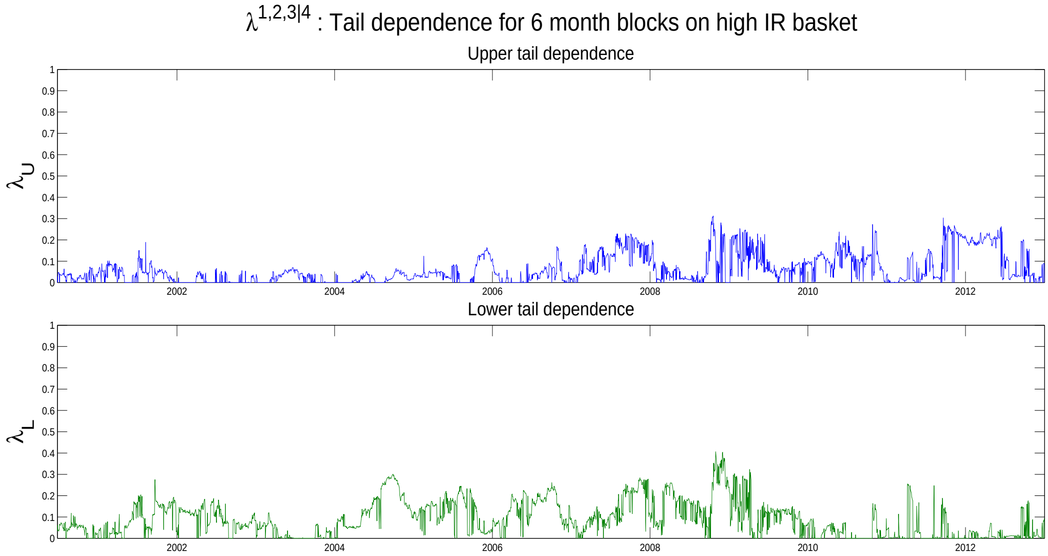


Figure 5.22: $\lambda^{123|4}$: 6 month blocks on high interest rate basket.

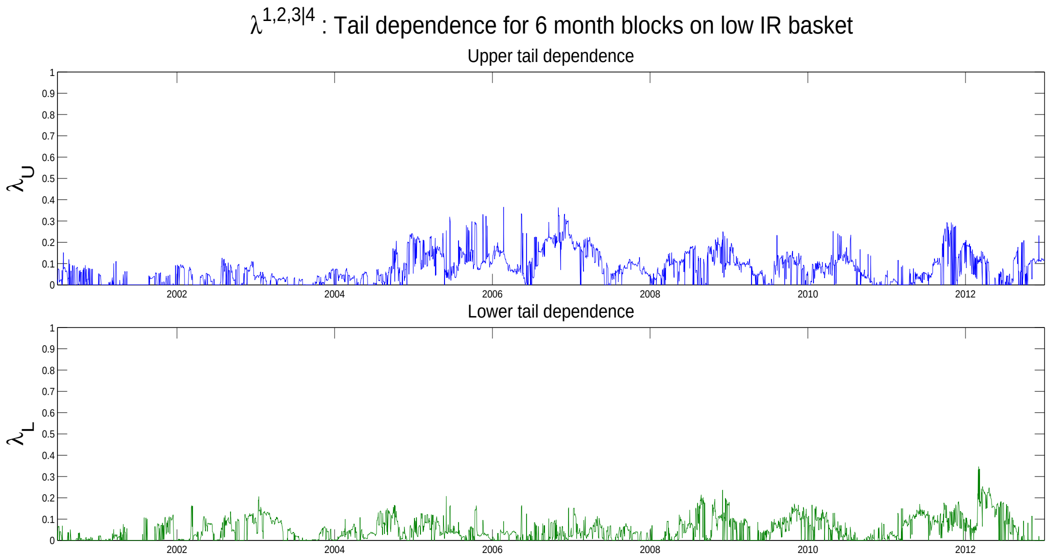


Figure 5.23: $\lambda^{123|4}$: 6 month blocks on low interest rate basket.

et al. [2008]; Menkhoff et al. [2012a].

The black lines plotted in Figures 5.24 and 5.25 furthermore display the mean tail dependence before and after August 2007 (which corresponds to the beginning of the global financial crisis). The data shows a large increase in upper tail dependence in the high interest rate basket after the crisis, as well as a smaller decrease in lower tail dependence. Interestingly there is very little difference in the mean tail dependence before and after the crisis for the low interest rate basket. The carry trade portfolios were particularly impacted by the sub-prime crisis as most of these currency positions were implemented and held by financial institutions which faced sudden difficulties to finance the leverage of their positions. Furthermore, another interesting point that can be made from the analysis of these two figures is the higher level of lower tail dependence before the financing crisis, especially between 2004 and 2007. The fact that during this three year period the VIX index was noticeably and continuously decreasing it is possible to imagine that this increase of the lower tail dependence results from lower risk aversion and the resulting tendency of investors to accordingly increase their leverage on risky positions, such as currency carry trades.

5.6 Pairwise Decomposition of Basket Tail Dependence

In the above analysis of model based parametric upper and lower tail dependence I focus on the joint extreme deviations in both the highest and the lowest interest rates currencies baskets. It is also informative to understand which pairs of currencies within a given currency basket contribute significantly to the downside or upside risks of the overall currency basket. In the class of Archimedean based mixtures considered in this thesis, the feature of exchangeability precludes decompositions of the total basket downside and upside risks into individual currency specific components. Here, I perform a decomposition of the risks within a basket, e.g. the downside risk of the funding basket, into contributions from each pair of currencies in the basket. This is achieved via a simple linear projection onto particular subsets of currencies in the portfolio that are of interest, which

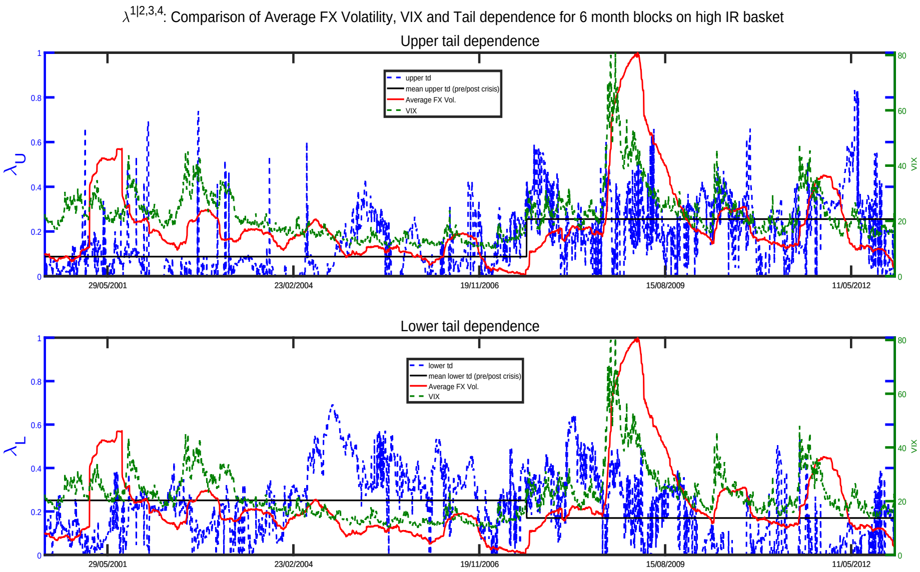


Figure 5.24: Comparison of Average FX volatility and Equity Volatility Index (VIX) with upper and lower tail dependence of the high interest rate basket.

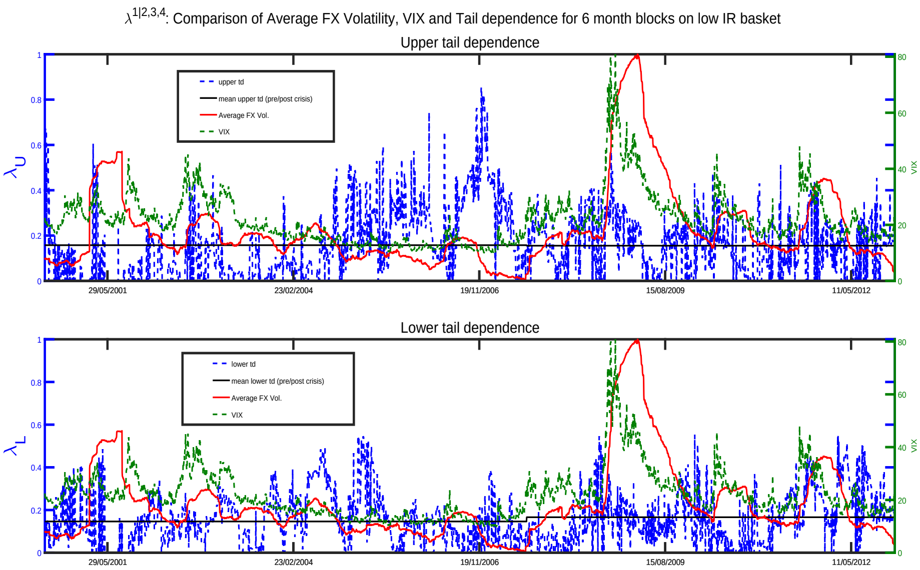


Figure 5.25: Comparison of Average FX volatility and Equity Volatility Index (VIX) with upper and lower tail dependence of the low interest rate basket.

leads for example to the following expression:

$$\mathbb{E}\left[\hat{\lambda}_u^{i|1,2,\ldots,i-1,i+1,\ldots,n}\middle|\hat{\lambda}_u^{2|1},\hat{\lambda}_u^{3|1},\hat{\lambda}_u^{3|2},\ldots,\hat{\lambda}_u^{n|n-1}\right]=\alpha_0+\sum_{i\neq j}^{n}\alpha_{ij}\hat{\lambda}_u^{i|j}, \tag{5–11}$$

where $\hat{\lambda}_u^{i|j}$ is the pairwise non-parametric tail dependence between currency i and currency j, and $\hat{\lambda}_u^{i|1,2,\ldots,i-1,i+1,\ldots,n}$ is a random variable since it is based on parameters of the mixture copula model which are themselves functions of the data and therefore random variables. Since the value of the tail dependence is bounded, i.e. $0 \leq \hat{\lambda}_u^{i|j} \leq 1$, this regression could be performed via a generalised linear model (glm) with a logit link function. However, in the empirical investigation here it was found that the restriction wasn't an issue. Such a simple linear projection will then allow one to interpret directly the marginal linear contributions to the upside or downside risk exposure of the basket obtained from the model, according to particular pairs of currencies in the basket by considering the coefficients α_{ij}, i.e. the projection weights. To perform this analysis it is necessary to obtain estimates of the pairwise tail dependences in the upside and downside risk exposures $\hat{\lambda}_u^{i|j}$ and $\hat{\lambda}_l^{i|j}$ for each pair of currencies $i, j \in \{1, 2, \ldots, n\}$. I obtain this through a non-parametric (model-free) estimator as discussed in Chapter 3. In particular, I will focus on the estimator presented in equation 3–24.

5.6.1 Non-Parametric Tail Dependence Results

In order to examine the contribution of each pair of currencies to the overall n-dimensional basket tail dependence I calculated the corresponding non-parametric pairwise tail dependences for each pair of currencies. In Figure 5.26 the average upper and lower non-parametric tail dependence for each pair of currencies during the 2008 Credit crisis can be seen, with the 3 currencies most frequently in the high interest rate and the low interest rate baskets labelled accordingly. The lower triangle represents the non-parametric pairwise lower tail dependence and the upper triangle represents the non-parametric pairwise upper tail dependence. Similarly, in Figure 5.27 the pairwise non-parametric tail dependences averaged over the last 12 months (01/02/2013 to 29/01/2014) can be seen. Comparing this heat map to the heat map during the Credit crisis (Figure 5.26) it can be seen

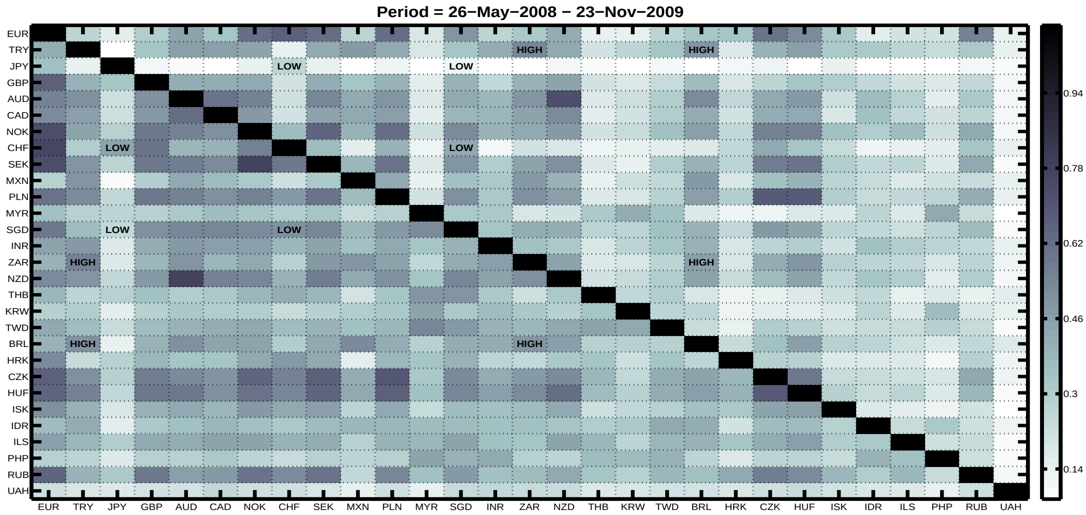


Figure 5.26: Heat map showing the strength of non-parametric tail dependence between each pair of currencies averaged over the 2008 Credit crisis period. Lower tail dependence is shown in the lower triangle and upper tail dependence is shown in the upper triangle. The 3 currencies most frequently in the high interest rate and the low interest rate baskets are labelled.

that in general there are lower values of tail dependence amongst the currency pairs.

Remark 5.6.1. *If one was trying to optimise their currency portfolio with respect to the tail risk exposures, i.e. to minimise negative tail risk exposure and maximise positive tail risk exposure, then one would sell short currencies with high upper tail dependence and low lower tail dependence whilst buying currencies with low upper tail dependence and high lower tail dependence.*

I then performed linear regression of the pairwise non-parametric tail dependence on the respective basket tail dependence for the period 01/02/2013 to 29/01/2014 for the days on which the 3 currencies all appeared in the basket (224 out of 250 for the lower interest rate basket and 223 out of 250 for the high interest rate basket). The regression coefficients and R^2 values can be seen in Table 5.3. This can be interpreted as the relative contribution of each of the 3 currency pairs to the overall basket tail dependence. It can be noted that for the low interest rate lower tail dependence and for the high interest rate upper tail

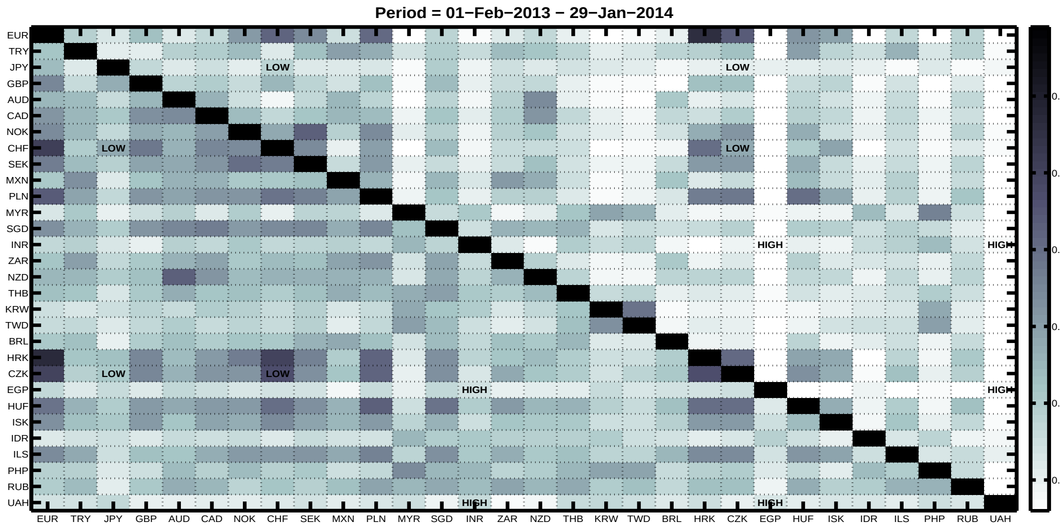


Figure 5.27: Heat map showing the strength of non-parametric tail dependence between each pair of currencies averaged over the last 12 months (01/02/2013 to 29/01/2014). Lower tail dependence is shown in the lower triangle and upper tail dependence is shown in the upper triangle. The 3 currencies most frequently in the high interest rate and the low interest rate baskets are labelled.

dependence there is a significant degree of cointegration between the currency pair covariates and hence it may be possible to use a single covariate due to the presence of a common stochastic trend.

5.7 Understanding the Tail Exposure Associated with the Carry Trade and Its Role in the UIP Puzzle

As was discussed in Section 5.3, the tail exposures associated with a currency carry trade strategy can be broken down into the upside and downside tail exposures within each of the long and short carry trade baskets. In order to assess the potential impact of these tail exposures on portfolio returns it is interesting to explore a risk adjustment approach. The downside relative exposure adjusted returns are obtained by multiplying the monthly portfolio returns by one minus

Table 5.3: Pairwise non-parametric tail dependence regressed on respective basket tail dependence for the period 01/02/2013 to 29/01/2014 (standard errors are shown in parentheses). The 3 currencies most frequently in the respective baskets are used as independent variables.

Low IR Basket	**Constant**	**CHF JPY**	**CZK CHF**	**CZK JPY**	R^2
Upper TD	0.22 (0.01)	0.02 (0.03)	0.18 (0.02)	0.38 (0.05)	0.57
Lower TD	0.71 (0.17)	-0.62 (0.25)	-0.38 (0.26)	0.23 (0.32)	0.28

High IR Basket	**Constant**	**EGP INR**	**UAH EGP**	**UAH INR**	R^2
Upper TD	0.07 (0.01)	-0.06 (0.33)	0.59 (0.08)	2.37 (0.42)	0.40
Lower TD	0.10 (0.02)	0.56 (0.05)	0.44 (0.08)	-0.40 (0.07)	0.44

the upper and the lower tail dependence values calculated using a lookback period of six months, i.e. there is no forward looking bias here, present respectively in the high interest rate basket and the low interest rate basket at the corresponding dates. This adjustment leads to higher discount of the returns associated to the high interest rates basket when the upper tail dependence of this basket is more important and conversely for the low interest rate basket and the related lower tail dependence. The upside relative exposure adjusted returns are obtained by multiplying the monthly portfolio returns by one plus the lower and upper tail dependence present respectively in the high interest rate basket and the low interest rate basket at the corresponding dates. Note that I refer to these as relative exposure adjustments only for the tail exposures since I do not quantify a market price per unit of tail risk. However, this is still informative as it shows a decomposition of the relative exposures from the long and short baskets with regard to extreme events.

As can be seen in Figure 5.28, the relative adjustment to the absolute cumulative returns for each type of downside exposure is greatest for the low interest

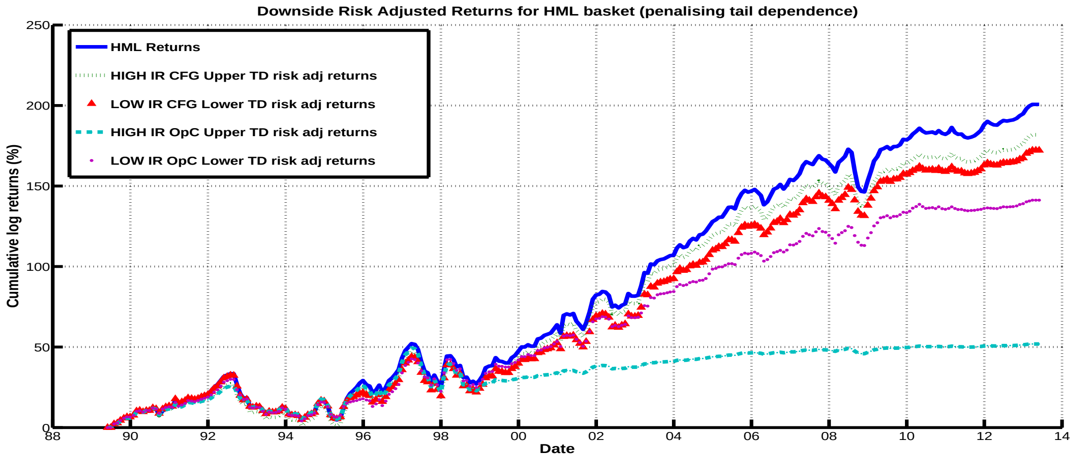


Figure 5.28: Cumulative log returns of the carry trade portfolio (HML = High interest rate basket Minus Low interest rate basket). Downside exposure adjusted cumulative log returns using upper/lower tail dependence in the high/low interest rate basket for the CFG copula and the OpC copula are shown for comparison.

rate basket, except under the OpC model, but this is due to the very poor fit of this model to baskets containing more than 2 currencies, which thus transfers to financial risk exposures. This is interesting because intuitively one would expect the high interest rate basket to be the largest source of tail exposure. However, one should be careful when interpreting this plot, since it is the extremal tail exposure. The analysis may change if one considered the intermediate tail risk exposure, where the marginal effects become significant. Similarly, Figure 5.29 shows the relative adjustment to the absolute cumulative returns for each type of upside exposure is greatest for the low interest rate basket. The same interpretation as for the downside relative exposure adjustments can be made here for upside relative exposure adjustments.

5.8 Conclusions

In this part of the thesis, I have investigated one of the most robust puzzles in international finance, namely the currency carry trade. This market phenomenon is particularly interesting from a theoretical standpoint as well as for the under-

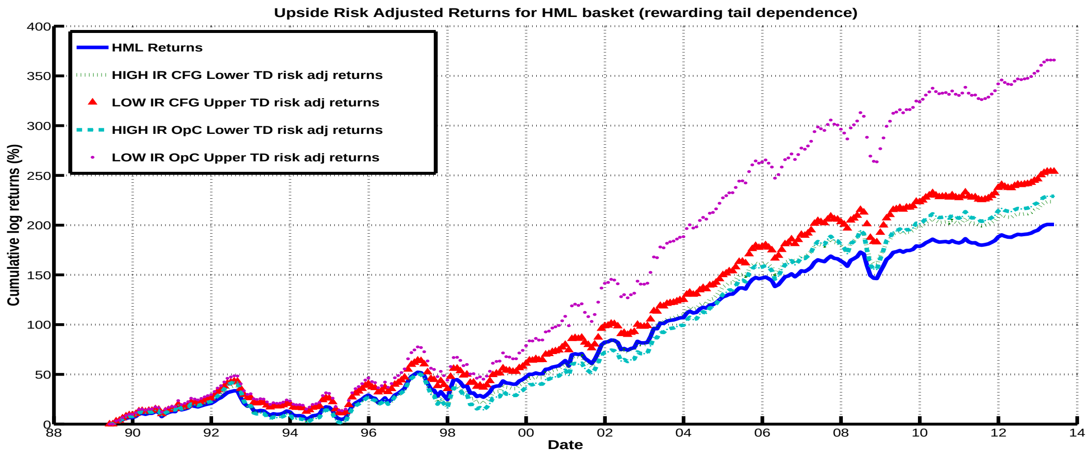


Figure 5.29: Cumulative log returns of the carry trade portfolio (HML = High interest rate basket Minus Low interest rate basket). Upside exposure adjusted cumulative log returns using lower/upper tail dependence in the high/low interest rate basket for the CFG copula and the OpC copula are shown for comparison.

standing of financial market mechanisms. It has been demonstrated empirically that the currency markets were violating a fundamental relation in finance connecting the currency exchange rates and the interest rates associated with two different countries.

The main contribution of this part of the thesis has been to propose a rigorous statistical modelling approach, which captures the specific statistical features of both the individual currency log-return distributions as well as the joint features such as the dependence structures prevailing between all the exchange rates.

In achieving this goal, I first assessed the marginal statistical features of each of the 20 currencies on an assumed locally stationary sliding window of six months, over all the trading days in the period 04/01/2000 to 02/01/2013. I found that a simple log-normal marginal distribution would not produce a suitable statistical fit for some of the key currencies that are regularly present in the high interest rate basket throughout this period. As detailed in the results section this was notably the case in unstable economies such as developing countries (for instance Turkey, Brazil or South Africa) where political stability or default risk create sudden and violent adjustments to their currency exchange rates with other countries. It can be noted that these currencies are still of direct significance to the study of global currency carry trade strategies since according to the modern portfolio theory this intrinsic country specific risk borne by an investor in these currencies can be diversified and mitigated by adding to the considered portfolio other currencies which depend themselves on different sources of intrinsic country specific risk. This would effectively establish a diversified portfolio of currencies violating the UIP hypothesis and would thus provide a very attractive average return for a very limited risk which has been the conclusion of several recent empirical studies in the finance literature.

The conclusion of this is that in this analysis these currencies are not excluded from the high interest rate basket analysis, even though they may demonstrate attributes resulting primarily from significant changes in their countries political and financial structure. One can also consider a restricted set of only developed market currencies, e.g. G10 currencies, as is the case in Part II of this thesis due to the open interest rate data being limited to only a subset of developed market currencies. As a result I needed to obtain more flexible marginal models to capture the features of these currencies more adequately. Consequently I modelled each currency exchange rate return marginally via a flexible three parameter parametric model which offers a wide range of skew-kurtosis relationships as well as the possibility of light exponential tails and heavier sub-exponential tail behaviours such as the log-normal member. The parametric family of distributions I selected for this purpose was the log-generalized gamma distribution.

Having modelled the marginal attributes of the high and low interest rate currency baskets over time adequately, the main emphasis was then to assess the multivariate dependence features of the currency baskets. In particular how this may change over time within a given basket, where I was particularly interested in the effect of the composition of the basket over time, and the response of the multivariate dependence features of the modelled basket and how it may respond in periods of heightened market volatility versus more stable periods. In addition to this within basket temporal analysis, from the perspective of undertaking a currency carry trade strategy, one would need to consider the relative relationships between the temporal dependence features of the high interest rate and low interest rate currency baskets. I demonstrate several interesting features from the model fits relating to asymmetries between the high and low interest rate baskets over time, especially during periods of high volatility in global markets. One way I ascertained such periods was through a comparison of the VIX versus features of the multivariate dependence relationships I modelled. Importantly I found substantial evidence to support arguments for time varying behaviours in the structural dependence hypotheses posed about the currency baskets, as captured by the relevant contributing copula components to the multivariate mixture model. In addition, substantial evidence was found for significant upper and lower tail dependences features in both the high and low interest rate baskets, which again displayed interesting asymmetries between both baskets over time.

The financial interpretation of the significance of these findings is related to the fact that it demonstrates that historically average rewards from a currency carry trade portfolio can be exposed to a significant risk of large losses arising from joint adverse movements in the currencies that would typically comprise the high and low interest rate baskets that an investor would go long and short on when trading. Hence, I conclude that our second contribution to the literature has been to rigorously demonstrate that such assertions relating to the profitability of the currency carry trade (based on Sharpe ratio for instance) are failing to appropriately take into consideration an important component of the risk which characterizes these types of portfolios of currencies named carry trade portfolios.

I conclude that indeed the copula theory employed in this thesis allows me to demonstrate statistically that beyond the intrinsic risk associated to high interest rate countries (which are generally paying higher interest rates to compensate for a higher risk) typically studied in the literature from a marginal perspective, another source of risk plays an important role. This second source of risk is related to the dependence structures linking these high interest rate currencies, more specifically the significant tail dependence features observed in this model analysis. Through an Archimedean copulae mixture model the significant presence of tail dependence among high interest rate currencies is shown. This tail dependence could have dramatic consequences on the carry trade portfolio's risk profile when appropriately accounted for in risk reward analysis. As a matter of fact, the tail dependence directly influences the diversification of the assets during stress periods and thus reduces the appealing convergence property stated by the modern portfolio theory.

In other words, this copula based probabilistic modelling approach allows me to demonstrate that besides the intrinsic risk associated to each particular high interest rate currency, another factor constitutes a determining source of risk which turns out to be the level of risk aversion prevailing in the market. It was demonstrated in this analysis that both upper and lower tail dependence features displayed significant association and asymmetries with each other between the high and low interest rate baskets during periods of relative financial stability versus periods of heightened market volatility.

These tail dependence features in the high interest rate basket were significantly increasing during crisis periods leading to an increased amount of risk associated with utilising such currency baskets (which were no longer diversified due to the presence of significant tail dependence features) in a carry trade. That being said, a rational portfolio manager's natural risk aversion tells them that they should receive an additional remuneration in order to offset any additional sources of risk associated to an investment. Therefore, to properly assess the profitability of the currency carry trade, such tail dependence features should be incorporated into the analysis of such risk-rewards when developing a trading strategy. To conclude, this investigation rigorously tempers the too often claimed attractiveness of the currency carry trade and provides to investors a risk management tool in order to control and monitor the risk contained in such positions.

Part II

Covariance Factor Modelling Contributions

Chapter 6

Part II Overview

In the second part of this thesis, the focus shifts away from the multivariate tail dependence modelling approaches presented in Part I, in which currencies are given equal weights in the portfolio (as is sometimes the case in the asset management industry). In order to provide a risk management solution, in which currency weights are considered, a covariance regression framework is introduced and its novel application to investigate how observable and interpretable explanatory factors influence the covariance structure of currency returns is presented.

Chapter 7 reviews the techniques available in the literature for modelling and forecasting covariance. The multivariate GARCH modelling approach and its many extensions are presented along with a discussion of the advantages and disadvantages of the various models. Then an approach that facilitates the incorporation of observable factors into the conditional covariance of the standard factor model is introduced, thus allowing for heteroskedastic unconditional covariance. Furthermore, a method by which the covariance factor models can be utilised in combination with time series models for the factors in order to forecast heteroskedastic covariance in an interpretable manner is detailed

Chapter 8 investigates how the behaviour of speculative traders impacts the dependence structure of currency carry trade baskets. Speculative trading volume factors are introduced into the covariance factor models presented in Chapter 7 allowing one to forecast portfolio covariance and thus perform risk based asset allocation in currency carry trades.

Chapter 7

Covariance Forecasting

In this chapter, the techniques available in the literature for modelling and forecasting covariance are reviewed. Accurate covariance forecasts are of key importance in providing a robust risk management approach to currency weight allocation in carry trade portfolios. The ARIMA models considered for the explanatory covariates are first presented. Building on this model, the multivariate GARCH modelling approach and its many extensions are presented along with a discussion of the advantages and disadvantages of the various models. Then an approach that facilitates the incorporation of observable factors into the conditional covariance of the standard factor model is introduced, thus allowing for heteroskedastic unconditional covariance. Furthermore, a method by which the covariance factor models can be utilised in combination with time series models for the factors in order to forecast covariance is detailed.

7.1 Univariate Time Series Models

In this section, the univariate autoregressive integrated moving average (ARIMA) models considered for modelling and forecasting the explanatory covariates are first presented. Following this, the extension to consider time-varying volatility in the form of ARCH models and furthermore the generalisation to GARCH models is detailed. This univariate GARCH model then forms the base of the multivariate models discussed in Section 7.2.

7.1.1 Univariate ARIMA Model

Autoregressive moving average models were introduced in the thesis of Peter Whittle, see Whittle [1951]. These models became popular following the work of Box and Jenkins, see Box et al. [2015]. An ARIMA model expresses the conditional mean of a time series Y_t as a function of both past observations and past innovations. Here Y_t is the time series resulting from first differencing the time series, X_t, d times.

Definition 38. ***ARIMA(p,d,q) Model***

$$Y_t = (1 - L)^d X_t \tag{7–1}$$

$$Y_t = c + \epsilon_t + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \tag{7–2}$$

where $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ are parameters, c is a constant, ϵ_t is a white noise process, and L is the lag operator, i.e. $LX_t = X_{t-1}$.

While the ARIMA model is a popular time series model in practice due to its parsimony, it is assumed that the error process is homoskedastic over time. However, it is a stylized fact that financial time series data contain heteroskedastic error processes.

7.1.2 Univariate ARCH Model

The ARIMA time series models for the mean are extended in the seminal paper of Engle [1982] to produce autoregressive conditional heteroskedastic (ARCH) models, which allow for time-varying volatility. These models assume the variance of the current error term is a function of the squares of the previous error terms. ARCH models are commonly employed in modelling financial time series that exhibit time-varying volatility clustering.

Definition 39. ***ARCH(q) Model***

$$\epsilon_t = \sigma_t z_t \tag{7–3}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \tag{7–4}$$

where z_t is a strong white noise process, $\alpha_0 > 0, \alpha_i \geq 0$ and $i > 0$.

7.1.3 Univariate GARCH Model

An extension of the ARCH model to allow past conditional variances to appear in the current conditional variance equation is proposed in Bollerslev [1986]. These generalized autoregressive conditional heteroskedastic (GARCH) models assume an ARMA model for the error covariance and thus allow a more flexible lag structure and in many cases permit a more parsimonious model.

Definition 40. ***GARCH(p,q) Model***

$$\epsilon_t = \sigma_t z_t \tag{7–5}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{7–6}$$

where z_t is a strong white noise process, $\alpha_0 > 0, \alpha_i \geq 0$, $\beta_j \geq 0$ and $i > 0$.

An introduction to ARCH and GARCH Models can be found in Engle [2001]. Furthermore, a comprehensive overview of these models can be seen in the book, Brooks [2014].

Following the introduction of the GARCH model there have been many extensions proposed in the literature:

1. Non-linear GARCH (NGARCH) Engle and Ng [1993] allows negative returns to increase future volatility by a larger amount than positive returns of the same magnitude.

2. Integrated GARCH (IGARCH) Engle and Bollerslev [1986] is a restricted version of the GARCH model, where the persistent parameters sum up to one, and imports a unit root in the GARCH process.

3. Exponential GARCH (EGARCH) Nelson [1991] models the log of the variance.

4. GARCH-in-mean (GARCH-M) Engle et al. [1987] incorporates the effect of the volatility of the series on the mean.

5. Quadratic GARCH (QGARCH) Sentana [1995] models asymmetric effects of positive and negative shocks.

6. Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) Glosten et al. [1993] allows for differing effects of negative and positive shocks, taking into account the leverage phenomenon.

7. Threshold GARCH (TGARCH) Zakoian [1994] which is similar to the GJR-GARCH model but instead models the standard deviation.

8. Family Garch (fGARCH) Hentschel [1995] is an omnibus model that nests a variety of other popular symmetric and asymmetric GARCH models.

9. Continuous-time Garch (COGARCH) Klüppelberg et al. [2004] is a continuous-time generalization of the discrete-time GARCH(1,1) process.

For synopses of GARCH model extensions in the univariate setting see Bollerslev et al. [1994]; Hentschel [1995]; Palm [1996]; Shephard [1996].

Remark 7.1.1. *The Exponentially Weighted Moving Average (EWMA) model is a simple practical alternative to fitting a GARCH model (in fact it is a GARCH(1,1) model with no long-run average volatility term). However, the subjective specification of λ, the volatility persistence parameter, is required. The RiskMetrics approach, see J. P. Morgan [1996], uses the EWMA model with $\lambda = 0.94$ for daily data and $\lambda = 0.97$ for weekly data.*

Definition 41. ***EWMA Model***

$$\sigma_t^2 = (1 - \lambda) r_{t-1}^2 + \lambda \sigma_{t-1}^2 \tag{7–7}$$

where $0 \leq \lambda \leq 1$.

7.2 Multivariate GARCH Framework

In this section, the key flavours of multivariate GARCH models will be introduced along with a discussion of the assumptions necessary and the advantages and disadvantages of the various models. For a comprehensive survey of multivariate GARCH models see Bauwens et al. [2006]; Silvennoinen and Teräsvirta [2009].

The multivariate extension of the GARCH framework is as follows: Consider a vector stochastic process $\boldsymbol{y}_t$ of dimension $N \times 1$ conditioned on the sigma field $\mathcal{F}_{t-1}$. Let $\boldsymbol{\theta} \in \mathbb{R}^d$ denote a finite vector of parameters associated with the mean and covariance of the process. Here, $\boldsymbol{H}_t(\boldsymbol{\theta})$ represents the conditional covariance matrix at time t. Both $\boldsymbol{H}_t(\boldsymbol{\theta})$ and $\boldsymbol{\mu}_t(\boldsymbol{\theta})$ depend on the unknown parameter vector $\boldsymbol{\theta}$, which can in most cases be split into two disjoint parts. Hereafter, for readability $\boldsymbol{H}_t(\boldsymbol{\theta})$ will be denoted by $\boldsymbol{H}_t$ and $\boldsymbol{\mu}_t(\boldsymbol{\theta})$ will be denoted by $\boldsymbol{\mu}_t$.

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t \ , \tag{7–8}$$

where $\boldsymbol{\mu}_t$ is the conditional mean vector and

$$\boldsymbol{\epsilon}_t = \boldsymbol{H}_t^{1/2} \boldsymbol{z}_t \ , \tag{7–9}$$

where $\boldsymbol{H}_t^{1/2}$ is a $N \times N$ positive definite matrix and $\boldsymbol{z}_t$ is a $N \times 1$ random vector such that:

$$\mathbb{E}(\boldsymbol{z}_t) = 0 \tag{7–10}$$

$$\text{and} \quad \mathrm{Var}(\boldsymbol{z}_t) = \boldsymbol{I}_N \ , \tag{7–11}$$

where $\boldsymbol{I}_N$ is the identity matrix of order N.
Thus it can be seen that:

$$\mathrm{Var}(\boldsymbol{y}_t | \mathcal{F}_{t-1}) = \mathrm{Var}_{t-1}(\boldsymbol{y}_t) = \mathrm{Var}(\boldsymbol{\epsilon}_t) \tag{7–12}$$

$$= \boldsymbol{H}_t^{1/2} \, \mathrm{Var}_{t-1}(\boldsymbol{z}_t) (\boldsymbol{H}_t^{1/2})' \tag{7–13}$$

where $\mathcal{F}_{t-1}$ is the natural filtration of $\boldsymbol{y}_{t-1}$.

In extending the GARCH framework to the multivariate level and hence facilitating the modelling of the time-varying covariances of variables a number of issues come into consideration. The following are important characteristics of such a model:

- flexibility to capture the dynamics of conditional variances and covariances;
- parsimony to allow robust estimation and interpretation;
- positive definiteness of the covariance matrix.

The literature consists of a wide range of proposed multivariate GARCH models due in part to the difficulty in combining these characteristics. A summary of some key MGARCH models follows below. The VEC model proposed by Bollerslev et al. [1988] was the first GARCH model and is the root of the very prolific literature about multivariate GARCH.

7.2.1 VEC-GARCH Model

The Vectorised GARCH (VEC) model, see Bollerslev et al. [1988], presents the most general extension of the univariate GARCH model in that each element of $\boldsymbol{H}_t$ is a linear function of the lagged squared errors and cross-products of errors and lagged values of the elements of $\boldsymbol{H}_t$.

Definition 42. ***VEC(p,q) Model***

$$vech(\boldsymbol{H}_t) = \boldsymbol{c} + \sum_{i=1}^{q} \boldsymbol{A}_i vech(\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}'_{t-i}) + \sum_{j=1}^{p} \boldsymbol{B}_j vech(\boldsymbol{H}_{t-j}) \tag{7–14}$$

where $vech(\cdot)$ is an operator that stacks the columns of the lower triangular part of its argument square matrix, c is an $N(N+1)/2 \times 1$ vector, and $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are $N(N+1)/2 \times N(N+1)/2$ parameter matrices.

As mentioned above, the VEC model is the most general MGARCH model and thus is very flexible. However, to ensure the positive definiteness of $\boldsymbol{H}_t$ a set of restrictive sufficient conditions must be satisfied, see Gouriéroux [2012] for details. Furthermore, the number of model parameters is very large unless N is small, e.g. for a VEC(1,1) model of a 5 dimensional time series the number of parameters, N_p, is given by $(p+q)(N(N+1)/2)^2 + N(N+1)/2 = 465$.

In addition the estimation of the parameters in the VEC model is computationally demanding since $\boldsymbol{H}_t$ needs to be inverted for every t in the likelihood at each iteration. The likelihood function of such a large number of parameters becomes very flat, and so convergence problems can be a serious issue in the optimisation routine.

Thus it is clear that some simplifying assumptions and structure has to be enforced on the covariance matrix. A quite severe restriction is considered in Bollerslev et al. [1988] in which the parameter matrices $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are diagonal matrices. This model contains $(p+q+1)N(N+1)/2$, e.g. for a Diagonal-VEC(1,1) model of a 5 dimensional time series the number of parameters is equal to 45. However, this model allows no interaction between the different conditional variances and covariances.

7.2.2 BEKK Model

Due to the difficulty of ensuring the positive definiteness of $\boldsymbol{H}_t$ without imposing strong restrictions on the parameters, Engle and Kroner [1995] introduced a new approach, named the Baba, Engle, Kraft and Kroner (BEKK) Model, that provides positive definiteness by construction.

Definition 43. ***BEKK(p,q,K) Model***

$$\boldsymbol{H}_t = \boldsymbol{C}\boldsymbol{C}' + \sum_{i=1}^{q}\sum_{k=1}^{K} \boldsymbol{A}'_{ki}\boldsymbol{\epsilon}_{t-i}\boldsymbol{\epsilon}'_{t-i}\boldsymbol{A}_{ki} + \sum_{j=1}^{p}\sum_{k=1}^{K} \boldsymbol{B}'_{kj}\boldsymbol{H}_{t-j}\boldsymbol{B}_{kj} \tag{7–15}$$

where $\boldsymbol{C}$, $\boldsymbol{A}_i$ and $\boldsymbol{B}_j$ are $N \times N$ parameter matrices. The summation limit K determines the generality of the process. Note $\boldsymbol{C}$ is lower triangular to ensure the positive definiteness of $\boldsymbol{H}_t$.

Whenever $K > 1$ there is an identification problem since there are several possible parameterizations that yield the same representation of the model, see Engle and Kroner [1995] for details. Estimation of a BEKK model is also computationally expensive due to a number of matrix inversions. The number of parameters in the BEKK model is still quite substantial, e.g. for a BEKK(1,1,1) model of a 5 dimensional time series the number of parameters, N_p, is given by $(p+q)KN^2 + N(N+1)/2 = 65$.

Remark 7.2.1. *Since VEC models and BEKK models contain such a high number of unknown parameters, even after imposing restrictions, they are rarely used for $N > 3$. In the literature this issue has been notably addressed by imposing a common dynamic structure on the elements of $\boldsymbol{H}_t$ via factor models or orthogonal models.*

7.2.3 Factor-GARCH Model

In order to address the curse of dimensionality issue a factor based approach was proposed in Engle et al. [1990]. Here, the authors assume that the observations are generated by a *small number* of factors that are conditionally heteroskedastic and have a GARCH structure.

Definition 44. ***Factor GARCH Model***

$$\boldsymbol{H}_t = \boldsymbol{\Omega} + \sum_{k=1}^{K} \boldsymbol{\omega}_k \boldsymbol{\omega}'_k f_{k,t} \tag{7–16}$$

where $\boldsymbol{\Omega}$ is an $N \times N$ positive semi-definite matrix, $\boldsymbol{\omega}_k$, $k = 1, \dots K$, are linearly independent $N \times 1$ vectors of factor weights, and $f_{k,t}$ are the factors. It is assumed that these factors have a first-order GARCH structure:

$$f_{k,t} = \xi_k + \alpha_k(\boldsymbol{\gamma}'_k \epsilon^2_{t-1}) + \beta_k f_{k,t-1} \tag{7–17}$$

where ξ_k, α_k, and β_k are scalars and $\boldsymbol{\gamma}_k$ is an $N \times 1$ vector of weights.

A two-step estimation procedure using maximum likelihood is described in Engle et al. [1990]. This approach is shown to be consistent but not efficient.

Remark 7.2.2. *The factor GARCH model makes the restriction that the factors are first order GARCH processes, which means that if multiple correlated factors are incorporated then a multivariate GARCH model is required. All factor GARCH models can be written as special BEKK models. This has prompted the proposal of orthogonal factor GARCH models.*

7.2.4 Orthogonal-GARCH Model

The Orthogonal GARCH (O-GARCH) model is introduced in Alexander and Chibumba [1997] to allow the conditional covariance matrix to be generated by a *small number* of orthogonal univariate GARCH factors.

Definition 45. ***Orthogonal GARCH(1,1,m) Model***

$$\boldsymbol{V}^{-1/2}\boldsymbol{\epsilon}_t = u_t = \boldsymbol{\Lambda}_m \boldsymbol{f}_t \tag{7–18}$$

where $\boldsymbol{V} = diag(v_1, \ldots, v_N)$*, with* v_i *the population variance of* ϵ_{it}*, and* $\boldsymbol{\Lambda}_m$ *is a matrix of dimension* $N \times m$ *given by:*

$$\boldsymbol{\Lambda}_m = \boldsymbol{P}_m diag(l_1^{1/2}, \ldots, l_m^{1/2}) \tag{7–19}$$

$l_1 \geq \ldots \geq l_m > 0$ *being the* m *largest eigenvalues of the population correlation matrix of* u_t*, and* $\boldsymbol{P}_m$ *the* $N \times m$ *matrix of associated (mutually orthogonal) eigenvectors. The vector* $f_t = (f_{1,t}, \ldots, f_{mt})'$ *is a random process such that the conditional expectation,* $\mathbb{E}_{t-1}$*, and conditional variance,* Var_{t-1}*, at time* $t-1$ *are as follows:*

$$\mathbb{E}_{t-1}(f_t) = 0 \tag{7–20}$$

$$\mathrm{Var}_{t-1}(f_t) = \boldsymbol{\Sigma}_t = diag(\sigma^2_{f_{1t}}, \ldots, \sigma^2_{f_{mt}}) \tag{7–21}$$

$$\sigma^2_{f_{1t}} = (1 - \alpha_i - \beta_i) + \alpha_i f^2_{i,t-1} + \beta_i \sigma^2_{f_{i,t-1}} \quad i = 1, \ldots, m \tag{7–22}$$

Thus,

$$\boldsymbol{H}_t = \mathrm{Var}_{t-1}(\boldsymbol{\epsilon}_t) = \boldsymbol{V}^{1/2}\boldsymbol{V}_t\boldsymbol{V}^{1/2} \tag{7–23}$$

where

$$\boldsymbol{V}_t = \mathrm{Var}_{t-1}(u_t) = \boldsymbol{\Lambda}_m\boldsymbol{\Sigma}_t\boldsymbol{\Lambda}_m' \tag{7–24}$$

The parameters of the model are $\boldsymbol{V}$, $\boldsymbol{\Lambda}_m$ and the parameters of the GARCH factors (α_i's and β_i's). Hence the number of parameters is $N(N+5)/2$ (if $m = N$). However, in practice $\boldsymbol{V}$ and $\boldsymbol{\Lambda}_m$ are replaced by their sample counterparts, and m is chosen by principal component analysis applied to the standardized residuals

$\hat{\boldsymbol{u}}_t$. A primer on the Orthogonal GARCH model can be seen in Alexander [2000]. It is important to note that the conditional variance matrix has reduced rank (if $m < N$), which may be a problem for applications and for diagnostic tests which depend on the inverse of $\boldsymbol{H}_t$.

7.2.5 GO-GARCH Model

Van der Weide [2002] relax the assumption of orthogonality by only assuming the matrix $\boldsymbol{\Lambda}$ is square and invertible. Thus producing the following generalisation of the O-GARCH model:

Definition 46. ***Generalised Orthogonal GARCH(1,1) Model***

The implied conditional correlation matrix of $\boldsymbol{\epsilon}_t$ can be expressed as:

$$\boldsymbol{R}_t = \boldsymbol{J}_t^{-1}\boldsymbol{V}_t\boldsymbol{J}_t^{-1} \tag{7–25}$$

where $\boldsymbol{J}_t = (\boldsymbol{V}_t\boldsymbol{I}_m)^{1/2}$, $\boldsymbol{V}_t = \boldsymbol{\Lambda}\boldsymbol{\Sigma}_t\boldsymbol{\Lambda}'$ and $\boldsymbol{\Sigma}_t := Var(\boldsymbol{\epsilon}_t)$

The singular value decomposition of the matrix $\boldsymbol{\Lambda}$ is used as a parametrization, i.e. $\boldsymbol{\Lambda} = \boldsymbol{P}\boldsymbol{L}^{1/2}\boldsymbol{U}$, where the matrix $\boldsymbol{U}$ is orthogonal, and $\boldsymbol{P}$ and $\boldsymbol{L}$ are the eigenvector and eigenvalue matrices. Note that the O-GARCH model (when $m = N$) corresponds to the particular choice $\boldsymbol{U} = \boldsymbol{I}_N$.

In order to estimate the model Van der Weide [2002] first replace $\boldsymbol{P}$ and $\boldsymbol{L}$ by their sample counterparts and then the remaining parameters, i.e. $\boldsymbol{U}$ are estimated together with the parameters of the GARCH factors in a second step.

Remark 7.2.3. *Lanne and Saikkonen [2007] also propose a Generalised Orthogonal Factor model, which allows some of the factors to be homoskedastic.*

7.2.6 FF-GARCH Model

A Full Factor GARCH model is introduced in Vrontos et al. [2003], in which the $\boldsymbol{W}$ matrix is restricted to be triangular.

Definition 47. ***FF-GARCH Model***

$$\boldsymbol{H}_t = \boldsymbol{W}\Sigma_t\boldsymbol{W}' \tag{7–26}$$

where W is a $N \times N$ triangular parameter matrix with ones on the diagonal and the matrix $\boldsymbol{\Sigma}_t = diag(\sigma^2_{1,t}, \ldots, \sigma^2_{N,t})$ where $\sigma^2_{i,t}$ is the conditional variance of the i-th factor, i.e. the i-th element of $W^{-1}\epsilon_t$, which can be separately defined as any univariate GARCH model.

Furthermore, the parameters in W are estimated directly using conditional information only.

Remark 7.2.4. *Dellaportas and Vrontos [2007] propose a class of multivariate threshold GARCH models to capture volatility asymmetries in financial time series. The approach is based on the idea of a binary tree where every terminal node parametrizes a (local) multivariate GARCH model for a specific partition of the data. Giannikis et al. [2008] introduce a class of flexible threshold normal mixture GARCH models to accommodate the stylized facts that appear in many financial time series. The authors develop a Bayesian stochastic method for the analysis of the proposed model allowing for automatic model determination and estimation of the thresholds and their unknown number.*

Zhang and Chan [2009] introduce three factor GARCH models in the framework of GO-GARCH and furthermore present a convenient two-step method for estimating these models. The three models are as follows:

1. Independent-factor GARCH model, which exploits factors that are statistically as independent as possible, as measured by mutual information.

2. Best-factor GARCH model, which contains factors that have the largest autocorrelation in their squared values, such that their volatilities could be forecast well by univariate GARCH.

3. Conditional-decorrelation GARCH model, in which the factors are as conditionally as uncorrelated as possible.

Factor-DCC models are then proposed as an extension to the factor GARCH models with dynamic conditional correlation (DCC) modelling the remaining conditional correlations between factors.

7.2.7 CCC Model

A more parsimonious approach to multivariate GARCH modelling can be attained by separately modelling the individual conditional variances and then the conditional correlation matrix. This results in a non-linear combination of univariate GARCH models.

The Constant Conditional Correlation (CCC) model, introduced in Bollerslev [1990], considers the conditional correlation matrix to be time-invariant. Thus the conditional covariance matrix can be expressed as follows:

Definition 48. ***Constant Conditional Correlation (CCC) Model***

$$\boldsymbol{H}_t = \boldsymbol{D}_t \tilde{\boldsymbol{R}} \boldsymbol{D}_t \tag{7–27}$$

where $\boldsymbol{D}_t = diag(h_{11,t}^{1/2}, \ldots, h_{NN,t}^{1/2})$, $h_{ii,t}$ *can be defined as any univariate GARCH model and the correlation between the returns is assumed to be constant over time and given by matrix* $\tilde{\boldsymbol{R}}$*. Note that the* $\tilde{\boldsymbol{R}}$ *is used here to distinguish between the usage of* $\boldsymbol{R}$ *in the next section to denote returns.*

The CCC model contains $N(N+5)/2$ parameters. $\boldsymbol{H}_t$ is positive definite if and only if all the N conditional variances are positive and $\tilde{\boldsymbol{R}}$ is positive definite. The unconditional covariances are difficult to calculate because of the non-linearity in equation 7–27.

7.2.8 DCC Model

Under the DCC model proposed by Engle [2002]; Christodoulakis and Satchell [2002]; Tse and Tsui [2002] the correlation is specified to be dynamically evolving in time. According to the model of Engle [2002] the conditional covariance matrix is specified as follows:

Definition 49. ***Dynamic Conditional Correlation (DCC) Model***

$$\boldsymbol{H_t} = \boldsymbol{D}_t \tilde{\boldsymbol{R}}_t \boldsymbol{D}_t \tag{7–28}$$

where $\boldsymbol{D}_t = diag(h_{11,t}^{1/2}, \ldots, h_{NN,t}^{1/2})$*,* $h_{ii,t}$ *can be defined as any univariate GARCH model and the dynamic of the conditional correlation is expressed according to the relationship:*

$$\tilde{\boldsymbol{R}}_{\boldsymbol{t}} = diag(q_{11,t}^{-1/2}, \ldots, q_{NN,t}^{-1/2}) \boldsymbol{Q_t} diag(q_{11,t}^{-1/2}, \ldots, q_{NN,t}^{-1/2}) \tag{7–29}$$

where the $N \times N$ *symmetric positive definite matrix* $\boldsymbol{Q_t} = (q_{ij,t})$ *is given by:*

$$\boldsymbol{Q_t} = (1 - \alpha - \beta)\bar{\boldsymbol{Q}} + \alpha u_{t-1} u_{t-1}^T + \beta \boldsymbol{Q_{t-1}} \tag{7–30}$$

with $u_{it} = \epsilon_{it}/\sqrt{h_{ii,t}}$*.* $\bar{\boldsymbol{Q}}$ *is the* $N \times N$ *unconditional variance matrix of* u_t*, and* α *and* β *are non-negative scalar parameters satisfying* $\alpha + \beta < 1$*.*

The elements of $\bar{\boldsymbol{Q}}$ can be estimated jointly with the other model parameters or can be set to the sample estimate to reduce the number of parameters and hence simplify the procedure.

The DCC model has been extensively studied in the literature, see Engle and Colacito [2006]; Aielli [2013], and is a popular proposal to model the conditional variances and correlations. An interesting paper on the properties of the DCC model is Caporin and McAleer [2013]. Extensions to the DCC model are the asymmetric conditional correlation model of Cappiello et al. [2006], the consistent DCC of Aielli [2013] and the sequential DCC model of Palandri [2009].

7.3 Covariance Factor Models

The broad financial literature on asset price dynamics has proposed various solutions to model the expected returns and the covariance between financial assets. While the multivariate GARCH models cope with a salient feature of financial asset prices, namely the heteroskedasticity of the variance and covariance of returns, they do not provide a direct interpretation of the factors intervening in the dynamic of the drift and the conditional variance or covariance. For instance, the GO-GARCH model proposed by Van der Weide [2002] or Lanne and Saikkonen [2007] relies upon latent independent, but not necessarily orthogonal, covariates which are not directly economically interpretable after going through the factor construction. The Factor GARCH model of Engle et al. [1990] does not allow for the incorporation of multiple interpretable non-independent factors, e.g. observed economic variables, without creating a multivariate GARCH model on the factors (and hence having to deal with the many issues associated with this).

In this section I will first present the standard multi-factor model and then introduce a class of generalised factor models (GFM) based around an explicit covariance regression model, first devised in the statistics literature in Hoff and Niu [2012]. This GFM model has the highly desirable property of being able to naturally incorporate interpretable covariates into the covariance dynamic. Thus, interpretable covariance forecasts can be achieved by modelling and forecasting these covariates. Furthermore, I will then give a detailed presentation of the estimation procedure considered for this model via a random-effects representation and expectation maximization (EM) algorithm that is numerically robust and efficient to implement in this context.

I first define two sets of information filtration, which will be used in this modelling framework:

Definition 50. ***Natural Filtration***
Let $\mathcal{F}_t$ denote the natural filtration of the observed portfolio vector valued returns, i.e. in the t-th window it would correspond to the observed σ-algebra generated by the asset returns, $\boldsymbol{R}_t$:

$$\mathcal{F}_t = \sigma\left(\boldsymbol{R}_t, \boldsymbol{R}_{t-1}, \ldots, \boldsymbol{R}_{t-T}\right) \tag{7–31}$$

obtained looking over a sliding window of length T.

The second filtration I will define is based on exogenous independent explanatory variables or factors $\boldsymbol{X}_t$ and will be denoted by $\mathcal{G}_t$.

Definition 51. ***Natural Covariate Filtration***
This filtration is the natural filtration of the observed covariates vector values, i.e.

in the t-th window it would correspond to the observed σ-algebra generated by

$$\mathcal{G}_t = \sigma\left(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-T}\right). \tag{7–32}$$

It is important to distinguish these two information sets as they will produce different ways of studying and constructing the portfolio.

Furthermore, when talking about population versus sample realizations of different model estimators it will also be useful to define the extended filtration $\widetilde{\mathcal{G}}_t$:

Definition 52. ***Extended Covariate Filtration***

$$\widetilde{\mathcal{G}}_t = \cup_{i=1}^{t} \mathcal{G}_i \tag{7–33}$$

i.e. all of the covariate information up until time t.

7.3.1 Standard Factor Model

In standard Multi-factor models, see Green and Hollifield [1992], Chan et al. [1999], the error terms $\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}}$ are assumed to be independent, identically distributed and importantly having a homoskedastic covariance over time, i.e. $\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}} \overset{iid}{\sim} \mathcal{WN}(\mathbf{0}, diag(\sigma_1^2, \ldots, \sigma_N^2))$ for some zero mean homoskedastic white noise driving vector, typically selected to be a multivariate normal distribution. The multi-factor model can be expressed as the standard multi-variate linear regression model displayed below:

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{X}_t + \tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}} \ , \tag{7–34}$$

where $N :=$ number of assets and $K :=$ number of covariates,
$\boldsymbol{R}_t :=$ N-dimensional vector of log asset returns at week t,
$\boldsymbol{\alpha} :=$ N-dimensional vector constant,
$\boldsymbol{\beta} :=$ N-by-K-dimensional matrix of mean covariate loadings,
$\boldsymbol{X}_t :=$ K-dimensional vector of covariate values at week t,
$\tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}} \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, diag(\sigma_1^2, \ldots, \sigma_N^2))$ are the N-dimensional errors at week t,

Extensions of such models sometimes incorporate lagged dependent variables, such as in multivariate (Autoregressive Distributed Lag) ARDL models, cointegration Error correction (ECM) models and more recently translation invariant copula models, in which the vector valued innovation error distribution is specified through marginals and a copula structure which may also be dynamic itself, see Hafner and Manner [2010]; Salvatierra and Patton [2015]. Under the simple version of this multi-factor model in Equation 7–34 the unconditional covariance matrix (population estimator) is easily obtained according to the following terms:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t) = \boldsymbol{\beta} Cov(\boldsymbol{X}_t|\widetilde{\mathcal{G}}_t)\boldsymbol{\beta}^T + diag(\sigma_1^2, \ldots, \sigma_N^2)\ . \tag{7–35}$$

One can think of this unconditional, in the sense of filtration $\mathcal{G}_t$, as being a population based realisation. Then there is the conditional covariance matrix given the covariates values $\boldsymbol{X}_t$ according to the expression:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_t) = diag(\sigma_1^2, \ldots, \sigma_N^2)\ . \tag{7–36}$$

This conditional covariance is to be understood in the sense of the conditioning on the realization of the exogenous covariates realized values and not the population variability. It can be seen from these two covariance specifications that this standard multi-factor model of the returns $\boldsymbol{R}_t$ indeed assumes that these random vectors are independent and homoskedastic given the covariates $\boldsymbol{X}_t$.

7.3.2 Generalised Multi-Factor Model Specification

The covariance regression model introduced in Ames et al. [2015b] and termed the Generalised Multi-Factor model is developed below to extend the traditional Multi-Factor model by allowing the factors to appear in the covariance of the idiosyncratic error terms and thus produce a more flexible model that is capable of capturing heteroskedasticity in the error terms and hence in both the unconditional and conditional covariance matrices. Furthermore, I will demonstrate how this model can be fit using an Expectation-Maximisation (EM) algorithm utilising a reformulation of the covariance regression structure under a specifically designed random-effects representation to produce a closed form E-step and a least squares solution for the M-step, as will be discussed in Section 7.3.3.

In order to capture the heteroskedastic effects of the covariates on the covariance of the returns, $\boldsymbol{R}_t$, the following model is proposed:

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{X}_t + \boldsymbol{e_t} \ , \tag{7–37}$$

where $N :=$ is the number of assets,
$K :=$ is the number of covariates,
$\boldsymbol{R}_t :=$ is the N-dimensional asset log returns,
$\boldsymbol{\alpha} :=$ is the N-dimensional constant,
$\boldsymbol{\beta} :=$ is the N-by-K-dimensional matrix of mean covariate loadings,
$\boldsymbol{X}_t :=$ is the K-dimensional vector of covariate values,
$\boldsymbol{e_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}\boldsymbol{X_t}\boldsymbol{X_t^T}\boldsymbol{C^T} + \boldsymbol{\Psi})$ are the N-dimensional errors,
$\boldsymbol{C} :=$ is the N-by-K matrix of covariate loadings,
$\boldsymbol{\Psi} :=$ is the N-by-N baseline covariance of the errors $\boldsymbol{e_t}$.

It is trivial to derive the unconditional covariance matrix as follows:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t) = \boldsymbol{\beta}Cov(\boldsymbol{X}_t|\widetilde{\mathcal{G}}_t)\boldsymbol{\beta}^T + \boldsymbol{C}\mathbb{E}(\boldsymbol{X}_t\boldsymbol{X}_t^T|\widetilde{\mathcal{G}}_t)\boldsymbol{C^T} + \boldsymbol{\Psi} \ , \tag{7–38}$$

where in this case the observed factors $\boldsymbol{X}_t$ are also assumed to be random vectors and therefore to admit a covariance structure that is locally stationary. The conditional covariance matrix, given the factors, of this multi-factor model is as follows:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_t \cup \mathcal{G}_t) = \boldsymbol{C}\boldsymbol{X_t}\boldsymbol{X_t^T}\boldsymbol{C^T} + \boldsymbol{\Psi} \ , \tag{7–39}$$

where the conditional covariance will be specified according to two terms, the baseline covariance structure Ψ that is present throughout all time, and a separate symmetric strictly positive definite covariance component that captures the relationship between the factors and the heteroskedasticity in the returns over time. From this it can be seen that the heteroskedasticity in the conditional covariance is given by the covariance of the error terms $\boldsymbol{e_t}$. One can note here that the difference between equation 7–38 and equation 7–39 is due to the conditioning of the covariance on the covariate values X_t only over the lookback period L.

7.3.3 Generalised Multi-Factor Model: Covariance Regression Model Estimation via Random-Effects Representation

To perform estimation it is convenient to formulate the covariance regression model as a special type of random-effects model, see Hoff and Niu [2012], for observed data $\boldsymbol{R_1}, \dots \boldsymbol{R_T}$ (N-dimensional returns of length T).

$$\begin{aligned}
&\boldsymbol{R_t} = \alpha + \boldsymbol{\beta X_t} + \gamma_t \times \boldsymbol{C X_t} + \boldsymbol{\epsilon_t} \ , \\
&\mathbb{E}[\boldsymbol{\epsilon_t}] = 0 \quad , \quad Cov(\boldsymbol{\epsilon_t}) = \boldsymbol{\Psi} \ , \\
&\mathbb{E}[\gamma_t] = 0 \quad , \quad Var[\gamma_t] = 1 \quad , \quad \mathbb{E}[\gamma_t \times \boldsymbol{\epsilon_t}] = 0.
\end{aligned} \tag{7–40}$$

Step 1: Mean De-trending of Returns.

The first step is to perform the mean-regression, via in this case a standard linear regression model. This will produce zero-mean residuals $\boldsymbol{\hat{e}_t}$, given by $\boldsymbol{\hat{e}_t} = \boldsymbol{R_t} - \boldsymbol{\hat{\alpha}} - \boldsymbol{\hat{\beta} X_t}$, where $\boldsymbol{\hat{\beta}}$ is the vector of mean regression loading estimates and the covariate vector is denoted by $\boldsymbol{X_t}$.

Step 2: Covariance Regression of Mean-Detrended Returns.

Next, perform the covariance regression of these residuals on the factors, using the random-effects representation:

$$\begin{aligned}
&\boldsymbol{\hat{e}_t} = \gamma_t \times \boldsymbol{C X_t} + \boldsymbol{\epsilon_t} \ , \\
&\mathbb{E}[\boldsymbol{\epsilon_t}] = 0 \quad , \quad Cov(\boldsymbol{\epsilon_t}) = \boldsymbol{\Psi} \ , \\
&\mathbb{E}[\gamma_t] = 0 \quad , \quad Var[\gamma_t] = 1 \quad , \quad \mathbb{E}[\gamma_t \times \boldsymbol{\epsilon_t}] = 0.
\end{aligned} \tag{7–41}$$

The resulting covariance matrix for $\boldsymbol{\hat{e}_t} = \boldsymbol{R_t} - \hat{\alpha} - \boldsymbol{\beta X_t}$, conditional on $\boldsymbol{X_t}$ is then given by,

$$\begin{aligned}
\Sigma_{\boldsymbol{X_t}} &:= \mathbb{E}[\hat{e}_t \hat{e}_t^T | \boldsymbol{X_t}] \\
&= \mathbb{E}[\gamma_t^2 \boldsymbol{C X_t X_t^T C^T} + \gamma_t(\boldsymbol{C X_t \epsilon_t^T} + \boldsymbol{\epsilon_t X_t^T C^T}) + \boldsymbol{\epsilon_t \epsilon_t^T} | \boldsymbol{X_t}] \\
&= \boldsymbol{C X_t X_t^T C^T} + \boldsymbol{\Psi}.
\end{aligned} \tag{7–42}$$

This random-effects model allows maximum likelihood parameter estimation of the coefficients, $\boldsymbol{C}$ and $\boldsymbol{\Psi}$, to be performed via the Expectation Maximization (EM) algorithm. I proceed by iteratively maximising the complete data log-likelihood of $\hat{\boldsymbol{E}} = \hat{\boldsymbol{e}}_{\boldsymbol{1}}, \ldots, \hat{\boldsymbol{e}}_{\boldsymbol{T}}$ denoted $l(\boldsymbol{C}, \boldsymbol{\Psi}) = \log p(\hat{\boldsymbol{E}}|\boldsymbol{C}, \boldsymbol{\Psi}, \boldsymbol{X}, \boldsymbol{\gamma})$, which is obtained from the multivariate normal density given by:

$$\begin{aligned} -2l(\boldsymbol{C}, \boldsymbol{\Psi}) =& TN\log(2\pi) + T\log|\boldsymbol{\Psi}| + \\ & \sum_{t=1}^{T} (\hat{\boldsymbol{e}}_{\boldsymbol{t}} - \gamma_t \boldsymbol{C}\boldsymbol{X}_{\boldsymbol{t}})^T \boldsymbol{\Psi}^{-1} (\hat{\boldsymbol{e}}_{\boldsymbol{t}} - \gamma_t \boldsymbol{C}\boldsymbol{X}_{\boldsymbol{t}}). \end{aligned} \tag{7–43}$$

Note that the conditional distribution of the random effects given the data and covariates is then conveniently given by a normal distribution in each element according to $\{\gamma_t|\hat{\boldsymbol{E}}, \boldsymbol{X}, \boldsymbol{\Psi}, \boldsymbol{C}\} = \mathcal{N}(m_t, v_t)$ with mean $m_t = v_t(\hat{\boldsymbol{e}}_{\boldsymbol{t}}^T \boldsymbol{\Psi}^{-1} \boldsymbol{C}\boldsymbol{X}_{\boldsymbol{t}})$ and variance $v_t = (1 + \boldsymbol{X}_{\boldsymbol{t}}^T \boldsymbol{C}^T \boldsymbol{\Psi}^{-1} \boldsymbol{C}\boldsymbol{X}_{\boldsymbol{t}})^{-1}$. The advantage of this random effects specification of the covariance regression is that taking the conditional expectation of the complete data log likelihood, with respect to the conditional distribution of the random effect parameters γ_t, one obtains a closed form expression for the Expectation E-step. In addition, expressions for the maximization step (m-step) are also attainable in closed form, see details in Hoff and Niu [2012].

7.4 Covariates and Covariance Forecasting

In this section, I present the method utilised to obtain forecasts of the returns covariance matrix under the Generalised Multi-Factor Model (GFM) framework that was developed in Section 7.3.2. In order to obtain forecasts of the covariance the covariates vector $\boldsymbol{X}_t$ needs to be forecast.

7.4.1 Big Data Time Series Forecasting

The traditional approach to modelling and forecasting a given time series is the well known Box-Jenkins method, see Box et al. [2015]. However, this can be prohibitively time consuming if there are many time series to model and forecast, which can be the case for sliding window analyses for example. The Box-Jenkins method is outlined below, followed by an automatic approach in order to overcome this big data issue.

7.4.1.1 Box-Jenkins Method

The Box-Jenkins methodology, see Box et al. [2015], is a five-step process for identifying, selecting, and assessing autoregressive integrated moving average (ARIMA) models for time series data. The steps are as follows:

Algorithm 4 Box-Jenkins methodology

1. Establish the stationarity of the time series and if there is any significant seasonality that needs to be modelled. If the series is not stationary, successively difference the series to attain stationarity. The sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of a stationary series decay exponentially (or cut off completely after a few lags).

2. Identify a (stationary) conditional mean model for the data. The sample ACF and PACF functions can help with this selection. For an autoregressive (AR) process, the sample ACF decays gradually, but the sample PACF cuts off after a few lags. Conversely, for a moving average (MA) process, the sample ACF cuts off after a few lags, but the sample PACF decays gradually. If both the ACF and PACF decay gradually, consider an ARMA model.

3. Specify the model, and estimate the model parameters via maximum likelihood estimation or non-linear least-squares estimation.

4. Check the goodness-of-fit of the proposed model. Residuals should be uncorrelated, homoskedastic, and normally distributed with constant mean and variance. Plotting the mean and variance of residuals over time and performing a Ljung-Box test or plotting autocorrelation and partial autocorrelation of the residuals are helpful to identify misspecification. If the residuals are not normally distributed, a Student's t distribution can be considered for the innovations. Return to step one if the model is inadequate.

5. After choosing a model and checking its fit and in-sample and out-of-sample forecasting ability the model can be used to forecast or generate Monte Carlo simulations over a future time horizon.

7.4.1.2 Automatic Covariate Forecasting

In order to obtain forecasts of the covariance of the returns, following the GFM model in equation 7–39, the covariates $\boldsymbol{X}_t$ must be forecast. If the number of time series to be modelled is large it will be important to consider an automatic procedure. Here, I present the Hyndman-Khandakar algorithm for automatic seasonal autoregressive integrated moving average (SARIMA) modelling as implemented in the auto.arima function in the R forecast package Hyndman [2015], see Hyndman and Khandakar [2008] for details.

The outline of the algorithm to fit the ARIMA model to each covariate time series is as follows:

Remark 7.4.1. *Such an automated procedure for model selection is particularly relevant in the context of the modelling in this thesis, since in Chapter 8 I carry out a sliding window analysis. I have K time series to be fit for every sliding window and there is one sliding window for each trading day over the entire length of data analysed. This is equivalent to $270 \times (15 + 11)$ total number of models to be fit. Thus, with this many models I need an automatic and efficient procedure.*

7.4.1.3 Covariate Forecasting Accuracy

To assess the suitability of a fitted ARIMA model it is important to analyse the accuracy of the forecasts. A well known measure of forecast accuracy is the Mean Absolute Percentage Error (MAPE) criterion.

Definition 53. *Mean Absolute Percentage Error (MAPE)*

$$MAPE_\tau = 100 \times \frac{1}{\tau} \sum_{t=1}^{\tau} \left(\frac{|X_t - \hat{X}_t|}{|X_t|} \right) \tag{7–44}$$

where the numerator is the forecast error at time t.

In addition to the MAPE criterion I also consider the Mean absolute scaled error (MASE) as given in Definition 54 and introduced in Hyndman and Koehler [2006]. The MASE measure scales the error based on the in-sample MAE from the naïve (random walk) forecast method and thus allows the comparison of time series on different scales and is also robust to values close to zero.

Algorithm 5 Automatic ARIMA model selection

1. The number of differences d is determined using repeated Kwiatkowski-Phillips-Schmidt-Shin KPSS hypothesis tests. This is a family of hypothesis tests for a time series that is assumed to be represented as the linear combination of a deterministic trend, a random walk, and a stationary error. Then the test statistic is formed from the Lagrange multiplier test of the hypothesis that the random walk has zero variance. Such a test is capable of testing both the unit root hypothesis and the stationarity hypothesis.

2. The values of p and q are then chosen by minimizing the AICc after differencing the data d times. Rather than considering every possible combination of p and q, the algorithm uses a stepwise search to traverse the model space.

   (a) The best model (with smallest AICc) is selected from the following four:
   ARIMA(2,d,2), ARIMA(0,d,0), ARIMA(1,d,0), ARIMA(0,d,1).
   If $d = 0$ then the constant c is included;
   if $d \geq 1$ then the constant c is set to zero. This is called the "current model".

   (b) Variations on the current model are considered:

      i. vary p and/or q from the current model by ± 1;
      ii. include/exclude c from the current model.

   The best model considered so far (either the current model, or one of these variations) becomes the new current model.

   (c) Repeat Step 2(b) until no lower AICc can be found.

Definition 54. *Mean Absolute Scaled Error (MASE)*

$$MASE_\tau = \frac{1}{\tau}\sum_{t=1}^{\tau}\left(\frac{|\tilde{e}_t|}{\frac{1}{n-1}\sum_{i=2}^{n}|X_i - X_{i-1}|}\right) \tag{7–45}$$

where the numerator $\tilde{e}_t$ is the forecast error at time t, defined as the actual value (X_t) minus the forecast value ($\hat{X}_t$) for that period, i.e. $\tilde{e}_t = X_t - \hat{X}_t$, and the denominator is the average in-sample forecast error over n data points of the one-step naïve (random walk) forecast method, which uses the actual value from the prior period as the forecast, i.e. $\hat{X}_t = X_{t-1}$.

Remark 7.4.2. *It is interesting to note that when the time series under consideration is very close to being white noise then the MASE measure will be close to 1 and thus can potentially mislead the statistician into thinking the model is useful for forecasting.*

7.4.2 Forecasting Covariance via Factor Models

Given forecasts of the covariate time series, the τ-step ahead unconditional covariance matrix can be forecast via the following procedure:

Algorithm 6 Covariance forecasting utilising GFM model and covariates forecasts

1. Fit Generalised Multi-Factor Model to the data period $[t - L : t]$ via the method in Section 7.3.3 to obtain parameter estimates $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\Psi}}$ and $\hat{\boldsymbol{C}}$. L is the lookback period: for example $L = 125$ data points.
2. Forecast τ-step ahead covariate values, $\hat{\boldsymbol{X}}_{t+\tau}$ for each covariate individually, as described by the SARIMA forecasting method in Hyndman and Khandakar [2008].
3. The τ-step ahead covariance matrix is calculated as: $\widehat{Cov}(\boldsymbol{R}_{t+\tau|t}|\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t) = \hat{\boldsymbol{\beta}} Cov(\hat{\boldsymbol{X}}_{t+\tau|t}|\widetilde{\mathcal{G}}_t)\hat{\boldsymbol{\beta}}^T + \hat{\boldsymbol{C}}\mathbb{E}(\hat{\boldsymbol{X}}_{t+\tau|t}\hat{\boldsymbol{X}}^T_{t+\tau|t}|\widetilde{\mathcal{G}}_t)\hat{\boldsymbol{C}}^T + \hat{\boldsymbol{\Psi}}$.
4. The τ-step ahead conditional covariance matrix forecast is given by:

$$\widehat{Cov}(\boldsymbol{R}_{t+\tau|t}|\hat{\boldsymbol{X}}_{t+\tau|t}, \mathcal{F}_t \cup \mathcal{G}_t) = \hat{\boldsymbol{C}}\hat{\boldsymbol{X}}_{t+\tau|t}\hat{\boldsymbol{X}}^T_{t+\tau|t}\hat{\boldsymbol{C}}^T + \hat{\boldsymbol{\Psi}} .$$

In the next chapter, I contribute to the currency multivariate dynamic modelling literature through the application of the proposed GFM model to currency covariance modelling. In doing so I propose to include a set of covariates which combine factors commonly found in currency modelling and other factors relating to speculative trading volumes.

Chapter 8

Covariance Forecasting Factor Models in Currency Carry Trades

In this chapter, I will investigate how the behaviour of speculative traders impacts the dependence structure of currency carry trade baskets. Speculative trading volume factors are introduced into the covariance factor models presented in Chapter 7 allowing one to forecast portfolio covariance and thus perform risk based asset allocation in currency carry trades. Furthermore, the impact of this speculator behaviour on the tail dependence of carry trade baskets is analysed.

8.1 Research Contribution: Speculative Trading Behaviour and Dependence Structure of Currency Returns

This chapter explores the impact of speculative trading behaviour on the dependence structure of currency returns. The ratio of speculative open interest (net non-commercial positions) to total open interest, termed the $SPEC$ factor, is shown to provide a good proxy to the behaviour of carry trade investors via a PCA analysis.

A covariance regression modelling approach whereby the influence of observed covariates on the covariance of the multivariate returns of a basket of assets is proposed. In particular, the impact of speculative trading behaviour, i.e. the $SPEC$ factors, on the covariance of carry currencies is investigated. These $SPEC$ factors are shown to hold several orders of magnitude more explanatory power than the price index factors, DOL and HML_{FX}, previously suggested in the literature. Furthermore, it is demonstrated that the time series for the DOL and HML_{FX} factors are very close to white noise and as such are essentially unforecastable. The suggested speculative open interest factors are shown to be amenable to ARIMA model fits and so produce reasonable forecast accuracy.

Thus, time series models for these covariates of interest are built and hence forecasts of the covariance of a basket of currencies can be obtained. Therefore, the inherent heteroskedasticity of the covariance of a basket of currencies can be modelled and forecast whilst maintaining the desirable property of interpretability of the model. As shown in this thesis, this forecasting ability is then useful for risk management, portfolio optimisation and trading strategy development.

A sensitivity analysis of the covariance to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. In addition, a regression of the tail dependence measures, obtained from the mixture copula modelling approach, on the $SPEC$ factors illustrates the influence of carry trade speculative behaviour on the extremal joint currency returns. The DOL and HML_{FX} are shown to hold little explanatory power in the joint tails.

8.2 Currency Data and Currency Factors Description

Here, I consider two sets of currency baskets typically associated with a currency carry trade strategy. One portfolio consisting of a long basket and a second portfolio consisting of a short basket. The long basket contains four major "investment" currencies, namely United Kingdom (GBP), Australia (AUD), Canada (CAD) and New Zealand (NZD), while the short basket contains three major "funding" currencies, namely Euro (EUR), Japan (JPY) and Switzerland (CHF). These baskets were chosen to replicate the pre-2009 NBER recession study carried out in Brunnermeier et al. [2008], and hence to highlight that the skewness relationships presented by the authors are found to be substantially different post-2009 NBER recession. In addition, the limited availability of the open interest rate data to only 7 developed market currencies didn't allow for the use of the same baskets considered in Part I of this thesis. I have considered daily settlement prices for each currency exchange rate as well as the daily settlement price for the associated 1 month forward contract in order to derive the weekly carry trade mark-to-market returns, $\boldsymbol{R}_t$. The daily time series analysed were obtained from Bloomberg and range from 04/01/1999 to 29/01/2014. As I am working on the trading volume based covariance modelling I choose 1st April 1999, i.e. the date of the introduction of the Euro, as the starting date of the sample.

For the explanatory factors in the currency analysis I consider a range of different factors that I motivate in this section from an economic perspective as well as a quantitative perspective. In a similar vein to the famous three stock-market factors and the two bond-market factors proposed by Fama and French [1993] to explain bonds and equities returns, Lustig et al. [2011] propose a factor decomposition of the currencies returns. Such models are built upon one of the cornerstones of financial theory which is the risk premium. These yields implicitly stored within asset returns would thus be received by investors willing to bear the associated sources of risk. Lustig et al. [2011] demonstrate with the help of a principal component analysis that two linearly independent factors could explain most of the variability in the cross section of the international exchange rates. The first factor would correspond to a level factor, named "dollar risk factor" or DOL, which is essentially the average relative value change of a foreign currency basket against the dollar[1]. The second factor embodies the market induced risk premium associated to the currencies with the highest differential of interest rates relative to the others and is accordingly named in the literature the High-Minus-Low risk factor or HML_{FX}.

Definition 55. *Dollar (DOL) Factor*

[1]When an American investor is considered. However it is asserted in the same article that similar results are obtained when we retain the Japanese, British or Swiss investor's point of view.

The DOL factor is defined as:

$$DOL = \frac{1}{N}\sum_{i=1}^{N} R_i\,, \tag{8–1}$$

where R_i is the log return of currency i and N is the total number of currencies.

Definition 56. ***High-Minus-Low (HML_{FX}) Factor***

The HML_{FX} factor is defined as:

$$HML_{FX} = \frac{1}{P}\sum_{j=1}^{P} R_j - \frac{1}{Q}\sum_{k=1}^{Q} R_k\,, \tag{8–2}$$

where R_j is the log return of currency j in the high interest rate basket, R_k is the log return of currency k in the low interest rate basket, and P and Q are the number of currencies in the high interest rate basket and the low interest rate basket respectively.

Lustig et al. [2011] show that over time higher interest rate currencies have a tendency to load more on the latter than low interest rate currencies. The explanatory power of the HML_{FX} factor is indeed significant when characterizing the intertemporal presence of the cross-sectional variation on average exchange rates among high and low interest rate currencies. This last statement justifies the inclusion of these market risk premium in the set of factors retained for the covariance regression model. Moreover, I also take into consideration the respective factor's volatility σ_{DOL}, σ_{HML} as well as the covariance between the factors $\sigma_{DOL,HML}$.

Ames et al. [2015a] recently demonstrated that on top of these price-based data sets another set of covariates is significant in explaining the joint dynamic between currencies. This additional set of covariates encompasses all the speculative net positions held by the non-commercial investors in the futures market. Leaning on a very rich academic literature, the relation between assets variance or covariance and trading volume has been recurrently demonstrated by academics. Among the seminal papers in this domain George E. Tauchen [1983] proposed the theoretical foundations with the Mixture-of-Distributions Hypothesis (MDH), which has been extended to the multivariate case recently by He and Velu [2014]. A parallel can be drawn between this branch of the literature and the empirical works concerning the influence of the speculative volumes upon financial assets joint and marginal dynamics Brunnermeier et al. [2008]; Brunnermeier and Pedersen [2009]; Anzuini and Fornari [2012]; Hutchison and Sushko [2013]; Fong [2013]; Ames et al. [2015a]. Therefore, I augment the price-based covariance regression model with speculative volume information provided in a weekly report published by the CFTC, see CFTC [2015]. In doing so, I assume the financial inflows and outflows resulting from the adjustments of the speculative long or short positions generate and thus could help explain dependences between international exchange rates, as demonstrated in Ames et al. [2015a]. This approach allows the market trading volumes resulting from the adjustments of the speculative long or short positions to be utilised to help explain the dependences between international exchange rates, as demonstrated in Ames et al. [2015a]. The weekly report provides the Commitments of Traders: showing the open positions, long and short, on the currency future contracts traded on the Chicago Mercantile Exchange and breaks them down into commercial, non-commercial and non-reportable positions. The group of the non-commercial investors, i.e. the category of speculative traders, are not holding the futures positions until expiry and will in general be more nimble and prone to build carry positions in the market. On the contrary, the commercial traders are using the futures market to hedge an existing exposure on the underlying asset, see Galati et al. [2007]; Fong [2013].

I will use the ratio of the net non-commercial futures position to the total open interest as a proxy of the speculative currency carry trade position for each currency. I will refer to this as the $SPEC$ ratio. This $SPEC$ factor is suggested as a good proxy for carry trade activity in a number of recent publications (see Galati et al. [2007]; Brunnermeier et al. [2008]; Cecchetti et al. [2010]; Anzuini and Fornari [2012]; Hutchison and Sushko [2013]; Fong [2013]).

Definition 57. ***Speculative Open Interest (SPEC) Ratio***

The $SPEC$ ratio for currency i is defined as:

$$SPEC_i = \frac{net\ non\text{-}commercial\ futures\ position_i}{total\ open\ interest_i} \tag{8–3}$$

where $i = 1, \ldots, N$ and N is the number of currencies.

Remark 8.2.1. *Note that the volume data provided by the CFTC for the Norwegian Krone are not long enough for a robust analysis. Thus it was necessary to exclude this currency from the analysis. I indeed retained for the volume analysis a period of time which spans 20/06/2006 to 29/01/2014*[1]*. The starting point corresponds to the date when the New Zealand Dollar (NZD) contract started to be liquid enough to be included in the analysis*[2]*. Furthermore, given that the CFTC data needed to run the regression analyses are available on a weekly basis, I build the corresponding weekly carry returns based on the daily settlement prices available.*

It is important to note that currencies are more heavily traded via Over-the-Counter (OTC) forward contracts rather than futures contracts. However, this forward trading volume data is not available. The PCA analysis presented in Section 8.3.1 is therefore key as it demonstrates the informational content, in terms of carry trade behaviour, of the futures trading volume data.

8.2.1 Data Preparation

In order to perform the empirical analyses considered in this chapter a substantial amount of effort and time was invested into collecting, cleaning and preparing the data. In particular, the following key steps were performed:

1. Collect daily currency spot price data: closing price, bid and ask price.

2. Collect daily currency forward price data - at maturities of one week, two weeks, three weeks and 1 month: closing price, bid and ask price.

3. Pre-process the price data to deal with missing data, i.e. if data is missing copy previous day's price.

[1] Since the 20th of June 2006 the Commitments of Traders report provides the commercial and non-commercial positions for the following currencies against the dollar: AUD, CAD, CHF, EUR, GBP, JPY, NZD.

[2] Before this date the NZD contract open interest was mostly equal to zero, which means that no positions were open during this time on the future markets, thus justifying the exclusion from the volume analysis the data until this date.

4. Match one month forward contracts with closing spot price on the correct date of delivery for the contract.

5. Calculate the forward premium (interest rate proxy).

6. Collect currency futures price open interest data: broken down into net commercial (hedgers) and net non-commercial futures positions (speculators).

7. Match open interest rate data to synchronous currency price data.

8.3 Exploring Intertemporal Cross-Sectional Volatility-Volume Relations

There has been a growing interest in the literature to study the effects of various macroeconomic and microeconomic factors on the mean and volatility dynamics of individual currency exchange rates. For instance, in Christiansen et al. [2012] the authors study the return volatility of exchange rates with respect to different functions of macroeconomic variables such as: equity market variables and risk factors such as the dividend price ratio, the earnings price ratio, equity market returns for leverage effects, Fama and French [1993] risk factors; interest rates, spreads and bond market factors such as T-bill rates, term spreads and other factors related to term structure forward rates discussed in Cochrane and Piazzesi [2002]; foreign exchange rate variables and risk factors such as the average forward discount for capturing counter cyclical FX risk premia, the DOL risk factor and the HML_{FX} carry factors of Lustig and Verdelhan [2007]; Lustig et al. [2011]; liquidity and credit risk factors such as yield spreads between BAA and AAA rated bonds (i.e. default spreads), TED spreads for LIBOR rate and T-Bill rates discussed in funding liquidity in Brunnermeier et al. [2008] as well as aggregate measures of bid-ask spreads in foreign exchange markets such as those discussed in Menkhoff et al. [2012a]; and macro-economic variables. In this study of Christiansen et al. [2012] they concentrate on explaining carefully individual exchange rates through Bayesian model averaging, however they neglect to study the joint relationships between multiple exchange rates and the influence of the proposed factors.

In this analysis, the intention is to generalize these types of studies to investigate joint behaviours in multiple exchange rates and focus on a few important factors principally related to the exchange rate market dynamics. In particular, I highlight the importance of factors based on speculative order flows in influencing the joint appreciation and depreciation dynamics of baskets of multiple exchange rates. This is interesting since there is mounting evidence that speculative order flows in the markets have a substantial impact on the dynamics of certain financial assets. An extended literature is documenting the empirical relation between the trading activity and its impact upon the asset drift (Hong and Yogo [2012]; Singleton [2013]) or price innovations relative to a benchmark (Henderson et al. [2014]) or even the volatility (Gallant et al. [1992]; Ané and Geman [2000]) of a given asset. The challenge with such findings is that speculator behaviours consist of an evolutionary process, which is naturally a function of the current market conditions, but also of the economic environment, which could be more or less prone to the growth of speculative inflows, for instance to satisfy the hedging needs from the non-financial sphere.

Following this literature trend I set out to demonstrate the existence of a dual relation between high and low interest rate differential currency baskets and the associated dependences, by comparing them with the amount of speculative inflows and outflows on the available funding and investing currency futures. While numerous authors have emphasized the relation between the volume traded on a specific asset and its volatility (Gallant et al. [1992]; Ané and Geman [2000]), I propose hereafter to focus more particularly on the speculative flows which allows one to broaden the analysis and in so doing investigate the relation between the non-commercial traders net positions, commonly considered as speculators, and the dependence between financial assets. Hence, it is assumed that setting up a carry position in the currency market will synchronously impact all the currency prices and thus result in higher price dependences and consequently a less sparse returns covariance matrix. While several articles, such as Hasbrouck and Seppi [2001]; Bernhardt and Taub [2008]; He and Velu [2014] investigate the relation between the volume commonalities and the price commonalities none of them have focused on the speculative volumes, nor have they studied the currency markets. Moreover, the analysis presented in this thesis contributes to the literature as

I propose in this section a new approach, namely the covariance regression, to study this relation between volume and asset prices, while in the following section I will investigate this relation for extremal return commonalities.

Through different price and volume based factors I thus explore in the next section the effect of the speculator behaviour on the first two moments of the cross-sectional currency returns. Firstly, I study the informational content of the speculative volume time series. I then present a mean regression of the individual currency returns followed by a covariance regression of the multivariate basket returns given the explanatory factors.

8.3.1 Informational Content of Speculative Trading Volumes

Before proceeding with the volume-volatility analysis it is important to justify the use of the CFTC non-commercial open positions as a proxy of the carry trade speculative positions. As a matter of fact, these time series provided by the CFTC do not distinguish the open positions resulting from the carry trade or the other potential motivations for the speculator. For instance, the open position will definitely be impacted by the setting up of a dollar position or a relative value trade between the European currencies and the others. In order to discern the common factors impacting the individual currencies percentage of non-commercial traders among futures open positions I run a principal component analysis on the net speculative positions published by the CFTC. It is found that 49.5% of the variance associated to this set of currencies is explained by the first principal component. Moreover, the currency loadings associated to this first factor are interestingly (almost) monotonically decreasing according to their respective average differential of interest rates with the US Dollar. It can also be observed that the associated eigenvector (see Figure 8.1) displays a positive sign for the principal financing currencies such as JPY and CHF whereas the investing currencies, such as NZD and AUD show a negative relation with the first component. These results confirm that a large part of the net speculative positions in futures is directly following from the carry trade strategy.

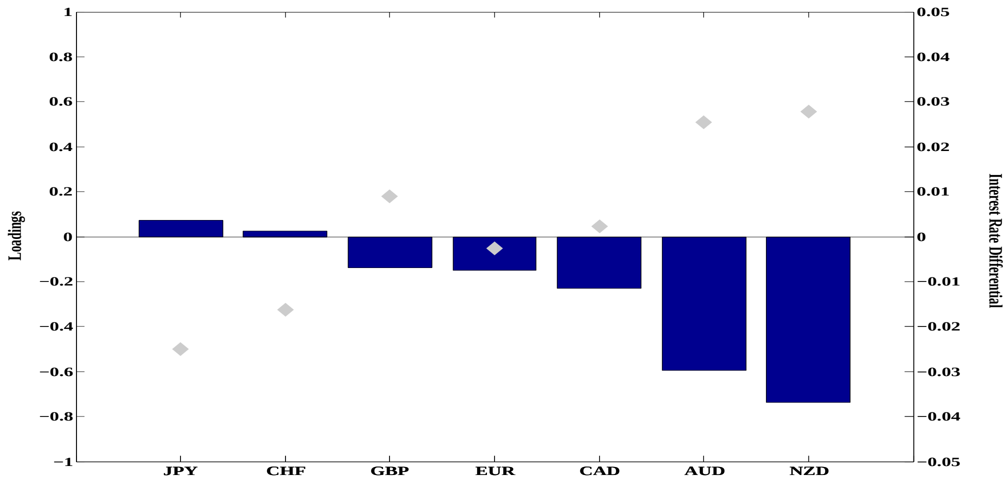


Figure 8.1: Loadings of the First Principal Component of Developed Countries Speculative Percentage. The bars (left axis) represent the loadings values on the speculative percentages first principal component while the grey diamonds (right axis) depicts the level of interest rate differential with the 1 month US interest rates.

8.3.2 Currency Mean Dynamic Decomposition

In order to understand, relative to the price based information flows, that the speculative trading volumes are distinctly influencing the international exchange rates, I first assess in this section if the cross-sectional ratios of speculator net positions on the market open interest, i.e. $SPEC$ ratios, have a significant impact on the mean dynamic of individual currencies once I have accounted for the variability explained by the common price based FX market factors described in Lustig and Verdelhan [2007], namely the dollar factor DOL and the carry high minus low factor HML_{FX}.

To achieve the analysis of the currency conditional mean dynamic, I consider a regression of each of the individual currency returns time series, for currencies utilised in the carry trade, onto explanatory risk factors given by the time series of the covariates DOL and HML_{FX}, as well as the ratios of the net speculative position relative to open interest, i.e. $SPEC$ ratios, for all the available currencies. As mentioned earlier, this regression is performed on weekly data, considering the period spanning from 20/06/2006 to 29/01/2014 and excluding as a consequence the Norwegian krone from the analysis. It can be seen in Table 8.1 that even though speculative volumes seem to contribute to the variability of a couple of currency returns, it turns out that their contribution to the explanatory power of the regression model remains ancillary. Indeed it is observed that the adjusted R^2 is very marginally improved once we include the speculative open interest covariates. This statement leads to the conclusion that the dollar and the high minus low factors proposed by Lustig and Verdelhan [2007] clearly prevail over the speculative volumes variables as far as the mean dynamic of the cross-sectional currency returns is concerned. This assertion corroborates the microeconomic literature regarding the variability-volume theory, wherein it has been demonstrated that volumes traded in equity markets mainly relate to the variance of the very same equities and not their price or even their return dynamics.

	AUD	CAD	CHF	EUR	GBP	JPY	NZD
$Constant$	0.0000	0.0020	-0.0008	-0.0009	0.0009	0.0011	-0.0007
	(0.256)	(0.376)	(0.444)	(0.975)	(0.061)	(0.348)	(0.546)
DOL	-0.5099**	-0.3130**	-0.3466**	-0.3834**	-0.3187**	-0.0585**	-0.5225**
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
HML_{FX}	0.3087**	0.2028**	-0.5919**	-0.2876**	-0.1095**	-0.5727**	0.3142**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$AUD\,SPEC$	-0.0014	0.0019	-0.0027	0.0022	-0.0002	0.0013	0.0004
	(0.264)	(0.586)	(0.921)	(0.424)	(0.366)	(0.115)	(0.826)
$CAD\,SPEC$	0.0009	-0.0056	-0.0014	0.0003	0.0015**	0.0006	0.0018
	(0.863)	(0.761)	(0.337)	(0.514)	(0.006)	(0.403)	(0.295)
$CHF\,SPEC$	0.0004	0.0047	0.0003	-0.0011	-0.0022	0.0017	-0.0014
	(0.561)	(0.471)	(0.387)	(0.794)	(0.065)	(0.882)	(0.471)
$EUR\,SPEC$	-0.0013	-0.0021	-0.0001**	-0.0007	0.0056	-0.0022	0.0002
	(0.756)	(0.478)	(0.032)	(0.496)	(0.390)	(0.963)	(0.939)
$GBP\,SPEC$	0.0013	0.0055	0.0012	-0.0037	-0.0041**	0.0038	-0.0007
	(0.065)	(0.169)	(0.128)	(0.493)	(0.024)	(0.514)	(0.769)
$JPY\,SPEC$	-0.0041	-0.0012**	0.0018**	0.0006**	0.0054	-0.0052	0.0010
	(0.765)	(0.031)	(0.022)	(0.005)	(0.505)	(0.331)	(0.577)
$NZD\,SPEC$	0.0004	-0.0026	0.004	0.0011	-0.0011	-0.0049**	-0.0008
	(0.561)	(0.088)	(0.590)	(0.805)	(0.272)	(0.045)	(0.712)
R^2 (DOL, HML_{FX})	92%	68%	80%	81%	60%	57%	90%
R^2 (DOL, HML_{FX}, $SPEC$)	92%	69%	81%	81%	61%	59%	90%

Table 8.1: Regression of the individual currency returns on the DOL index, HML_{FX} index and the $SPEC$ ratio (the ratio of each currency future speculative net position to the total future open interest, as provided by the CFTC), as well as cross relations among them. The open interest data provided by the CFTC as well as the computed DOL and HML_{FX} indexes are weekly data. The period of time considered for this analysis spans from 20th June 2006 to 29th January 2014 and corresponds to the longest available overlapping sample for all the currencies considered. Numbers in parentheses show Newey and West (1987) HAC p-values. All the possible cross effects among the currencies are not significantly contributing to the regression (HAC p-value below 5%)

8.3.3 A Covariance Regression Model Considering DOL, HML_{FX} and $SPEC$ Factors

In the previous section I considered regression on the mean structure looking at whether the market price factor DOL, carry factor HML_{FX} and speculative volume factors $SPEC$ provided statistically significant explanatory power in describing the trend in the returns dynamic of individual currencies and currency baskets constructed from ordering of interest rate differentials. In this section, I extend these mean-regressions of carry trade basket returns to study how these factors load directly on the regression against the covariance structure of the assets in the currency baskets. This will reveal the proportion of covariance, between the currencies in each basket, that can be explained by the DOL, HML_{FX} and $SPEC$ factors. I will investigate two sets of covariates for each of the high and low interest rate baskets. Firstly, I will examine the power of the DOL and HML_{FX} factors in explaining the covariance. Secondly I will analyse what explanatory power is contributed by the $SPEC$ factors and the first order cross terms between the $SPEC$ factors. The first order cross terms are included in the model to allow interactions between the speculative trading behaviour across pairs of currencies, as multiple currencies are utilised to construct the baskets.

To perform this study I formulate the covariance regression model as a special type of random-effects model, as described in Chapter 7 of this thesis (see Hoff and Niu [2012]), for the observed asset returns data $\boldsymbol{y_1}, \ldots \boldsymbol{y_T}$ (d-dimensional high or low interest rate basket weekly log returns for a time block of length T). The interest in utilising this covariance regression model here is to examine the contribution of the explanatory factors to the conditional covariance matrix of the currency log returns relative to the associated interest rate differentials, and thus the basket each currency belongs to.

The summary measure described in Equation 8–4 shows the proportion of covariation in the covariance regression attributed to these factors relative to the total second order explanatory power of the covariance regression on each 125 week sliding window. This measure focuses on the covariance explained when $\boldsymbol{X}_t$

takes its median value, denoted $\boldsymbol{X}_{\mathbf{(0.5)}}$.

$$\text{Non-Baseline Variance } \% = 100 \times \frac{trace(\boldsymbol{C}\boldsymbol{X}_{\mathbf{(0.5)}}\boldsymbol{X}_{\mathbf{(0.5)}}^T\boldsymbol{C}^T)}{trace(\boldsymbol{C}\boldsymbol{X}_{\mathbf{(0.5)}}\boldsymbol{X}_{\mathbf{(0.5)}}^T\boldsymbol{C}^T) + trace(\boldsymbol{\Psi})} \tag{8–4}$$

where $\boldsymbol{C}$ is the matrix of covariate loadings and $\boldsymbol{X}_t$ is the vector of covariate values.

Figure 8.2 shows this result for both the high interest rate basket and the low interest rate basket. It can be seen that the explanatory power of the two factors, DOL and HML_{FX}, is time-varying, but that there is very little power in explaining the linear second order co-movements of currencies in either the high or low interest rate baskets as captured by the covariance structure intertemporally. This observation further strengthens the hypothesis that one must look at co-currency movements using flexible models that capture appropriate concordance/discordance relationships such as tail dependence, as explored in Chapter 5 and also later in this chapter. The large increase in explanatory power obtained by using the full $DOL + HML_{FX} + SPEC + SPEC \times SPEC$ model over the two factor $DOL + HML_{FX}$ model can be seen for both the high interest rate basket and the low interest rate basket in Figure 8.3. Here, I plot the log of the ratio of the percentage explanatory power of the full model divided by the percentage explanatory power of the two factor $DOL + HML_{FX}$ model. It is observed that the explanatory power of the model incorporating the $SPEC$ factors and its crosses is several orders of magnitude greater than the two factor $DOL + HML_{FX}$ model.

It is worth highlighting that for each sliding window I estimate the regression relationships with covariate vector $\boldsymbol{x}_t = [1, x_{1,t}, \ldots, x_{q,t}]$ given by the factors discussed and the resulting covariance matrix parameter vectors defined by row vectors of the $(d \times q)$ matrix $\boldsymbol{C}$[1], which is given by $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_d\}$ vectors, which

[1] Here the dimensions of the matrix $\boldsymbol{C}$ are linked to the portfolio considered. In the high interest rate basket covariance regression, I analyse the de-trended returns variability of GBP, AUD, CAD and NZD relative to 12 covariates (which are the DOL, the HML_{FX}, the four associated currency $SPEC$ factors and six cross $SPEC \times SPEC$ factors) which leads to (4×12) matrix $\boldsymbol{C}$. Likewise, in the low interest rate basket the dimensions of $\boldsymbol{C}$ are (3×8) as I focus on the variability of EUR, JPY and CHF relative to 8 covariates (which are the DOL, the HML_{FX}, the three associated currency $SPEC$ factors and three cross $SPEC \times SPEC$ factors).

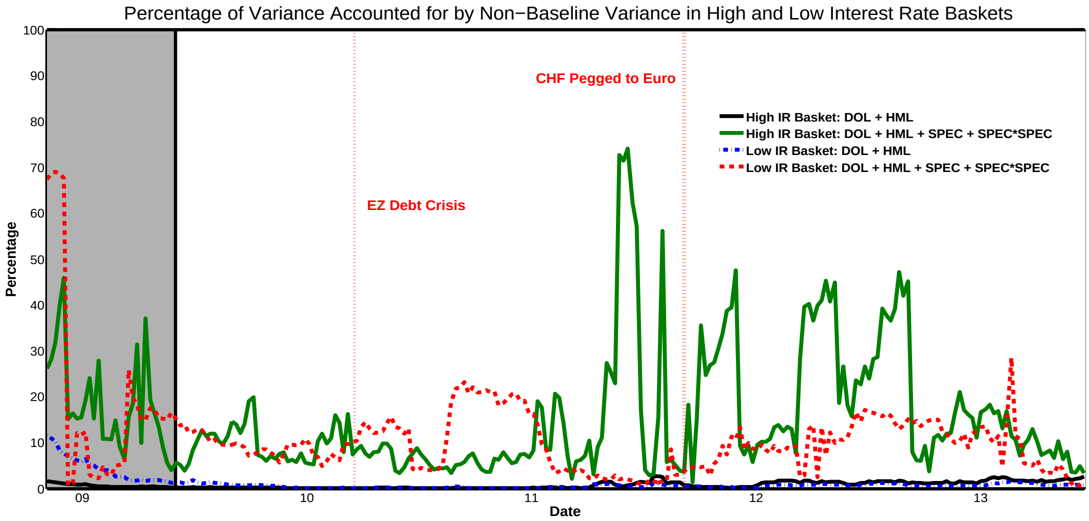


Figure 8.2: High interest rate and Low interest rate basket. $DOL + HML_{FX}$ vs $DOL + HML_{FX} + SPEC + SPEC \times SPEC$. 125 week lookback periods.

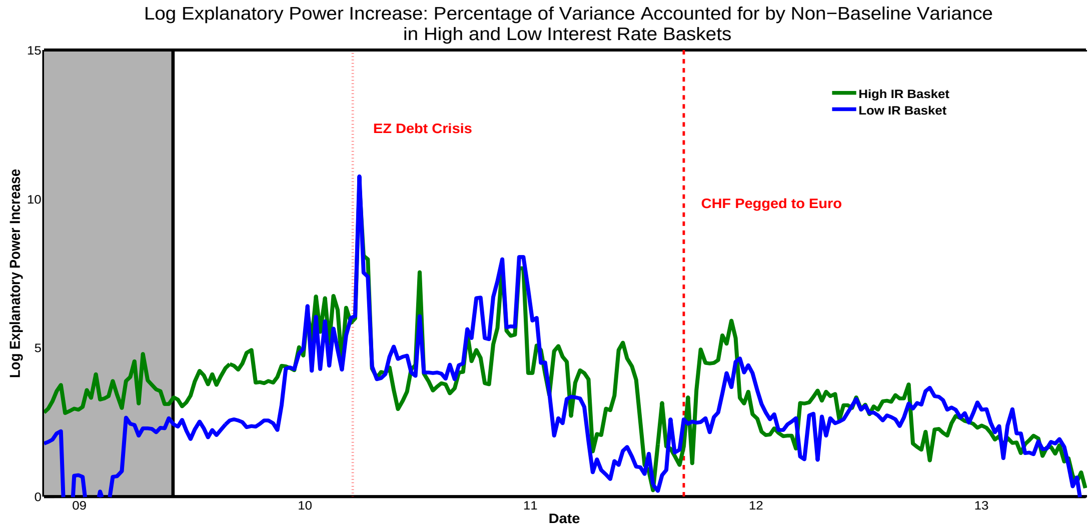


Figure 8.3: Log Explanatory Power Increase: High interest rate and Low interest rate basket. $DOL + HML_{FX}$ vs $DOL + HML_{FX} + SPEC + SPEC \times SPEC$. 125 week lookback periods. Here, we plot the log of the ratio of the percentage explanatory power of the full model divided by the percentage explanatory power of the two factor $DOL + HML_{FX}$ model.

then result in the regression models for each element j of the covariance matrix of the currency basket return residuals $\boldsymbol{e}_t = \boldsymbol{y}_t - \boldsymbol{\beta}\boldsymbol{x}_t$ being given by:

$$\begin{aligned}\mathbb{V}\text{ar}\left[\boldsymbol{e}_{t,j}\right|\boldsymbol{x}_t] &= \Psi_{j,j} + \boldsymbol{c}_j\boldsymbol{x}_t\boldsymbol{x}_t^T\boldsymbol{c}_j^T \\ &= \Psi_{j,j} + \sum_{s=1}^{q} c_{j,s}x_{s,t}x_{s,t}^T c_{j,s}^T\,,\end{aligned} \tag{8–5}$$

$$\begin{aligned}\mathbb{C}\text{ov}\left[(\boldsymbol{e}_{t,j}, \boldsymbol{e}_{t,k})\right|\boldsymbol{x}_t] &= \Psi_{j,k} + \boldsymbol{c}_j\boldsymbol{x}_t\boldsymbol{x}_t^T\boldsymbol{c}_k^T \\ &= \Psi_{j,k} + \sum_{s=1}^{q} c_{j,s}x_{s,t}x_{s,t}^T c_{k,s}^T.\end{aligned} \tag{8–6}$$

This estimation is performed on a weekly sliding window, whereby for each sliding window period point estimators are obtained for the $\boldsymbol{C}$ matrix parameters. Therefore as the window slides different realizations are obtained based on the data fits for the estimated parameter relationships. I summarise these by constructing box plots of the parameter estimates for the full model containing $DOL + HML_{FX} + SPEC + SPEC \times SPEC$ covariates for both the high interest rate basket, which can be seen in Figure 8.4, and also for the low interest basket which is provided in Figure 8.5. In addition, for each sliding window the statistical significance of the estimated coefficient can be tested, where the null hypothesis would be that the parameter is zero versus an alternative that it is non-zero. A description of the test statistic is provided in Appendix B. The results of the test on each sliding window are indicated by adjusting the width of each box so that it is equal to the proportion of the sliding windows for which this parameter was significant, i.e. its confidence intervals did not cross zero. The baseline width of the box, i.e. if the parameter was significant on every sliding window, is given by the horizontal distance between two adjacent box midpoints. For details of the calculation of the confidence intervals for the parameters see Appendix B and more generally Hoff and Niu [2012].

To aid in the interpretation of this analysis I partition the results for the $\boldsymbol{C}$ matrix parameter estimation according to the loadings on each currency for a given factor, for instance in the high interest rate analysis there are four currencies considered. Thus for each factor in Figure 8.4, such as $DOL, HML_{FX}, SPEC$ and $SPEC \times SPEC$, as separated by the vertical red dotted lines and labelled at the top of the plot, there are four boxes, one for each of the time-evolving parameter estimate loadings for each currency in the order GBP, AUD, CAD and NZD. Similarly, the low interest rate basket contains three boxes per factor since only EUR, JPY and CHF are contained in this basket.

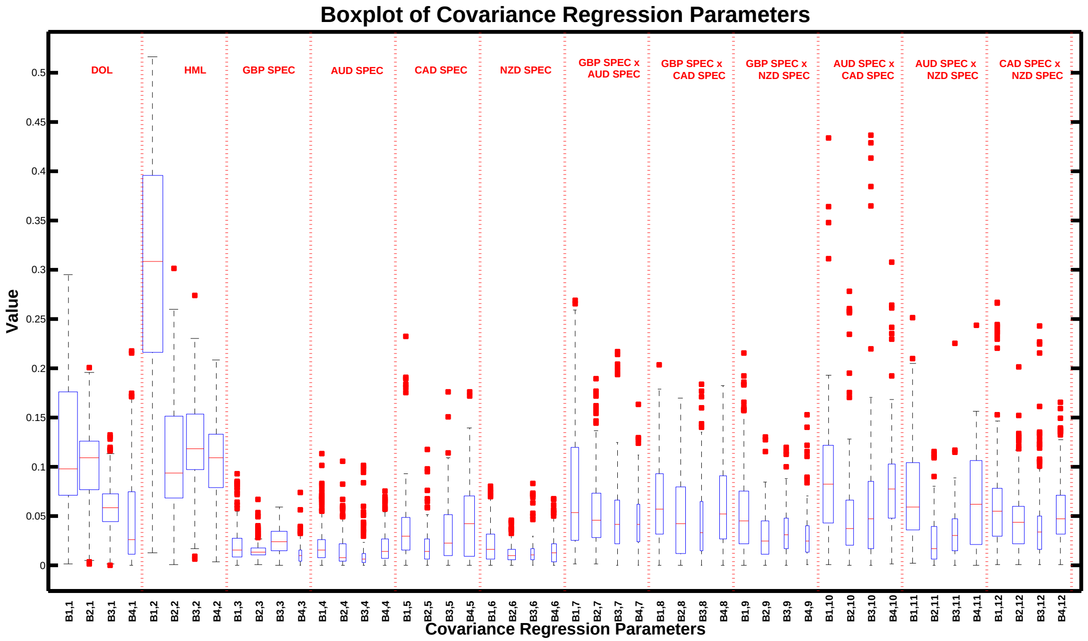


Figure 8.4: High interest rate basket parameter boxplot: $DOL + HML_{FX} + SPEC + SPEC \times SPEC$. 125 week lookback periods. The 4 currencies in the high interest rate basket are ordered as (GBP; AUD; CAD; NZD). The width of each box is equal to the proportion of the sliding windows for which this parameter was significant, i.e. 95% confidence intervals not crossing zero. The baseline width of the box, i.e. if the parameter was significant on every sliding window, is given by the horizontal distance between two adjacent box midpoints.

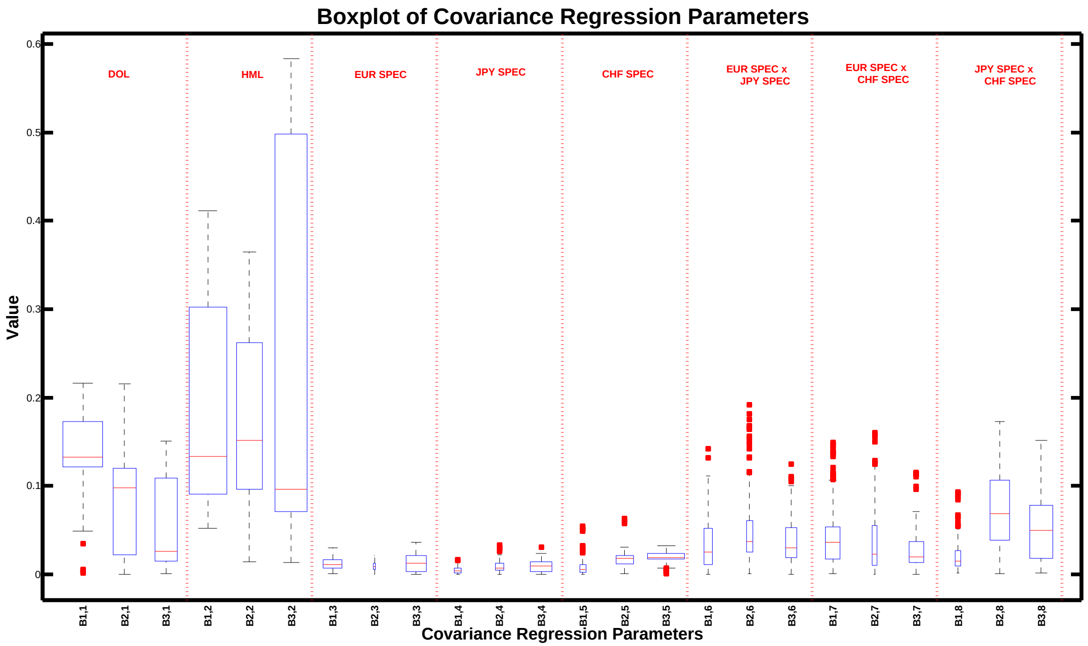


Figure 8.5: Low interest rate basket parameter boxplot: $DOL + HML_{FX} + SPEC + SPEC \times SPEC$. 125 week lookback periods. The 3 currencies in the low interest rate basket are ordered as (EUR; JPY; CHF). The width of each box is equal to the proportion of the sliding windows for which this parameter was significant, i.e. 95% confidence intervals not crossing zero. The baseline width of the box, i.e. if the parameter was significant on every sliding window, is given by the horizontal distance between two adjacent box midpoints.

These factor loadings can be interpreted as the proportion of change one would expect in the covariance relationships between each currency in the basket given a unit change in the factor. The utility of the covariance regression model lies in the additional variability $\mathbf{C x_i x_i^T C^T}$, which is randomly added to the baseline variability, $\boldsymbol{\Psi}$, of the de-trended data. Thus, for the low interest rate basket the set of vectors $\{\mathbf{c_{CHF}}, \mathbf{c_{JPY}}, \mathbf{c_{EUR}}\}$ associated to each currency should be interpreted as how the additional variability or heteroskedasticity manifests across the covariates. In other words the components of each vector, for instance $\mathbf{c_{CHF}} = [\mathrm{c_{CHF,DOL}}, \mathrm{c_{CHF,HML_{FX}}}, \dots]$, correspond to the sensitivities of the Swiss franc de-trended returns variability to the set of factors $\{DOL, HML_{FX}, \dots\}$. Thus, a high norm of the vector $\mathbf{c_{CHF}}$ should accordingly be associated with a high heteroskedasticity in the Swiss franc de-trended returns. Another interesting aspect of this analysis is the cross analysis of the vectors $\{\mathbf{c_{CHF}}, \mathbf{c_{JPY}}, \mathbf{c_{EUR}}\}$, since the pair of significantly different from zero vectors $\{\mathbf{c_{CHF}}, \mathbf{c_{JPY}}\}$ means that the Swiss franc and the Japanese yen become more correlated when their variances increase. It is possible to augment this analysis and interpret the highly significant pairs of vector components $\{c_{CHF,DOL}, c_{JPY,DOL}\}$ and $\{c_{CHF,HML_{FX}}, c_{JPY,HML_{FX}}\}$ as the significant effect of the DOL and HML_{FX} factors upon the covariance between the Swiss franc and the Japanese yen. Said differently, when the DOL and the HML_{FX} factors change, the covariance of the Swiss franc and the Japanese yen consequently increases. Furthermore, features can be seen which indicate that for the joint currency covariance structure the DOL factor is significant on all three currencies, but loads more substantially on the variance of the EUR compared to the JPY and CHF, indicating the heightened sensitivity of the volatility of the EUR to the DOL factor than the other currencies. This last reasoning from the covariance regression is particularly relevant for the analysis of the impact of the speculative positions under consideration in this thesis.

In Figure 8.5, the width of the boxes associated to the Swiss franc speculative interest covariate shows a significant increase in the covariance between the Swiss franc and the Japanese yen returns when the speculative interests on the Swiss franc increase. Whereas the width of the boxes associated to the Euro speculative covariate show that when the Euro speculative interests increase, a significant increase of the correlation between the Euro and the Swiss franc is noticeable.

These results demonstrate that speculator interventions on low interest rate currencies systematically influence the covariances among these currencies and that volume based information provides complementary indications about asset price dynamics and dependences.

Likewise, it is apparent in the high interest rate basket results that the DOL and the HML_{FX} factors load significantly on the covariance relationships for the GBP, more substantially than any of the relationships for the other currencies in this basket. This is especially the case for the HML_{FX} factor and in the majority of cases they are statistically significant loadings at 5% significance. Moreover, there seems to be an asymmetry in the factor loadings for the speculative open interest of one currency exchange on another currency exchange. For instance, whilst the impact of the speculative open interest of the GBP $SPEC$ is predominantly significant in all fits for the GBP, AUD and CAD exchange rates (not for the NZD), it can be seen that this speculative open interest factor loads much more substantially historically on the CAD than it does on the AUD or GBP. Conversely, an asymmetry arises in the speculative open interest on the CAD since the most dominant loading of this speculative open interest is on the GBP, CAD and NZD with the majority of fitted time periods being statistically significant.

Another very interesting point worth noting from this covariance regression is the relation between the GBP variance and the speculative volumes. It can be observed that for each $SPEC$ covariate the average width of the first box, which corresponds to the statistical significance of each covariate effect upon the GBP variance, is conspicuously higher than for the other currencies. Finally, the last two covariate results displayed in Figure 8.4 shed light on: the relation existing between the AUD and NZD speculator inflows correlation; the covariance between the GBP and NZD exchange rates; the synchronicity relation between the CAD and NZD speculator inflows; and the correlation between GBP, AUD and NZD exchange rates. These interesting outcomes substantiate the assertion in this thesis about the impact of the speculative open interest changes upon the individual variability and the dependence structure among the high interest rate currencies as well as the low interest rate currencies.

8.4 Skewness of Cross-Sectional Currency Returns: Pre and Post-Crisis Analysis

Standard linear measurements of association or concordance fail to provide any measurement of the asymmetric extreme relations exhibited by exchange rates. Here, I extend the approach proposed in Brunnermeier et al. [2008], where the authors compare the individual skewnesses of a set of currencies once they have been ranked as a function of their interest rate differential, to include more recent data. I present a similar chart to that of the authors representing the skewness and interest rate differentials[1] for developed countries, but I also extend the analysis by considering a combination of developed and developing countries[2]. Furthermore, I divided the data sample into three distinct periods, before, during and after the 2008 financial crisis, i.e. 29th July 1999 to 30 June 2007; 1st July 2007 to 30th June 2009; and 1st July 2009 to 29th January 2014.

It can be observed that whatever the basket of exchange rates under scrutiny, (Figures 8.6 and 8.7 or Figures 8.8 and 8.9), between 01/01/1999 and 29/01/2014 the skewnesses of the highest interest rate countries were clearly positive, which means that the depreciations of these currencies were asymmetrically more important than their appreciations. Likewise it can be observed that the currencies with the lowest interest rate differentials display a negative skewness, which shows a significant asymmetry with higher interest rate differential currencies. That being said, Figures 8.7 and 8.9, as well as the associated Tables 8.2 to 8.4 show

[1]Please note that in the analysis presented in this thesis the data used corresponds to the inverse of the exchange rates considered in the article of Brunnermeier et al. [2008]. Indeed the amount of foreign currency per unit of dollar is utilised, whereas Brunnermeier et al. [2008] consider the amount of dollars per foreign currency. Thus, a decrease in the exchange rates means in my case a depreciation of the dollar relative to the foreign currency, while an increase in the exchange rate represents an appreciation of the dollar. The slope of the regression of the skewness on the interest rate differentials is accordingly of opposite sign in the analysis presented in this thesis. It should also be noted that a regression of X on Y will not have the exact same inverse relationship that $1/X$ would have on Y.

[2]To choose these currencies, I consider all the developed and developing currencies available and look at the average interest rates differential with the US local interest rates over time periods spanning from 01/01/1999 to the 29/01/2014 and then rank them. I retain the five currencies that are most often present among the five highest interest rates differentials (namely TRY, BRL, ZAR, INR, MXN) and the five currencies that are most often present among the five lowest interest rates differentials (namely JPY, TWD, CHF, SGD, EUR).

contradictory information between 30/06/2009 (which corresponds to the end of the most recent recession according to the NBER statistics[1]) and 29/01/2014 given that both the high and the low interest rate currency marginal distributions are suggesting a positive skewness over this period of time. One can also note that the six-month period at the beginning of 2012 could potentially have a substantial effect on these regression estimates. During the crisis period, see Table 8.3, the explanatory power of the interest rate differential in the regression is substantially lower for both the developed countries regression and the developed and developing countries regression. In addition, when considering the developed countries alone the regression is not statistically significant. The evolution of the rolling high and low interest rate exchange rates cross-sectional average skewness across time (Figure 8.10) confirms this finding and shows that since the end of the financial crisis a different average asymmetry dynamic of the respective currency marginal distributions has prevailed. It seems indeed that the cross-sectional average skewness of the two sets of currencies are, over the past five years, showing noticeable synchronicity, which was not necessarily the case formerly. Furthermore, as can be seen in Figure 8.11, the low interest rate basket components recently display significant positive skewness, which reflects the European debt crisis and the monetary policy decisions made by the Japanese and Swiss central banks during this period.

Thus, as far as the speculator impact on market prices is concerned, the linear cross-sectional relation between marginal skewness and the interest rates differential pointed out by Brunnermeier et al. [2008] seems to be non-stationary over time and thus puts into question the cross-sectional relation between carry trade speculative flows and the currency dynamics described in the same article. Here, I argue that even though a peculiar event could marginally impact a specific currency, the construction of the carry trade portfolio by speculators should on the contrary simultaneously affect several currencies as a function of the associated interest rate differential. That being said, such confined events can still impact the carry trade performance and thus lead to position unwinding, which should

[1]The National Bureau of Economic Research dates economic recessions in the USA, by defining them as a persistent decline in several economic variables such as real GDP, real income, employment, industrial production, and wholesale-retail sales. For more information you can access these data on the NBER's website (http://www.nber.org/cycles.html).

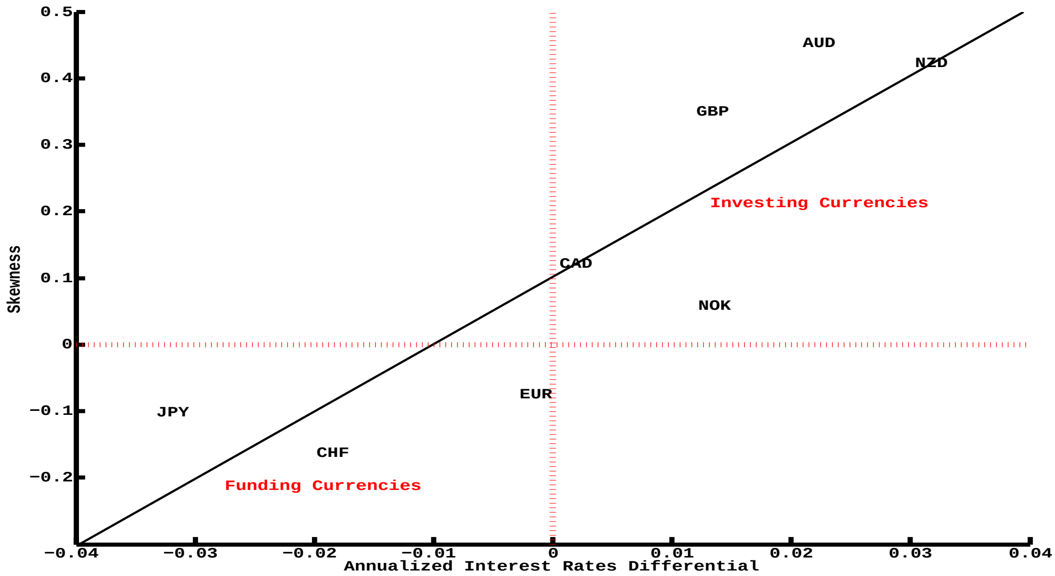


Figure 8.6: Developed Countries Before July 2007: Skewness vs Interest Rate Differential. Before July 2007.

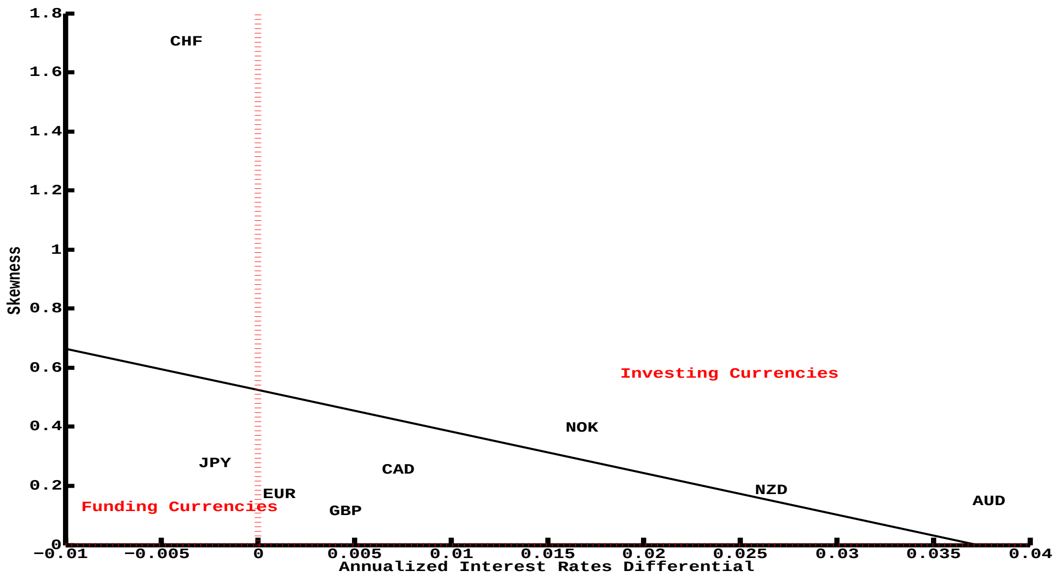


Figure 8.7: Developed Countries: Skewness vs Interest Rate Differential. After June 2009.

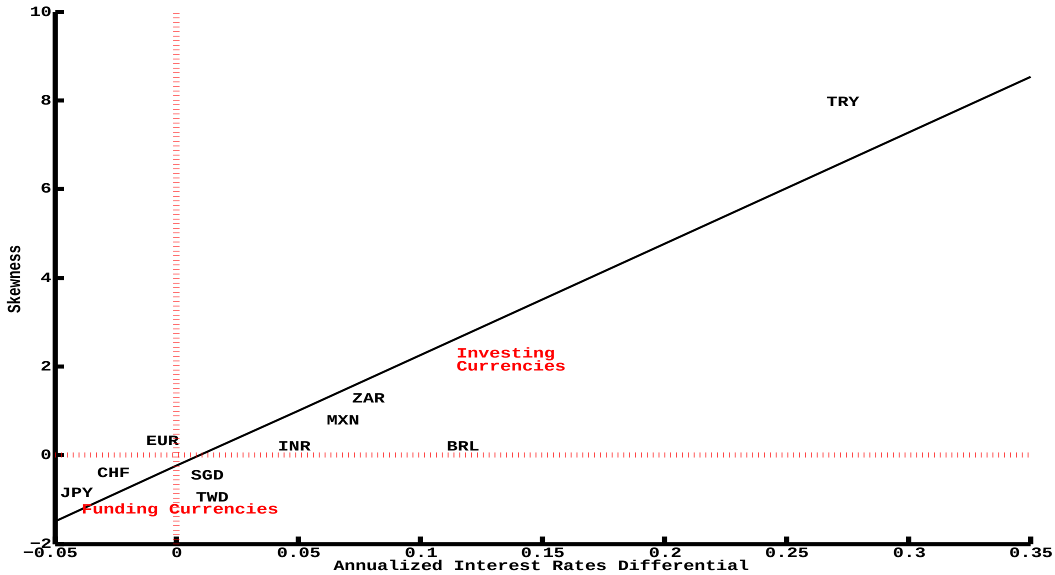


Figure 8.8: Developed and Developing Countries Before July 2007: Skewness vs Interest Rate Differential.

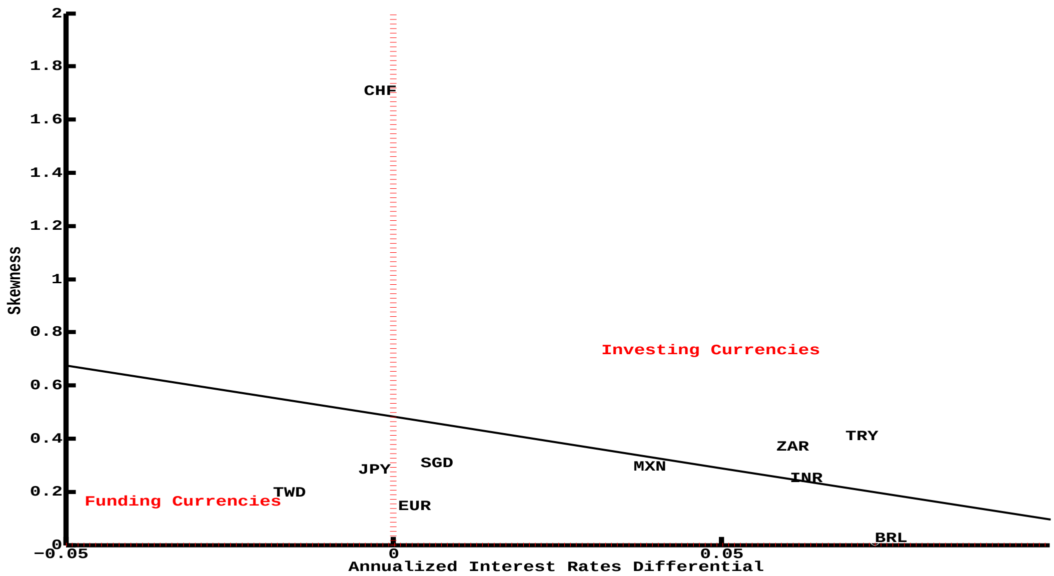


Figure 8.9: Developed and Developing Countries After June 2009: Skewness vs Interest Rate Differential.

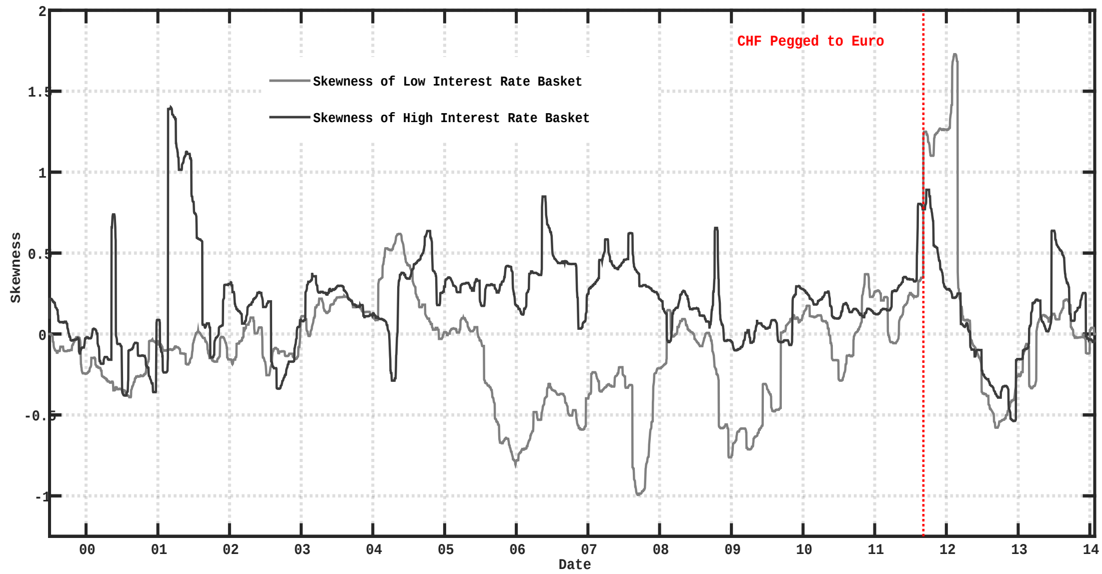


Figure 8.10: 6-month rolling average individual skewness of high interest rate developed countries (averaged over each individual currency: GBP, AUD, CAD, NOK, NZD) compared to rolling average individual skewness of low interest rate developed countries (averaged over each individual currency: JPY, CHF, EUR).

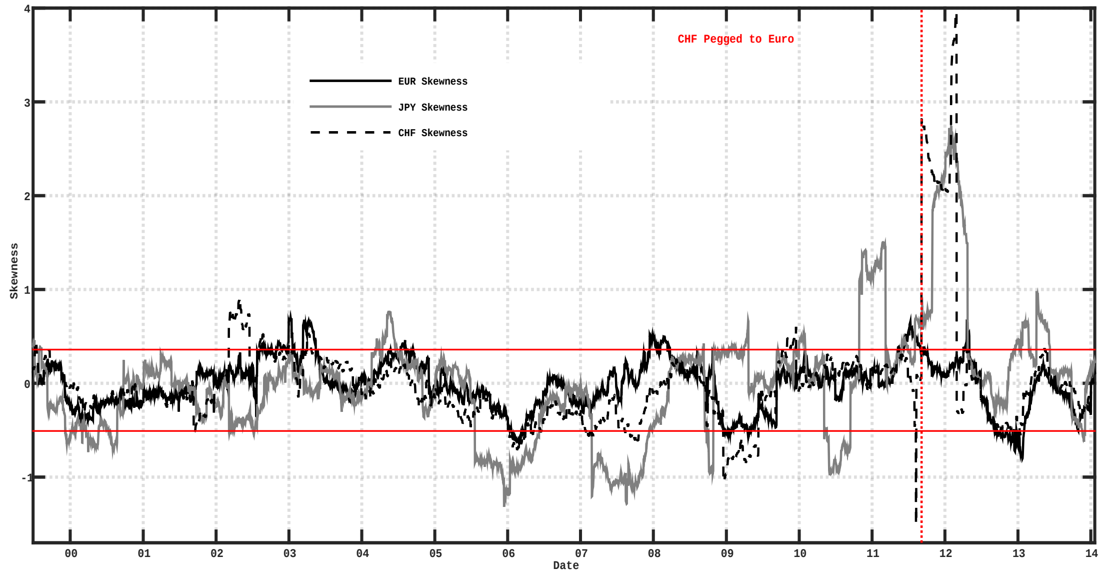


Figure 8.11: 6-month rolling average individual skewness of low interest rate developed countries (namely JPY, CHF, EUR) with upper and lower confidence intervals.

	Before Crisis (29-Jul-1999 / 30-Jun-2007)	
	Developed Countries	Developed and Developing Countries
Intersect	0.044	-0.337
Slope	7.911	26.23
R^2	0.847	0.7856
t-stat	5.773	6.902
P-value	0.001	1.01×10^{-4}

Table 8.2: Before July 2007: cross-sectional regression of the skewness on the interest rates differential for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

	Before Crisis (01-July-2007 / 30-June-2009)	
	Developed Countries	Developed and Developing Countries
Intersect	0.026	-0.109
Slope	7.370	8.948
R^2	0.295	0.456
t-stat	1.587	3.285
P-value	0.163	0.005

Table 8.3: During credit crisis: cross-sectional regression of the skewness on the interest rates differential for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

	After Crisis (01-Jul-2009 / 29- Jan-2014)	
	Developed Countries	Developed and Developing Countries
Intersect	0.524	0.482
Slope	-14.018	-3.853
R^2	0.151	0.082
t-stat	-1.034	-0.843
P-value	0.341	0.424

Table 8.4: After June 2009: cross-sectional regression of the skewness on the interest rates differential for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

again simultaneously impact the currencies composing the carry portfolio. Said differently, the features of the marginal distributions, such as the individual skewnesses retained by Brunnermeier et al. [2008], are not necessarily taking into account the joint distribution characteristics, such as high and low interest rate tail dependences. Furthermore, I assert that by selling the funding currencies and buying the investing currencies, speculators should asymmetrically influence the upper and lower extremal currency joint behaviour. In such a context, the dynamic of the respective upper and lower tail dependences characterizing the high and the low interest rate currency baskets are investigated in this thesis. To ensure that the speculative flows influence the extreme joint behaviour of the exchange rates it is necessary to first understand how theoretically the building and the unwinding of a dynamic carry trade strategy is impacting the high and low interest rate currencies. As detailed in Chapter 4, in order to benefit from the UIP violations a speculator will buy the high interest rate currencies while selling the low interest rate currencies relative to a reference currency, which here is the US dollar. As a result, when the international exchange rates system receives speculative inflows it should be possible to perceive an increase of the low interest rates basket upper tail dependence (evidence of significant sales of the basket currencies against the US dollar) while the high interest rates currencies will simultaneously display an increasing lower tail dependence (evidence of significant purchases of the basket currencies against the US dollar). It is assumed that at the same time no carry trade position will be unwound, which is represented by a low upper tail dependence within the high interest rate basket combined with the converse in the funding basket, i.e. a low lower tail dependence of the low interest rates currencies. Conversely, when the international exchange rates system faces speculative outflows, high interest currencies will be simultaneously sold in order to buy low interest rate currencies closing existing carry trade positions. It is assumed that no reverse carry trade positions are permitted in this economy, even though this would not dampen the conclusions. The outcome of this financial operation is naturally an increase of the high interest rate basket upper tail dependence and simultaneously an increase of the low interest rate lower tail dependence. Provided that investors are closing their positions I assume that no carry trade inflows are taking place at this point, hence it should be possible to observe a decrease of the high interest rate basket lower tail dependence as well as a decrease of the low interest rate upper tail dependence or at least observe low levels for these two dependence measures.

In this chapter the covariance regression model proposed in Chapter 7 has been utilised to investigate the effect of covariates, including speculative trading volumes, on the covariance structure. In addition, the skewness of the cross-sectional currency returns has been explored. In the next chapter, this analysis is extended to further understand the relationship between speculator behaviour and the dependence of currency returns via the average volatility and the tail dependences.

Chapter 9

Speculative Behaviour and Tail Dependence of Currency Returns

In this section, I focus mainly on the interpretation of the copula mixture estimation of exchange rate time series ranked relative to their respective level of local interest rate. The mixture copula parameters estimated provide, through the combination of the copula mixture components (Definition 36) and the associated upper or lower tail dependence expressions (Equations (3–33) and (3–34)), a parametric estimation of the upper or lower tail dependences, which quantify the level of upper or lower extremal dependence among the high interest rate and low interest rate sets of currencies. More precisely, the resulting estimation of the respective upper and lower tail dependences characterizing each basket of currencies reveals the complex non-linear relations existing between currencies, which remain totally imperceptible when one only considers either marginal characteristics of individual exchange rates or any linear central measure of dependence, such as covariance or correlation.

9.1 Extremal Carry Trade Behaviour and Average Currency Volatility

The financial literature about the carry trade states that carry positions will be sensitive to the risk aversion and the level of volatility in the foreign exchange markets (Brunnermeier et al. [2008]; Farhi and Gabaix [2008]; Clarida et al. [2009]; Menkhoff et al. [2012a]). Therefore it could be postulated that speculators tend to close their carry positions when the foreign exchange market volatility is increasing or at a high level (embodying an increasing uncertainty and thus investor risk aversion), whereas they will build their carry positions when the foreign exchange market volatility is decreasing or at a low level. This first assumption about speculator behaviour improves the ability to detect the potential relation between tail dependence and the propensity of a speculator to build or unwind carry trades. According to the investigation on the respective basket average skewness presented in Section 8.4, I split the data set into three very distinct sub-periods. Firstly, I estimate the low and high interest rate basket tail dependences on the pre-financial crisis period, which runs from 01/01/1999 to 30/06/2007. Secondly, I estimate on the during crisis period, which runs from 01/07/2007 to 30/06/2009. Finally, I estimate on the post-financial crisis period, which runs from 01/07/2009 to 29/01/2014. Then, I individually regress the tail dependence time series upon the average foreign exchange market volatility as shown in Equation 9–1.

$$\hat{\lambda}^i_{j,t} = \beta^{i,j}\sigma^{FX}_t + \epsilon^i_{j,t} \quad , \quad i = \{H, L\}\, , \, j = \{u, l\}. \tag{9–1}$$

where $\sigma^{FX}_t = \frac{1}{N}\sum_{n=1}^{N} |r_{n,t}|$ and $r_{n,t}$ is the log return of currency n on day t. The i index here corresponds to the high interest rate basket or low interest rate basket. The j index corresponds to the upper tail dependence or lower tail dependence measure. Since the value of the tail dependence is bounded, i.e. $0 \leq \hat{\lambda}^i_{j,t} \leq 1$, this regression could be performed via a generalised linear model (glm) with a logit link function. However, it was found that in the empirical investigation there were no incidences of boundary problems with the estimated values of $\hat{\lambda}^i_{j,t}$, i.e. the estimated values were not limited by the hard constraints of $[0, 1]$.

The results of this regression before the crisis in Table 9.1 demonstrate that the high interest rate lower tail dependence is negatively sensitive to the average volatility of foreign exchange markets, except for the developing countries. Likewise, the upper tail dependence of the low interest rate basket displays a negative significant relation with the average volatility in currency markets. Contrary to this, the low interest rate lower tail dependence is positively reacting to market volatility. Note that during the pre-crisis period the high interest rate upper tail dependence regression coefficients were less significant, which could be explained by the fact that a large part of this sample was characterized by a low volatility environment during which the carry trade strategies performed very well (see Lustig and Verdelhan [2007]). This is a well-known time of heightened carry trade construction, hence the most important sensitivities to consider during this period were the low interest rate upper tail dependence and the high interest rate lower tail dependences. During the crisis, see Table 9.2, the regression relationships are mixed for the low interest rate countries. However, for the high interest rate countries (both developed and developing) one can see that the upper tail dependence and the average FX volatility is negatively related, which one would expect during this period. Now turning to the post-crisis period, shown in Table 9.3, it can be observed that in such a high volatility environment the high interest rate upper tail dependence is significantly positively related to the market volatility and the high interest rate lower tail dependence is conversely significantly negatively affected by volatility changes (except for the developing countries, but this is not statistically significant). These results validate the proposed model and the associated hypothesis about the impact of carry trade speculative flows upon the extreme joint behaviour of international exchange rates relative to their level of short-term interest rates. It is also particularly interesting to notice that during the post-financial crisis the low interest rate upper and lower tail dependences remain significantly sensitive to the level of volatility in the foreign exchange markets and thus, according to the hypothesis, to the carry trade flows. This statistical stability has to be weighed against the switching behaviour of the average skewness identified in the previous chapter.

9.2 Extremal Carry Trade Behaviour and Currency Speculative Open Positions

To validate the assertion about the influence of carry trade speculative flows on currency extremal joint behaviour, I propose in this second model to consider the same covariates as for the covariance regression model in Section 8.3.3, namely

	Before Crisis (29-Jul-1999 / 30-Jun-2007)	
	Upper TD	Lower TD
Low Int. Rates Developed Countries	-2.914 (0.000)	1.539 (0.000)
High Int. Rates Developed Countries	0.651 (0.377)	-1.147 (0.047)
High Int. Rates All Countries	-1.012 (0.061)	2.107 (0.000)

Table 9.1: Before July 2007: Regression of the high interest rate upper and lower tail dependences time series ($\hat{\lambda}^H_{u,t}$, $\hat{\lambda}^H_{l,t}$) and the low interest rate upper and lower tail dependences time series ($\hat{\lambda}^L_{u,t}$, $\hat{\lambda}^L_{l,t}$) on the average volatility for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

	During Crisis (01-Jul-2007 / 30-Jun-2009)	
	Upper TD	Lower TD
Low Int. Rates Developed Countries	-0.508 (0.008)	-1.400 (0.000)
High Int. Rates Developed Countries	1.101 (0.019)	-0.901 (0.011)
High Int. Rates All Countries	1.245 (0.003)	-0.120 (0.727)

Table 9.2: During credit crisis: Regression of the high interest rate upper and lower tail dependences time series ($\hat{\lambda}^H_{u,t}$, $\hat{\lambda}^H_{l,t}$) and the low interest rate upper and lower tail dependences time series ($\hat{\lambda}^L_{u,t}$, $\hat{\lambda}^L_{l,t}$) on the average volatility for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

	After Crisis (01-Jul-2009 / 29-Jan-2014)	
	Upper TD	Lower TD
Low Int. Rates Developed Countries	-1.771 (0.000)	2.067 (0.000)
High Int. Rates Developed Countries	0.657 (0.094)	-1.344 (0.000)
High Int. Rates All Countries	2.279 (0.000)	0.231 (0.565)

Table 9.3: After June 2009: Regression of the high interest rate upper and lower tail dependences time series ($\hat{\lambda}^H_{u,t}$, $\hat{\lambda}^H_{l,t}$) and the low interest rate upper and lower tail dependences time series ($\hat{\lambda}^L_{u,t}$, $\hat{\lambda}^L_{l,t}$) on the average volatility for developed (AUD, CAD, JPY, NZD, NOK, CHF, GBP and EUR) and developing countries (SGD, TWD, INR, MXN, ZAR, BRL and TRY).

the DOL and the HML_{FX} factors combined with the $SPEC$ factor. It was observed in Section 8.3.3 of Chapter 8 that speculative positions in the markets have a substantial impact on the covariance dynamics of international exchange rates. To complete this analysis I demonstrate in this section the existence of the previously stated dual relation between on one hand the high and low interest rate currency baskets upper and lower tail dependences, and on the other hand the size of the speculative positions associated to these funding and investing currencies futures. In this section, the empirical study consists thus in investigating the relation between the non-commercial traders net position (long $-$ short) and the extreme environment dependence measure, namely the tail dependence. The base idea is to assume that while speculators set up or unwind a carry position in the currency market, this will synchronously impact all the currency prices increasing accordingly certain tail dependences among high and low interest rates currencies. Furthermore, a synchronous change in the net open position of the speculators should be observable. As a first example, it can be seen from Figure 9.1 that there is a negative relationship between the net position of speculators on the Swiss franc (one of the main financing currencies) and the upper tail dependence

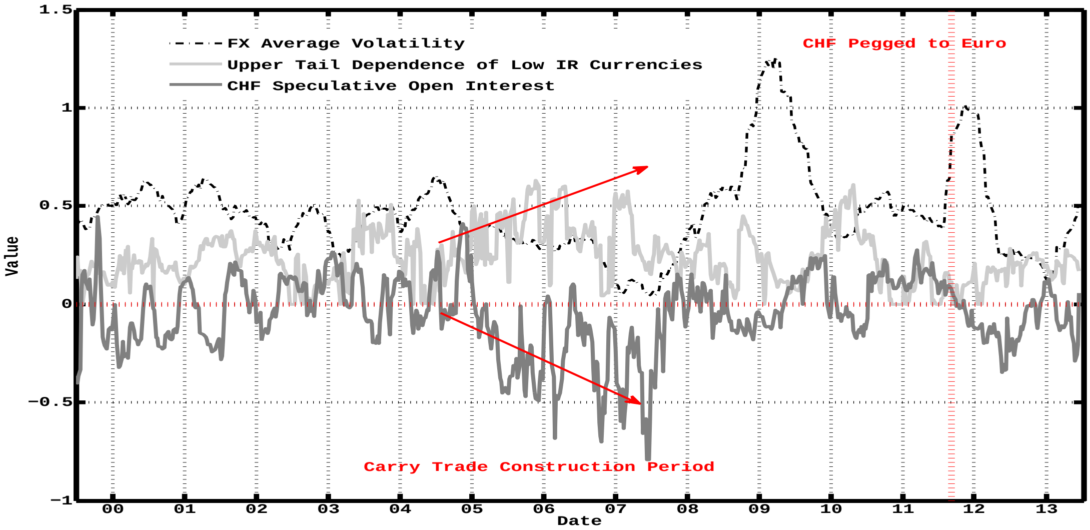


Figure 9.1: 6-month rolling upper tail dependence of low interest rate developed countries (namely JPY, CHF, EUR) compared to net open position of the Swiss franc future contract traded on the CME. The black line corresponds to the average 6-month rolling historical volatility computed from the low interest rate basket.

associated to the low interest rates basket. Since 1999 the Swiss franc has indeed always been one of the lowest interest rate currency relative to the US Dollar and thus always used by the speculator as a financing currency.

In order to verify this assertion I model the four tail dependences as a function of the *SPEC* factor, i.e. the ratio of the non-commercial net positions at the end of each week divided by the total number of futures contracts still open in the market at the end of each week. This will then act as a factor to help explain how much of the currency extremal dependence can be explained by the speculative positions. The first problem to deal with is the homogeneous impact that the dollar can have on the common behaviour of the currency open positions.

When the dollar index, defined as a basket of currencies against the dollar, increases the tail dependences could potentially be modified too. To extract the linear effects associated to this component I follow the analysis carried out in Lustig and Verdelhan [2007], who demonstrated the effect of the dollar index through a principal component analysis in which they interpret the first principle component as the dollar index.

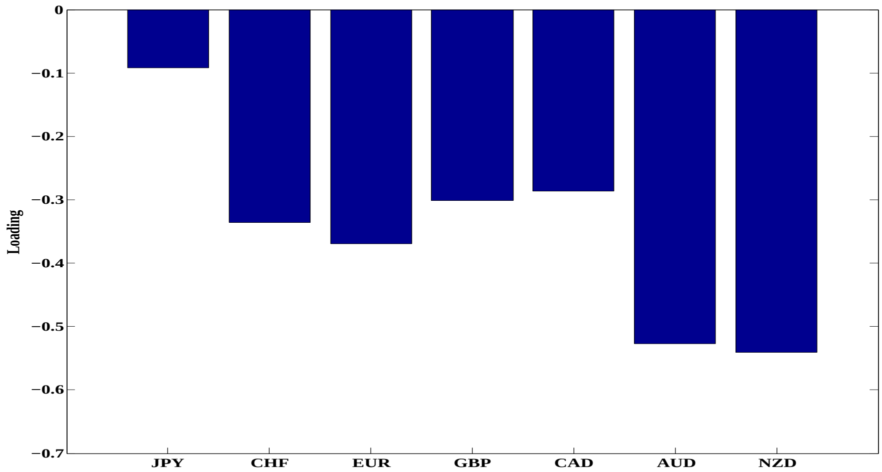


Figure 9.2: Loadings of the First Principal Component of Developed Countries Currency Returns.

To achieve this, I first extend the monthly PCA analysis of Lustig and Verdelhan [2007] to a daily frequency, motivating the construction of daily DOL and HML_{FX} factors, which I then use to compute the weekly factors. To construct these factors at the daily frequency, it is necessary to calculate the daily carry returns via an interpolation on the 1 month forward curve using the following market price data: overnight rate, one week rate, two week rate, three week rate and the one month rate. The details of this interpolation procedure can be seen in Appendix C, along with an example of an interpolated curve for one particular day. From this interpolated forward curve it is possible to construct a daily time series of carry returns for each of the 7 currencies via a mark to market of the forward contract that would be held if one was continuously rolling one month forward contracts at the end of each month, as in Lustig and Verdelhan [2007]. These individual currency carry returns can then be used to compute the covariance matrix that is used for the principal component analysis.

Instead of applying the principal component analysis to a set of portfolios, I used directly the seven currencies for which the CFTC open interest data is available. Figure 9.2 shows that the daily analysis replicates the results of Lustig and Verdelhan [2007], where all the currencies are negatively impacted by

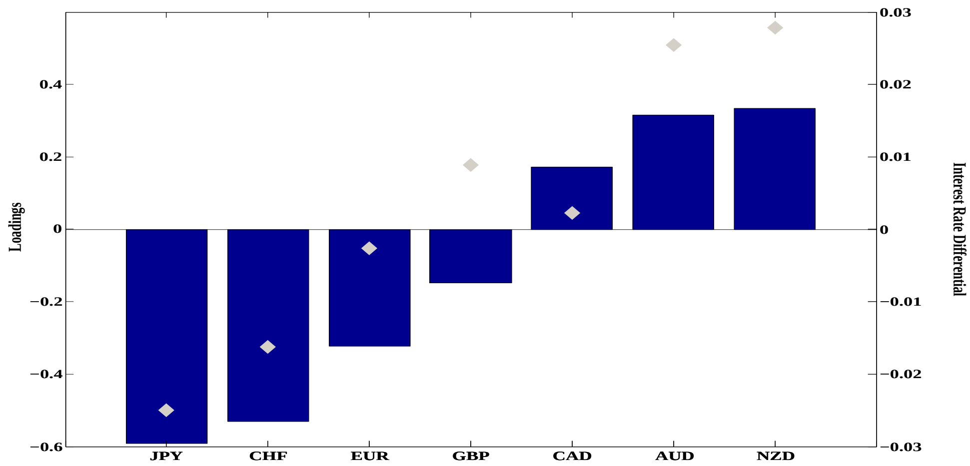


Figure 9.3: Loadings of the Second Principal Component of Developed Countries Currency Returns. The bars (left axis) represent the loadings values on the returns second principal component while the grey diamonds (right axis) depicts the level of interest rate differential with the one month US interest rate.

the first component (see Figure 9.2) which represents the dollar effect (DOL), whereas the second component is (almost) monotonically increasing with the rate differential (see Figure 9.2), which is analysed as the high minus low effect (HML_{FX}). It is found that over 76% of the variation of the daily carry returns can be explained in the first two principal components. It is worth emphasizing that these two projections of the currency returns to linear combinations of the currency returns with the largest unconditional variances are not necessarily related to the conditional variance model described in Section 8.3 of Chapter 8. In the remainder of this section I consider the tail dependence regression and thus use the first two principal components time series (since the PCA analysis is run on a sliding window basis) as independent variables for the regression to cancel out the effect of these two price effects related to DOL and HML_{FX}.

More formally, I perform linear regression on the upper and lower tail dependences of the high and the low interest rate sets of currencies (respectively $\hat{\lambda}_u^H, \hat{\lambda}_l^H, \hat{\lambda}_u^L, \hat{\lambda}_l^L$) using three models. The first model contains only the DOL and HML_{FX} factors as well as their volatilities and covariance. The second model includes the $SPEC$ factors[1]. Finally the third model further allows the cross terms between the $SPEC$ factors:

$$\begin{aligned}\hat{\lambda}^{i}_{j,t} &= \alpha^{i,j} + \underbrace{\beta^{i,j}_{DOL} DOL_t + \beta^{i,j}_{HML_{FX}} HML_{FX_t}}_{\text{Dollar and Carry Factors}} \\ &+ \underbrace{\beta^{i,j}_{\sigma_{DOL_t}} \sigma_{DOL_t} + \beta^{i,j}_{\sigma_{HML_{FX_t}}} \sigma_{HML_{FX_t}} + \beta^{i,j}_{\mathrm{Cov}_{DOL_t,HML_{FX_t}}} \mathrm{Cov}_{DOL_t,HML_{FX_t}}}_{\text{Dollar and Carry Factor Volatilities and Covariance}} \\ &+ \underbrace{\sum_{k=1}^{N} \beta^{i,j}_{k} SPEC^{k}_{t}}_{\text{Speculative Volume Factors}} + \underbrace{\sum_{k=1}^{N}\sum_{l>k} \beta^{i,j}_{k,l} SPEC^{k}_{t} \times SPEC^{l}_{t}}_{\text{Speculative Volume Cross Terms}} + \epsilon^{i,j}_{t} \quad , \qquad (9\text{–}2)\end{aligned}$$

where $i = \{H, L\}$ and $j = \{u, l\}$.

[1] Utilising $\Delta SPEC$ as the covariate was investigated but found to not hold as much explanatory power in the covariance regression modelling framework utilised here.

	$\hat{\lambda}_u^H$			$\hat{\lambda}_u^L$			$\hat{\lambda}_l^H$			$\hat{\lambda}_l^L$		
$Constant$	0.096	0.057	0.080	0.370	0.337	0.270	0.350	0.433	0.339	0.121	0.086	0.178
	(0.122)	(0.435)	(0.268)	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.045)*	(0.190)	(0.012)**
DOL	0.083	-0.023	-0.162	-0.319	-0.264	-0.107	0.143	0.134	0.014	0.249	-0.006	-0.061
	(0.707)	(0.913)	(0.317)	(0.089)	(0.079)	(0.418)	(0.321)	(0.390)	(0.920)	(0.223)	(0.972)	(0.707)
HML_{FX}	0.303	0.030	-0.032	-0.205	0.056	-0.172	-0.020	0.135	-0.058	0.410	0.155	-0.041
	(0.380)	(0.922)	(0.921)	(0.445)	(0.807)	(0.398)	(0.945)	(0.609)	(0.785)	(0.227)	(0.590)	(0.859)
σ_{DOL}	-8.696	-2.343	0.625	-8.733	-7.866	-8.883	3.166	-0.335	1.640	-2.268	-1.606	2.691
	(0.051)	(0.553)	(0.863)	(0.000)**	(0.001)**	(0.000)**	(0.324)	(0.920)	(0.503)	(0.609)	(0.664)	(0.510)
$\sigma_{HML_{FX}}$	26.700	23.177	16.159	-8.918	-10.300	4.468	-0.246	-1.715	-2.267	31.495	31.478	16.269
	(0.003)**	(0.009)**	(0.058)	(0.098)	(0.186)	(0.516)	(0.976)	(0.869)	(0.794)	(0.001)**	(0.000)**	(0.065)
$\sigma_{DOL,HML_{FX}}$	-258.50	-13.83	-201.89	-1088.66	-958.77	-710.73	577.01	239.62	333.00	1103.72	868.16	776.10
	(0.488)	(0.966)	(0.576)	(0.000)**	(0.000)**	(0.004)**	(0.072)	(0.378)	(0.173)	(0.008)**	(0.007)**	(0.023)*
$AUD\,SPEC$		-0.129	-0.173		0.089	-0.049		0.083	0.272		-0.083	-0.105
		(0.017)*	(0.059)		(0.052)	(0.403)		(0.066)	(0.000)**		(0.154)	(0.160)
$CAD\,SPEC$		0.126	0.281		0.025	-0.017		-0.053	0.114		0.054	-0.013
		(0.008)**	(0.011)*		(0.663)	(0.804)		(0.160)	(0.204)		(0.236)	(0.899)
$CHF\,SPEC$		-0.023	-0.014		-0.136	-0.130		0.173	0.036		0.186	0.219
		(0.679)	(0.900)		(0.002)**	(0.028)*		(0.000)**	(0.687)		(0.001)**	(0.027)*
$EUR\,SPEC$		0.154	0.232		0.129	0.149		-0.084	-0.109		0.002	-0.031
		(0.014)*	(0.042)*		(0.012)*	(0.049)*		(0.101)	(0.125)		(0.975)	(0.812)
$GBP\,SPEC$		0.046	0.133		-0.162	0.035		-0.070	-0.290		-0.011	-0.124
		(0.404)	(0.203)		(0.004)**	(0.637)		(0.214)	(0.006)**		(0.855)	(0.214)
$JPY\,SPEC$		-0.023	-0.053		-0.040	0.016		-0.128	-0.137		-0.054	-0.109
		(0.631)	(0.534)		(0.370)	(0.771)		(0.002)**	(0.009)**		(0.278)	(0.084)
$NZD\,SPEC$		0.008	-0.020		-0.027	-0.064		-0.104	0.025		0.122	0.129
		(0.878)	(0.745)		(0.379)	(0.141)		(0.007)**	(0.675)		(0.017)*	(0.024)*
$CROSS\,SPEC$			$Cross1$			$Cross2$			$Cross3$			$Cross4$
R^2	14.5%	27.1%	39.4%	15.6%	30.9%	53.7%	5.2%	23.6%	46.0%	11.7%	31.7%	48.5%

Table 9.4: Regression of the high and low interest rate respective tail dependences on the DOL index, HML_{FX} index, DOL index volatility, HML_{FX} index volatility, DOL and HML_{FX} indices covariance and the $SPEC$ ratio (the ratio of each currency future speculative net positions to the total future open interest, as provided by the CFTC) as well as cross relations among them. The open interest data provided by the CFTC as well as the computed DOL and HML_{FX} indexes are weekly data while the respective tail dependence measurement corresponds to the average value over each week. The period of time considered for this analysis spans from June 20th 2006 to January 28th 2014 and corresponds to the longest overlapping sample for all the currencies considered and available. Numbers in parentheses show Newey and West [1987] HAC p-values. Cross1 corresponds to all the possible cross effects among which the following are statistically significant (below 5%): AUD/EUR, EUR/GBP, EUR/NZD. For Cross2 the following cross effects are statistically significant (below 5%): AUD/JPY, CAD/EUR, EUR/GBP. For Cross3 the following cross effects are statistically significant (below 5%): AUD/GBP, EUR/GBP, GBP/JPY. For Cross4 the following statistically crosses effects are statistically significant (below 5%): AUD/EUR, AUD/GBP, CHF/EUR, EUR/JPY, JPY/NZD.

The first observation to make on the results displayed in Table 9.4 is that the speculator activity contributes significantly to the explanatory power of the tail dependences regression whereas the DOL and HML_{FX} factor variances and covariance do not systematically explain with significance this currency extremal joint behaviour. Furthermore, it should also be highlighted that the adjusted R^2 is noticeably increased once the variables related to the speculator positions in the market are included.

From the graphical analysis of the Swiss franc speculative open interests, seen in Figure 9.1, and Table 9.4 it can be seen that this common financing currency is significantly related to the upper and the lower tail dependences of the low interest rates basket and that the sign associated is also corroborating the hypothesis of this thesis. As a matter of fact, when the Swiss franc is primarily sold by the speculators, who are building their carry portfolio, the upper tail dependence of the low interest rate currencies is increasing, as seen in Figure 9.1 during the known carry trade construction period of 2004 to 2007. Conversely, when the Swiss franc is primarily bought by the speculators, who are unwinding their carry portfolio, the lower tail dependence of the low interest rates basket tends to increase.

Interestingly enough, the Australian dollar is playing exactly the same role for the high interest rate basket. This typical investing currency is indeed significantly contributing to the upper and the lower tail dependences of the high interest rate basket. Furthermore, the two signs associated to the regression coefficients of the Australian currency (as seen in columns 8 and 9 of the $AUDSPEC$ row of Table 9.4) also validate the theory proposed in this thesis, since the purchase of the Australian dollar, following from the construction of a carry trade position by speculators is leading to an increase in the lower tail dependence among the high interest rate currencies. On the contrary, when speculators sell the Australian currency in order to reduce their carry trade exposure an increase in the upper tail dependence among the high interest rate currencies is observed (as seen in columns 2 and 3 of the $AUDSPEC$ row of Table 9.4). Finally, it can also be seen that the speculator positions on the Euro, the British pound and the New Zealand dollar have informative power about the extremal joint behaviour of the international exchange rates.

Part III

Currency Portfolio Optimisation Contributions

Chapter 10

Part III Overview

In the third part of this thesis, portfolio optimisation techniques are introduced and then the novel covariance forecasting approach developed in Part II is utilised in order to investigate portfolio optimisation in currency carry portfolios. The complementary "upstream" and "downstream" approaches associated with the challenges of portfolio optimisation are considered.

Chapter 11 reviews the literature on portfolio optimisation. To begin, the traditional Markowitz mean-variance approach to portfolio optimisation is introduced and then the more recently popular risk-based portfolio allocation techniques are presented. Furthermore, the sensitivity of the portfolio weights to the covariance factors introduced in Part II is discussed.

Chapter 12 explores the utility of the proposed GFM model in covariance forecasting and portfolio optimisation. In particular, the added performance of incorporating the $SPEC$ factors into the model for the covariance is analysed. The details of the procedure followed for the comparison of the covariance forecasting models is explained and finally the resulting performance of the proposed approach is discussed.

Chapter 11

Portfolio Optimisation

In this chapter, the literature on portfolio optimisation is reviewed. To begin, the traditional Markowitz mean-variance approach to portfolio optimisation is introduced and then the more recently popular risk-based portfolio allocation techniques are presented. Furthermore, the sensitivity of the portfolio weights to the covariance factors introduced in Part II is discussed.

11.1 Introduction

The advent of modern portfolio theory, with the seminal mean-variance model proposed by Markowitz [1952], forged new frontiers for a large area of finance literature and contributed to significant developments within the asset management industry. Nevertheless, the performance of such models and more importantly the validity of the accompanying statistical assumptions underpinning the application of such models to portfolio selection has been questioned. This is due to widely documented observed inconsistencies in the model assumptions and the practical applications. This has resulted in numerous interrogations about the practical implementation of this seminal model and subsequent model extensions to the original framework to address such issues.

The task of portfolio allocation can be divided into four stylized non-independent stages as follows:

1. Statistical model estimation and model selection of the portfolio constituent

multivariate return processes historically.

2. Some form of forecasting under the estimated model selected.

3. Selection and estimation of a risk measure on which to measure performance of the portfolio.

4. An optimization criterion upon which to perform portfolio allocation based on the portfolio forecast risk measure.

Under the classical mean-variance based models several challenging model prerequisites for such a framework arise. Most notably these include the estimation of essential but unknown parameters such as each portfolio component's drift and diffusion terms as well as the dependence structure between them, as measured often through correlation and covariance relationships, but sometimes also through other concordance measures such as tail dependence. When such statistical models are then utilized for stage two, the forecasting, and subsequently stages three and four in the portfolio selection, it is important to study the influence of the model assumptions, the model choice, the model estimation and the model forecast accuracy on the performance of the portfolio allocation method in stages three and four. In this regard, several works have undertaken analysis of such considerations in terms of considering the sensitivity of the mean-variance optimal portfolio behaviour, examples can be found in both the academic and practitioners' literature, see Jobson and Korkie [1981]; Frost and Savarino [1988]; Michaud [1989]; Chopra and Ziemba [1993]; Broadie [1993]. For instance it has been shown that the basic mean-variance quadratic program happens to be highly sensitive to models that fail to account for heteroskedascity in the covariance, and such models have been shown to be equivalent to a re-expression of the estimation problem as a measurement error maximization program, further highlighting the importance of this covariance modelling feature, see Michaud [1989]; Nawrocki [1996].

Several of these studies have demonstrated that indeed one of the most important features to capture in the real portfolio data returns is in fact the trend structure of the portfolio returns, but perhaps even more importantly the heteroskedastic nature of the portfolio covariance structure over time. Whilst trend is widely considered to be notoriously difficult and unpredictable even with the most carefully developed models, the heteroskedastic nature of the covariance structure is considered to be more reliably predictable and amenable to model development. Not only have these features been shown to be important model components to capture accurately in stages one and two of the process, but in addition since the portfolio allocation and subsequently portfolio performance in terms of returns and risk performance is highly sensitive to the ability of the model to correctly capture these dynamic features over time, they also directly affect stages three and four.

Therefore, several approaches have subsequently been developed in the academic literature to address these problems and generally they can be split into two categories. The categorisation depends on which aspect of the four stages they modify to try to address the above identified issues, particularly on heteroskedasticity of the portfolio covariance: i.e. at the modelling stage, the forecasting stage, the risk measure specification stage or in the portfolio optimization program objective function in stage four.

Stages one and two are referred to as "Upstream" approaches, which focus on:

1. Improving the model development and forecasting framework, i.e. the input estimation that produces the risk measure of the portfolio and acts as input to portfolio optimization.

2. Reducing the noise on the input sources.

Stages three and four are referred to as "Downstream" approaches. These approaches consider the input noise as an inexorable feature of financial market data and accordingly focus on:

4. The risk measure.

5. Adjusting the optimization program through reformulation of the loss function or through refinement of optimization constraints in order to restrain the estimator bias and its effect upon the optimal portfolio allocation solution.

In summary, one could make more robust model estimations and forecasts and utilise existing portfolio allocation methods, or alternatively one can make more resilient and constrained portfolio allocation methods to account for weaker models in stages one and two.

From the quantitative finance perspective, it has been more popular in the academic literature to address the challenges highlighted through refinement of the upstream aspects. In this context there exists an abundant literature wherein three different approaches are particularly worth discussing in the context of this thesis. Firstly, factor models (Green and Hollifield [1992]; Chan et al. [1999] or more recently Santos and Moura [2014]) have been proposed in which the potential portfolio assets have their conditional covariance matrix and drift modelled based on considerations of a constructed value-weighted market index. This is akin to the approach adopted in single factor models such as Sharpe [1963] or the augmented multi-factors models devised by Fama and French [1993] or Carhart [1997]. In the same vein the latent factor models instead promote a transformation and dimension reduction approach based on constructing factors that are orthogonal and typically obtained based on principle component analysis (PCA) based decompositions. This is achieved at the expense of the economic interpretation that would have been offered by the non-transformed factors, see Han [2006]; Zhang and Chan [2009].

The second approach in the literature to tackle issues with the upstream modelling involves development of models that attempt to capture the portfolio assets price time series heteroskedasticity through time series model structures. Typically this includes the modelling of correlation and volatility time variability under some variant of a multivariate GARCH model such as the widely considered class of Dynamic Conditional Correlation (DCC) models, see Engle [2002]; Engle and Colacito [2006]; Aielli [2013], where the heteroskedasticity is only temporal and does not depend on economic factors. The class of DCC models has been a focus in the literature since they calculate the covariance between the asset returns as a function of their past volatility and the correlations between them. The relationship between the DCC models and GARCH models means that a DCC model typically utilises recent past information in the estimation of the present correlation between series, thereby implicitly filtering or down weighting historical returns over some horizon. Such models involve the estimation of the covariance matrix which can be made either directly, as in the vectorised

GARCH (VEC) formulations developed in Bollerslev [1990] and the diagonal VEC (DVEC) and restricted VEC (often called BEKK) models, see Engle and Kroner [1995], or indirectly using conditional correlations as in CCC, DCC or STCC (Smooth Transition Conditional Correlations) models. Then there are also dimension reduction based versions of such models such as the orthogonal GARCH (O-GARCH) proposed by Alexander [2000], which develops the model as linear combinations of uncorrelated factors. In this manner it is akin to the approach of principal component analysis dimension reduction. However, it has been observed that in cases in which the portfolio returns are weakly correlated, or the portfolio components have similar unconditional variance, then it is likely that problems in the estimation of O-GARCH will occur and manifest typically in numerical instability of the fits and forecasts and therefore of the overall portfolio allocation framework that results. Consequently, this O-GARCH framework was further refined to the generalised version GO-GARCH of Van der Weide [2002]. In addition to these classes of DCC models there are also models known as time varying correlation (TVS) models, see Christodoulakis and Satchell [2002].

Finally, the third approach involves Bayesian methods, which have been proposed to reduce the variance of the input estimator. The technique of shrinkage was originally applied to the mean parameter estimation by Jorion [1985, 1986] and then subsequently extended by incorporating qualitative inputs with the Black-Litterman model, see Black and Litterman [1991]. This approach was extended to the covariance matrix by Ledoit and Wolf [2003] and more recently to the inverse of the same covariance matrix by Kourtis et al. [2012]. This technique optimally combines two existing estimators, such as the sample estimator (respectively for the expected value or the covariance matrix) and for instance a factor model based estimator. More recently Garlappi et al. [2006]; Boyle et al. [2012]; Branger et al. [2013] utilise a multi-prior model to take account of the investor's aversion to ambiguity or model mis-specifications in the optimal portfolio.

From the "downstream" viewpoint, but not so far from the aforementioned Bayesian approach, another stream of literature focuses on the optimization program objective function and constraint specifications. Often termed robust portfolio theory, it proposes to deal with the upstream input estimator's lack of precision or noise by an extreme value of the portfolio variance minimization given a preset uncertainty around inputs, which could take the shape of percentile based intervals (Tutuncu and Koenig [2004]) or ellipsoidal sets (Goldfarb and Iyengar [2003]). Close to this concept, another determining contribution has been to irrevocably admit the presence of noise within the inputs and as a consequence to constrain voluntarily and pragmatically the portfolio weights in order to limit the uncertainty hanging over the portfolio risk exposure, see Frost and Savarino [1988]; Jagannathan and Ma [2003] and more recently DeMiguel et al. [2009a]. Interestingly enough, it has been demonstrated that these last two methods can be reformulated using Bayesian shrinkage of the covariance matrix, see Scherer [2007]; DeMiguel et al. [2009a]. It is clear to see that whatever the angle considered, both noise-reduction alternatives are closely related.

In this context, the contribution to modern portfolio theory literature presented in this thesis contains multiple aspects:

- A novel "upstream" model is proposed for the portfolio optimization inputs at the crossroads of the time series and multi-factor models.
- Considering a conditional mean and covariance regression model the heteroskedastic component of the covariance is expressed as a function of a set of economically relevant and known factors, which are also potentially intervening in the drift dynamic.
- Furthermore, the influence of heteroskedasticity within the covariance matrix upon the efficient frontier and the optimal mean-variance portfolio weights is demonstrated.
- A stress testing framework is developed based on the GFM model to assess the most influential factors in the portfolio allocation and hence the resulting performance.

Whilst there are vast individual literatures on portfolio optimisation and the currency carry trade, there are relatively few papers addressing the challenge of portfolio optimisation in currency carry trades. Barroso and Santa-Clara [2015] test the relevance of technical and fundamental variables in forming currency portfolios. In addition to the carry strategy the authors combine momentum and reversal in order to optimise portfolio performance. The resulting optimal portfolio is found to outperform the carry trade and other naive benchmarks in an extensive 16 year out-of-sample test. Its returns are not explained by risk and are valuable to diversified investors holding stocks and bonds. Exposure to currencies increases the Sharpe ratio of diversified portfolios by 0.5 on average, while reducing crash risk. Furthermore, the authors argue that currency returns are an anomaly which is gradually being corrected as hedge fund capital increases.

Daniel et al. [2014] examine carry trade returns formed from the G10 currencies. The authors find that performance attributes depend on the base currency and that dynamically spread-weighting and risk-rebalancing positions improves performance. it is demonstrated that equity, bond, FX, volatility, and downside equity risks cannot explain profitability. Dollar-neutral carry trades are shown to exhibit insignificant abnormal returns, while the dollar exposure part of the carry trade earns significant abnormal returns with little skewness. Downside equity market betas of the constructed carry trades are found to be not significantly different from unconditional betas, while hedging with options reduces but does not eliminate abnormal returns. Furthermore, the distributions of drawdowns and maximum losses from daily data indicate the importance of time-varying autocorrelation in determining the negative skewness of longer horizon returns.

Ackermann et al. [2016] demonstrate that the key difference between the currency markets setting and the equity markets setting is that in currency markets interest rates provide a predictor of future returns that is free of estimation error, which permits the application of mean-variance analysis. The authors show that over the last 26 years, a mean-variance efficient portfolio constructed in this fashion has a Sharpe ratio of 0.91, versus only 0.15 for the equally weighted portfolio.

11.2 Markowitz Mean-Variance Approach

The traditional mean-variance optimization approach proposed by Markowitz [1952] is a framework for portfolio allocation that maximises the expected return for a given level of risk, defined as variance. The key insight in this seminal paper is that an asset's risk and return should not be assessed individually, but by how it contributes to a portfolio's overall risk and return. Thus an investor can reduce their portfolio risk simply by holding combinations of instruments that are not perfectly positively correlated.

The general closed-form Markowitz framework for calculating the optimal portfolio weights in the unconstrained case, i.e. when weights $\boldsymbol{w}$ are allowed to be negative, is presented below. The idea here is that an investor specifies a target level of portfolio volatility and then calculates the asset weights so as to achieve the maximum level of portfolio return, i.e. an efficient portfolio. Markowitz [1952] showed that there is an equivalent dual problem in which an investor specifies a target level of portfolio return and then calculates the asset weights so as to achieve the minimum level of portfolio volatility (standard deviation). In practice, in order to find efficient portfolios of risky assets the dual problem in equation 11–1 is most often solved due to computational efficiency.

Definition 58. ***Markowitz Mean-Variance Optimisation***

An investor seeks to solve the following unconstrained optimisation problem:

$$\min_{\boldsymbol{w}} \sigma^2_{p,\boldsymbol{w}} = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\Sigma} \boldsymbol{w} \quad s.t. \tag{11–1}$$

$$\mu_p = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\mu} = \mu_{p,0} \quad and \quad \boldsymbol{w}^{\boldsymbol{T}} \mathbf{1} = 1 \ .$$

where $\boldsymbol{w}$ are the weights of the assets in the portfolio, $\boldsymbol{\Sigma}$ is the associated covariance matrix, $\boldsymbol{\mu}$ is the mean returns vector, μ_p is the portfolio return, $\sigma^2_{p,\boldsymbol{w}}$ is the portfolio variance, and $\mathbf{1}$ is a vector of ones.

To solve the unconstrained minimization problem equation 11–1, first it is necessary to form the Lagrangian function

$$L(w, \lambda_1, \lambda_2) = \mathbf{w^T \Sigma w} + \lambda_1(\mathbf{w^T}\mu - \mu_{p,0}) + \lambda_2(\mathbf{w^T 1} - 1) \ . \tag{11–2}$$

The first order conditions (FOCs) for a minimum are thus the linear equations

$$\frac{\partial L(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \mathbf{w}} = 2\mathbf{\Sigma w} + \lambda_1 \boldsymbol{\mu} + \lambda_2 \mathbf{1} = \mathbf{0} \ , \tag{11–3}$$

$$\frac{\partial L(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \lambda_{\mathbf{1}}} = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\mu} - \mu_{p,0} = 0 \ , \tag{11–4}$$

$$\frac{\partial L(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \boldsymbol{\lambda_2}} = \boldsymbol{w^T 1} - 1 = 0 \; . \tag{11–5}$$

The system of linear equations can be represented using matrix algebra as

$$\begin{pmatrix} 2\boldsymbol{\Sigma} & \boldsymbol{\mu} & \boldsymbol{1} \\ \boldsymbol{\mu}^T & 0 & 0 \\ \boldsymbol{1}^T & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \mu_{p,0} \\ 1 \end{pmatrix} \; ,$$

or

$$\boldsymbol{A} \boldsymbol{z_w} = \boldsymbol{b_0} \; , \tag{11–6}$$

where

$$\boldsymbol{A} = \begin{pmatrix} 2\boldsymbol{\Sigma} & \boldsymbol{\mu} & \boldsymbol{1} \\ \boldsymbol{\mu}^T & 0 & 0 \\ \boldsymbol{1}^T & 0 & 0 \end{pmatrix} \; , \; \boldsymbol{z_w} = \begin{pmatrix} \boldsymbol{w} \\ \lambda_1 \\ \lambda_2 \end{pmatrix} \; \text{and} \; \boldsymbol{b_0} = \begin{pmatrix} \boldsymbol{0} \\ \mu_{p,0} \\ 1 \end{pmatrix} \; . \tag{11–7}$$

The solution for $\boldsymbol{z_w}$ is then

$$\boldsymbol{z_w} = \boldsymbol{A}^{-1} \boldsymbol{b_0} \; . \tag{11–8}$$

Note that the first d elements of $\boldsymbol{z_w}$ are the optimal portfolio weights $\boldsymbol{w} = (w_1, \dots, w_d)$ for the minimum variance portfolio with expected return $\mu_{p,w} = \mu_{p,0}$. If $\mu_{p,0}$ is greater than or equal to the expected return on the global minimum variance portfolio then $\boldsymbol{w}$ is an efficient portfolio.

Remark 11.2.1 (Long basket and short basket constraint)**.** *In the currency carry trade portfolio examples studied in Chapter 12 of this thesis it will be necessary to constrain the weights of the currencies to be positive in the long basket and negative in the short basket in order to enforce the long/short nature of the strategy. For this constrained case, there is no closed form solution available, however a quadratic programming approach can be utilised. See Boyd and Vandenberghe [2004]; Palomar and Eldar [2010] for detailed references.*

One particular choice of portfolio weights that is of key interest is the portfolio with the smallest possible volatility, i.e. the global minimum variance portfolio.

Definition 59. ***Global Minimum Variance Portfolio***

The unconstrained Global Minimum Variance (GMV) portfolio satisfies the following minimisation problem:

$$\min_{\boldsymbol{w}} \sigma^2_{p,\boldsymbol{w}} = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\Sigma} \boldsymbol{w} \;\; s.t. \;\; \boldsymbol{w}^{\boldsymbol{T}} \mathbf{1} = 1 \; . \tag{11–9}$$

which has the solution

$$\boldsymbol{w}^{\star}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}} \tag{11–10}$$

where $\boldsymbol{w}$ are the weights of the assets in the portfolio, $\boldsymbol{\Sigma}$ is the associated covariance matrix, and $\sigma^2_{p,\boldsymbol{w}}$ is the portfolio variance.

However, if the weights are constrained to be positive ($\boldsymbol{w} > \mathbf{0}$) then there no longer exists a closed form solution. A simple quadratic programming approach can be applied in this case.

The GMV portfolio does not take the expected returns of the assets into consideration, but seeks to minimise portfolio volatility and as such provides an anchor at the leftmost end point of the efficient frontier.

Definition 60. ***Efficient Portfolio Frontier***

The efficient portfolio frontier is a graph of μ_p versus σ_p values for the set of efficient portfolios generated by solving equation 11–1 for all possible target expected return levels $\mu_{p,0}$ above the expected return on the global minimum variance portfolio. This is equivalent to solving the following minimisation problem, where $q \geq 0$ is a "risk tolerance" parameter ($q = 0$ results in the Global Minimum Variance (GMV) portfolio):

$$\min_{\boldsymbol{w}} \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\Sigma} \boldsymbol{w} - q \times \boldsymbol{\mu}^T \boldsymbol{w} \;\; s.t. \tag{11–11}$$

$$\mu_p = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\mu} = \mu_{p,0} \; and \;\; \boldsymbol{w}^{\boldsymbol{T}} \mathbf{1} = 1 \; .$$

where $\boldsymbol{w}$ are the weights of the assets in the portfolio, $\boldsymbol{\Sigma}$ is the associated covariance matrix, $\boldsymbol{\mu}$ is the mean returns vector, μ_p is the portfolio return, $\sigma^2_{p,\boldsymbol{w}}$ is the portfolio *variance.*

The efficient frontier therefore provides an infinite number of potential portfolios, all of which are optimal in the sense that there exists no other portfolio with a higher expected return for equal or less risk (here defined as portfolio volatility). Among this collection of available portfolios it will be interesting to focus on a number of particular points on the frontier in addition to the GMV portfolio. Firstly, the Markowitz portfolio with the maximum Sharpe ratio is of key importance in traditional portfolio optimisation.

Definition 61. ***Sharpe Ratio***

The Sharpe ratio is defined as the portfolio mean divided by the portfolio volatility, i.e.

$$\textit{Sharpe Ratio} = \frac{\mu_p}{\sigma_p} \tag{11–12}$$

where μ_p is the portfolio mean, and σ_p is the portfolio volatility.

Definition 62. ***Maximum Sharpe Ratio (MSR) Portfolio***

The Maximum Sharpe Portfolio (Tangency Portfolio) is a portfolio on the efficient frontier at the point where the line drawn from the point (0, risk-free rate) is tangent to the efficient frontier. The risk-free rate is the rate of return earned on an asset assumed to have zero risk, e.g. a short dated treasury bill.

$$\textit{Max Sharpe Ratio} = \max_{\boldsymbol{w}} \frac{\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\mu}}{\sqrt{\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Sigma}\boldsymbol{w}}} \tag{11–13}$$

which in the unconstrained case has the solution

$$\boldsymbol{w}^{\star}_{MSR} = \frac{\boldsymbol{\Sigma}\boldsymbol{\mu}^{-1}}{\mathbf{1}^T\boldsymbol{\Sigma}\boldsymbol{\mu}^{-1}} \tag{11–14}$$

Remark 11.2.2. *In the general case, finding the Maximum Sharpe Portfolio requires a non-linear solver since the Sharpe Ratio is a non-linear function of $\boldsymbol{w}$. However, as long as all constraints are homogeneous of degree 0, i.e. if $\boldsymbol{w}$ is multiplied by a number the constraint is unchanged, a quadratic solver can be used to find the Maximum Sharpe Portfolio weights.*

Secondly, it will also be interesting to explore the concept of the most diversified portfolio, introduced by Choueifaty and Coignard [2008] as a risk based approach, but here I will further restrict the optimisation to the set of portfolios on the efficient frontier. Choueifaty et al. [2013] provides an interesting discussion of the properties of the diversification ratio and most diversified portfolio and indeed the intuition behind it.

Definition 63. ***Diversification Ratio***

The diversification ratio is the ratio of the weighted average of the asset volatilities divided by the portfolio volatility, i.e.

$$\textit{Diversification Ratio} = \frac{\boldsymbol{w^T diag(\Sigma)}}{\sqrt{\boldsymbol{w^T \Sigma w}}} \tag{11–15}$$

Definition 64. ***Most Diversified Efficient (MDE) Portfolio***

The Most Diversified Efficient Portfolio is the portfolio lying on the efficient frontier which maximises the diversification ratio, i.e. the MDE portfolio is a constrained version of the well known Most Diversified Portfolio (MDP).

$$\boldsymbol{w}^{\star}_{MDE} = \underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w^T diag(\Sigma)}}{\sqrt{\boldsymbol{w^T \Sigma w}}} \;\; \textit{s.t.} \;\; \boldsymbol{w}^{\star}_{MDE} \textit{ is efficient.} \tag{11–16}$$

The solution to this optimisation problem can be found by numerically evaluating the diversification ratio over a grid of portfolios along the efficient fronter, and selecting the portfolio with the maximum diversification ratio.

An example plot of an efficient frontier can be seen in Figure 11.1. The trade-off between risk and return can be observed in the various portfolios plotted. At the leftmost point on the efficient frontier one can observe the constrained GMV portfolio, whereas on the rightmost end point one can observe the portfolio with 100% allocation to Asset 4. Furthermore, the asset weights for the three example portfolio allocation methods can be seen in Figure 11.2. It can be seen here that only the constrained MDE portfolio actually allocates some weight to each of the assets, i.e. the constrained GMV portfolio and the constrained Markowitz Maximum Sharpe portfolio concentrates the allocation into only three assets.

Remark 11.2.3. *In practice, investors can utilise the pragmatic approach of requiring a minimum holding in each of the assets considered for the portfolio.*

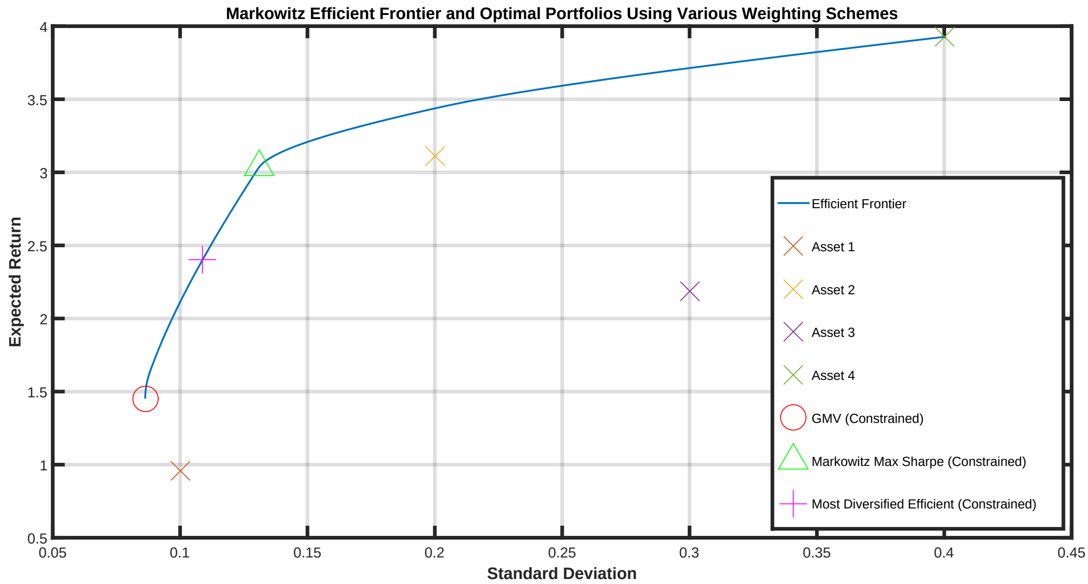


Figure 11.1: Illustrative example: efficient frontier and some key Markowitz portfolios.

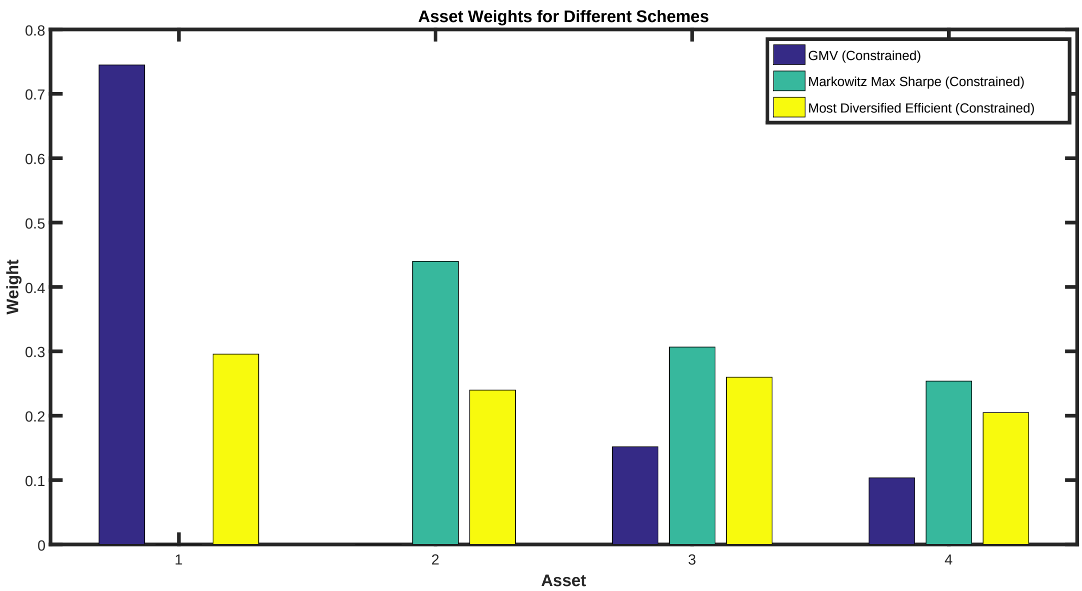


Figure 11.2: Illustrative example: bar plot of asset weights for some key Markowitz portfolios.

This can alleviate problems with portfolios becoming too concentrated in a small number of particular assets.

11.3 Risk Based Approaches

It is known that the mean-variance portfolio optimisation approach can be highly sensitive to the input parameters, and in particular to the expected returns, see Merton [1980]. Therefore, risk-based techniques have arisen as an alternative, see Roncalli [2013]. In focusing on just the risk of the portfolio the investor is admitting that he has no useful knowledge of expected returns, and thus effectively assuming that all potential assets under consideration have equal expected returns. Of course it is possible to go further in this direction and assume no knowledge of asset covariances. In this case an investor can use equal asset weights. Indeed DeMiguel et al. [2009b] present empirical evidence to suggest that the gain from optimal diversification is more than offset by estimation error when considering various weighting schemes versus the equal weight approach.

Definition 65. ***Equal Weight (EW) Portfolio***
The equal weight portfolio is simply defined as:

$$w^{\star}_{EW} = \frac{1}{N} \tag{11–17}$$

where N is the number of assets in the portfolio.

Arguably the next simplest approach to asset weighting is achieved by weighting each asset based on its relative volatility, thus it is assumed each pair of assets has an equal correlation. Here, I refer to this as naïve risk parity, see Anderson et al. [2012]; Asness et al. [2012].

Definition 66. ***Naïve Risk Parity (NRP) Portfolio***
The naïve risk parity portfolio is simply defined as:

$$w^{\star}_{NRP}(i) = \frac{1}{\sigma_i \sum_{j=1}^{N} \frac{1}{\sigma_j}} \tag{11–18}$$

where N is the number of assets in the portfolio and σ_i is the volatility of asset i.

A key contribution to the risk-based portfolio literature is provided by the Equal Risk Contribution (ERC) approach of Maillard et al. [2010]. The aim of this approach is to find the portfolio in which each asset contributes equally to the total portfolio volatility.

Definition 67. ***Equal Risk Contribution (ERC) Portfolio***

The optimal asset weights are defined as:

$$\boldsymbol{w}^{\star}_{\boldsymbol{ERC}} = \{\boldsymbol{w} \in [0,1]^N : \sum w_i = 1, w_i \times \partial_{w_i}\sigma(\boldsymbol{w}) = w_j \times \partial_{w_j}\sigma(\boldsymbol{w}) \ \forall i,j\} \quad (11\text{–}19)$$

where $\partial_{w_i}\sigma(\boldsymbol{w}) = \frac{\partial\sigma(\boldsymbol{w})}{\partial w_i}$ *are the marginal risk contributions, i.e. the impact of an infinitesimal increase in an asset's weight on the total portfolio volatility, and* $\sigma(\boldsymbol{w}) = \sqrt{\boldsymbol{w}^T\Sigma\boldsymbol{w}}$ *is the portfolio variance.*

Note that $\partial_{w_i}\sigma(\boldsymbol{w}) \propto (\Sigma\boldsymbol{w})_i$, where $(\Sigma\boldsymbol{w})_i$ denotes the i-th row of the vector issued from the product of $\boldsymbol{\Sigma}$ with $\boldsymbol{w}$. This reduces the optimisation problem to the following:

$$\boldsymbol{w}^{\star}_{\boldsymbol{ERC}} = \{\boldsymbol{w} \in [0,1]^N : \sum w_i = 1, w_i \times (\Sigma\boldsymbol{w})_i = w_j \times (\Sigma\boldsymbol{w})_j \ \forall i,j\} \quad (11\text{–}20)$$

In order to find the optimal weights a Sequential Quadratic Programming (SQP) algorithm can be applied to solve the following problem (see details in Maillard et al. [2010] or Chaves et al. [2012] for alternative algorithms):

$$\boldsymbol{w}^{\star}_{\boldsymbol{ERC}} = \text{argmin}\, f(\boldsymbol{w}) \ \ s.t. \ \ \boldsymbol{1}^T\boldsymbol{w} = 1 \ \text{ and } \ \boldsymbol{0} \geq \boldsymbol{w} \leq \boldsymbol{1}. \quad (11\text{–}21)$$

where

$$f(\boldsymbol{w}) = \sum_{i=1}^{N}\sum_{j=1}^{N} (w_i(\Sigma\boldsymbol{w})_i) - w_j(\Sigma\boldsymbol{w})_j)^2 \quad (11\text{–}22)$$

Note that the existence of the ERC portfolio is ensured only when the condition $f(\boldsymbol{w}^{\star}) = 0$ is verified, i.e. $w_i(\Sigma\boldsymbol{w})_i - w_j(\Sigma\boldsymbol{w})_j$ for all i, j. Thus, it can be seen that the program minimises the variance of the rescaled risk contributions. For alternative solutions see Lee [2011]; Kaya and Lee [2012]; Baltas and Kosowski [2015].

A more general approach to risk-based portfolio construction can be seen in

Jurczenko et al. [2015]. The authors show that all risk-based approaches can be mapped on a plane defined with only two parameters. The first parameter is a regularization parameter which implies differences in sensitivity to covariance estimates. Thus the GMV portfolio, which is highly sensitive to the covariance matrix, and the EW portfolio, which is totally independent from it, represent both extremes of the spectrum. The ERC portfolio lies in between, i.e. it is sensitive to risk parameters but less so than the GMV portfolio.

The second parameter gives the tolerance for individual total risks. The GMV portfolio, which is the most averse to risk, is at one end of the spectrum. At the other end of the spectrum is what the authors term 'the Most Concentrated (MC) portfolio', which is only invested in the riskiest asset. The Most Diversified (MD) portfolio lies in between, being more diversified than the MC portfolio but less focused on individual total risks than the GMV portfolio. Implicitly, the correlation is the definitive input for the MD portfolio.

Definition 68. *Generalized Risk Based Portfolio*

The optimal weights for any risk-based portfolio can thus be found by solving the following minimisation problem with the two parameters γ and δ:

$$\boldsymbol{w}^{\star} = \operatorname{argmin} D(f(w_i;\gamma,\delta)) \ \ s.t. \ \ \mathbf{1}^T\boldsymbol{w} = 1. \tag{11–23}$$

where

$$f(w_i;\gamma,\delta) = \frac{w_i^{\gamma}}{\sigma_i^{\delta}} \times MRC_i \tag{11–24}$$

where $\gamma \geq 0$ and $\delta \geq 0$ are the two key parameters that define the plane of all risk based portfolios. $D(.)$ is a dispersion matrix and MRC_i is the marginal risk contribution of asset i.

This general optimisation problem can be solved using a number of methods with various strengths and weaknesses, for example via a Newton-Raphson algorithm (see Maillard et al. [2010]).

An interesting approach is presented in Stefanovits et al. [2014] where the authors introduce a measure of model risk and then show that under the assumption of known constraints and unbiased estimators, optimal portfolios are on average negatively affected by model risk. The analytical results in the paper show that mean-variance optimization is seriously compromised by model uncertainty, in particular, for non-Gaussian data and small sample sizes. In order to mitigate these shortcomings, a method is proposed in which the sample covariance matrix is adjusted in order to reduce model risk.

11.4 Portfolio Weights Sensitivity to Factors

In this section, I provide expressions for the sensitivity of the conditional covariance and the optimal Markowitz portfolio weights to the explanatory exogenous factors that make up the filtration $\mathcal{G}_t$ introduced in Chapter 7. These sensitivity results can be utilised to study the stress testing of the portfolio to variations in the factors. This gives an indication of the robustness of portfolio performance to variations in the driving factors and also an indication of the influence that such model based reactivity will have on risk based performance.

11.4.1 Conditional Covariance Sensitivity to Covariates

To begin, the sensitivity of the conditional covariance of the portfolio under the GFM model to each factor in the model is derived as follows:

$$\begin{aligned} \Sigma_{X_t} &= \mathbb{E}[e_t e_t^T | \mathcal{F}_t \cup \mathcal{G}_t] \\ &= \Psi + B X_t X_t^T B^T \ . \end{aligned} \tag{11–25}$$

$$\mathrm{Cov}(e_m, e_n | \mathcal{F}_t \cup \mathcal{G}_t) = \Sigma_{X_t}^{m,n} = \Psi^{m,n} + (B_{m,:} X_t) \times (B_{n,:} X_t) \ , \tag{11–26}$$

where $m = 1, \ldots, d$, $n = 1, \ldots, d$ and d is the number of assets in the portfolio. Differentiating Σ_{X_t} w.r.t covariate $X_{k,t}$ gives:

$$\frac{\partial \Sigma_{X_t}^{m,n}}{\partial X_{k,t}} = (B_{m,k} \times (B_{n,:} X_t)) + (B_{n,k} \times (B_{m,:} X_t)) \ . \tag{11–27}$$

These results can be utilised to study the influence that each factor has on the portfolio allocation and performance in the SFM and GFM frameworks. In addition these results can be used to study the effect of the factors on the forecast covariance performance.

11.4.2 Optimal Markowitz Weights Sensitivity to Covariates

Having obtained the sensitivity of the conditional covariance of the portfolio under the GFM model to each factor in the model, this can now be extended to study the sensitivity of the allocation weights selected for the portfolio to the factors.

$$
\begin{aligned}
\frac{\partial \boldsymbol{z_w}}{\partial X_{k,t}} &= \left(\frac{\partial \boldsymbol{A}^{-1}}{\partial X_{k,t}} \times \boldsymbol{b_0}\right) \quad + \quad \underbrace{\left(\frac{\partial \boldsymbol{b_0}}{\partial X_{k,t}} \times \boldsymbol{A}^{-1}\right)}_{=0,\text{ since } \boldsymbol{b_0} \text{ doesn't depend on } X_{k,t}} \\
&= \left(-\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial X_{k,t}}\boldsymbol{A}^{-1}\right)\boldsymbol{b_0} \\
&= \left(-\boldsymbol{A}^{-1}\begin{bmatrix} 2\left((\boldsymbol{B}_{n,:}X_t)\boldsymbol{B}_{m,k} + (\boldsymbol{B}_{m,:}X_t)\boldsymbol{B}_{n,k}\right) & \boldsymbol{\beta}_{:,k} & 0 \\ \boldsymbol{\beta}^T_{:,k} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\boldsymbol{A}^{-1}\right)\boldsymbol{b_0}\,.
\end{aligned}
\tag{11–28}
$$

In the next chapter, the portfolio optimisation techniques presented above will be utilised to analyse the accuracy and the associated performance of the covariance models introduced in Part II.

Chapter 12

Investigating Optimal Currency Portfolios via Generalised Factor Model Covariance Forecasting

In this chapter, the utility of the proposed GFM model in covariance forecasting and portfolio optimisation is explored. In particular, the added performance of incorporating the $SPEC$ factors into the model for the covariance is analysed. The details of the procedure followed for the comparison of the covariance forecasting models is explained and finally the resulting performance of the proposed approach is discussed.

12.1 Covariance Forecasting Accuracy

In order to compare the accuracy of the covariance forecasts following from the heteroskedastic Generalised Multi-Factor model (GFM) with those generated through the Standard Multi-Factor model (SFM) and the DCC model I consider a set of allocation approaches which are unevenly impacted by the input sources of uncertainty. These include: the variability and information content of the conditioning filtrations in the model estimation; the variability of the forecasts of the conditional and unconditional covariance matrices for the portfolio under each model; the sensitivity of the model estimation; and the stress of the model relative to variability in the explanatory factors in the filtration $\mathcal{G}_t$ in each window. I will present different comparative results to study each of these factors both in terms of the allocation and in terms of portfolio performance.

I will first consider two naive methods, which are the equal weighted method studied recently by DeMiguel et al. [2009b] as well as the naïve risk parity approach (Asness et al. [2012]; Anderson et al. [2012]) where it is assumed that the correlations across all pairs of assets are equal such that only the variances should be considered for the risk parity allocation. The former is naturally insensitive to the input mismeasurement while the latter is impacted by poor forecasts of the asset volatilities.

In addition to these I also consider the classical approach of the mean-variance optimization program proposed by Markowitz [1952]. The prerequisites for this method are expected value and covariance assessments. This method is as a consequence highly sensitive to the measurement error affecting the expected value and to a lower extent the covariance forecasts errors (Jobson and Korkie [1981]; Frost and Savarino [1988]; Michaud [1989]; Chopra and Ziemba [1993]; Broadie [1993]; Nawrocki [1996]). I then complete this analysis with several other risk based approaches to portfolio allocation that have been more recently proposed. In this regard I focus on the portfolio allocation methods displaying more or less sensitivity to the covariance measurement error. Among them I implemented the equal risk contribution (ERC) approach proposed by Maillard et al. [2010], the minimum variance portfolio proposed by Haugen and Baker [1991], as well as the maximum diversification devised by Choueifaty and Coignard [2008]. Interestingly, it has been recently pointed out by Jurczenko et al. [2015] that the measurement error on the covariance used as input is particularly influencing the optimal weights calculated through these techniques. This effect is more pronounced for a minimum variance portfolio, which amounts to the Markowitz mean-variance model but with equal expected returns for all the portfolio components, than for the ERC model[1].

[1]Please refer to Jurczenko et al. [2015] for a more detailed review.

12.2 Currency Data and Currency Factors Description

The dataset considered in this chapter is as in Chapter 8 of this thesis. For readability, this is briefly recapped here. I consider two sets of currency baskets typically associated with a currency carry trade strategy. One portfolio consisting of a long basket and a second portfolio consisting of a short basket. The long basket contains four major "investment" currencies, namely United Kingdom (GBP), Australia (AUD), Canada (CAD) and New Zealand (NZD), while the short basket contains three major "funding" currencies, as in Brunnermeier et al. [2008], namely Euro (EUR), Japan (JPY) and Switzerland (CHF). I have considered daily settlement prices for each currency exchange rate as well as the daily settlement price for the associated 1 month forward contract in order to derive the weekly carry trade mark-to-market returns, $\boldsymbol{R}_t$. The daily time series analysed were obtained from Bloomberg and range from 04/01/1999 to 29/01/2014. As I am working on the trading volume based covariance modelling I chose 1st April 1999, i.e. the date of the introduction of the Euro, as the starting date of the sample.

For the explanatory factors in the currency analysis I consider a range of different factors that I motivate in this section from an economic perspective as well as a quantitative perspective. In a similar vein to the famous three stock-market factors and the two bond-market factors proposed by Fama and French [1993] to explain bonds and equities returns, Lustig et al. [2011] propose a factor decomposition of the currencies returns. Such models are built upon one of the cornerstones of financial theory which is the risk premium. These yields implicitly stored within asset returns would thus be received by investors willing to bear the associated sources of risk. Lustig et al. [2011] demonstrate with the help of a principal component analysis that two linearly independent factors could explain most of the variability in the cross section of the international exchange rates. The first factor would correspond to a level factor, named "dollar risk factor" or DOL, which is essentially the average relative value change of a foreign currency basket against the dollar[1]. The second factor embodies the market induced risk premium associated to the currencies with the highest differential of interest rates relative to the others and is accordingly named in the literature the High-Minus-Low risk factor or HML_{FX}.

12.2.1 Data Preparation

In order to perform the empirical analyses considered in this chapter a substantial amount of effort and time was invested into collecting, cleaning and preparing the data. In particular, the following key steps were performed:

1. Collect daily currency spot price data: closing price, bid and ask price.
2. Collect daily currency forward price data - at maturities of one week, two weeks, three weeks and 1 month: closing price, bid and ask price.
3. Pre-process the price data to deal with missing data, i.e. if data is missing copy previous day's price.
4. Match one month forward contracts with closing spot price on the correct date of delivery for the contract.
5. Calculate the forward premium (interest rate proxy).
6. Collect currency futures price open interest data: broken down into net commercial (hedgers) and net non-commercial futures positions (speculators).
7. Match open interest rate data to synchronous currency price data.

12.2.2 Covariate SARIMA Forecast Results

Before analysing the covariance forecasts it is important to examine the accuracy of the SARIMA forecasting models for the individual covariates. Here, I utilise the MASE and the MAPE measures, as discussed in Section 7.4.1.3 of Chapter 7.

[1]When an American investor is considered. However it is asserted in the same article that similar results are obtained when we retain the Japanese, British or Swiss investors point of view.

The MASE forecast accuracy results, shown as a time series in Figure 12.1 and as a boxplot summary in Figure 12.2, suggest that on average all of the models constructed from the ARIMA automated fitting procedure described in Section 7.4.1.2 of Chapter 7 behave as expected. The covariates with non-trivial ARIMA model structures produce reasonably accurate forecast performance over the one month forecast horizon, which is required for the applications to carry trade strategies considered in this thesis.

It is important to note that the DOL and HML_{FX} covariates are risk premia and therefore shouldn't be expected to be forecastable, since otherwise there is no risk to be compensated for. Indeed these covariates seem to be generated from models that are close to white noise and hence the naïve in-sample forecasting method in these cases can be very poor. Thus, looking at the boxplot summary of the Mean Absolute Percentage Error (MAPE) forecast accuracy results in Figure 12.3 it can be seen that the DOL and HML_{FX} have median MAPEs of 100%, i.e. often these covariates are forecasted as zero.

If instead of just considering DOL and HML_{FX}, additional covariates are considered based on the volatility in these factors and there covariance, then the fitted models for the factors σ_{DOL}, σ_{HML} and $\text{Cov}_{DOL,HML}$ demonstrate a much more accurate forecast performance: having median MAPEs of 11%, 12% and 21% respectively. The accuracy of the forecast performance in these covariates is even more accurate outside the period of poor forecast performance corresponding to the 2008 Financial Crisis, which is not unexpected. Furthermore, the $SPEC$ and cross $SPEC$ covariates for the low interest rate currencies have median MAPEs of 38%, 56%, 50%, 70%, 60% and 78% respectively. The speculative volume covariates for the high interest rate currencies show similar forecasting accuracy.

An important contribution of the studies in this section is to demonstrate a feature not previously discussed in the literature on carry trade portfolio analysis that will have practical significance in the actual carry trade portfolio construction. In previous studies, as described in Section 8.3.2 of Chapter 7, the factors known as DOL and HML_{FX} were shown to have strong explanatory power of the carry trade portfolio returns when studied from an in-sample analysis via PCA. However, as demonstrated in this section, this has not carried forward to good forecast performance for the models fitted for these DOL and HML_{FX} factors. The

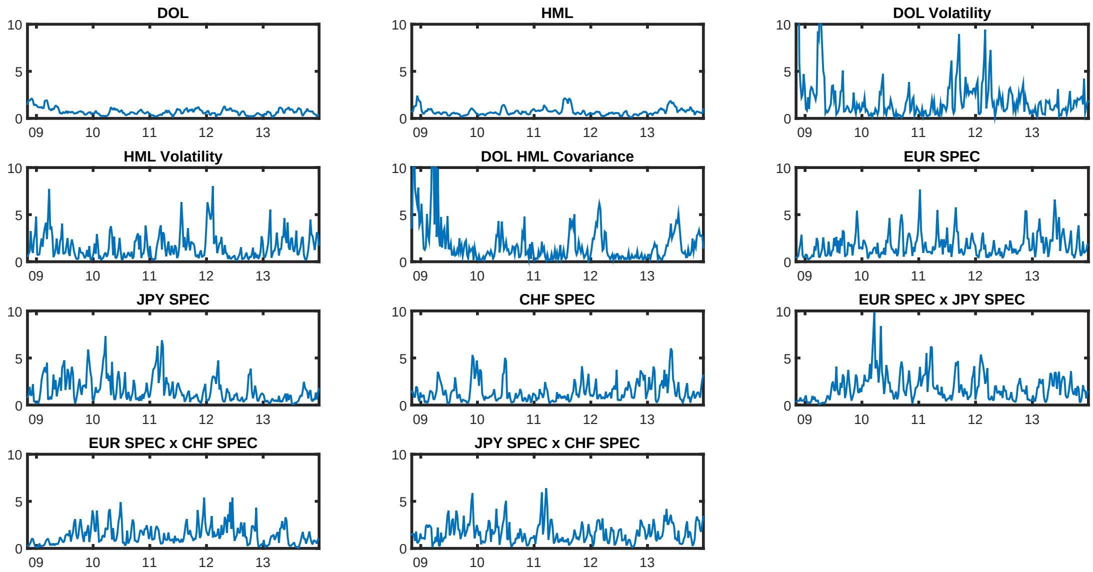


Figure 12.1: Mean Absolute Scaled Errors (MASE) for Low Interest Rate Basket Covariate Forecasts.

reason for this is explained by the fact that the models fit to these factors tend to demonstrate that they behave historically in a similar manner to white-noise which naturally therefore results in high forecast errors under the MAPE criterion.

In fact these findings further strengthen the arguments presented in this thesis that one must include other explanatory factors such as the speculative open interest volume covariates into the currency carry trade portfolio descriptions. These factors were found to have both good in-sample explanatory power in the covariance regression structures as well as good out-of-sample forecast performance under the ARIMA models selected for these factors. This means that such factors can be both significant in interpreting inter-temporal variation in carry returns as well as instrumental in improving covariance model forecasts in the proposed GFM model and therefore may contribute to improving the portfolio performance that results from such a model. I will investigate this second aspect further in the studies contained in the remaining sections.

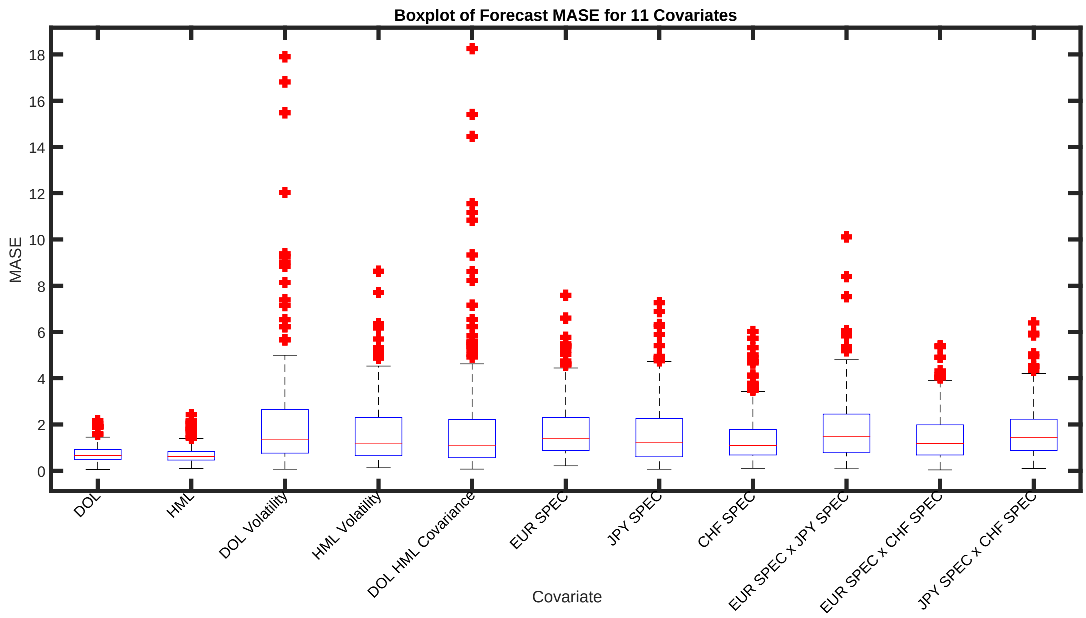


Figure 12.2: Boxplots of Mean Absolute Scaled Errors (MASE) for Low Interest Rate Basket Covariate Forecasts.

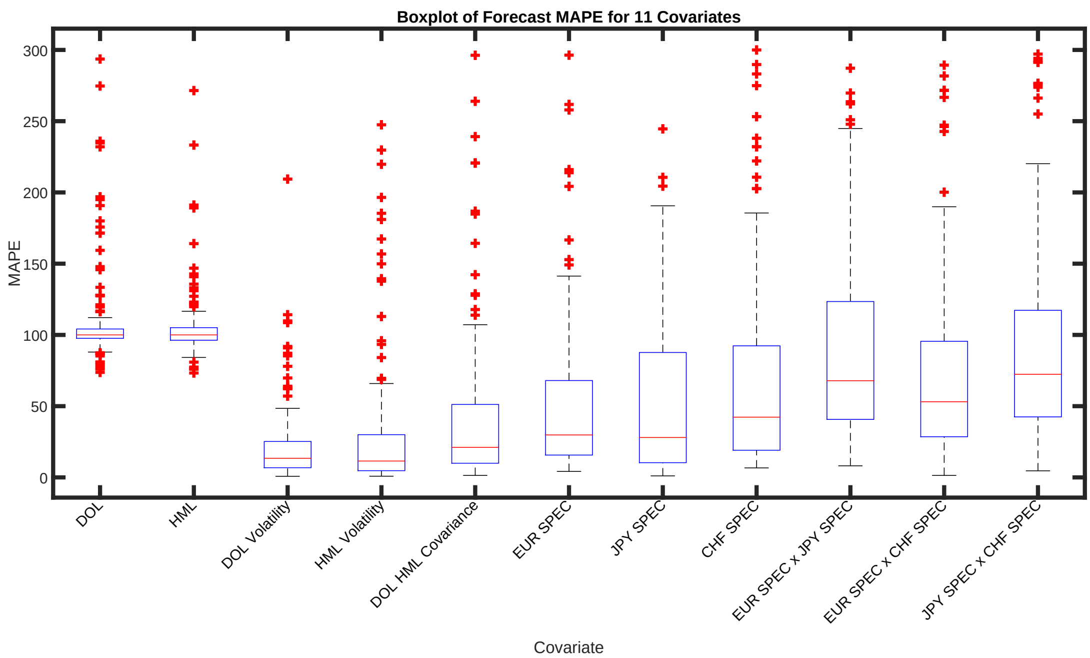


Figure 12.3: Mean Absolute Percentage Errors (MAPE) for Low Interest Rate Basket Covariate Forecasts.

12.2.3 Covariance Dynamics and Forecasting Accuracy

This section aims to study two important aspects of the models that have been described for the portfolio returns. The first is how the stage 1 and stage 2 model forecast covariance structures described in Section 8.3.3 of Chapter 7 for the SFM, GFM and DCC models behave under the different conditional assumptions with respect to the previously defined filtrations $\mathcal{F}_t$, $\mathcal{G}_t$ and $\widetilde{\mathcal{G}}_t$. In particular, I demonstrate that each model's forecast covariance produces significantly different behaviours over time in both the information content captured and more importantly in the reactivity of the covariance model forecasts to inter-temporal variation in the information content contained in the filtrations $\mathcal{F}_t$ and $\widetilde{\mathcal{G}}_t$. The second aspect of this analysis is to assess the downstream portfolio performance of the covariance regression models as a result of the propagation of the forecasts of the covariates/currency factors and the resulting covariance forecasts when used in the portfolio allocation, as described in Section 7.4 of Chapter 7. Performing these studies can be achieved in a number of different ways. The approach presented below is based on a similar type of analysis performed in Engle and Colacito [2006].

The first study highlights the distinctive features and benefits of using the proposed GFM model versus the SFM and DCC models. Demonstration of the differences in the second order modelled information content is achieved through analysis of the forecast covariance matrix. Here, two measures are used: the trace to study the variation and reactivity of each model forecast to marginal volatility fluctuations; and the maximum eigenvalue of the covariance matrix forecasts over time to summarise additional second order covariance structure in off-diagonal dependence structure information content captured by each model and to observe its reactivity over time.

The second set of studies performed considers the accuracy of the forecast covariance models as measured through the portfolio ex-post performances. For sake of comparison between all models, and to remove the influence that the mean prediction of returns plays on the portfolio selection, the global minimum variance (GMV) portfolio allocation framework is considered to undertake the studies in this section. This is largely due to the widely acknowledged fact that forecasting the mean return can be highly challenging, whereas one may expect much better performance when considering the second order information in the volatility and covariance, see discussions on this in Chapter 11.

Furthermore, this second aspect of the study of accuracy of the model forecasts, as measured through the global minimum variance portfolio performances, is based around the type of analysis performed in Engle and Colacito [2006], modified for the context of the models in this thesis. This required the use of a bootstrap procedure, over each sliding window, in order to obtain a time series of estimators of the realized portfolio performance variance (population portfolio volatility). I will denote this time series of estimators as the "ex-post" portfolio volatility that the different models will be trying to achieve with their portfolios constructed from the different covariance forecast structures in a global minimum variance allocation framework. The bootstrap procedure takes 21 days (one trading month) of out-of-sample daily carry returns, selects a random start day uniformly between 1 and 16 and then calculates the one week portfolio volatility from the selected weights of the model and the sums of the next 5 days synchronised daily carry returns for each currency. I draw 1000 bootstrap replicate samples and then calculate the covariance of these bootstrapped weekly portfolio volatilities. The ex-post portfolio volatility obtained is compared to each of the forecasts and resultant global minimum variance portfolios constructed using each of the forecast covariance models for the SFM, GFM and DCC. However, as noted in Chapter 7 there are several variants of these models which contain different sources of conditional information. For instance some versions of these models have information coming from filtrations $\mathcal{F}_t$, $\mathcal{G}_t$ and $\widetilde{\mathcal{G}}_t$, depending on whether they contain factors and whether they are population based estimations such as for the SFM and GFM models in Equations (7–35) and (7–38) respectively, or locally adapted conditional estimations as in the SFM and GFM models in Equations (7–36) and (7–39) respectively. In the context considered here, $\mathcal{F}_t$ contains the currency returns over a lookback period of length T until time $t-1$, $\mathcal{G}_t$ contains the covariate information over a lookback period of length T until time t, and $\widetilde{\mathcal{G}}_t$ contains all of the historical covariate information up until time t.

To interpret the comparison between the "ex-post" portfolio volatility and each of the SFM, GFM and DCC model forecast results a great deal of care is required.

I shall undertake this comparison under the following statistical assumptions: I assume that the population based covariance estimate for the model factors, that are constructed from the filtration $\widetilde{\mathcal{G}}_t$, form an unbiased and consistent estimator of a stationary population based covariance. Furthermore, since the filtration $\widetilde{\mathcal{G}}_t$ is comprised of a time series of length tT whereas the filtration $\mathcal{F}_t$ is of length T for each sliding window in t, I will assume that for comparison purposes the contribution to the unconditional covariance, for the SFM and GFM models in Equations (7–35) and (7–38) respectively is approximately "exact". To be more precise, I assume the convergence rate of the second order moments of $\boldsymbol{X}_t$, which are constructed based on $\widetilde{\mathcal{G}}_t$, are a function of tT and as such, I will assume that as T and t go to infinity, asymptotically only the leading contribution is observed to the portfolio volatility from the SFM and GFM unconditional covariance models, which is arising from local (in the current t-th sliding window) variability due to the filtration $\mathcal{F}_t$. In this sense it is then possible to compare the models for the SFM and GFM, which are based on $\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t$ with the version of the DCC model which is based only on $\mathcal{F}_t$. If this were not the case, the results are still valid but direct comparison between model performance would be less obvious.

An alternative approach would be to extend the bootstrap procedure to also sample multiple realisations of the factors $\boldsymbol{X}_t$ that make up the filtration $\widetilde{\mathcal{G}}_t$. These sampled bootstrap replicates could then be used to numerically average out the variability due to the realisation of the factors attributed to the terms such as $Cov(\boldsymbol{X}_t|\widetilde{\mathcal{G}}_t)$ and $\mathbb{E}(\boldsymbol{X}_t\boldsymbol{X}_t^T|\widetilde{\mathcal{G}}_t)$ in the SFM and GFM models when considering the unconditional covariance, in order to isolate the influence on portfolio volatility attributed to $\mathcal{F}_t$.

It is demonstrated that a key difference among the set of models described earlier lies in the conditioning filtrations considered. Furthermore, another distinguishing feature involves the choice of conditional variance and covariance dynamic considered, which means in this case either heteroskedastic or homoskedastic models in the SFM, GFM and DCC models.

In the following, the differences in reactivity among the set of covariance models under scrutiny is emphasised and it is demonstrated that not only does the conditional dynamic of the dependence structure have a role to play, but in addition the filtration utilised in constructing the portfolio variance also has an important role to play in determining how fast each estimator can adapt to abrupt changes of environment. Therefore, it is interesting to then study whether if a particular model is found to be more reactive to the local environment, as will be shown with a version of the GFM model, does this necessarily translate into better portfolio performance and in what sense?

To this end, I distinguish the reactivity for each model in adjusting the average conditional variances behaviour for the associated marginal distributions and the dependence structures behaviour for the multivariate component. It can be seen in the upper panels of Figures 12.4 and 12.5 that the traces of the covariance matrices resulting from the GFM model are more reactive than those generated by the SFM model or the historical covariance matrix model even though the amplitude of the adjustment stayed restrained with respect to the DCC. It should be noted here that there is a trade-off between being reactive enough to capture changes in covariance and the trading costs associated with increasing portfolio turnover. While the trace embodies the average variability of the matrix diagonal elements, i.e. the vector of asset variances, the relative importance of the first eigenvalue displays on the contrary a higher reactivity and absolute amplitude of adjustment for the GFM model as shown by the lower panels of Figures 12.4 and 12.5.

These two study results lead to the conclusion that the DCC model accompanied by the marginal GARCH dynamics tend to be particularly sensitive to the changes occurring at the marginal volatility level whereas the GFM model is more sensitive to the changes occurring at the asset dependence level. Said differently, the heteroskedasticity seems to be more influential at the covariance level of the GFM model generated covariance matrices while the DCC generated matrices react more significantly to the variance heteroskedasticity component.

As discussed, to further the comparison between the GFM and the DCC models I propose to assess the forecasting accuracy of the two models by comparing the difference between the model based volatility forecast for the next month and the bootstrapped realized volatility of the optimal portfolio over the same period. This graph should indicate the accuracy with which each model anticipates the joint and marginal behaviours of the assets composing the portfolio. As shown in Figures 12.6 and 12.7 the accuracy of the two methods is quite similar and

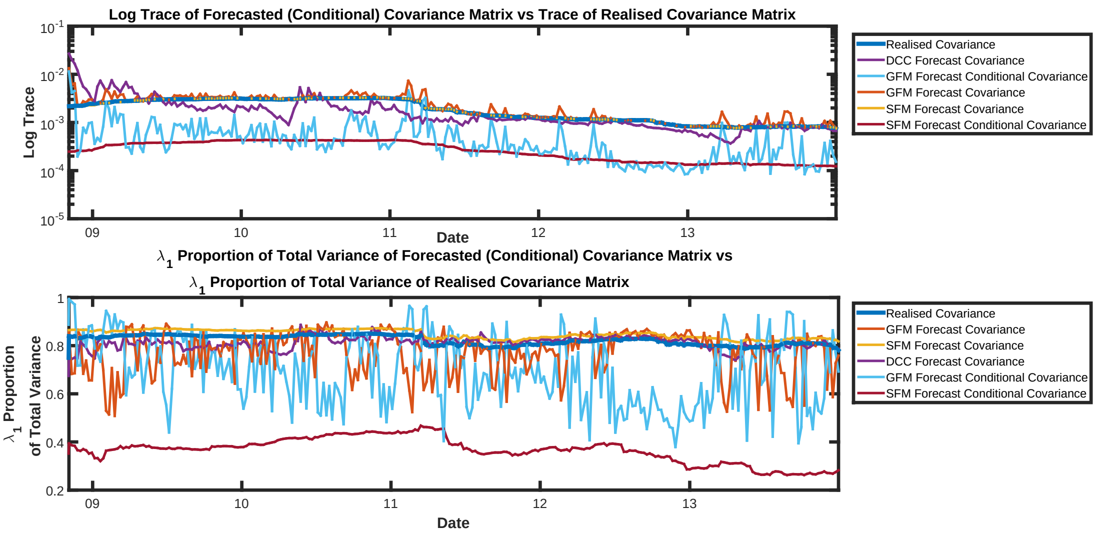


Figure 12.4: High interest rate basket. Upper panel: Trace of covariance matrix.
Lower panel: Proportion of variance explained by first principal component.

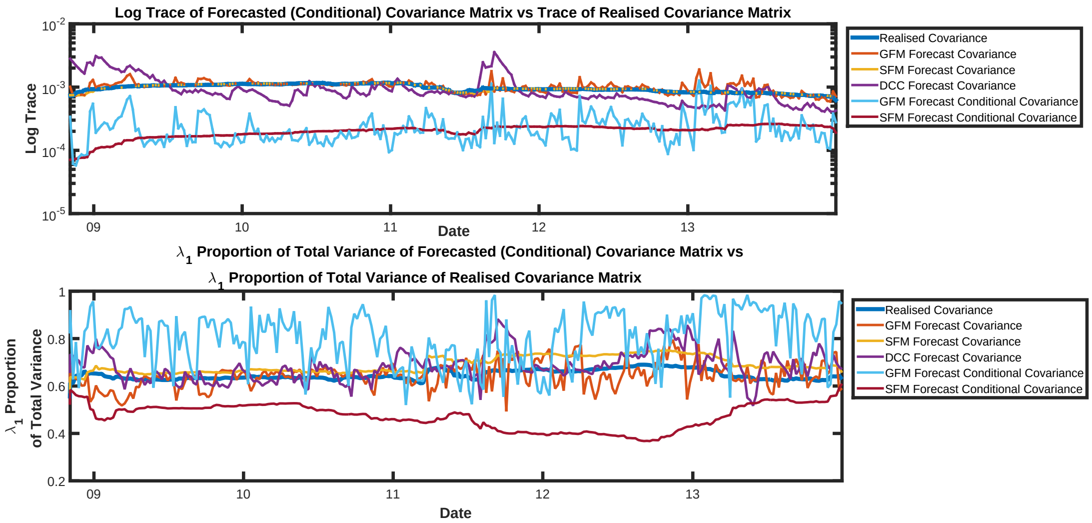


Figure 12.5: Low interest rate basket. Upper panel: Trace of covariance matrix.
Lower panel: Proportion of variance explained by first principal component.

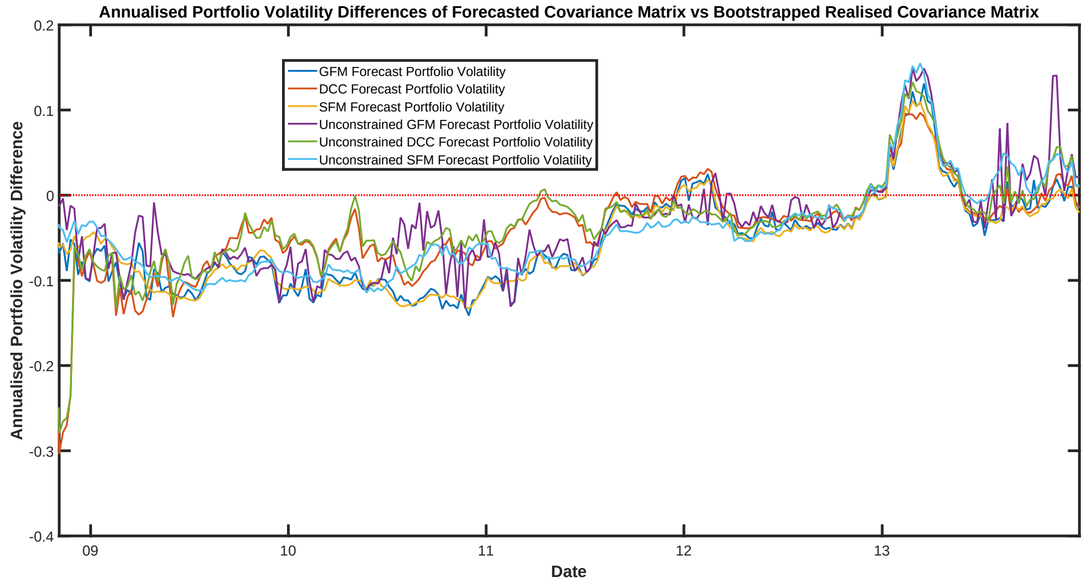


Figure 12.6: High interest rate basket. Annualised portfolio volatility differences between forecast covariance matrix and realised bootstrapped covariance matrix for different covariance forecasting models.

remains within the $+/-15\%$ annualised portfolio volatility bounds. This shows that the GFM and the DCC, while depending on different filtrations and thus leading to different estimator sensitivity to innovations in the data process, still display quite similar accuracy in forecasting the future covariance matrices.

12.3 Portfolio Performance and Conditioning of The Covariance Matrix

In this section, I explore the influence played by portfolio optimisation methods that consider portfolio weight constraints versus those that are unconstrained. It was shown in the innovative paper of Jagannathan and Ma [2003] that such constraints can result in a form of regularization or shrinkage effect implicitly induced on the portfolio variance through the optimization routine and not directly through the stage one or stage two statistical model estimations. This is particularly interesting to consider in the context of the models studied in this thesis for the SFM, GFM and DCC covariance forecast models.

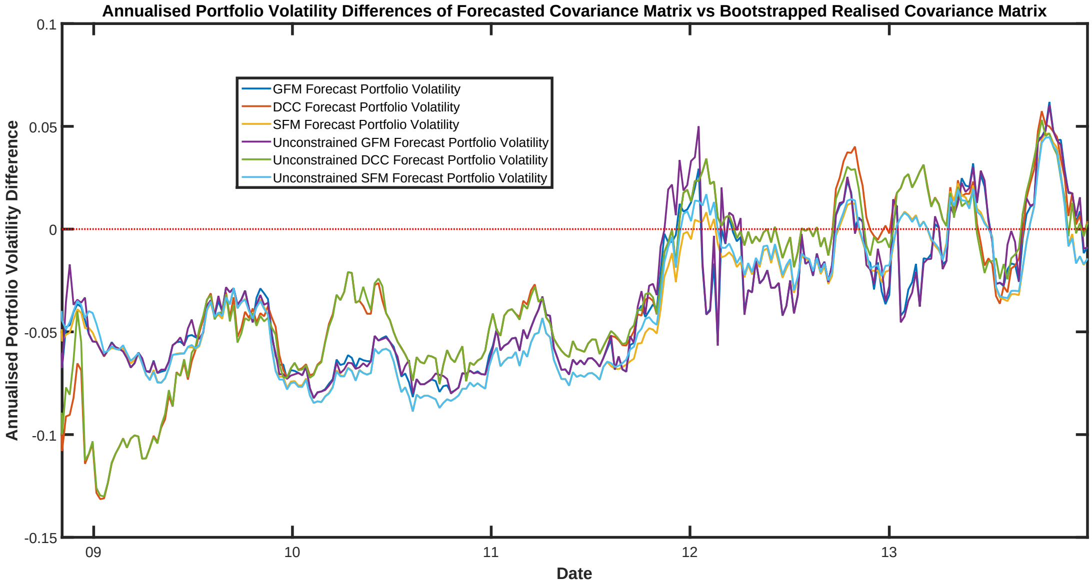


Figure 12.7: Low interest rate basket. Annualised portfolio volatility differences between forecast covariance matrix and realised bootstrapped covariance matrix for different covariance forecasting models.

Therefore, I investigate the consequences of the weight constraints upon the characteristics of the global minimum variance portfolio. As mentioned earlier, the carry trade strategy presumes that an investor is long the high interest rate currencies while financing this position through short positions on the low interest rate currencies. This implies that the weights are constrained to be positive in the high interest rate currencies basket optimization program, whereas the weights are enforced to be negative in the low interest rates currencies basket. As a result of this supposedly slight modification of the global minimum variance optimization program the input covariance is accordingly affected, and more precisely an implicit form of shrinkage occurs on the matrix. For instance, it can be seen that the objective function for the global minimum variance will contain, in the resulting constrained Lagrangian, a form of 'penalty' term given by $(\lambda 1^T - 1\lambda^T)$, where λ corresponds to the Lagrange multipliers column vector for the non-negativity constraints, see details in Jagannathan and Ma [2003]. Furthermore, Jagannathan and Ma [2003] argues that such an ex-post alteration of the input covariance matrix used for the portfolio optimization naturally lowers the contribution of any estimator improvement technique. That is, it regularizes to some extent the resulting contribution one may obtain by trying to improve the model forecast performance in stages one and two of the upstream model improvements. More precisely, it can be shown that the explicit 'penalty' term that results from the weight constraint takes the form of $\lambda_i + \lambda_j$ which acts to reduce the joint covariance between the returns for currency i and currency j.

While Jagannathan and Ma [2003] demonstrate that the ex-post average return and volatility associated to a set of global minimum variance portfolios optimized with various sample estimators of the covariance matrix are almost indistinguishable once the positivity constraint is affixed, the plot of the 12-month rolling Sharpe ratios for the various estimators analysed in this thesis goes in the same direction. Indeed it can be noticed that the risk return profiles associated to the global minimum variance portfolios built on various estimators are barely distinguishable when the positivity constraint is enforced, as shown in Figure 12.8.

Contrary to this, the unconstrained results, plotted in Figure 12.9, show that the differences among estimators are clearly noticeable on a rolling window basis when these constraints are not imposed. It is important to emphasise that as a result of this statement, the carry trade optimal portfolio, being constrained on the sign of the positions for the high and the low interest rates basket, is likely to be largely independent of the covariance estimator choice.

This is true as far as the filtrations $\widetilde{\mathcal{G}}_t$ or $\mathcal{F}_t$ are considered. However, the GFM family of models explored in this thesis also enables the conditioning of the covariance estimator upon a different combination of filtrations such as the union of the data and covariates sample filtrations, $\mathcal{G}_t$ and $\mathcal{F}_t$, according to the equation 7–39 derived earlier. Thus, while the DCC, the GFM unconditional covariance matrix and the SFM unconditional covariance matrix models are based respectively upon the following filtrations $\mathcal{F}_t$, $(\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t)$and $(\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t)$, alternatively the GFM conditional covariance matrix is instead conditioned upon the sample filtration $(\mathcal{F}_t \cup \mathcal{G}_t)$, which respectively contain the currency returns over a lookback period of length T until time $t-1$, and the covariate information over a lookback period of length T until time t. As such this conditional covariance model thus allows different high and low interest rates currency optimal portfolios to be constructed, which will represent in a sense the non-diagonal heteroskedastic

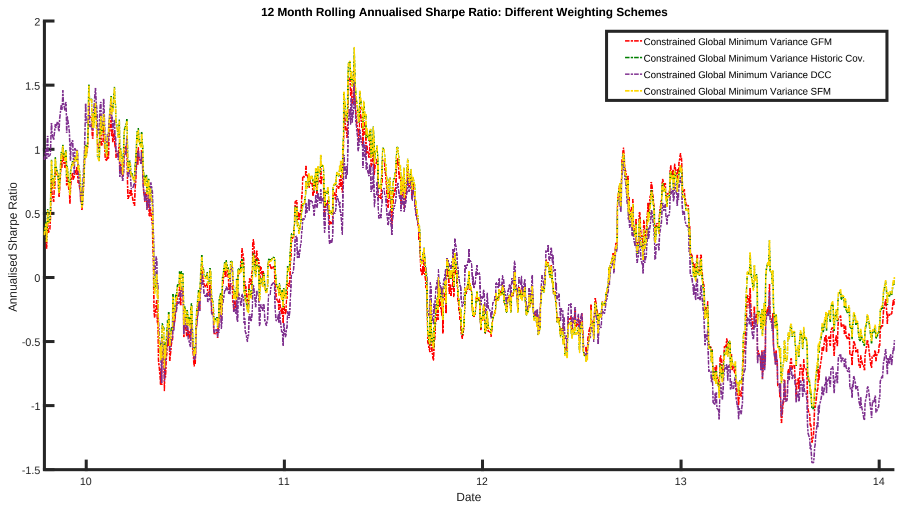


Figure 12.8: High interest rate basket. Constrained GMV 12 month rolling Sharpe ratio comparison.

component of the diffusion and the dependence structure characterised separately from the related high and low interest rates sets of currencies. The conditioning of this covariance on the covariate values over only the most recent lookback period allows for a measure which is more reactive to recent changes in macroeconomic variables. If a comparison is performed between the 12-month rolling Sharpe ratio of the constrained minimum variance portfolio based on the GFM conditional covariance matrix with the unconditional GFM and the conditional DCC models, all of them being conditioned on different filtrations or combinations of filtrations, the former strikingly stands out from the two others.

Figures 12.10 and 12.11 display a substantially different behaviour of the rolling Sharpe ratio for the portfolios based on the GFM conditional covariance matrix even though the non-negativity constraint is affixed to the global minimum variance optimization program. It can also be observed that the combination of the GFM conditional minimum variance optimal long positions on the high interest rates currencies and the GFM conditional minimum variance optimal short positions on the low interest rates currencies basket leads to a noticeable improvement of the strategy Sharpe ratio as demonstrated in Section 12.5.

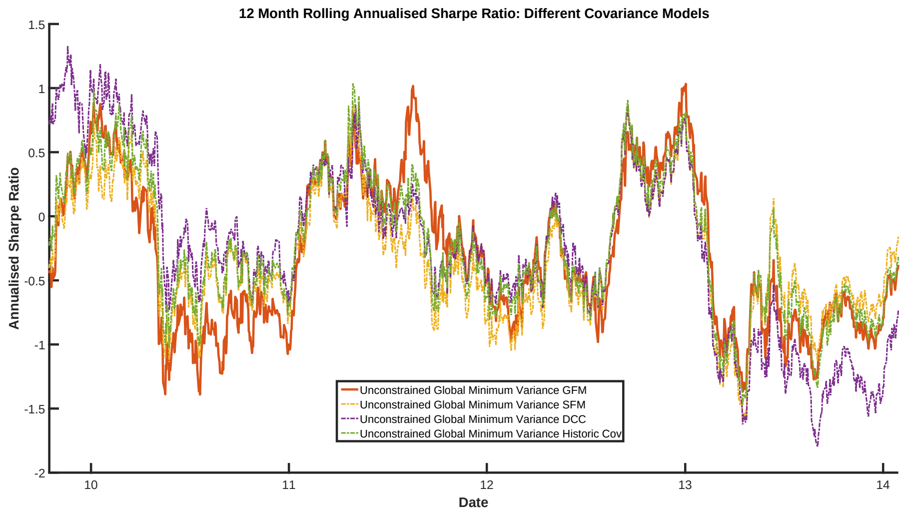


Figure 12.9: High interest rate basket. Unconstrained GMV 12 month rolling Sharpe ratio comparison.

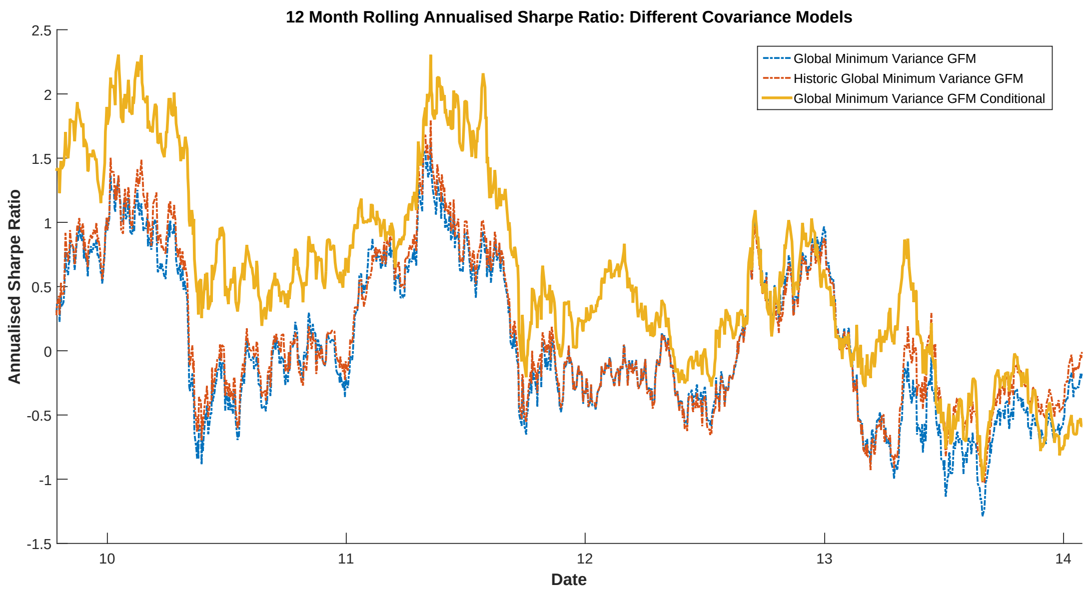


Figure 12.10: High interest rate basket. 12 month annualised rolling Sharpe ratio. Comparison of Conditional GFM and Unconditional GFM.

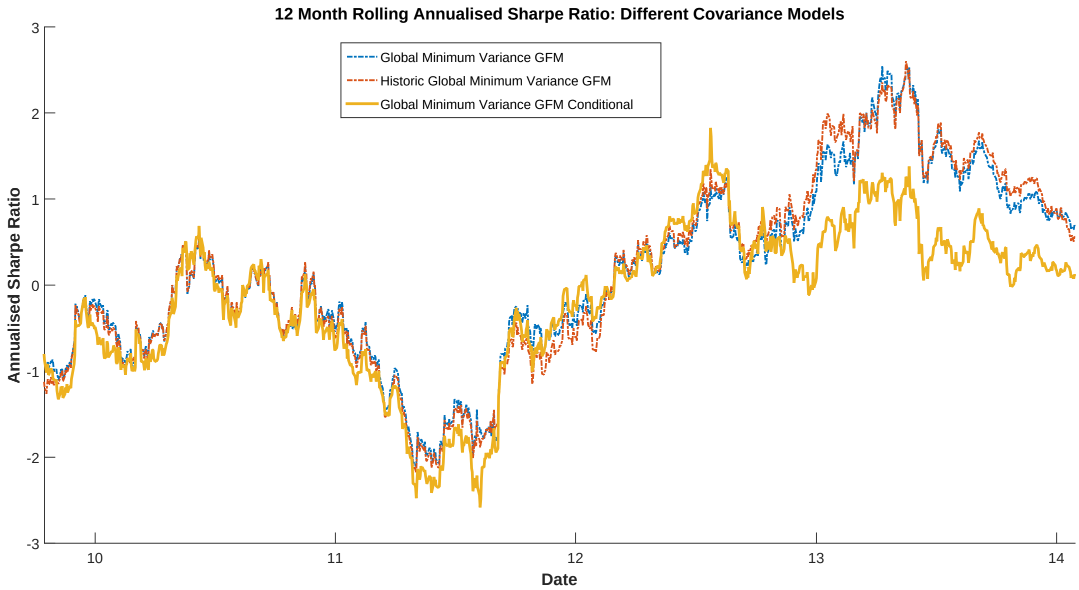


Figure 12.11: Low interest rate basket. 12 month annualised rolling Sharpe ratio. Comparison of Conditional GFM and Unconditional GFM.

12.4 Sensitivity Analysis

In this section, I study the sensitivity of the global minimum variance portfolio obtained using the GFM models based on both the unconditional and conditional covariance models, formed from filtrations $\mathcal{F}_t \cup \widetilde{\mathcal{G}}_t$ and $\mathcal{F}_t \cup \mathcal{G}_t$ respectively. Then to perform this study I systematically vary each individual covariate, one-by-one, from the set of currency factors considered in Section 8.2. The amount of variation considered was to increase and decrease each factor systematically by their inter-quartile ranges, i.e. the quantiles of 25% and 75% respectively. These new perturbed factor values on each day were then fed into the estimated covariance regression model for each sliding window and the global minimum variance portfolio re-estimated. I then summarise the behaviour through portfolio based metrics of the perturbation effect of each covariate. This makes it possible to study which covariates are most influential in driving the portfolio performance and which covariates are likely to result in the largest sensitivity of results. Such an analysis is easily undertaken due to the specific model structure developed in this thesis for the GFM structure where the covariates enter explicitly into the covariance matrix.

This enlightening robustness analysis allows one to estimate a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. The formula in equation 11–27, derived earlier, is thus plugged into the optimization program for various percentile values of each covariate to subsequently determine the effect of a given variation of the independent variables upon the ex-post variance of the GFM unconditional minimum variance portfolio. Figures 12.12 and 12.13 show that some covariate changes can lead to a larger effect on the structure of dependence among assets and their respective marginal features leading accordingly to a large modification of the optimal portfolio volatility. As an example, it can be seen in Figure 12.12 that the GBP speculative open interest has a larger impact on the variance of the global minimum variance portfolio while the uncertainty over the DOL factor has a more limited impact. This limited informative content of the DOL and HML_{FX} factors should be added to the limited forecasting quality highlighted earlier. Furthermore, it can be globally noticed that the global minimum variance portfolio volatility for high interest rates currencies is less sensitive to the price based information represented by the DOL and HML_{FX} factors as well as their respective volatility and the covariance between them. This is in contrast to the speculative volume based data, for which changes lead to larger modifications of the high interest rates global minimum variance portfolio ex-post volatility. This statement demonstrates the interest of understanding and investigating the relation existing between the speculative volumes and the dependence structure among financial assets or at least currency crosses. Another interesting point these two graphs reveal is the asymmetric effect that an increase of certain covariates can have upon the optimal portfolio relative to a decrease of the very same covariates. This phenomenon looms from the inversion of the covariance matrix which will give different results for a positive or a negative modification of a given covariance matrix entry.

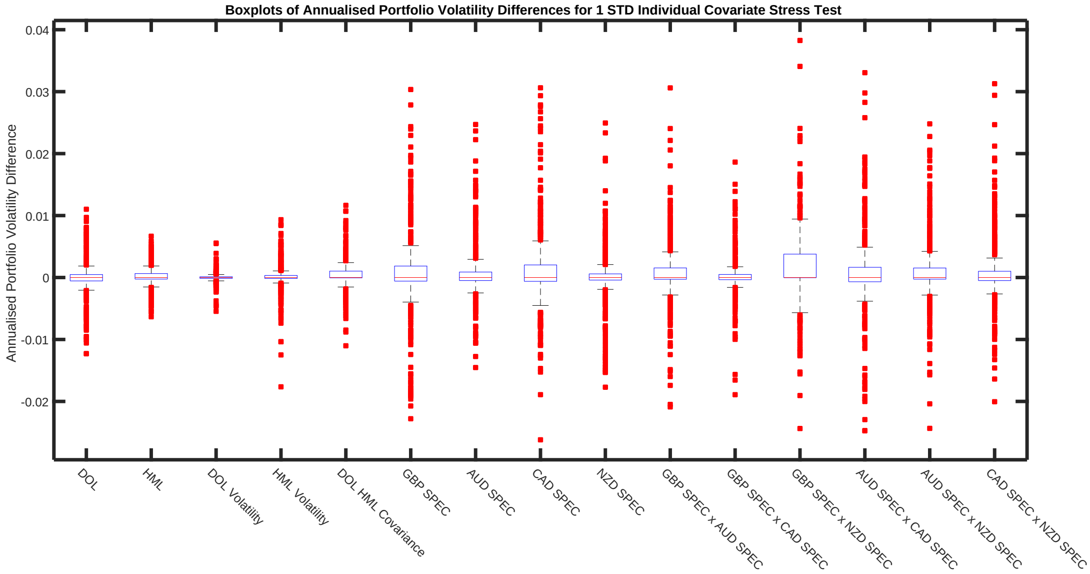


Figure 12.12: High interest rate basket. Boxplot of annualised portfolio volatility differences resulting from one standard deviation individual perturbation of each covariate for GFM model with GMV weights.

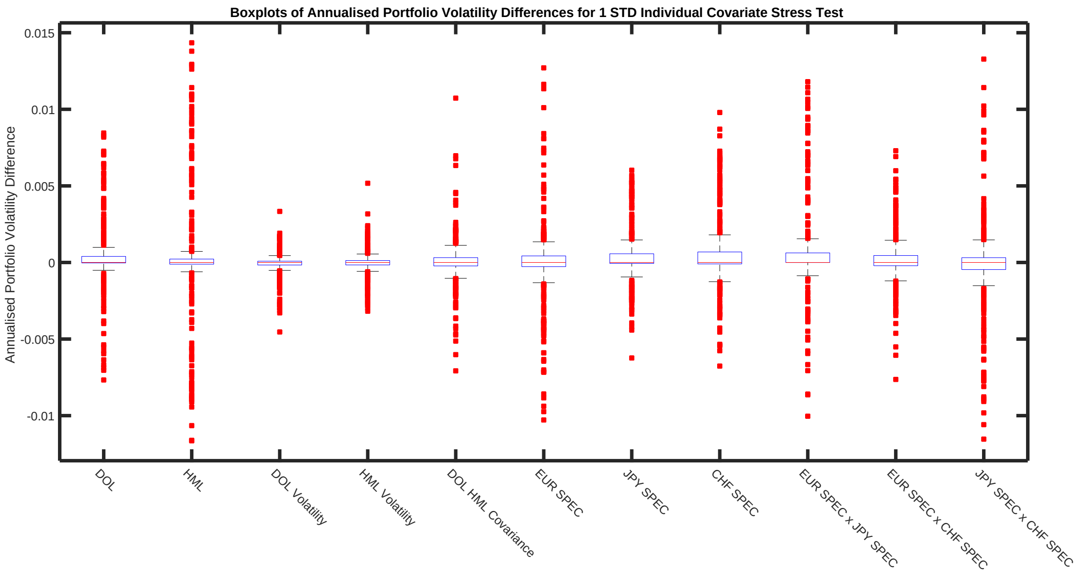


Figure 12.13: Low interest rate basket. Boxplot of annualised portfolio volatility differences resulting from one standard deviation individual perturbation of each covariate for GFM model with GMV weights.

12.5 The Carry Trade Portfolio

This section is devoted to the performance analysis of the combination of the optimal high interest rates currencies basket and the optimal low interest rates currencies basket considering different covariance estimators under the SFM, GFM and DCC models. It is important to stress the fact that this could not to be considered as the optimal carry trade portfolio as I split the optimization into two optimisation subprograms conditionally on different sets of filtrations $\mathcal{G}_t^{\text{high}}$, $\mathcal{F}_t^{\text{high}}$ and $\mathcal{G}_t^{\text{low}}$, $\mathcal{F}_t^{\text{low}}$, associated to the high and the low interest rates basket models respectively. It is worth mentioning that this two-step procedure was motivated by the noticeably different dependence structure behaviours for the high and the low baskets, as demonstrated in previous studies in Ames et al. [2015a,c]. Considering the carry trade portfolio configuration, I focus in this section on the constrained version of the global minimum variance optimiser for the high and the low interest rates currency portfolios. As expected, the non-negativity (equivalently the non-positivity) weight constraint for the high interest rates basket (the low interest rates currencies basket) results in a very similar 12-month rolling Sharpe for the DCC, the GFM unconditional and the SFM unconditional estimator. Nevertheless, in Figures 12.14 and 12.15 the difference of behaviour of the carry trade portfolio optimised using the GFM conditional estimator is observable. With the exception of the second half of 2013 this portfolio has always shown a significantly higher Sharpe ratio on a 12-month rolling basis, which demonstrates the robustness of the improvement. During this period the conditional information contained in the GFM covariance regression models utilising the DOL, HML_{FX} and the associated volatilities and covariance resulted in a markedly different weighting allocation under the GMV approach. Furthermore, Tables 12.1 and 12.2 underpin this argument and shows a noticeable improvement of the Sharpe ratio but without deteriorating the Calmar ratio, which estimates the extreme risk associated (measured by the sample maximum drawdown) to a given strategy relative to its average annualised returns. In addition, the downside volatility is not penalizing the GFM conditional estimator, as demonstrated by the Sortino ratio. Thus, it may be concluded from this analysis that the conditional covariance estimator developed in this thesis under the GFM model family is clearly displaying

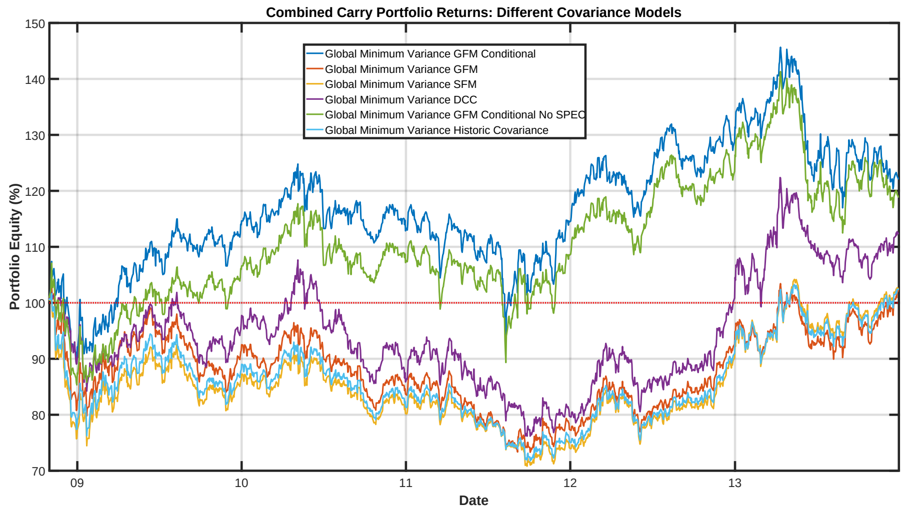


Figure 12.14: Carry trade portfolio performance. I re-optimise the portfolios on a monthly basis using an annual portfolio volatility target of 15% and hence scale the monthly returns according to the expected portfolio volatility for each method. I assume that we initially capitalise the strategy to the value of the unleveraged baskets.

interesting properties: such as its lower sensitivity to the shrinkage effect resulting from the weight constraints and also on the resulting improvements of the global minimum variance portfolio risk and return profile.

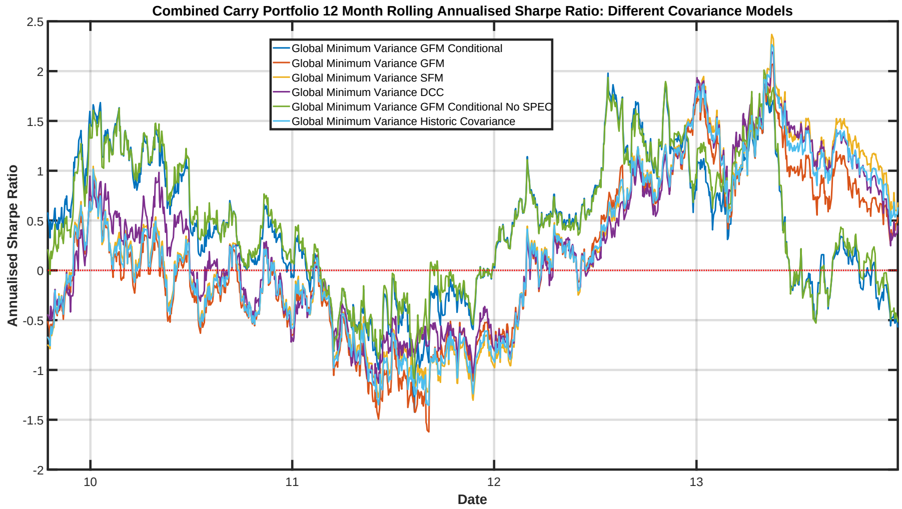


Figure 12.15: Carry trade portfolio 12 month annualised rolling Sharpe ratio. The Sharpe ratio is defined as return divided by volatility.

Table 12.1: Carry trade portfolio risk measures for different covariance forecasting techniques. The Sharpe ratio is defined as return divided by volatility. The Sortino ratio is the return divided by downside volatility. The Omega ratio is the probability weighted ratio of gains versus losses for some threshold return target (we use 0). Max DD is the maximum decline from historical peak.

Risk Measure	GFM Cond.	GFM Cond. (No SPEC)
Sharpe	**0.31**	0.27
Sortino	**0.43**	0.38
Omega	**1.87**	1.78
Max DD	31.71	**28.82**
Calmar	**0.15**	0.15

Table 12.2: Carry trade portfolio risk measures for different covariance forecasting techniques (2). Here, the GFM is the unconditional model. The Sharpe ratio is defined as return divided by volatility. The Sortino ratio is the return divided by downside volatility. The Omega ratio is the probability weighted ratio of gains versus losses for some threshold return target (we use 0). Max DD is the maximum decline from historical peak.

Risk Measure	GFM	SFM	DCC
Sharpe	0.11	0.11	**0.21**
Sortino	0.14	0.14	**0.30**
Omega	1.27	1.28	**1.58**
Max DD	30.3	**30.04**	31.50
Calmar	0.05	0.05	**0.10**

12.6 Conclusions

The Standard Multi-Factor model (SFM) family discussed in this thesis has been widely used in the finance and econometrics literature primarily because of the readily available economic interpretation it offers when linking exogenous factors to the portfolio returns. In addition it is efficient with regard to estimation due to its model based parsimony. However, the standard form of this multi-factor model is known to fail to account for an important feature displayed by financial assets real data returns, which is the heteroskedastic nature of the assets covariance structure over time. In this thesis, a Generalised version of the Multi-Factor model, the GFM family of models, is developed. The main purpose of this extension is to address the short-comings offered under the SFM family whilst preserving the direct interpretation of factors in the model and their influence on explaining the portfolio returns. By introducing such a model it has been demonstrated that it is possible to fill this gap by proposing a generalized version of the multi-factor model which incorporates the factors into the covariance of the idiosyncratic error term and hence allows for heteroskedastic unconditional and conditional covariance based model structures. I show that the GFM model is directly interpretable in terms of how it depends on the assets return filtration but also on the σ-algebra generated by the covariates or selected explanatory factors. The use of the GFM model in applications involving portfolio allocation requires the ability to easily and efficiently forecast the future value of the covariance matrix assuming the stationarity of the trend and the covariance regression parameters. The GFM model developed makes it possible to devise and estimate robust forecasting models for the set of independent variables selected. This is demonstrated in numerous different studies on the forecast performance of the GFM family of models as well as the performance of the optimal portfolios under a global minimum variance portfolio allocation framework.

Another contribution of this thesis involves the selection of meaningful econometric factors that have both explanatory power in-sample as well as good forecast performance when used to develop a portfolio covariance forecast. I demonstrate that in the currency portfolio studies performed, whilst two well known factors studied in the literature, the DOL and HML_{FX}, are providing strong in-sample explanatory power, their out-of-sample forecast performance is very poor. It is important to note that the DOL and HML_{FX} covariates are risk premia and therefore shouldn't be expected to be forecastable, since otherwise there is no risk to be compensated for. This makes them difficult to utilise in portfolio selection frameworks which require the forecast portfolio trend and covariance. In this thesis, I have obtained additional volume based explanatory factors that admit both strong in-sample explanatory power as well as providing reasonable forecasting performance, making them directly useful in the portfolio allocation problem. Furthermore, the factors considered are directly interpretable and it is possible to relate there attributes to an established literature in economics, relating returns to volume and liquidity of an asset, see Ames et al. [2015a]. This established relation between the speculative positions and the asset returns dependence structure means it is possible to better capture the heteroskedasticity prevailing in the asset returns, notably in high volatility environment. Through this empirical application to the currency market I demonstrated that the conditional formulation of the covariance proposed outperforms, on a risk-return basis, several widely implemented models, such as the DCC or the single factor models, even though non-negativity and non-positivity constraints are necessarily appended to the high and low interest rates baskets optimization in order to build self-financing portfolios.

Part IV

Hybrid Multi-Factor State Space Modelling Contributions

Chapter 13

Part IV Overview

The previous three parts of this thesis have focused on currencies, in particular dependence modelling and optimisation of multiple-currency baskets in the currency carry trade strategy. However, it can be noted that many of the currencies typically utilised in the high interest rate basket of the carry trade are heavily linked to commodity prices, see Ready et al. [2017]. These currencies are known as 'commodity currencies'. There is a growing strand of literature surrounding this link between commodity price fluctuations and currencies price dependences. Therefore, it is important to understand the dynamics of commodity prices in order to further understand the dependence dynamics in currencies. In the fourth part of this thesis, a novel Hybrid Multi-Factor Stochastic Differential Equation framework is introduced. This state-space modelling framework is utilised to investigate the influence of observable exogenous covariates on the behaviour of commodity prices.

Chapter 14 introduces the traditional approaches utilised in the literature to model commodity prices and then details the proposed novel Hybrid Multi-Factor (HMF) state-space modelling framework. The flaws inherent in the traditional two-stage approaches to analysing the influence of covariates on commodity prices are discussed. Furthermore, the benefits of the HMF framework are presented.

Chapter 15 utilises the novel Hybrid Multi-Factor (HMF) model developed in Chapter 14 to investigate commodity futures and spot price dynamics in terms of interpretable observable factors that influence speculators and hedgers heterogeneously. In particular, the focus is on understanding the macroeconomic and microeconomic factors influencing the behaviour of oil prices.

Chapter 14

Hybrid Multi-Factor Modelling Framework

In this chapter, the traditional approaches utilised in the literature to model commodity prices will be introduced and then the novel Hybrid Multi-Factor (HMF) state-space modelling framework proposed in this thesis will be detailed.

14.1 Model

In this section, I describe the Hybrid Multi-Factor (HMF) SDE model, which is a genuine, statistically robust and consistent approach to incorporation of both stochastic latent factor interpretation of unobserved spot price dynamics from futures panel dynamics as well as the incorporation of important influential and informative explanatory covariates that are observed in the global macro and micro economy and commodity markets. Furthermore, it enables one to differentiate the impact that certain observable exogenous macro, micro and fundamental variables can have upon the dynamic of this commodity and how significant they can be in explaining the short term and the long term dynamic of a given commodity market. This HMF model is particularly interesting for regression analysis as it avoids the common two-stage regression generally proposed in the literature (Dempster et al. [2012], Prokopczuk and Wu [2013]) with extraction of the latent factors and then the regression upon the independent variables chosen by the econometricians.

Detailed discussions of the inconsistency of such two stage approaches from a statistical estimation, model selection and testing as well as forecasting perspective can be found in Ames et al. [2016].

The estimation and inference in the proposed joint HMF stochastic models still allow convenient, robust and statistically optimal state space modelling estimation procedures to be adopted on closed form risk neutral analytic futures price dynamics without the violation of inconsistent statistical modelling assumptions. The inference of this model consists of a one block estimation of the parameters, thus avoiding any misspecification of the residuals dynamic as pointed out in Ames et al. [2016]. The other appeal of this approach lies in the capacity of separating the impact of a given covariate on the various parameters of a latent factor dynamic. In this thesis, I investigate the short term/long term model proposed by Schwartz and Smith [2000] and analyse the impact of several macroeconomic as well as microeconomic variables upon the respective stochastic latent factors dynamics. Furthermore and as will be demonstrated, this model also allows one to incorporate different features observed in the market such as the correlation between the spot price and the stochastic convenience yield as proposed in Casassus and Collin-Dufresne [2005] but also the inventory and scarcity effects upon the commodity price dynamic.

14.1.1 Gibson-Schwartz Stochastic Convenience Yield Model

Here, the model introduced in Gibson and Schwartz [1990] is reviewed, with the notation as adopted in Schwartz [1997]. This stochastic convenience yield model and its extensions to include a third latent factor are very popular in the literature as it allows one to model the convenience yield as one of the latent factors. Hence it is straightforward to perform regressions of observable covariates, such as inventories and production, on the filtered convenience yield factor, as presented in Dempster et al. [2012]; Prokopczuk and Wu [2013].

The real world dynamics of the two-factor stochastic convenience yield model of Gibson and Schwartz [1990] are expressed as follows:

Definition 14.1.1. *Gibson-Schwartz 1990 (GS90) Model*

$$X_t = lnS_t \tag{14–1}$$

$$dX_t = (\mu - \delta_t - \frac{1}{2}\sigma_1^2)dt + \sigma_1 dZ_t^1 \tag{14–2}$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dZ_t^2 \tag{14–3}$$

$$\textit{where} \quad \mathbb{E}\left[dZ_t^1 dZ_t^2\right] = \rho dt \tag{14–4}$$

where S_t is the spot price at time t, μ is the equilibrium spot price level, δ_t is the convenience yield, α is the equilibrium level of the convenience yield, κ is the speed of mean reversion of the convenience yield, σ_1^2 and σ_2^2 are the volatilities of the brownian increments of the log spot price and the convenience yield respectively, and dZ_t^1 and dZ_t^2 are increments of standard Brownian motion.

The risk-neutral formulation of this model can be obtained in the standard fashion (adjusting the drift terms) as:

$$X_t = lnS_t \tag{14–5}$$

$$dX_t = (r - \delta_t - \frac{1}{2}\sigma_1^2)dt + \sigma_1 d\tilde{Z}_t^1 \tag{14–6}$$

$$d\delta_t = \kappa(\alpha - \lambda - \delta_t)dt + \sigma_2 d\tilde{Z}_t^2 \tag{14–7}$$

$$\textit{where} \quad \mathbb{E}\left[dZ_t^1 dZ_t^2\right] = \rho dt \textit{ and } r \textit{ is the risk-free rate.} \tag{14–8}$$

Prokopczuk and Wu [2013] employ a third latent factor to model the stochastic interest rate. Dempster et al. [2012] allow the third latent factor to act as a medium term factor in order to capture business cycles, net oil demand and trading variables. Thus an affine combination of two of the factors combine to model the convenience yield.

14.1.2 Schwartz-Smith 2000 (SS2000) Model

The two factor long-term/short-term model introduced in Schwartz and Smith [2000] is equivalent to the Gibson and Schwartz [1990] model, as shown in Section 14.1.3, but also comes with a number of advantages, as described in Schwartz and Smith [2000]: "While many find the notion of convenience yields elusive, the idea of stochastically evolving short-term deviations and equilibrium prices seems more natural and intuitive. Moreover, these factors are more "orthogonal" in their dynamics, which leads to analytic results that are more transparent and allow us to simplify the analysis of many long-term investments."

The real world dynamics of the two-factor long-term/short-term model of Schwartz and Smith [2000] are expressed as follows:

Definition 14.1.2. *Schwartz-Smith 2000 (SS2000) Model*

Real Process:

$$X_t = ln(S_t) = \chi_t + \xi_t \tag{14–9}$$

$$d\chi_t = -\beta\chi_t dt + \sigma_\chi dZ_t^\chi \tag{14–10}$$

$$d\xi_t = \mu_\xi dt + \sigma_\xi dZ_t^\xi \tag{14–11}$$

$$\mathbb{E}\left[dZ_t^\chi dZ_t^\xi\right] = \rho_{\chi\xi} dt \tag{14–12}$$

where S_t is the spot price at time t, χ_t is the short term dynamics latent factor, ξ_t is the long term dynamics latent factor, β is the short term mean reversion parameter, μ_ξ is the long term equilibrium parameter, σ_χ^2 and σ_ξ^2 are the volatilities of the brownian increments, and dZ_t^χ and dZ_t^ξ are increments of standard Brownian motion.

The risk-neutral formulation of this model can be obtained in the standard fashion (adjusting the drift terms) as:

Risk-Neutral Process:

$$d\chi_t = (-\beta\chi_t - \lambda_\chi)dt + \sigma_\chi d\tilde{Z}_t^\chi \tag{14–13}$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi d\tilde{Z}_t^\xi \tag{14–14}$$

where S_t is the spot price at time t, χ_t is the short term dynamics latent factor, ξ is the long term dynamics latent factor, and dZ_t^χ and dZ_t^ξ are increments of standard Brownian motion.

14.1.3 Equivalence of Schwartz-Smith 2000 Model and Gibson-Schwartz Stochastic Convenience Yield Model

The equivalence between the Schwartz-Smith 2000 model and Gibson-Schwartz stochastic convenience yield model is shown in Table 14.1. The factors in each model can be represented as linear combinations of the factors in the other model.

Table 14.1: The Relationships Between Parameters in the Long-Term/Short-Term Model and the Stochastic Convenience Model of Gibson and Schwartz [1990].

Long-Term/Short-Term Model Parameter		
Symbol	Description	Definition in Terms of Stochastic Convenience Yield Model
β	Short-term mean-reversion rate	κ
σ_χ	Short-term volatility	σ_2/κ
dZ_χ	Short-term process increments	dZ_2
μ_ξ	Equilibrium drift rate	$\mu - \alpha - \frac{1}{2}\sigma_1^2$
σ_ξ	Equilibrium volatility	$(\sigma_1^2 + \sigma_2^2/\kappa^2 - 2\rho\sigma_1\sigma_2/\kappa)^{1/2}$
dZ_ξ	Equilibrium process increments	$(\sigma_1 dZ_1 - (\sigma_2/\kappa)dZ_2)(\sigma_1^2 + \sigma_2^2/\kappa^2 - 2\rho\sigma_1\sigma_2/\kappa)^{-1/2}$
$\rho_{\xi\chi}$	Correlation in increments	$(\rho\sigma_1 - \sigma_2/\kappa)(\sigma_1^2 + \sigma_2^2/\kappa^2 - 2\rho\sigma_1\sigma_2/\kappa)^{-1/2}$
λ_χ	Short-term risk premium	λ/κ
λ_ξ	Equilibrium risk premium	$\mu - r - \lambda/\kappa$

14.1.4 Extension to Schwartz-Smith 2000 Model: SSX Model

The Schwartz and Smith [2000] model can be extended to allow for mean reversion in the long term drift component, which is desirable since it is a stylized fact that commodity prices mean revert in the long term. Such a feature is first introduced in Peters et al. [2013] Section 3.1 and Binkowski et al. [2009], although here I study this feature in significantly more detail in the novel class of HMF model structures introduced in this chapter. In particular, Chapter 15 demonstrates the statistically significant gain in model fit such an extended Schwartz-Smith 2000 (hereafter abbreviated as SS2000) model presents when used to explain inter-temporal variation in oil futures prices.

Definition 14.1.3. *Schwartz-Smith 2000 Extended (SSX) Model*
The real world and risk-neutral dynamics of the Schwartz-Smith 2000 Extended (hereafter abbreviated as SSX) model can be expressed as follows:

Real Process

$$X_t = ln(S_t) = \chi_t + \xi_t \tag{14–15}$$

$$d\chi_t = -\beta\chi_t dt + \sigma_\chi dZ_t^\chi \tag{14–16}$$

$$d\xi_t = (\mu_\xi - \gamma\xi_t)dt + \sigma_\xi dZ_t^\xi \tag{14–17}$$

$$\mathbb{E}\left[dZ_t^\chi dZ_t^\xi\right] = \rho_{\chi\xi}dt \tag{14–18}$$

where S_t is the spot price at time t, χ_t is the short term dynamics latent factor, ξ_t is the long term dynamics latent factor, β is the short term mean reversion parameter, μ_ξ is the long term equilibrium parameter, γ is the long term mean reversion parameter, σ_χ^2 and σ_ξ^2 are the volatilities of the brownian increments, and dZ_t^χ and dZ_t^ξ are increments of standard Brownian motion.

Risk-Neutral Process

$$dchi_t = (-\beta\chi_t - \lambda_\chi)dt + \sigma_\chi d\tilde{Z}_t^\chi \tag{14–19}$$

$$d\xi_t = (\mu_\xi - \lambda_\xi - \gamma\xi_t)dt + \sigma_\xi d\tilde{Z}_t^\xi \tag{14–20}$$

where $\mathbb{E}[d\tilde{Z}_t^{\chi}, d\tilde{Z}_t^{\xi}] = \rho_{\chi\xi}dt$ *and it is assumed that there are constant, deterministic unknown risk premia for compensation of the drift in the short-term and long-term dynamics of the latent stochastic spot price. Such assumptions on risk premia are standard in the literature and are mostly made for convenience to aid in derivation of a closed form expression for the futures prices.*

Remark 14.1.4. *One can note that so far these models are purely stochastic (mathematical) models, in that the factors utilised to explain the futures curve dynamics are stylized latent stochastic processes and not constructed based on independent observable covariates that not only stochastically explain cross-sectional and serial correlation and stochastic variations in the observed futures panels, but also lead to greater economic insight and interpretability of the models. At present, the current literature tries to achieve this extended goal of interpretability of these latent factor models with exogenous covariates added in an ad hoc, non-statistically consistent two stage set of procedures. Ames et al. [2016] explains in more detail some of the challenges with such two stage procedures. In the following sections, I explain how to consistently perform calibration and estimation of a Hybrid Multi-Factor stochastic differential equation (s.d.e.) model that structurally incorporates exogenous covariates as explanatory factors, whilst admitting efficient and statistically consistent estimation procedures.*

14.1.5 The Hybrid Multi-Factor (HMF) Model

The Hybrid Multi-Factor (hereafter abbreviated as HMF) is referred to as a hybrid model since it combines the latent factor modelling approach and the observable factor linear regression modelling approach into a model which allows for consistent estimation. The model structure presented below allows for several nested subclasses of model to be developed, which includes linear regression predictors for incorporation of exogenous covariates through a link function to the stochastic latent spot price dynamic factors. The link function relating the linear predictors to the latent s.d.e. model factors can be achieved in a number of structurally interpretable approaches in the drift function and the volatility function, effectively allowing one to develop generalised diffusion dynamics for the multi-factor s.d.e. commodity model whilst still incorporating a closed form analytic risk neutral futures price. This can be achieved in the long term equilibrium price and the rates of mean reversion in the short and long term latent spot dynamics, with structurally different effects as well as differing interpretation. Furthermore, the latent factors in this model can be easily incorporated in a statistically consistent manner with lagged exogenous covariates, instantaneous effects and even forward looking, smoothing based information models.

Definition 14.1.5. *Hybrid Multi-Factor (HMF) Model*

$$X_t = ln(S_t) = \chi_t + \xi_t \tag{14–21}$$

$$d\chi_t = - \underbrace{\beta_t}_{\psi_{c1} + \sum_{j=1}^{J} \sum_{k=-K}^{K'} \psi_{1,j} m_{t+k,j}} \chi_t dt + \sigma_\chi dZ_t^\chi \tag{14–22}$$

$$d\xi_t = (\underbrace{\mu_{\xi,t}}_{\psi_{c2} + \sum_{j=1}^{J} \sum_{k=-K}^{K'} \psi_{2,j} m_{t+k,j}} - \underbrace{\gamma_t}_{\psi_{c3} + \sum_{j=1}^{J} \sum_{k=-K}^{K'} \psi_{3,j} m_{t+k,j}} \xi_t) dt + \sigma_\xi dZ_t^\xi \tag{14–23}$$

$$\mathbb{E}\left[dZ_t^\chi dZ_t^\xi\right] = \rho_{\chi\xi} dt \tag{14–24}$$

where $m_{t,j}$ is the value of the observable covariate j at time t, J is the number of covariates considered, and K and K' determine the time period over which the covariates are summed.

Remark 14.1.6. *In this modelling framework, $m_{t,j}$ is assumed to be observable and known (or more formally, that $m_{t,j}$ is part of the filtration). It is not necessary to have knowledge of the process of $m_{t,j}$, since it is sufficient to be able to observe* *the value at any point t required. However, in reality $m_{t,j}$ is partially observed and thus some form of approximation is necessary at fixed time points. There are 3 approaches that can be considered:*

1. *Estimate $m_{t,j}$ as a fixed quantity, i.e. assume constant over time.*

2. *Utilise lagged values of the time series $m_{t,j}$.*

3. *Fit a time series model to the covariate $m_{t,j}$ in order to forecast future values.*

The empirical investigation presented in Chapter 15 adopts the first and second approaches. Furthermore, initial explorations of the third approach were undertaken, but are not contained in this thesis.

14.2 Deriving The Futures Price Expression

One can derive the futures price, $F_{t,T}$, for the HMF model using the Backward-Kolmogorov equation (BKE):

$$F_{t,T} = \tilde{\mathbb{E}}[S_T|S_t] = \tilde{\mathbb{E}}[e^{\chi_T+\xi_T}|\chi_t, \xi_t] \tag{14–25}$$

BKE:

$$\begin{aligned} \frac{\partial p}{\partial t} + \frac{1}{2}\sigma_\chi^2 \frac{\partial^2 p}{\partial \chi_t^2} + \frac{1}{2}\sigma_\xi^2 \frac{\partial^2 p}{\partial \xi_t^2} + (-\beta_t \chi_t - \lambda_\chi)\frac{\partial p}{\partial \chi_t} \\ + (\mu_{\xi,t} - \lambda_\xi - \gamma_t \xi_t)\frac{\partial p}{\partial \xi_t} + \rho_{\chi\xi}\sigma_\chi\sigma_\xi \frac{\partial^2 p}{\partial \chi_t \partial \xi_t} = 0 \end{aligned} \tag{14–26}$$

subject to boundary condition $p(X_T, T|X_t, t = T) = \tilde{\delta}(X_T - X_t)$.

Multiplying by e^{X_T} and then integrating w.r.t. X_T allows one to express each term in the pde with respect to the futures price:

$$\frac{\partial F}{\partial t}+\frac{1}{2}\sigma_\chi^2\frac{\partial^2 F}{\partial \chi_t^2}+\frac{1}{2}\sigma_\xi^2\frac{\partial^2 F}{\partial \xi_t^2}+(-\beta_t\chi_t-\lambda_\chi)\frac{\partial F}{\partial \chi_t}$$
$$+(\mu_{\xi,t}-\lambda_\xi-\gamma_t\xi_t)\frac{\partial F}{\partial \xi_t}+\rho_{\chi\xi}\sigma_\chi\sigma_\xi\frac{\partial^2 F}{\partial\chi_t\partial\xi_t}=0 \qquad (14\text{–}27)$$

subject to boundary condition $F(X_T,T|X_t,t)=e^{X_T}$.

Assume the solution of this backward Kolmogorov equation has an exponential affine form:

$$F_{t,T}=e^{B_{0,t}(\tau)+B_{1,t}(\tau)\chi_t+B_{2,t}(\tau)\xi_t}\quad , \qquad (14\text{–}28)$$

where $\tau=T-t$.

Now, since $F(t=T,T)=e^{X_T}$ we have $B_{0,t}(0)=0,\ B_{1,t}(0)=1,\ B_{2,t}(0)=1$.

Substituting this expression for the futures price into the BKE:

$$F_{t,T}\Big[\frac{\partial B_{0,t}(\tau)}{\partial t}+\frac{\partial B_{1,t}(\tau)}{\partial t}\chi_t+\frac{\partial B_{2,t}(\tau)}{\partial t}\xi_t\Big]+\frac{1}{2}\sigma_\chi^2B_{1,t}^2(\tau)F_{t,T}$$
$$+\frac{1}{2}\sigma_\xi^2B_{2,t}^2(\tau)F_{t,T}+(-\beta_t\chi_t-\lambda_\chi)B_{1,t}(\tau)F_{t,T}$$
$$+(\mu_{\xi,t}-\lambda_\xi-\gamma_t\xi_t)B_{2,t}(\tau)F_{t,T}+\rho_{\chi\xi}\sigma_\chi\sigma_\xi B_{1,t}(\tau)B_{2,t}(\tau)F_{t,T}=0 \qquad (14\text{–}29)$$

Note $\partial\tau=-\partial t$.

$$-F_{t,T}\Big[\frac{\partial B_{0,t}(\tau)}{\partial \tau}+\frac{\partial B_{1,t}(\tau)}{\partial \tau}\chi_t+\frac{\partial B_{2,t}(\tau)}{\partial \tau}\xi_t\Big]+\frac{1}{2}\sigma_\chi^2B_{1,t}^2(\tau)F_{t,T}$$
$$+\frac{1}{2}\sigma_\xi^2B_{2,t}^2(\tau)F_{t,T}+(-\beta_t\chi_t-\lambda_\chi)B_{1,t}(\tau)F_{t,T}$$
$$+(\mu_{\xi,t}-\lambda_\xi-\gamma_t\xi_t)B_{2,t}(\tau)F_{t,T}+\rho_{\chi\xi}\sigma_\chi\sigma_\xi B_{1,t}(\tau)B_{2,t}(\tau)F_{t,T}=0 \qquad (14\text{–}30)$$

Dividing by $F_{t,T}$ and re-arranging gives:

$$\begin{aligned}\frac{\partial B_{0,t}(\tau)}{\partial\tau}+\frac{\partial B_{1,t}(\tau)}{\partial\tau}\chi_t+\frac{\partial B_{2,t}(\tau)}{\partial\tau}\xi_t =&\frac{1}{2}\sigma_\chi^2 B_{1,t}^2(\tau)+\frac{1}{2}\sigma_\xi^2 B_{2,t}^2(\tau)\\ &+(-\beta_t\chi_t-\lambda_\chi)B_{1,t}(\tau)\\ &+(\mu_{\xi,t}-\lambda_\xi-\gamma_t\xi_t)B_{2,t}(\tau)\\ &+\rho_{\chi\xi}\sigma_\chi\sigma_\xi B_{1,t}(\tau)B_{2,t}(\tau)\end{aligned}\qquad(14\text{–}31)$$

One now has a system of 3 ODEs:

$$\frac{dB_{1,t}(\tau)}{d\tau}\chi_t=-\beta_t\chi_t B_{1,t}(\tau)\implies\frac{dB_{1,t}(\tau)}{d\tau}=-\beta_t B_{1,t}(\tau)\qquad(14\text{–}32)$$

$$\frac{dB_{2,t}(\tau)}{d\tau}\xi_t=-\gamma_t\xi_t B_{2,t}(\tau)\implies\frac{dB_{2,t}(\tau)}{d\tau}=-\gamma_t B_{2,t}(\tau)\qquad(14\text{–}33)$$

$$\begin{aligned}\frac{dB_{0,t}(\tau)}{d\tau}=&\frac{1}{2}\sigma_\chi^2 B_{1,t}^2(\tau)+\frac{1}{2}\sigma_\xi^2 B_{2,t}^2(\tau)-\lambda_\chi B_{1,t}(\tau)\\ &+(\mu_{\xi,t}-\lambda_\xi)B_{2,t}(\tau)+\rho_{\chi\xi}\sigma_\chi\sigma_\xi B_{1,t}(\tau)B_{2,t}(\tau)\end{aligned}\qquad(14\text{–}34)$$

with initial conditions: $B_{1,t}(0)=1, B_{2,t}(0)=1, B_{0,t}(0)=0$.

Solving this system of ODEs one obtains:

$$B_{1,t}(\tau)=e^{-\int\beta_t d\tau}\qquad(14\text{–}35)$$

$$B_{2,t}(\tau)=e^{-\int\gamma_t d\tau}\qquad(14\text{–}36)$$

$$B_{0,t}(\tau) = \int \Big[\frac{1}{2}\sigma_\chi^2 B_{1,t}^2(\tau) + \frac{1}{2}\sigma_\xi^2 B_{2,t}^2(\tau) - \lambda_\chi B_{1,t}(\tau) + (\mu_{\xi,t} - \lambda_\xi) B_{2,t}(\tau) + \rho_{\chi\xi}\sigma_\chi\sigma_\xi B_{1,t}(\tau) B_{2,t}(\tau)\Big] d\tau \qquad (14–37)$$

$$B_{0,t}(\tau) = \int \Big[\frac{1}{2}\sigma_\chi^2 e^{-2\int \beta_t d\tau} + \frac{1}{2}\sigma_\xi^2 e^{-2\int \gamma_t d\tau} - \lambda_\chi e^{-\int \beta_t d\tau} + (\mu_{\xi,t} - \lambda_\xi) e^{-\int \gamma_t d\tau} + \rho_{\chi\xi}\sigma_\chi\sigma_\xi e^{-\int \beta_t d\tau} e^{-\int \gamma_t d\tau}\Big] d\tau \qquad (14–38)$$

Using $B_{0,t}(0) = 0$ we see:

$$B_{0,t}(\tau) = -\frac{\sigma_\chi^2}{4\beta_t}(e^{-2\int \beta_t d\tau} - 1) - \frac{\sigma_\xi^2}{4\gamma_t}(e^{-2\int \gamma_t d\tau} - 1) + \frac{\lambda_\chi}{\beta_t}(e^{-\int \beta_t d\tau} - 1) - \frac{1}{\gamma_t}(\mu_{\xi,t} - \lambda_\xi)(e^{-\int \gamma_t d\tau} - 1) - \frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{(\beta_t + \gamma_t)}(e^{-\int (\beta_t + \gamma_t) d\tau} - 1) \qquad (14–39)$$

Thus one can express the futures price as

$$F_{t,T} = e^{B_{0,t}(\tau) + B_{1,t}(\tau)\chi_t + B_{2,t}(\tau)\xi_t} \qquad (14–40)$$

and hence one has the following expression for the log futures price

$$lnF_{t,T} = e^{-\int \beta_t d\tau}\chi_t + e^{-\int \gamma_t d\tau}\xi_t + B_{0,t}(\tau). \qquad (14–41)$$

14.3 State-Space Model Formulation

Having derived the futures price for the HMF model above, the state-space model is now formulated in terms of a measurement equation and transition equation as follows.

Measurement Equation:

Let $y_t(\tau) = lnF_t(\tau)$ and $\tau_i = T_i - t$, where $T_i, i = 1, \ldots, N$ are the maturities of the contract available at time t.

$$\begin{bmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{bmatrix} = \begin{bmatrix} e^{-\int \beta_t d\tau_1} & e^{-\int \gamma_t d\tau_1} \\ e^{-\int \beta_t d\tau_2} & e^{-\int \gamma_t d\tau_2} \\ \vdots & \vdots \\ e^{-\int \beta_t d\tau_N} & e^{-\int \gamma_t d\tau_N} \end{bmatrix} \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} + \begin{bmatrix} B_{0,t}(\tau_1) \\ B_{0,t}(\tau_2) \\ \vdots \\ B_{0,t}(\tau_N) \end{bmatrix} + \begin{bmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_N) \end{bmatrix} \tag{14–42}$$

$$y_t(\tau) = \Lambda_t(\tau) f_t + B_{0,t}(\tau) + \epsilon_t(\tau) \tag{14–43}$$

where $\epsilon_t(\tau)$ is the observation error at time t of contract with maturity τ.

Note: when using futures data $\Lambda_t(\tau)$ changes with time, since each day the time to maturity reduces by one day for each contract until expiry. However, in the practical application considered in Chapter 15 I interpolate a fixed maturity futures curve at each time step and so $\Lambda_t(\tau)$ is in fact constant.

Transition Equation:

$$\begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} 0 \\ \mu_{\xi,t}\Delta t \end{bmatrix} + \begin{bmatrix} e^{-\int \beta_t d\tau} & 0 \\ 0 & e^{-\int \gamma_t d\tau} \end{bmatrix} \begin{bmatrix} \chi_{t-1} \\ \xi_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t^{\chi} \\ \eta_t^{\xi} \end{bmatrix} \tag{14–44}$$

$$f_t = c_t + A_t f_{t-1} + \eta_t \tag{14–45}$$

with the error terms following a white noise (WN) distribution given by

$$\begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix} \sim WN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & H \end{bmatrix} \right) \tag{14–46}$$

where

$$Q = \begin{bmatrix} \sigma_\chi^2 \frac{1-e^{-2\int \beta_t d\tau}}{2\beta_t} & \rho_{\chi\xi}\sigma_\chi\sigma_\xi \frac{1-e^{-\int(\beta_t+\gamma_t)d\tau}}{\beta_t+\gamma_t} \\ \rho_{\chi\xi}\sigma_\chi\sigma_\xi \frac{1-e^{-\int(\beta_t+\gamma_t)d\tau}}{\beta_t+\gamma_t} & \sigma_\xi^2 \frac{1-e^{-2\int \gamma_t d\tau}}{2\gamma_t} \end{bmatrix} \tag{14–47}$$

$$H = \begin{bmatrix} s_1 & 0 & 0 & \dots & 0 \\ 0 & s_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & s_N \end{bmatrix} \tag{14–48}$$

and

$$\Lambda_t(\tau) = \begin{bmatrix} e^{-\int \beta_t d\tau_1} & e^{-\int \gamma_t d\tau_1} \\ e^{-\int \beta_t d\tau_2} & e^{-\int \gamma_t d\tau_2} \\ \vdots & \vdots \\ e^{-\int \beta_t d\tau_N} & e^{-\int \gamma_t d\tau_N} \end{bmatrix} \tag{14–49}$$

$$f_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} c_t = \begin{bmatrix} 0 \\ \mu_{\xi,t}\Delta t \end{bmatrix} A_t = \begin{bmatrix} e^{-\int \beta_t d\tau} & 0 \\ 0 & e^{-\int \gamma_t d\tau} \end{bmatrix} \tag{14–50}$$

14.4 Filtering and Parameter Estimation via Kalman Filter

Due to the way that the HMF model has been developed it is still possible to obtain optimal estimation of both the latent stochastic factors as well as all static model parameters in a statistically consistent manner via computationally efficient and widely utilised methods based on Kalman filtering followed by marginal likelihood estimation under recursive least squares estimation methods, which provide the best linear unbiased estimators of the model parameters and latent states, see discussions in Peters et al. [2013] as well as Schwartz and Smith [2000].

14.4.1 Kalman Filter

The Kalman filter component is introduced in this section. The filter equations can be split into a prediction step and a correction step. The prediction step consists of projecting the current state to obtain an a priori estimation of the latent factor which is then corrected during the update step, once the new measurement is taken into account.

Prediction stage:

$$\hat{f}_{t|t-1} = c_t + A_t \hat{f}_{t-1|t-1} \tag{14–51}$$

$$P_{t|t-1} = A_t P_{t-1|t-1} A_t^T + Q$$

Update stage:

$$\hat{f}_{t|t} = \hat{f}_{t|t-1} + K_t(y_t - \Lambda_t \hat{f}_{t|t-1} - B_{0,t}(\tau))$$

$$P_{t|t} = P_{t|t-1} - K_t \Lambda_t P_{t|t-1} \tag{14–52}$$

where the weighting function K_t is named the Kalman Gain and is equal to:

$$K_t = P_{t|t-1} \Lambda_t^T (\Lambda_t P_{t|t-1} \Lambda_t^T + H)^{-1} \tag{14–53}$$

The function K_t will place more or less weight on the prediction error.

14.4.2 Maximum Likelihood Parameter Estimation

To derive the maximum likelihood estimation one starts from the prediction error:

$$v_t = y_t - \hat{y}_{t|t-1} = y_t - \Lambda_t \hat{f}_{t|t-1} - B_{0,t}(\tau) \tag{14–54}$$

while the variance of this prediction error can be written as:

$$W_t = Var(v_t) = H + \Lambda_t P_{t|t-1} \Lambda_t^T \tag{14–55}$$

Then, since the prediction error is assumed to be Gaussian one has:

$$y_t | y_{t|t-1} \sim \mathcal{N}(\Lambda_t \hat{f}_{t|t-1} + B_{0,t}(\tau), W_t) \tag{14–56}$$

Based on this conditional distribution, one can now compute the log-likelihood function of $\Theta = \{\beta_t, \sigma_\chi, \lambda_\chi, \mu_\xi, \sigma_\xi, \gamma_t, \lambda_\xi, \rho_{\chi\xi}, s_1, \ldots, s_N\}$ by computing the joint density of $y_t|y_{t|t-1}$, $t = 1, 2, \ldots, T$.

$$l(\Theta) = -\frac{NT}{2} log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} log|W_t| - \frac{1}{2}\sum_{t=1}^{T} v_t^T W_t^{-1} v_t \tag{14–57}$$

This log likelihood function can then be maximised using an optimisation algorithm, for example the interior-point algorithm implementation in the MATLAB fmincon function.

14.4.3 Consistently Incorporating Exogenous Explanatory Covariates

To understand the interest of directly incorporating the exogenous covariates in the latent factors dynamic and the challenge that arises with the calibration in a two stage process I first present the current practice being adopted in the literature in works such as Dempster et al. [2012] and Prokopczuk and Wu [2013]. From the Kalman filter the optimally estimated state for long and short term latent stochastic factors (or convenience yield formulation) is obtained at each time t according to the estimator:

$$\mathbb{E}\left[f_t|\widehat{\boldsymbol{\theta}}^{(s1)}, \widehat{\boldsymbol{f}}_{t-1|t-1}, \mathcal{F}_t\right] = \widehat{\boldsymbol{c}}^{(s1)} + \widehat{\boldsymbol{A}}^{(s1)}\widehat{\boldsymbol{f}}_{t-1|t-1} \tag{14–58}$$

where $\mathcal{F}_t$ is the filtration generated by the futures panels $\{\mathbb{F}_1, \mathbb{F}_2, \ldots, \mathbb{F}_t\}$ with the observed futures prices on day t given by random vector $\mathbb{F}_t = (F_{t,1}, \ldots, F_{t,p_t})$. Note the number of contracts p_t may change over time depending on which contracts are observable and have sufficient traded volumes for incorporation in the model. $\widehat{\boldsymbol{\theta}}^{(s1)}$ denotes the static model parameters estimated at stage 1. Moreover these expectations are estimated via the Kalman filter as $\hat{f}_{t|t}\left(\widehat{\boldsymbol{\theta}}^{(s1)}\right)$.

Then in papers such as Dempster et al. [2012] and Prokopczuk and Wu [2013] these estimated states are used to perform simple multiple linear regressions based on simplified models under the following typically implicitly utilised statistical model assumptions:

- The observed values $\hat{f}_{t|t}\left(\widehat{\boldsymbol{\theta}}^{(s1)}\right)$ for times $t \in \{1, 2, \ldots, T\}$ are assumed to be i.i.d. realizations of "observations" of the latent factors, conditional on exogenous covariates $\boldsymbol{m}_t \in \mathbb{R}^d$. Note, these observations are already implicit functions of the static model parameters from stage 1 of estimation $\widehat{\boldsymbol{\theta}}^{(s1)}$.

- The mean regression model is assumed to be given by

$$\begin{aligned}
&\mathbb{E}\left[\widehat{\boldsymbol{f}}_{t|t}|\boldsymbol{m}_t\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\boldsymbol{f}_t|\widehat{\boldsymbol{\theta}}^{(s1)}, \widehat{\boldsymbol{f}}_{t-1|t-1}, \mathcal{F}_t\right]|\boldsymbol{m}_t\right] \\
&= \boldsymbol{\theta}_1^{(s2)} + \boldsymbol{\theta}_2^{(s2)}\boldsymbol{m}_t
\end{aligned} \tag{14–59}$$

  for d' exogenous observed covariates $\boldsymbol{m}_t \in \mathbb{R}^{d'}$ with unknown deterministic static parameters to be estimated in stage 2, denoted by $\boldsymbol{\theta}_1^{(s2)} \in \mathbb{R}^2$ and $\boldsymbol{\theta}_2^{(s2)} \in \mathbb{R}^2 \times \mathbb{R}^{d'}$.

- Furthermore, it is commonly assumed in the above cited works that the covariance is conditionally heteroskedastic and given by $\mathbb{V}\text{ar}\left[\hat{f}_{t|t}\left(\widehat{\boldsymbol{\theta}}^{(s1)}\right)|x\right] = \Omega^{(s2)}$, often with a diagonal covariance matrix.

Such a simple linear model can then be estimated via a generalized least squares procedure such that the stage two model parameters $\left[\boldsymbol{\theta}_1^{(s2)}, \boldsymbol{\theta}_2^{(s2)}\right]$ are obtained as the solution to the quadratic minimization:

$$\begin{aligned}
\left[\widehat{\boldsymbol{\theta}}_1^{(s2)}, \widehat{\boldsymbol{\theta}}_2^{(s2)}\right] = \arg\min &\Bigg[\left(\widehat{\boldsymbol{f}}_{1:T|1:T} - M\left[\boldsymbol{\theta}_1^{(s2)}, \boldsymbol{\theta}_2^{(s2)}\right]^T\right)^T \left(\Omega^{(s2)}\right)^{-1} \\
&\left(\widehat{\boldsymbol{f}}_{1:T|1:T} - M\left[\boldsymbol{\theta}_1^{(s2)}, \boldsymbol{\theta}_2^{(s2)}\right]^T\right)\Bigg],
\end{aligned} \tag{14–60}$$

where $\widehat{\boldsymbol{f}}_{1:T|1:T}$ denotes the set of "responses" in the regression given by $\left\{\widehat{\boldsymbol{f}}_{1|1}, \widehat{\boldsymbol{f}}_{2|2}, \ldots, \widehat{\boldsymbol{f}}_{T|T}\right\}$ and M is the design matrix of the exogenous covariates.

From this quadratic form, for a given response covariance matrix $\Omega^{(s2)}$ the Best Linear Unbiased Estimator for $\left[\boldsymbol{\theta}_1^{(s2)}, \boldsymbol{\theta}_2^{(s2)}\right]$ is given by:

$$\begin{aligned}
\left[\widehat{\boldsymbol{\theta}}_1^{(s2)}, \widehat{\boldsymbol{\theta}}_2^{(s2)}\right] &= \left(M^T\left(\Omega^{(s2)}\right)^{-1}M\right)^{-1} M^T\left(\Omega^{(s2)}\right)^{-1}\widehat{\boldsymbol{f}}_{1:T|1:T} \\
&= \left(M^T\left(\Omega^{(s2)}\right)^{-1}M\right)^{-1} M^T\left(\Omega^{(s2)}\right)^{-1} \oplus_{t=1}^T \left[\widehat{\boldsymbol{c}}^{(s1)} + \widehat{\boldsymbol{A}}^{(s1)}\widehat{\boldsymbol{f}}_{t-1|t-1}\right]
\end{aligned} \tag{14–61}$$

where $\oplus$ is the kronecker sum.

The challenge with this approach is that the aforementioned assumptions are typically not satisfied, making this form of regression both inefficient, due to the two stage parameter estimation performed, as well as biased and inaccurate in the model estimation and conclusions on model structure (for more detail please refer to Ames et al. [2016]). Conversely, the one-stage estimation framework proposed in this thesis allows a simultaneous inference of the latent factors dynamic as well as the covariates coefficients and thus to overcome this estimation error associated with the two-stage approach generally proposed in the literature.

Chapter 15

Investigating Cross-Sectional Dependence in Commodity Prices via Hybrid Multi-Factor State Space Models

In this chapter, the novel Hybrid Multi-Factor (HMF) model developed in Chapter 14 is utilised to investigate commodity futures and spot price dynamics in terms of interpretable observable factors that influence speculators and hedgers heterogeneously. In particular, the factors driving the behaviour of oil prices are analysed.

15.1 Introduction

Oil has historically been one of the most closely scrutinized commodities in the market. First and foremost, this is due to the important role this commodity plays in the worldwide economy and international relations, which gives it a prominent role, when compared to other energy, agricultural and metals commodities, in many aspects of the global economy and each country's specific macro, micro and monetary economic policy decisions.

The prominence of oil futures can be easily demonstrated for instance by the fact that its share of the global GDP was 4.8% in 2013 (Aguilera and Radetzki [2015]). In addition, as discussed in Backus and Crucini [2000] it has a significant influence over the respective balance of trade of consuming and producing countries and thus the resulting geopolitical interactions among them.

Historically, one has observed the importance that economies place on the price variation of oil and understanding the factors that affect such a dynamic in order to better understand the determinants of shocks and volatility regimes in the spot price, demand and supply.

Another determining reason lies in the frequent shocks affecting the supply and demand of the so called "black gold", giving birth to sudden and dramatic price movements such as during the 1973/74 oil crisis. The price of this exhaustible commodity has indeed been in the past heavily impacted by the discovery of new fields or the conflicts in oil-producing countries. On the other hand, the demand behaviour has generally been more influenced by the business cycles or even the evolution of the extracted oil inventories. That being said, according to the US Department of the Interior (DOI) as well as the US Energy Information Administration (EIA), the technology used for its extraction has recently been the main factor influencing the market supply. Over the last decade, advances in the application of horizontal drilling and hydraulic fracturing in shale have indeed drastically modified the international supply and demand equilibrium as well as the existing international relations by allowing the biggest oil consumer, namely the United States, to become over the same time period less and less dependent on its energy imports. According to the EIA, in 2015, 24% of the petroleum consumed in this country was imported, which corresponds to the lowest level since 1970.

From a modelling perspective, these features are significant and need to be incorporated into any interpretable and realistic commodity futures stochastic model. In addition, if the model is developed, as is the case with the class of Hybrid Multi-Factor Models (HMF) introduced in this thesis, to allow for clear closed form representations of structural features such as sensitivity, shock transient response and perturbation influence on the model parameters and the driving exogenous covariates characterizing the features just discussed, then such a class of models has the potential to significantly aid in the study of stochastic variation in oil futures prices and to aid in forecasting and policy decision. The main aim of this research is to provide such a class of models and demonstrate their utility in incorporating a range of exogenous covariates into different structural components that will clearly explain short term and long term speculator and hedger positions in oil futures and their influences.

Although one can obtain a coarse picture of the principal fundamental events affecting oil price dynamics throughout history, the modelling and the choice of explanatory variables for oil price dynamics is still fiercely debated in the academic literature. Several reasons for this have been put forward, among which is the microeconomic interactions between different types of agents who intervene in the market and who are generally classified into two distinct groups, labelled respectively hedgers and speculators. The pre-eminent role they can play in the price discovery process of the market has raised unanswered questions about the causality relationship existing between the future prices and the physical or spot price observed in the real economy. As a matter of fact, several papers have demonstrated that not just the speculators but also the commodity-index funds were so influential in the market that the future price was actually leading the spot price and thus disconnecting the oil price from the fundamentals, such as those mentioned earlier (Kaufmann and Ullman [2009], Silvrio and Szklo [2012], Kilian and Murphy [2014]). Following this strand of the literature, certain authors (Bessembinder [1992]; Acharya et al. [2013]; Etula [2013]; Adrian et al. [2014]) considered the limits-to-arbitrage as one of the main reasons for the inverted price discovery process. Through such analyses they were able to argue that this demonstrated that any market friction limiting the arbitrage capacity of the financial intermediaries was translating into limits to hedging for the producers and accordingly impacting the real sphere participants' behaviour as well as related variables such as the spot oil prices.

The fact that macro and micro-economic observable variables influence the determination of market price dynamics by directly influencing the decisions and behaviour of speculators and hedgers in the market has naturally led to an alternative proposition from academics consisting of modelling the oil price dynamic through state space models where the log-price can be represented as a combination of several latent processes, which can then be generically interpreted without being necessarily related to any fundamental or microeconomical variables (Gibson and Schwartz [1990]; Schwartz and Smith [2000]; Casassus and Collin-Dufresne [2005]). Among advocates for this approach, authors notably decomposed the future prices as a combination of short term and long term latent components while others have assumed equivalently that the latent process should be associated to the convenience yield and thus determine the basis level or said differently the price difference between the spot and the future contract. Kaldor [1939] explains that the inter-temporal difference between futures and thus between the future price and the spot price are linked to the cost of storage and also the so-called convenience yield which embodies the benefits accrued to the owner of the physical commodity by providing him with a certain flexibility with regards to his reaction in case of market shocks. Schwartz and Smith [2000] demonstrated through a change of variable the linear equivalence between modelling the convenience yield or the dynamic of a long and a short term latent factor in order to model the futures price curve. Another advantage in considering these models resides in the ease of financial change of measure to risk neutral formulations that admit closed form analytical futures prices in terms of stochastic factors assumed to explain the spot price stochastic unobserved dynamics. From this systematic model differentiation between macro, micro and latent factors and given also the fact that the storage cost or the convenience yield are both naturally related to fundamental elements such as the storage capacity in the market, followed several articles dissecting the behaviour of the latent processes relative to a set of fundamental and microeconomic variables (Dempster et al. [2012], Daskalaki et al. [2014]). On the contrary other academics focused on demonstrating that the fundamental factors were not marginally contributing to the explanation provided by the futures prices themselves, and thus the latent processes (Daskalaki et al. [2014]; Cummins et al. [2016]).

The research presented in this thesis reconciles two classes of model, the latent factor stochastic multi-factor s.d.e. models and the alternative class of observable regression econometric factor models, in a statistically consistent manner from interpretation and estimation perspectives. This is achieved with the novel class of stochastic HMF models that I develop, which allow for incorporation of exogenous covariate structures in a statistically rigorous manner.

One will notice that the proposed HMF stochastic models are a genuine combination of the two approaches and do not presume any prevalence from one approach or the other. The crux of the matter lies in building a model which allows a one-stage estimation with simultaneous inference of the latent factors dynamic and the covariates coefficients to overcome the estimation error associated to the two-stage approach generally proposed in the literature. In such a two-stage model (as in Dempster et al. [2012]), the authors recommend to first extract the latent factor estimates to later regress as a function of a set of covariates. This conditional estimation of the latent factor suffers from several flaws compared to the conditional estimates proposed in this thesis.

Furthermore, as detailed in formal statistical arguments in Ames et al. [2016] the current approaches proposed in the literature adopting such two stage estimation procedures to estimate latent stochastic factors followed by regression relationships in stage two for incorporation of exogenous covariates, often do so with inappropriate statistical assumptions and regression models. This makes claims and analysis coming from such models speculative at best, see discussion in detail in Ames et al. [2016].

First and foremost, conditioning on a set of macroeconomic and microeconomic variables commonly used in the literature leads to an undeniable improvement of the inference procedure relative to the two-stage method while I also show how the fundamental factors influence the different parameters of the latent factor models presented in the literature. For instance, the estimation method makes possible the distinction between the fundamental covariates which are impacting the mean reverting component of the latent factors and those which are influencing their respective trend.

The method utilised in this research allows one to consider covariate forecasts in order to extrapolate values for the futures prices while considering the confidence interval associated to this estimate. This is particularly convenient in risk management and commodity hedging as one needs to consider not only the amount to invest but also the uncertainty associated to this measurement.

Furthermore, the proposed model also copes with the topical problem of the marginal contribution of certain fundamental variables modelling relative to the latent process approaches. As a matter of fact, the results presented here show that adding a mean reversion component in the long term latent process and combining it with the mean reverting dynamic of the short term latent process devised by Schwartz and Smith [2000] model definitely improves the likelihood. Last but not least, I demonstrate through a likelihood ratio test how certain fundamental factors also consistently improve the inference of the state space model parameters showing that those covariates provide additional information not contained in the latent factors. Thus some adjustment in the latent factors models is required in order to take into consideration for instance the stochastic dynamic of the dollar variable. At present, there is no such general framework in the literature and this model could answer numerous questions still debated among academics about oil price dynamics.

15.2 Description of Price Data and Explanatory Covariates

In this section, a discussion of the oil futures price data and the explanatory covariate data used for the empirical analysis is presented. Furthermore, the choices made for the explanatory covariates investigated in this thesis are detailed.

15.2.1 Explanatory Covariates Data

To facilitate the empirical analysis one can distinguish between the different types of data sources among the covariates considered, as detailed in Table 15.1. The main distinction is naturally between macroeconomic (coming from the spot or physical sphere) and the microeconomic (coming from the financial markets sphere) variables. For instance in the physical sphere I considered the Baltic Dry Index (BDI) which represents an assessment of the freighting cost and is a composite of the daily quotes for various sized dry-vessels bookings across 23 different shipping routes and thus embodies an estimator for the price of moving the major raw materials by sea. The interest of this index has already been demonstrated in the literature (Bakshi et al. [2011],Geman and Smith [2012],Henderson et al. [2014]) and is due to the fact that the supply of cargo ships is quite inflexible and so the BDI index mainly fluctuates following the demand for raw materials. I naturally consider the US weekly crude oil Ending Stocks (excluding the strategic petroleum reserves) which corresponds to the number of barrels of oil in inventories at the end of each week in the United States (this data has been extensively used in the literature, for instance in Dempster et al. [2012] and Gorton et al. [2013]). Combining this data with the weekly refinery utilization rate, which measures the percentage of the operable crude oil distillation units utilized at this time (this indicator has been notably used in Kaufmann et al. [2008] and is also provided by the Energy Information Administration) and the US Field Production which represents the number of barrels of crude oil produced on a weekly basis in the US (this information provided by the EIA has also been considered in Dvir and Rogoff [2014]) allows one to take into account different fundamental information about the US physical market. First and foremost, it is commonly admitted in the literature that there is a negative relation between the convenience yield and the level of inventories (Victor K. Ng [1994], Milonas and Henker [2001]) while the freighting cost is directly related to the basis level, defined as the spread between the future and the spot prices (Geman and Smith [2012]). Moreover, the recent modification in the techniques used for extracting oil, in other words the advances in the application of horizontal drilling and hydraulic fracturing in shale should be quantified by the impressive growth of the US production over the last decade (Dvir and Rogoff [2014]). Provided that crude oil is priced in dollar, the level of the US currency relative to the other currencies is naturally affecting both the supply and the demand side of the market and thus the dynamic of the short and the long term latent variables. To measure the dollar fluctuations I retained the US Dollar Index (also used in Tang and Xiong [2012],Dempster et al. [2012]) which is a weighted geometric mean of the dollar's value relative to other selected currencies (which are the Euro, Japanese yen, British pound sterling, Canadian dollar, Swedish krona and the Swiss franc).

For the microeconomic or financial variables I retained two commonly mentioned indices in the limit to arbitrage literature, which are the speculative trading pressure (for more detail about the complementary value of this covariate, i.e. hedging pressure, please refer to Basu and Miffre [2013] and Acharya et al. [2013]), estimated as the ratio of net open speculative investor futures positions to the total open interest in the market, and the leverage ratio which represents the level of tightness of financial intermediaries' funding constraints, computed as the ratio of dealers' assets to liabilities (for more detail about this ratio please refer to Adrian et al. [2010], Adrian et al. [2014] and Daskalaki et al. [2014]). While the speculative trading pressure is computed from the daily commitment of traders reports published by the CFTC, the leverage factor on the contrary is only available on a quarterly basis and is computed using the amounts of financial assets and financial liabilities of security broker-dealers as published in Table L.129 of the Federal Reserve Flow of Funds. I also take into consideration two financial indices, which are the S&P500 and the Goldman Sachs Commodity Index (GSCI) which respectively represent the market capitalisation weighted index of the 500 largest public companies in the US and a weighted average of 24 commodities among which crude oil and other energy products represent about 64% of the index (both of these indices have been used in Daskalaki et al. [2014] and Büyükşahin and Robe [2014]).

Table 15.1: List of covariates (and their abbreviations) investigated in this modelling framework.

Covariate	Abbreviation
Baltic Dry Index	BDI
Dollar Index	DXY
Ending Stocks	End Stocks
Goldman Sachs Commodity Index	GSCI
Leverage Ratio	Lev Rat
Refinery Utilization	Ref Util
S&P 500 Index	S&P500
SPEC Ratio	SPEC
United States Field Production	US Prod

One challenge to resolve when working with such disparate and variable macro and micro economic data sources is the difference of publication frequency. In this research I match the frequency of all the fundamentals variables and take into account their date of publication to cope with the problem of mismatch between the data value date and the publication date. The weekly data published by the EIA containing information up to the previous Friday is released at 10:30 a.m. (Eastern Time) on Wednesdays and also the speculative positions data published by the CFTC containing information up to Tuesday is released at 3:30 p.m. (Eastern time) on Fridays. Thus, I take the closing price on Wednesdays as the weekly price data and use the latest available published fundamental data as the synchronous covariate value. One can also note that Wednesdays are affected by the least number of holidays.

I decided to consider in this study five different environments and periods as they were presumably impacted by different variables. I decided to look at equal sized samples, introducing no a priori bias, and considered period of five years as according to Postali and Picchetti [2006] the average long term cycle in the crude oil industry has been estimated to be 4-6 years. I first took the last five years where financialisation of the commodity market has been more pronounced according to several authors (Henderson et al. [2014], Büyükşahin and Robe [2014], Singleton [2014]). Then, I considered the period from 2006 up to 2011 which includes the financial crisis of 2008. Finally I looked at the period from 2000 to 2006 with the burst of the dot-com bubble, 1995-2000 with the LTCM collapse and finally the period going from 1990 to 1995 including the Iraqi Army's occupation of Kuwait in August 1990.

Standardised time series of the covariates considered in this analysis can be seen for the entire time period 1990 - 2016 in Figures 15.1 and 15.2. In these plots the individual covariates have been standardised using the approach proposed in Gelman [2008], i.e. by subtracting the mean and dividing by twice the standard deviation.

15.2.2 Crude Oil Futures Price Data

The crude oil price data considered in this research is the West Texas Intermediate (WTI) Crude oil futures prices traded on the New York Mercantile Exchange (NYMEX) obtained from Bloomberg for the empirical study in this thesis. I utilise the 1, 5, 9, 13 and 17 month expiry contracts as considered in Gibson and Schwartz [1990]; Schwartz [1997]; Prokopczuk and Wu [2013] since these contracts are sufficiently liquid. The Wednesday closing prices are retained at a weekly frequency in order to match with the weekly release of oil related data from the U.S. Energy Information Administration (EIA). The data sample covers

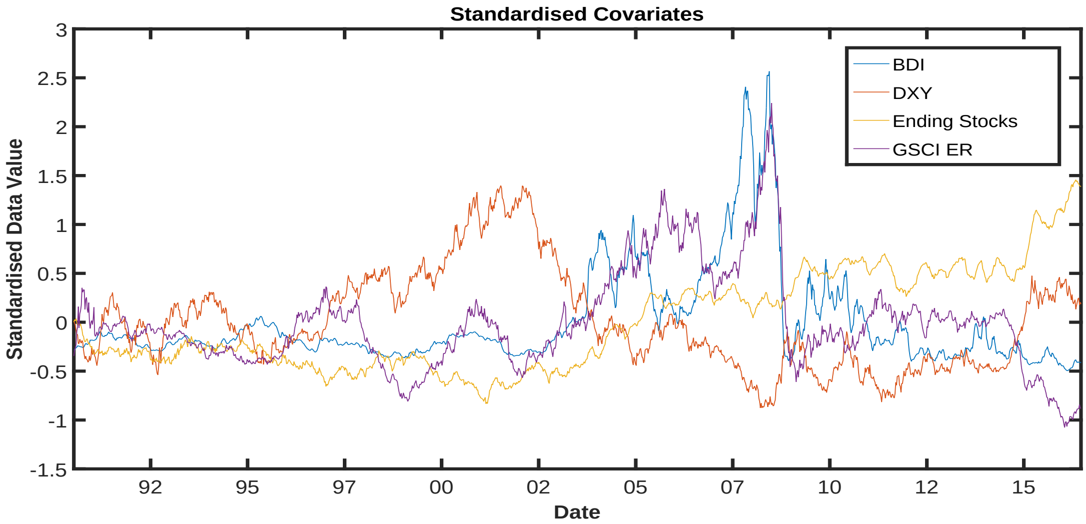


Figure 15.1: Standardised time series of the following covariates (using Gelman [2008] approach): BDI, DXY, Ending Stocks and GSCI Excess Returns.

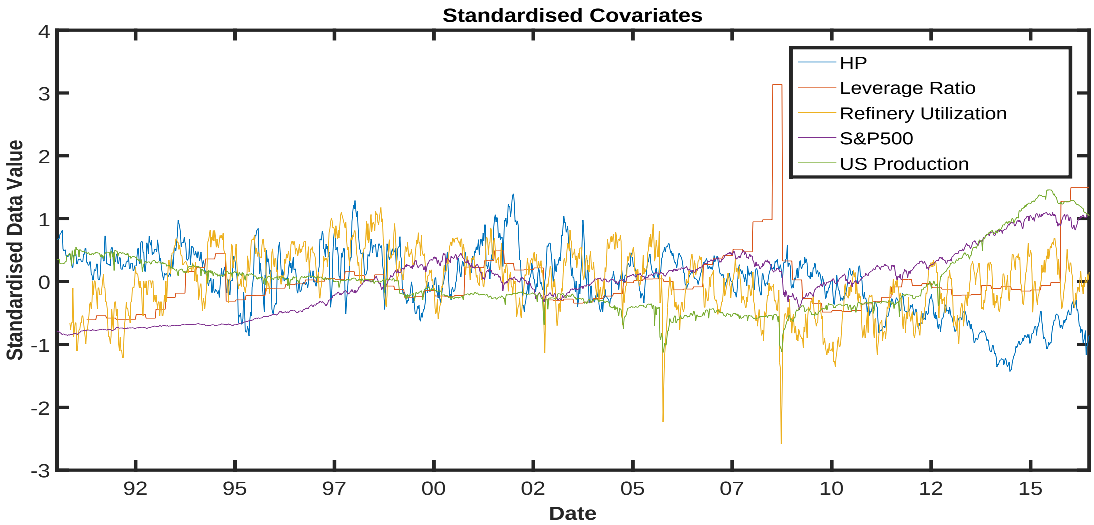


Figure 15.2: Standardised time series of the following covariates (using Gelman [2008] approach): Hedging Pressure, Leverage Ratio, Refinery Utilization, S&P500 and US Production.

the period 11th July 1990 to 22nd June 2016, i.e. 26 years or 1355 weeks. I divide the sample into five equal length blocks of roughly five years in order to provide a more detailed granular analysis.

In order to analyse the observable covariates at constant fixed points on the futures curve and hence produce comparable and interpretable coefficient values, a cubic spline interpolation approach is utilised to extract a fixed maturity futures curve from the actual futures data, i.e. maturities of 1 month, 5 months, 9 months, 13 months, 17 months are extracted from the raw futures data for which the days remaining until expiry of the contracts varies daily.

Summary statistics of the oil futures price time series data can be seen for each of the five periods in Tables 15.2 to 15.6 respectively.

Table 15.2: Descriptive statistics of WTI futures prices for the period 90-95. The mean, standard deviation, skewness, kurtosis, maximum and minimum of each futures maturity time series is presented. In addition, the average, maximum and minimum percentage backwardation is shown, where percentage backwardation is calculated as $100 \times (F_t(\tau_1) - F_t(\tau_5))/F_t(\tau_1)$.

Variable	$\boldsymbol{\tau_1}$	$\boldsymbol{\tau_2}$	$\boldsymbol{\tau_3}$	$\boldsymbol{\tau_4}$	$\boldsymbol{\tau_5}$
Mean	20.21	19.80	19.58	19.51	19.53
STD	3.97	2.95	2.33	2.00	1.80
Skew	2.13	1.70	1.30	1.11	1.00
Kurt	8.95	7.44	5.82	4.99	4.57
Max	38.46	32.70	28.81	27.02	26.08
Min	14.10	14.83	15.46	16.03	16.50
Avg. Bwd. %	-	1.30	0.75	0.16	-0.20
Max. Bwd. %	-	20.20	12.45	7.48	4.24
Min. Bwd. %	-	-7.79	-5.24	-4.25	-3.70

15.2.3 Data Preparation

In order to perform the empirical analyses considered in this chapter a substantial amount of effort and time was invested into collecting, cleaning and preparing the data. In particular, the following key steps were performed:

Table 15.3: Descriptive statistics of WTI futures prices for the period 95-00. The mean, standard deviation, skewness, kurtosis, maximum and minimum of each futures maturity time series is presented. In addition, the average, maximum and minimum percentage backwardation is shown, where percentage backwardation is calculated as $100 \times (F_t(\tau_1) - F_t(\tau_5))/F_t(\tau_1)$.

Variable	τ_1	τ_2	τ_3	τ_4	τ_5
Mean	20.85	20.06	19.53	19.16	18.91
STD	5.49	4.39	3.68	3.14	2.72
Skew	0.55	0.74	0.90	1.01	1.06
Kurt	2.80	3.27	3.73	4.08	4.30
Max	35.40	32.25	30.80	29.40	28.10
Min	11.34	12.18	12.70	13.14	13.54
Avg. Bwd. %	-	2.51	1.92	1.45	0.99
Max. Bwd. %	-	16.45	9.54	6.77	5.52
Min. Bwd. %	-	-17.07	-7.70	-6.00	-4.51

Table 15.4: Descriptive statistics of WTI futures prices for the period 00-06. The mean, standard deviation, skewness, kurtosis, maximum and minimum of each futures maturity time series is presented. In addition, the average, maximum and minimum percentage backwardation is shown, where percentage backwardation is calculated as $100 \times (F_t(\tau_1) - F_t(\tau_5))/F_t(\tau_1)$.

Variable	τ_1	τ_2	τ_3	τ_4	τ_5
Mean	36.54	35.74	34.75	33.93	33.27
STD	13.08	13.79	14.02	14.08	14.06
Skew	0.89	1.05	1.14	1.19	1.23
Kurt	2.60	2.76	2.91	3.02	3.09
Max	69.15	69.87	69.50	69.61	69.44
Min	18.63	19.45	19.83	20.07	20.25
Avg. Bwd. %	-	2.73	3.16	2.61	2.12
Max. Bwd. %	-	16.28	9.32	7.02	5.75
Min. Bwd. %	-	-4.96	-2.26	-1.50	-1.08

Table 15.5: Descriptive statistics of WTI futures prices for the period 06-11. The mean, standard deviation, skewness, kurtosis, maximum and minimum of each futures maturity time series is presented. In addition, the average, maximum and minimum percentage backwardation is shown, where percentage backwardation is calculated as $100 \times (F_t(\tau_1) - F_t(\tau_5))/F_t(\tau_1)$.

Variable	$\boldsymbol{\tau_1}$	$\boldsymbol{\tau_2}$	$\boldsymbol{\tau_3}$	$\boldsymbol{\tau_4}$	$\boldsymbol{\tau_5}$
Mean	77.76	79.91	80.92	81.49	81.80
STD	20.40	19.03	18.21	17.52	17.00
Skew	0.84	1.01	1.10	1.17	1.23
Kurt	3.83	4.19	4.43	4.60	4.74
Max	143.81	145.02	145.06	144.26	143.41
Min	35.67	42.95	46.61	49.25	51.49
Avg. Bwd. %	-	-3.55	-1.59	-0.95	-0.54
Max. Bwd. %	-	4.66	3.21	2.87	2.31
Min. Bwd. %	-	-24.90	-9.17	-7.37	-5.60

Table 15.6: Descriptive statistics of WTI futures prices for the period 11-16. The mean, standard deviation, skewness, kurtosis, maximum and minimum of each futures maturity time series is presented. In addition, the average, maximum and minimum percentage backwardation is shown, where percentage backwardation is calculated as $100 \times (F_t(\tau_1) - F_t(\tau_5))/F_t(\tau_1)$.

Variable	$\boldsymbol{\tau_1}$	$\boldsymbol{\tau_2}$	$\boldsymbol{\tau_3}$	$\boldsymbol{\tau_4}$	$\boldsymbol{\tau_5}$
Mean	81.04	81.61	81.43	81.05	80.59
STD	23.80	22.13	20.87	19.80	18.82
Skew	-0.77	-0.78	-0.78	-0.77	-0.77
Kurt	2.03	2.04	2.09	2.14	2.18
Max	112.89	113.70	113.27	112.30	110.95
Min	28.19	32.48	34.45	35.97	37.15
Avg. Bwd. %	-	-1.76	-0.38	0.02	0.21
Max. Bwd. %	-	6.04	4.92	3.81	3.16
Min. Bwd. %	-	-19.49	-7.80	-5.05	-3.88

1. Collect commodity futures price data at maturities of 1 month, 5 months, 9 months, 13 months and 17 months.

2. Collect exogenous covariate data. This is available at various frequencies depending upon the covariate being considered, e.g. the Baltic Dry Index (BDI) is reported daily whereas the US Field Production is reported weekly and indeed the Leverage Ratio is reported only quarterly.

3. Pre-process the commodity price data and the covariate data to deal with missing data, i.e. if data is missing copy previous day's price.

4. Match commodity price data to synchronous covariate data.

15.3 Results and Discussion

In order to appreciate the importance of each physical and financial covariate within the crude oil prices dynamic I first highlight in this section the rationale behind the model. To this end I first put in parallel the theoretical and empirical features described in the literature with the model components and the result obtained after calibration. I then show through an impulse response profile how the futures curves react to modifications of the parameters and demonstrate the impact of parameter changes upon the level and the slope of the crude oil futures curve. Finally, I study the role played by each covariate in the dynamic behaviour of the model parameters: first by looking at their statistical contribution to the log likelihood of the model and then through a stress scenario analysis by highlighting, conditionally on the past extreme values of a given covariate, their meaningfulness in the market behaviour.

15.3.1 Relevance of the long term mean reversion

One of the contributions of this research to the extensive literature about commodity price modelling is the addition of a long term mean reversion to the two factors model proposed by Schwartz and Smith [2000]. In Schwartz and Smith [2000] the authors assume that the supply elasticity to the price changes on the crude oil market is fitting into a short term time frame. They indeed consider that for a given rise of the crude oil barrel market price the profitability or internal rate of return of some producers may suddenly be positive and consequently entail an increase of the production on a short term basis. Reciprocally, if the price of the crude oil decreases, the following reduction of the highest cost producers margin will push them out of the market. On the other side the consumer demand elasticity to oil prices plays a symmetrical role and thus can in the short to mid-term constrain the fluctuations of the price. These mechanisms naturally entail a mean reversion of the oil price on a short period of time and has proved its efficacy in representing the commodity prices dynamic, see Cortazar et al. [2015]. I confirm this finding in Table 15.7 where the parameter β over several periods of time is always statistically significant. That being said, the speed of this mean reversion has been quite controversial in the literature and several authors have demonstrated the existence of a very long term mean reverting behaviour towards a long term equilibrium price (Bessembinder et al. [1995]; Pindyck [1999]; Postali and Picchetti [2006]; Maslyuk and Smyth [2008]). This phenomenon is mainly due to the long term horizon associated to drilling projects, which means investment decisions from oil producers fit into a longer term time frame. The proposed model thus presents the interest of coping with two horizons of mean reversion, the first one which could be linked to the short to mid-term reaction of consumers or producers having capacity of production or being capable to reduce rapidly their production combined with a longer term behaviour from other producers which start or suspend investment projects on a longer term basis, hence the supply adjustments and thus price reversion on a longer time horizon. For its part, the long term equilibrium price or marginal cost of production is directly influenced by the extraction technologies as well as the discovery of new fields. This will be represented by the value of the parameter μ_ξ. I also tested the presence of a short term stochastic trend to the mean reverting dynamic of the latent factor χ_t by estimating its marginal contribution to the log likelihood but it turns out this was not improving the model fit which means that there the long term equilibrium price taking into account the marginal cost of production is embodied by the parameter μ_ξ. As far as the model goodness of fit is concerned, I should point out through the comparison of Table 15.7 and Table 15.8 that the addition of the long term mean reversion has significantly contributed to the likelihood of the proposed

model. If one considers indeed that the Schwartz and Smith [2000] model is a particular model of the SSX model where the mean reversion parameter γ is equal to zero then we can compare the difference of likelihood between the two models with a chi-squared distribution with 1 degree of freedom, corresponding to the number of parameters that has been constrained. One can now consider the test statistics for a 99% critical value. Given that the 99th percentile of a chi-squared distribution with 1 degree of freedom is 6.635 one can see that except for the 2006-2011 sample all the other sample likelihood have been improved by adding the mean reversion component. This is confirming the findings in the literature about the long term mean reversion of commodity prices (Bessembinder et al. [1995], Schwartz [1997],Postali and Picchetti [2006]).
Furthermore, Table 15.8 also displays a noticeable volatility in the trend parameters under the historical probability as well as the risk premiums associated to each latent factor. This feature has been frequently highlighted in the literature (Cortazar et al. [2015] for instance) since the simultaneous estimation of the spot dynamic and the associated risk premium is particularly perilous when one only considers the futures prices. This set of prices indeed only contains information about the risk neutral dynamic. Moreover if one shifts the trend associated to each latent factor with a constant:

$$\chi_t = \chi_t + \frac{\Delta}{\beta_t \gamma_t} \tag{15–1}$$

$$\xi_t = \xi_t - \frac{\Delta}{\beta_t \gamma_t} \tag{15–2}$$

$$\tag{15–3}$$

and accordingly modifies the respective risk premium:

$$\lambda_{\chi_t} = \lambda_{\chi_t} - \frac{\Delta}{\gamma_t} \tag{15–4}$$

$$\lambda_{\xi_t} = \lambda_{\xi_t} + \frac{\Delta}{\beta_t} \tag{15–5}$$

$$\tag{15–6}$$

then no changes have been made to equation 14–40 and hence once has an equivalent expression. These two arguments explain why for all the data sub-samples, the estimation of the two risk premiums as well as the true trend obtained with the models look very large and chaotic. Nevertheless, if one computes now the sub-sample average risk neutral drifts, as the parameter one ultimately estimates during the inference procedure, one obtains a value approximately equals to 8% for all the sub-sample estimates, which is not an absurd value and is reasonably close to the results obtained by Schwartz and Smith [2000].

Table 15.7: Parameter estimates of Schwartz-Smith model (no covariates).

Variable	**90-95**	**95-00**	**00-06**	**06-11**	**11-16**
β	1.4651 (0.0329)	0.8619 (0.0245)	1.0460 (0.0473)	0.6604 (0.0265)	0.6702 (0.0237)
σ_χ	0.2976 (0.0146)	0.2891 (0.0126)	0.2663 (0.0136)	0.2744 (0.0130)	0.2378 (0.0121)
λ_χ	-0.0105 (0.0290)	0.0760 (0.0222)	-0.0183 (0.0251)	-0.4678 (0.0219)	-0.6038 (0.0236)
σ_ξ	0.1397 (0.0062)	0.1518 (0.0066)	0.1669 (0.0076)	0.2658 (0.0114)	0.1802 (0.0084)
λ_ξ	-0.0771 (0.0637)	0.0629 (0.0646)	0.2511 (0.0756)	0.1166 (0.1336)	-0.1900 (0.0817)
$\rho_{\chi\xi}$	0.2848 (0.0676)	-0.2631 (0.0617)	0.1217 (0.0728)	0.1119 (0.0692)	0.2779 (0.0766)
μ_ξ	-0.0602 (0.0636)	0.0686 (0.0646)	0.2048 (0.0761)	0.0506 (0.1329)	-0.2414 (0.0818)
Std Pricing error τ_1	0.0260 (0.0012)	0.0435 (0.0019)	0.0229 (0.0019)	0.0267 (0.0013)	0.0204 (0.0012)
Std Pricing error τ_2	0.0002 (0.0014)	0.0091 (0.0004)	0.0033 (0.0013)	0.0049 (0.0002)	0.0050 (0.0004)
Std Pricing error τ_3	0.0032 (0.0002)	0.0000 (0.0004)	0.0020 (0.0005)	0.0000 (0.0002)	0.0006 (0.0003)
Std Pricing error τ_4	0.0000 (0.0002)	0.0000 (0.0002)	0.0000 (0.0002)	0.0000 (0.0002)	0.0000 (0.0003)
Std Pricing error τ_5	0.0041 (0.0002)	0.0047 (0.0002)	0.0043 (0.0004)	0.0024 (0.0001)	0.0035 (0.0002)
NLL	-4318	-4108	-4273	-4437	-4444

Table 15.8: Parameter estimates of Extended Schwartz-Smith (SSX) model (no covariates).

Variable	**90-95**	**95-00**	**00-06**	**06-11**	**11-16**
β	0.1792 (0.0127)	1.1320 (0.0371)	0.8687 (0.0532)	0.0001 (0.0127)	0.0909 (0.0071)
σ_χ	0.1871 (0.0088)	0.3397 (0.0192)	0.2807 (0.0156)	0.2661 (0.0127)	0.2355 (0.0108)
λ_χ	-0.0046 (0.0045)	0.1712 (0.0424)	0.0370 (0.0265)	0.0655 (0.0092)	0.0083 (0.0044)
σ_ξ	0.2933 (0.0130)	0.2971 (0.0193)	0.1821 (0.0094)	0.2652 (0.0137)	0.1929 (0.0107)
γ	1.7901 (0.0438)	0.3782 (0.0173)	0.0422 (0.0085)	0.6753 (0.0245)	0.9611 (0.0356)
λ_ξ	0.1152 (0.1243)	0.1225 (0.1345)	0.2759 (0.0848)	-0.0740 (0.1182)	-0.0610 (0.0833)
$\rho_{\chi\xi}$	0.2010 (0.0612)	-0.5639 (0.0585)	-0.1006 (0.0947)	0.1221 (0.0666)	0.1107 (0.0740)
μ_ξ	5.4037 (0.1842)	1.2561 (0.1622)	0.3831 (0.0993)	2.3741 (0.1431)	3.6727 (0.1605)
Std Pricing error τ_1	0.0246 (0.0011)	0.0293 (0.0013)	0.0253 (0.0016)	0.0266 (0.0013)	0.0162 (0.0009)
Std Pricing error τ_2	0.0000 (0.0024)	0.0000 (0.0012)	0.0056 (0.0006)	0.0049 (0.0002)	0.0002 (0.0010)
Std Pricing error τ_3	0.0031 (0.0001)	0.0028 (0.0001)	0.0009 (0.0003)	0.0000 (0.0002)	0.0025 (0.0002)
Std Pricing error τ_4	0.0000 (0.0002)	0.0000 (0.0002)	0.0000 (0.0003)	0.0000 (0.0002)	0.0000 (0.0002)
Std Pricing error τ_5	0.0036 (0.0002)	0.0042 (0.0002)	0.0037 (0.0002)	0.0024 (0.0001)	0.0039 (0.0002)
NLL	-4381	-4326	-4298	-4443	-4501

15.3.2 Sensitivity analysis

In this section, I stress the financial meaning of each parameter estimated in the model and try to assess the impact or response over time of the crude oil futures curve following an instantaneous shock or impulse on one of the parameters of the model. In the next section, I propose to decompose the dynamic of each latent factors parameter as an affine function of the covariates described earlier. As a consequence, it is necessary to present beforehand the consequences on the curve level, its slope or even the crude oil convenience yield of a marginal parameter change. But for ease of understanding it is worth re-emphasizing before that the significance of each parameter in light of the calibration results. As one can notice in Table 15.8, the order associated to the two mean reverting parameters, namely γ and β, can from one sub-sample to another be interchangeable - meaning that a qualification of the two latent factors as being respectively a short and a long term component is not straightforward and needs some clarifications. If one takes the period 2011-2016 it can be seen that the values associated to the parameter β, (+0.0909) and γ, (+0.9611) corresponds respectively to equivalent half-lives of 3.31 and 0.31[1]. Thus on average it will take 3.31 years for the random variable χ_t to cross half of the distance existing at time t between the initial value χ_t and the long term average value μ_χ. Similarly, it will take 0.31 years to reach the midpoint between ξ_t and its long term expected value 0. As can be seen in this case, the speed of mean reversion for the process χ_t is below that of ξ_t, which means that contrary to the interpretation made in the paper Schwartz and Smith

[1] I define the half-life as the average time necessary for the process to revert half-way from the mean. In such a case I want to find the time $t^\star$ when the expected value of $\chi(t)$ will reach the middle point between $\chi(0)$ and the long term mean μ_χ. To compute this value one just needs to make the distance between the expected value of the mean reverting process $\chi(t)$ and its long term mean equal half of the distance between the initial value and the long term mean, which leads to the following equality:

$$t^\star \in \mathbb{R}^+ \quad s.t. \quad \mathbb{E}\left[\chi\left(t^\star\right) - \mu_\chi\right] = \frac{\chi\left(0\right) - \mu_\chi}{2}$$

which leads to the following solution:

$$\begin{aligned} e^{-\beta t^\star}\left[\chi\left(0\right) - \mu_\chi\right] &= \tfrac{\chi(0) - \mu_\chi}{2} \\ t^\star &= \tfrac{log2}{\beta} \end{aligned}$$

[2000], for the model proposed in this thesis and specifically for this sub-sample, the latent factor χ_t does not embody the short term factor but instead a long term dynamic. However, the second latent factor ξ_t for his part is the combination of a long term expected value or equilibrium price and a quite short cyclicality around it. Naturally, this does not tarnish the interest of the proposed model, but it is necessary to adjust the parameter interpretation accordingly. In order to plot the impulse response charts one needs to calculate the average backwardation of log prices, which is given by the following expression:

$$
\begin{aligned}
\text{Avg. Bwd of Log Prices}_t &= \mathbb{E}[lnF_t(\tau_1) - lnF_t(\tau_5)|\xi_0, \chi_0] \\
&= \left(e^{-\beta\tau_1} - e^{-\beta\tau_5}\right)\chi_0 e^{-\beta t} \\
&\quad + \left(e^{-\gamma\tau_1} - e^{-\gamma\tau_5}\right)\left(\xi_0 e^{-\gamma t} + \frac{\mu}{\gamma}\left(1 - e^{-\gamma t}\right)\right) \\
&\quad - \frac{\sigma_\chi^2}{4\beta}(e^{-2\beta\tau_1} - 1) - \frac{\sigma_\xi^2}{4\gamma}(e^{-2\gamma\tau_1} - 1) + \frac{\lambda_\chi}{\beta}(e^{-\beta\tau_1} - 1) \\
&\quad - \frac{1}{\gamma}(\mu_\xi - \lambda_\xi)(e^{-\gamma\tau_1} - 1) - \frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{(\beta+\gamma)}(e^{-(\beta+\gamma)\tau_1} - 1) \\
&\quad + \frac{\sigma_\chi^2}{4\beta}(e^{-2\beta\tau_5} - 1) + \frac{\sigma_\xi^2}{4\gamma}(e^{-2\gamma\tau_5} - 1) - \frac{\lambda_\chi}{\beta}(e^{-\beta\tau_5} - 1) \\
&\quad + \frac{1}{\gamma}(\mu_\xi - \lambda_\xi)(e^{-\gamma\tau_5} - 1) + \frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{(\beta+\gamma)}(e^{-(\beta+\gamma)\tau_5} - 1)
\end{aligned} \tag{15–7}
$$

Tables 15.9 and 15.10 show the instantaneous and asymptotic ($t \to \infty$) sensitivity of the average backwardation to shocks on μ, β or γ. Expressions for the sensitivity of the average backwardation of log prices at time t to shocks on μ, β or γ can be seen in Appendix E. From these tables one can see the time-varying nature of the instantaneous sensitivities over the five time periods. In addition, it can be seen that the equilibrium sensitivity of the backwardation to the μ parameter is zero. Furthermore, one can plot the impulse response charts over each 5-year period as shown in Figures 15.3 to 15.7. One can see in these figures the effects of increasing/decreasing the various parameter values by 25%. In particular, it can be noted that increasing the γ parameter results in a positive change to the level of backwardation, whereas decreasing the γ parameter results in a negative change. The converse of this is true for the μ and β parameters.

Table 15.9: Instantaneous Sensitivity of Average Backwardation.

	90-95	**95-00**	**00-06**	**06-11**	**11-16**
$\frac{\partial AB_t}{\partial \mu}\|_{t=0}$	-0.44	-1.01	-1.29	-0.83	-0.69
$\frac{\partial AB_t}{\partial \beta}\|_{t=0}$	0.15	-0.03	0.04	0.17	0.18
$\frac{\partial AB_t}{\partial \gamma}\|_{t=0}$	1.30	3.08	3.89	2.75	2.45

Table 15.10: Equilibrium Sensitivity of Average Backwardation.

	90-95	**95-00**	**00-06**	**06-11**	**11-16**
$\lim_{t\to\infty} \frac{\partial AB_t}{\partial \mu}$	0	0	0	0	0
$\lim_{t\to\infty} \frac{\partial AB_t}{\partial \beta}$	0.031	-0.060	-0.001	0.010	0.041
$\lim_{t\to\infty} \frac{\partial AB_t}{\partial \gamma}$	-0.017	-0.057	-0.235	0.066	0.035

15.3.3 Impact of Fundamental Variables Upon the Crude Oil Futures Term Structure

In order to analyse the informational content of the crude oil prices term structure I considered two different data samples. While the two samples include exactly the same covariates I decided to base the model selection procedure on the quality of the calibration when I was using backward looking data and simultaneous covariates data. As a consequence, to determine the model parameters in the first case I consider at time t the covariates values for the previous 8 weeks, whereas in the second case I consider the last available covariates value. The average likelihoods for the statistically significant covariates show that most of time the regression models for the three parameters γ, β, μ are noticeably improved when one considers only the latest data available for each covariates.

Table 15.11 shows the respective covariates impact upon the components of the crude oil spot dynamic following from the HMF model, namely the short-term and

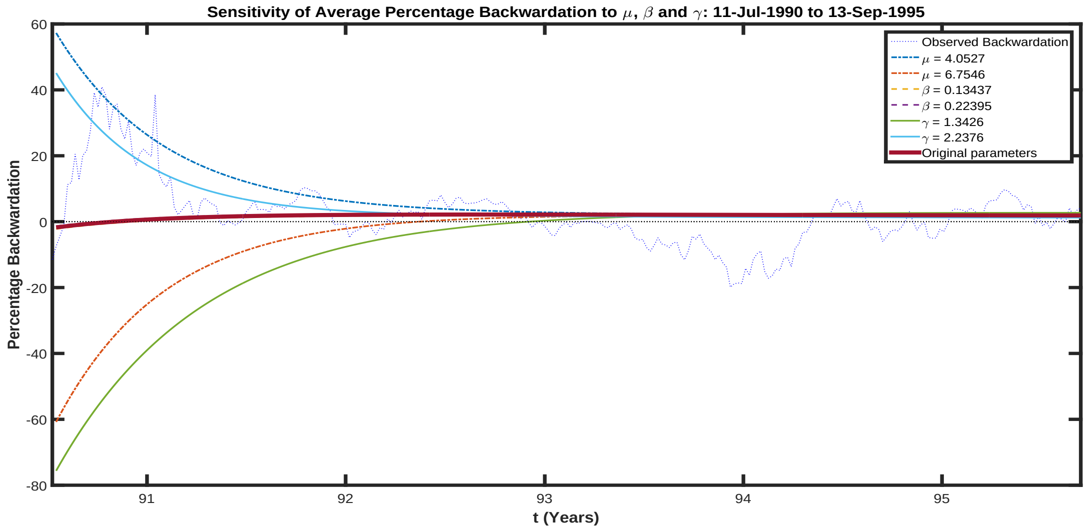


Figure 15.3: Sensitivity of Average Percentage Backwardation to μ, β and γ during the period 1990 to 1995.

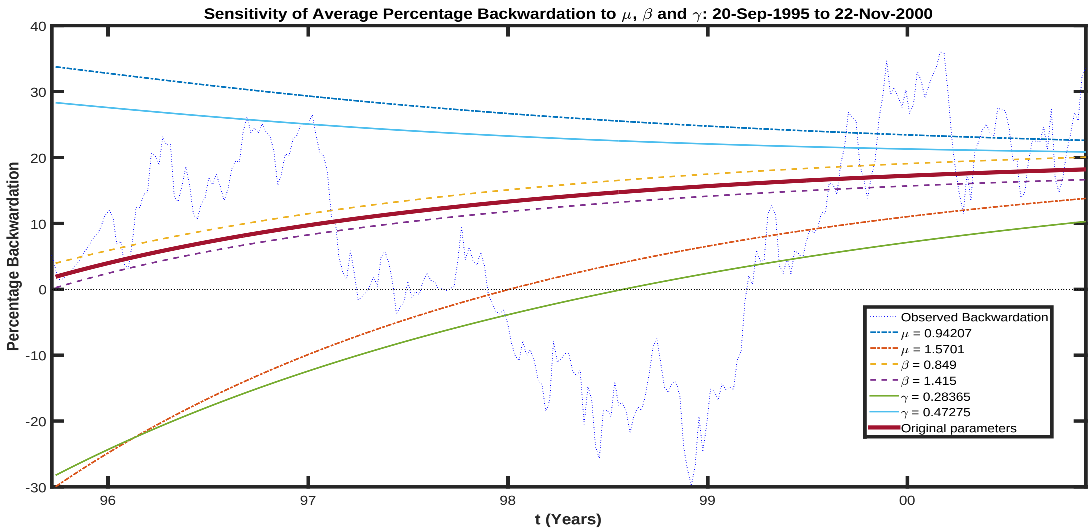


Figure 15.4: Sensitivity of Average Percentage Backwardation to μ, β and γ during the period 1995 to 2000.

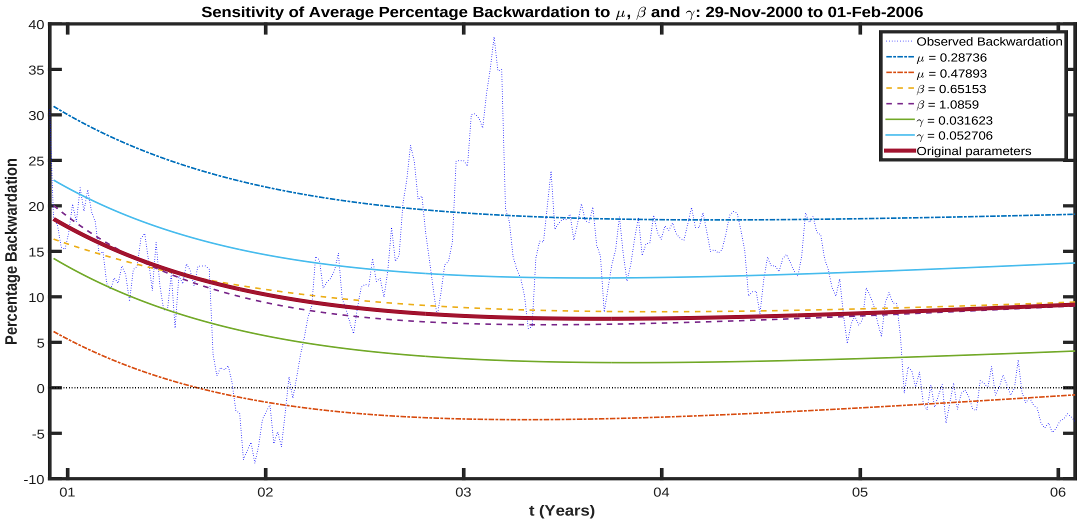


Figure 15.5: Sensitivity of Average Percentage Backwardation to μ, β and γ during the period 2000-2006.

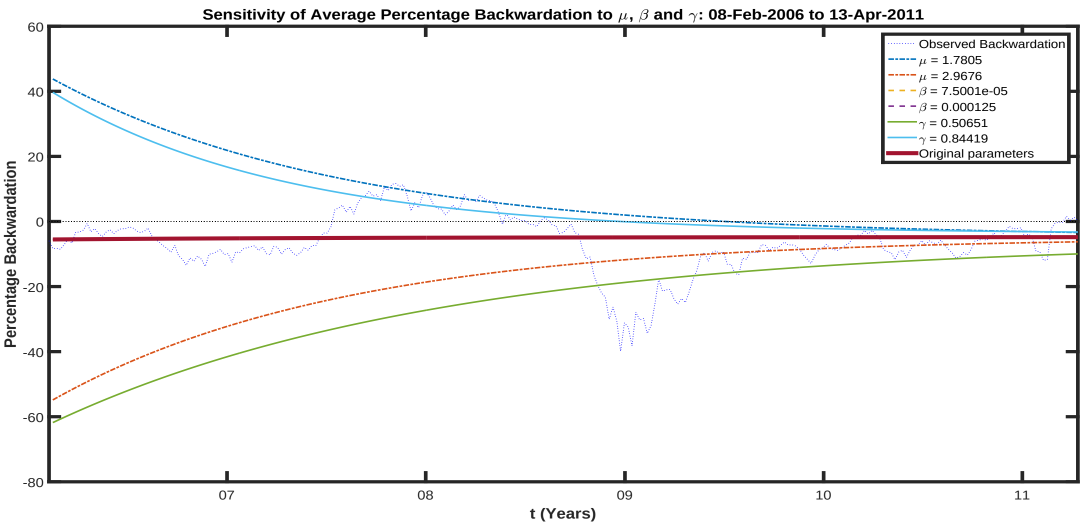


Figure 15.6: Sensitivity of Average Percentage Backwardation to μ, β and γ during the period 2006-2011.

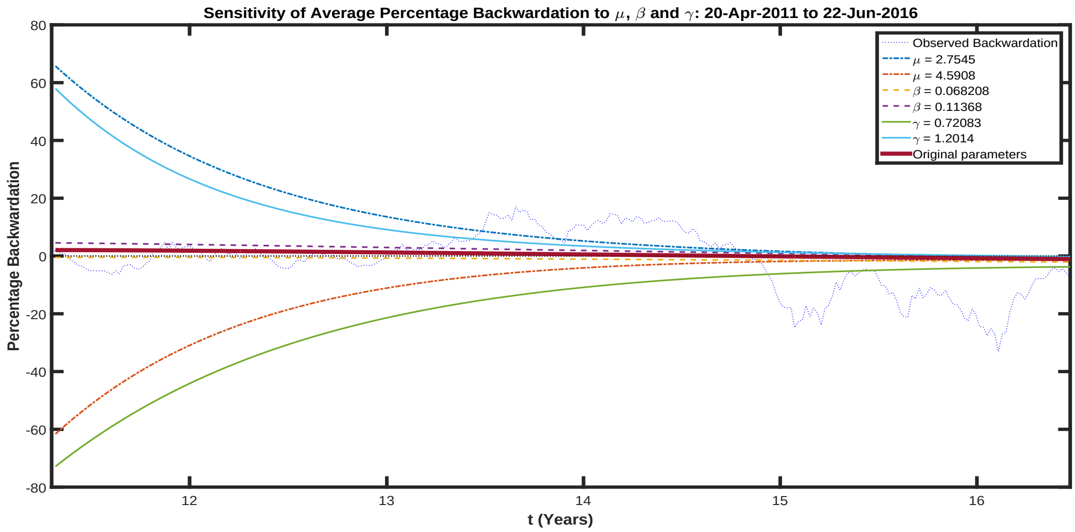


Figure 15.7: Sensitivity of Average Percentage Backwardation to μ, β and γ during the period 2011-2016.

the long-term mean reversion factors as well as the long term trend. The detailed model parameter estimates for each of the five periods can be seen in Appendix F. It can first and foremost be noticed that the relevant covariates are not necessarily the same among the three latent factors parameters and thus that each factor influences the dynamic of the crude oil price term structure in a different manner (as demonstrated also by Dempster et al. [2012]). Furthermore, it can be observed that this influence can evolve over the course of time. For instance, when one looks at the factors impacting the long term trend μ_ξ and the long term mean reversion parameters (which could be the β or the γ parameter according to the related half-life value as explained before) one notices that the dollar was one of the most important factors between 1995 and 2011 period during which the dollar index went conspicuously up until the burst of the dot-com bubble when it was almost always decreasing until the end of 2008. The negative sign of the μ_ξ regression coefficient associated to the Dollar Index shows a inverse relation between the dollar and the price of the crude oil which means that when the dollar is going up the price of the crude oil, expressed in this currency, tends to decrease. This relation has also been described in Akram [2009]. Nevertheless, this negative relation associated to the DXY index in the regression of the long term trend μ_ξ

and its impact on the other components of the crude oil dynamic have been less influential over the last five years (see Table 15.11). This statement is completing the results published by Reboredo [2012] and Reboredo et al. [2014] which analyse the relation between the WTI price and a set of currencies between the 4th of January 2000 and the 5th of May 2012 and point out that the intensity of these relations can fluctuate across time and notably reached their climax during the 2008 financial crisis.

These results not only demonstrate that the effect of the dollar on the dynamics of the crude oil price has, over the last five years, faded out but also that at the same time the US production of oil has recently weighed a lot more on the dynamic of oil prices while it was not so influential in the past as can be seen in Table 15.11. This result quantifies the recent impact of the advances in the application of horizontal drilling and hydraulic fracturing in shale which have obviously modified the international supply and demand equilibrium and thus the oil price dynamic itself (Outlook [2013] and Dvir and Rogoff [2014]). Interestingly enough one can also notice that this decrease of the US energy dependence has not impacted the long term trend value but more the shape of the curve and the rate of reversion to the equilibrium price. The latest negative sign indeed shows that the noticeable increase of the US oil production over the last five years has significantly pushed the futures curve into contango and thus explains in part this recent change in the crude oil basis sign. This increase of the US oil production has to be put in parallel with the increase of the refinery utilization rate that has been observed lately and which has also positively impacted the curve slope and the current contango situation in the crude oil market according to the results presented here. Furthermore, according to the competitive rational expectations model of storage (Pindyck [1994], Routledge et al. [2000] , Casassus and Collin-Dufresne [2005] and see Gorton et al. [2013] for a detailed review of the literature) one should also find a statistically significant negative relation between the inventories and the level of backwardation of the curve which is indeed mostly the case in this model if you consider the long term mean reversion impact on the curve slope and thus the convenience yield. This negative relation indeed means that when the US oil inventories decrease the long term mean reversion accelerates which consequently, and as demonstrated in the previous subsection (see Table 15.9), generally increases persistently the backwardation of the futures term structure. Although the sign of the inventories coefficients associated to the long term mean reversion is mostly negative and statistically significant, this effect has however not been as meaningful as the US oil production or the refinery utilization rate for the last five years (as shown by the AIC criterion ranking). This is once again confirming that in order to measure oil scarcity, the level of inventories in the US is maybe not enough and it is necessary nowadays to take into consideration the fact that the US are themselves producing a large part of their energy needs and can potentially accommodate shocks to demand by adjusting their production. This is also pointed out in Dvir and Rogoff [2014], who propose an extended commodity storage model where they assume that the supply can shift from a restricted to an unrestricted regime. The authors demonstrate that since the crisis of 1973, following the Organization of Petroleum Exporting Countries members embargo, there has been a restricted regime where supply does not react to shocks on the oil demand because of the production capacity constraints or structural limitations, e.g. refinery capacity, railroad infrastructure etc. Moreover, the results presented in this chapter also echo the conclusion of Dvir and Rogoff [2014] regarding the stability and potential shift of regime towards an unrestricted supply where the US production can satisfy the shocks on demand[1]. According to the same authors, in such a case the relation between the inventories and oil price, being a function of the flexibility of the production, should be negative instead of positive. This explains why I obtain a negative relation between the inventories and the oil price level for the last period (sign of the inventories coefficient relative to the trend component) while this relation was mostly positive in the past.

The proposed model also shows interesting results about the relation between stock markets and goods price or inflation. This financial economics topic which is linked to the Fisher hypothesis, according to which the long-run relation between equities and goods price should be positive, has been contradicted by several studies showing a negative relation between the equity indices and the goods prices (Bodie [1976]; Fama [1981]; Lee [2010]). This point has led to the conclusion

[1] Dvir and Rogoff [2014], p. 127: "...regarding the availability of shale oil in the U.S., and assuming that the industry will remain competitive, within a few years we may be again in a period where increased demand can be easily met by more production from U.S. sources. This development may well reverse the long-run relationship between inventories and price".

that the common stocks are a good hedge against inflation and thus commodities a good diversifier for equity portfolio risk. However as recently shown in the literature this relation may revert, weaken or at least not be consistently significant over time (Kilian and Park [2009]; Büyükşahin et al. [2009]; Büyükşahin and Robe [2014]). This concurs with the results presented here, as can be seen in Appendix F the S&P500 is first and foremost almost always significantly impacting the three elements of the crude oil spot price dynamic and is often (three periods out of five) one of the three most important factors, according to the AIC criterion ranking of the significant factor. Furthermore, the impact on the long term trend is for the three periods of time spanning from 1990 to 2006 always negative, substantiating the findings of the literature mentioned before. Nevertheless, over the last ten years the results in Table 15.11 show the direction of the relation is flipping around and becoming positive. One could conjecture that this change in the sign of the relation between oil and equity is also linked to the last decade's significant increase of the US supply capacity which has reduced the impact of the demand shocks for precautionary reasons. This argument is in line with the findings of Kilian and Park [2009] who demonstrate that only demand shocks for precautionary reasons such as those following political disturbances in the middle east (such as the Iraqi invasion of Kuwait in August 1990) can generate a significantly negative relation between equities and crude oil prices. The authors also add that there is no evidence that supply shocks will have the same outcome.

Finally, when one looks at the financial sphere impact the results show that the impact of the hedging pressure upon the trend of the crude oil price is not obvious and even insignificant over the last decade. Nevertheless, the influence of the hedgers seems to fall back on the two mean reversion components of the crude oil dynamic which are directly linked to the slope of the futures curve. When the hedging pressure is increasing, which means that there are more producers hedging their exposure than processors, a net short position is thus appearing and is subsequently offset by the speculators who accept to bear the risk in return for which they ask for a risk premium Keynes [1930]. This risk premium is actually materialized by a backwardation situation as explained by Hirshleifer [1990] and through the excess returns of long-short hedging pressure portfolios in Basu and Miffre [2013]. Table F.4 in Appendix F shows that during the 2008 financial crisis the hedging pressure was not at all altering the long term equilibrium price, namely μ_ξ, but instead increasing the speed of the long term mean reversion which then increased the backwardation of the curve validating accordingly the aforementioned literature. This result means also that during the financial crisis of 2008 and the related large fluctuations of the crude oil price the speculator has more assumed the role of an insurance provider under the increasing hedgers pressure but has not necessarily impacted directly the trend of this commodity. In line with recent articles raising questions about the potential impact of the commodity market financialisation (Kilian and Lee [2014],Goldstein and Yang [2015]), this statement sheds light on the complex role played by the speculators in the energy futures market during the 2008 crisis which should not be boiled down to the trend following strategies implemented by certain hedge funds as claimed in different articles (Kaufmann and Ullman [2009], Cifarelli and Paladino [2010], Kaufmann [2011]).

A final point should be made about the role of the financial institutions in the market for the period of 2006 to 2016, while the results show an insignificant influence of the hedging pressure on the curve dynamic from 2011 to 2016 it seems nonetheless that the leverage ratio associated to the financial intermediaries has significantly influenced the curve dynamic for the last five years and for the other periods under scrutiny. If one considers the 2008 spike and collapse and the 2015 spikes of this variable (see Figure 15.2), it seems that they impacted differently the curve slope as well as the crude oil long term equilibrium price. If one looks at the 2006 to 2011 period it can be seen that the crude oil collapse can be explained by the impact of the leverage factor on the long term trend (see Table F.4 in Appendix F) which confirms the findings of Acharya et al. [2013], who demonstrate that a limitation in the risk-taking capacity of financial institutions (which corresponds to a decrease of the financial institutions leverage ratio) could mechanically generate a hedging pressure on the futures curve and thus a backwardation situation as well as a decrease of the crude oil price, leading to a positive relation between the leverage ratio and the oil price and a negative relation between the same oil price and the backwardation. While the signs of the mean reversion and trend components obtained through the proposed model validate this theory for the 2006 to 2011 financial crisis it seems that the signs shift for the last five year invalidates it. According to the results for 2011 to 2016 it seems indeed that the recent spike (Q4 2015) of the financial intermediaries leverage factor has on the contrary negatively impacted the long term trend and has also increased the backwardation of the term structure.

	1990-1995			1995-2000			2000-2006			2006-2011			2011-2016		
Covariate	ST Mean Reversion	LT Mean Reversion	LT Trend	ST Mean Reversion	LT Mean Reversion	LT Trend	ST Mean Reversion	LT Mean Reversion	LT Trend	ST Mean Reversion	LT Mean Reversion	LT Trend	ST Mean Reversion	LT Mean Reversion	LT Trend
BDI										-0.009	-0.029				
DXY				-0.193	0.010	-0.032	-0.422	0.021	-0.070	0.009	0.032	-0.028			
End Stocks								-0.037	0.117						
GSCI	-0.042		0.127	0.177	-0.011	0.034						0.191	-0.042	0.054	0.168
Lev Rat	-0.020	0.210	0.053				-0.344	0.017	-0.054				0.011		-0.036
Ref Util		0.172												-0.019	
SP500	0.046	0.266	-0.129	-0.185	0.008	-0.025	-0.289				0.026	0.045	-0.017	-0.051	0.070
Hedging Pressure															
US Prod										0.002				-0.036	

Table 15.11: Three Highest AIC Criterion Contributors. In this table the three most important contributors to the log likelihood among the significant regressions are selected. To do so only the statistically significant factors are retained and their associated Akaike Information Criterion computed, then ranked in descending order and only the three highest values are retained. The value corresponds to the coefficient associated to the regression of the three model components on the macroeconomic and microeconomic independent variables.

15.3.4 Backwardation Changes Due to Perturbing Covariates: a Stress Scenario Analysis

Using the fitted HMF models one can assess the impact that a shock to each of the covariates will have on the futures curve. Figures 15.9 to 15.11 show the effect to the percentage backwardation of the futures curve resulting from a three[1] standard deviation increase to the covariate value during the period 2011 to 2016. Here, percentage backwardation is calculated as the percentage difference in futures price of the one month contract to the 5 month contract, i.e. $100 \times (F_t(\tau_1) - F_t(\tau_2))/F_t(\tau_1)$. It can be seen that such a stress test confirms the previous analysis based on the AIC criterion. In particular, the impact of the US production on the slope of the futures term structure over the last five years as shown in Figure 15.10 is more substantial than the effect of inventories and the other covariates taking into account a three standard deviation shock on the real values. As far as the short term mean reversion is concerned, namely γ for this period, one clearly sees that the GSCI as the average price of the commodities has a dominant impact on the contango of the crude oil futures curve.

15.4 Conclusions

This chapter contributes to the literature about commodity term structure dynamic modelling by proposing a model combining two mean reverting latent factors for which the stochastic dynamic can be expressed as a function of a set of macroeconomic covariates. Starting from the short term/long term model proposed by Schwartz and Smith [2000] the interest of adding a second mean reversion component with a higher half-life was demonstrated statistically and conceptually. Furthermore, this research contributes to the literature by proposing an innovative state-space framework which allows one to extract latent stochastic factors as well as all static model parameters in a statistically consistent manner. This model bridges the existing gap between the latent factor modelling literature

[1] The effects to the percentage backwardation resulting from one, two and three standard deviation increases to the covariate value were investigated here. Three standard deviations are presented in this section to aid in the interpretation of the plots, i.e. the lines are further apart and hence easier to see.

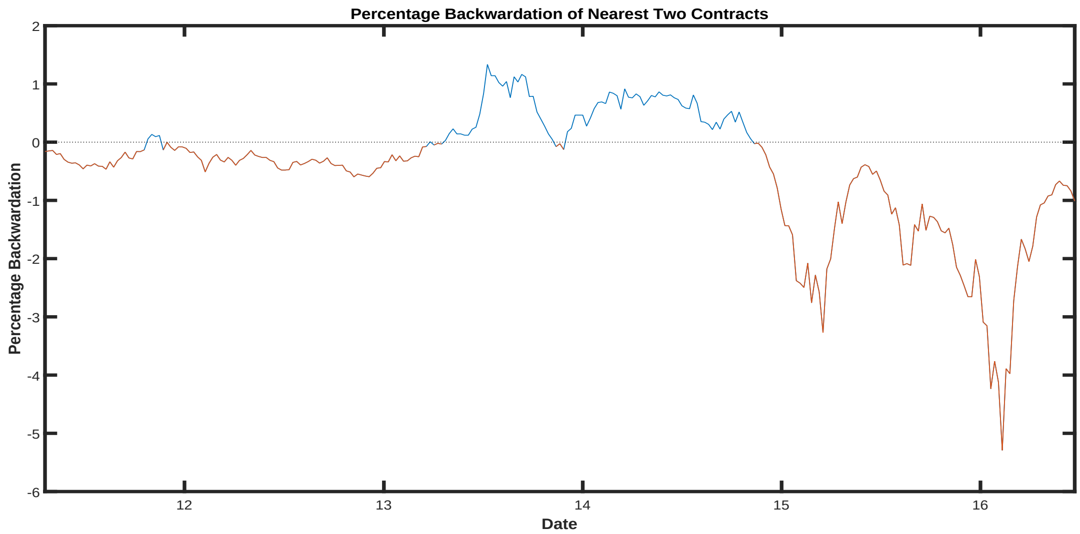


Figure 15.8: Percentage backwardation of the nearest two contracts during the period 2011-2016. The line is coloured blue when the the backwardation is positive and red when the backwardation is negative (i.e. contango).

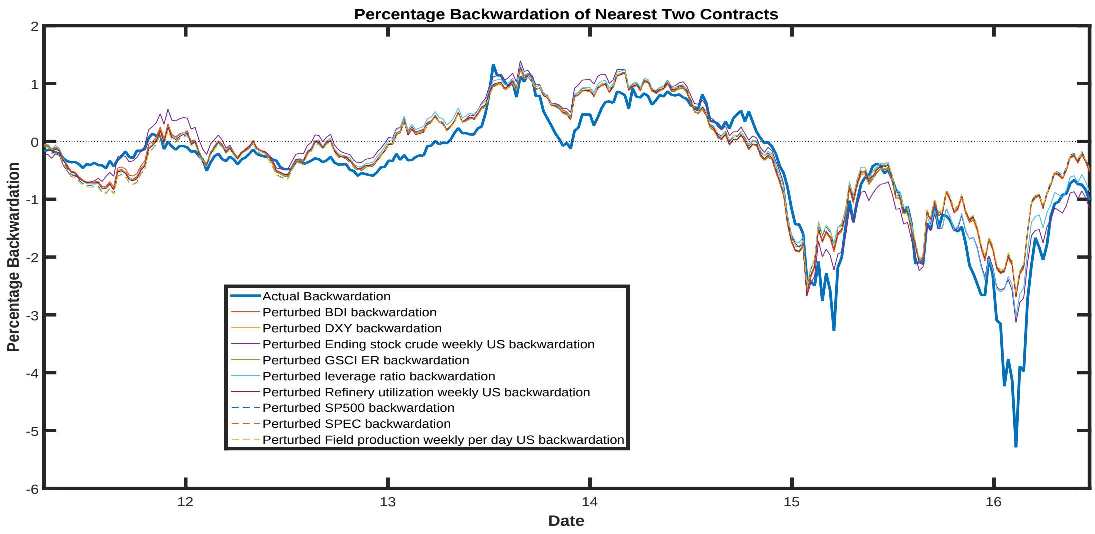


Figure 15.9: Percentage backwardation of the nearest two contracts resulting from a three standard deviation increase to the covariate value during the period 2011 to 2016. Here the fitted model links the covariate to the μ parameter.

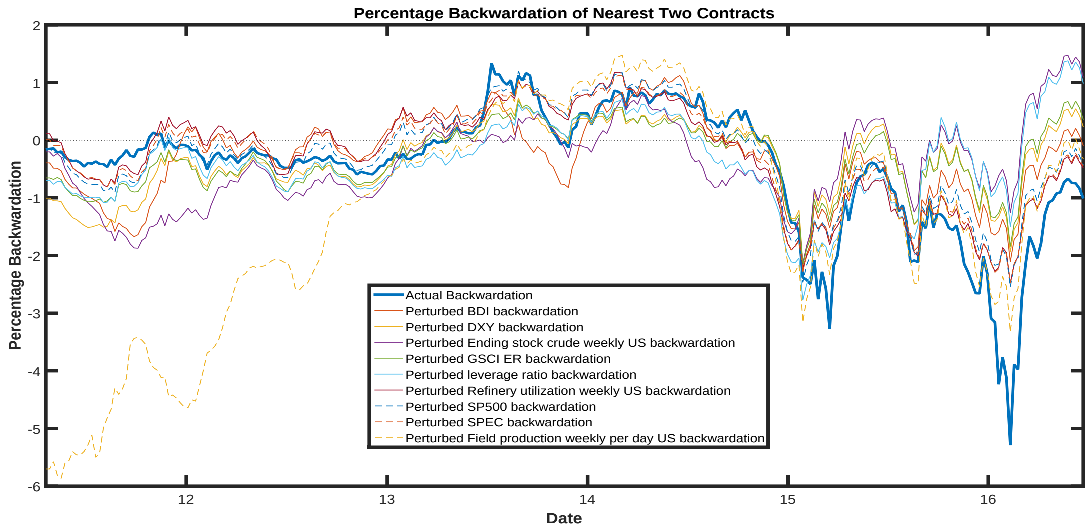


Figure 15.10: Percentage backwardation of the nearest two contracts resulting from a three standard deviation increase to the covariate value during the period 2011 to 2016. Here the fitted model links the covariate to the β parameter.

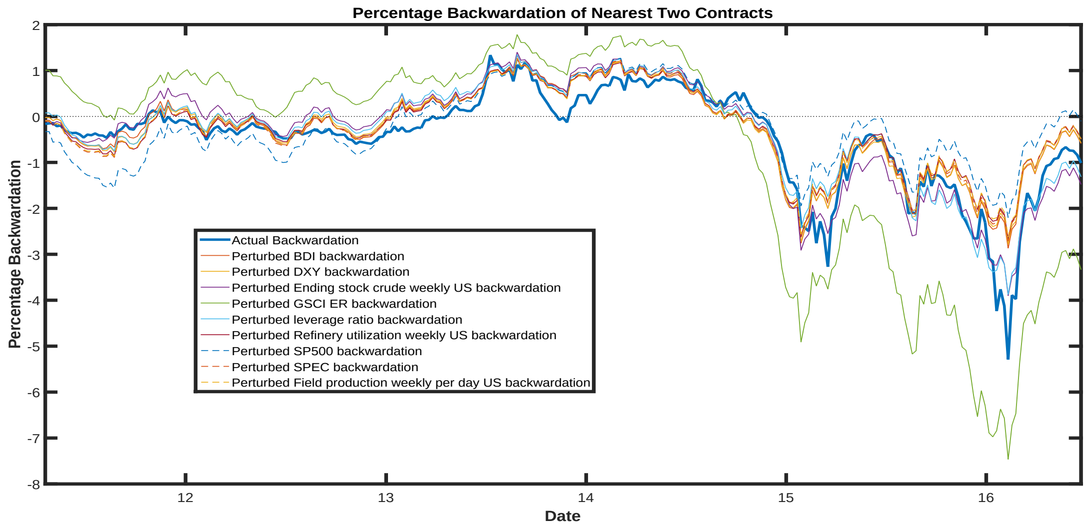


Figure 15.11: Percentage backwardation of the nearest two contracts resulting from a three standard deviation increase to the covariate value during the period 2011 to 2016. Here the fitted model links the covariate to the γ parameter.

and the two-step regression models generally proposed to explain the a priori estimated latent factors stochastic dynamics as functions of macroeconomic and microeconomic variables. Finally, the results presented here shed light upon several topical challenges raised in the literature about the relation between crude oil term structure behaviour and financial or physical information available in the market. Notably, one can conclude that the recent increase of the US oil production over the last decade has significantly influenced the behaviour of the crude oil long term equilibrium price but also the dynamic of the futures term structure emphasizing accordingly the interest of the extended commodity storage model proposed recently by Dvir and Rogoff [2014].

Chapter 16

Conclusions and Future Work

16.1 Summary

The research presented in this thesis constitutes an important and novel contribution towards modelling dependence in financial applications. In particular, contributing to the literature by further understanding the key factors driving the dynamic nature of such dependence. This thesis focuses on two key financial applications: modelling multiple-currency basket returns and modelling commodity prices.

16.2 Statistical Modelling and Estimation Contributions

Three complementary dependence modelling approaches are developed in this thesis. The first two approaches address the challenge of modelling the multivariate distribution of a portfolio of asset returns. The third approach developed concerns commodity price dependence modelling where the link between maturities through the term structure of futures prices is considered. Firstly, a parametric copula modelling approach is considered in order to capture the complex dependence structure present in such data. In particular, flexible mixture copula models, consisting of weighted Archimedean copula members such as Clayton, Frank and Gumbel components, are developed including additional structural flexibility via distortion transforms corresponding to inner and outer-transform variants estimated via the inference for margins method which consists of a two step fitting procedure for the marginal model and then the dependence structure. In addition, an expectation-maximisation method is considered.

Secondly, a covariance factor regression framework is utilised in order to understand the influence of observed covariates on the covariance of the multivariate distribution of a portfolio of asset returns. This framework provides a number of desirable properties. Crucially, the model is interpretable in a way that GARCH-type models are not and as such, forecasting the covariance matrix is straightforward and transparent. This is achieved by constructing time series models for the observed covariates and calculating forecasts, which are then used as inputs to the covariance matrix forecast. Furthermore, the estimation of the covariance factor model can be performed using a simple and efficient Expectation-Maximization (EM) algorithm. A sensitivity analysis of the covariance matrix to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression.

The resulting forecasts of the covariance matrix of asset returns can then be utilised in portfolio optimisation. In particular, this modelling framework allows one to calculate the sensitivity of the portfolio weights to the observable covariance factors and accordingly helps to devise a global and dynamic hedging strategy for portfolios of assets. Thus, the relationship between interpretable factors and the weightings of assets in a portfolio can be further understood.

Thirdly, a novel Hybrid Multi-Factor (HMF) state-space modelling framework is also proposed in order to understand the key factors driving the dependence structure among commodity futures prices along their term structure. A consistent estimation framework is developed, which builds on the familiar two-factor model of Schwartz and Smith (2000), to allow for an investigation of the influence of observable covariates on commodity prices. Using this novel Hybrid Multi-Factor (HMF) model, it is possible to obtain closed form futures prices under standard risk neutral pricing formulations, and one can incorporate state-space model estimation techniques to consistently estimate both the structural features related to the convenience yield and spot price dynamics (long and short term stochastic dynamics) and also the structural parameters that relate to the influence on the spot price of the observed exogenous covariates. Such models can then be utilised to gain significant insight into the futures and spot price dynamics in terms of interpretable observed factors that influence speculators and hedgers heterogeneously. This is not attainable with existing modelling approaches.

16.3 Novel Insights into Finance and Econometric Studies

This thesis also contributes to the literature by the application of the dependence structure modelling described above to two challenging financial modelling problems: modelling multiple-currency basket returns and modelling commodity futures price term structure. In order to perform the empirical analyses considered in this thesis a substantial amount of effort and time was invested into collecting, cleaning and preparing the data.

Multiple Currency Basket Modelling

Firstly, this thesis investigates the well-known financial puzzle of the currency carry trade, which is yet to be satisfactorily explained. It is one of the most robust financial puzzles in international finance and has attracted the attention of academics and practitioners alike for the past 25 years. The currency carry trade is the investment strategy that involves selling low interest rate currencies in order to purchase higher interest rate currencies, thus profiting from the interest rate differentials. Assuming foreign exchange risk is uninhibited and the markets have rational risk-neutral investors, then one would not expect profits from such strategies. That is uncovered interest rate parity (UIP); the parity condition in which exposure to foreign exchange risk, with unanticipated changes in exchange rates, should result in an outcome that changes in the exchange rate should offset the potential to profit from such interest rate differentials.

A dataset of daily closes on spot and one month forward contracts for 20 currencies from 2000 to 2013 was used to investigate the behaviour of carry portfolios, formed by sorting on the forward premium (a proxy to the interest rate differential to US dollar). A rigorous statistical modelling approach is proposed, which captures the specific statistical features of both the individual currency log-return distributions as well as the joint features, such as the dependence structures prevailing between the exchange rates.

The individual currency returns were transformed to standard uniform margins after fitting appropriately heavy tailed marginal models, namely log-normal and log generalised gamma models. To analyse the tail dependence present in the carry portfolios - mixture copula models, consisting of weighted Clayton, Frank and Gumbel components, were fitted on a rolling daily basis to the previous six months of transformed log returns. Extracting and interpreting the multivariate tail dependence present in the rolling daily baskets provided significant evidence that the average excess returns earned from the carry trade strategy can be attributed to compensation for not only individual currency tail risk, but also exposure to significant risk of large portfolio losses due to joint adverse movements.

A key contribution of this thesis is therefore to provide a rationale for the unintuitive excess returns seen empirically in the currency carry trade via the presence of multivariate tail dependence and therefore increased portfolio crash risk. This is a novel and promising approach. A further contribution of this research is the identification of significant periods of carry portfolio construction and unwinding through the analysis of multivariate tail dependence in mixture copula models.

From a fundamental perspective this thesis also explores the impact of speculative trading behaviour on the dependence structure of currency returns. The ratio of speculative open interest (net non-commercial positions) to total open interest, termed the $SPEC$ factor, is shown to provide a good proxy to the behaviour of carry trade investors via a PCA analysis and consequently the resulting complex nonlinear relation between international exchange rates.

To investigate this phenomenon, a covariance regression modelling approach whereby the influence of observed covariates on the covariance of the multivariate returns of a basket of assets is proposed. In particular, the impact of speculative trading behaviour, i.e. the $SPEC$ factors, on the covariance of carry currencies is investigated. These $SPEC$ factors are shown to hold several orders of magnitude more explanatory power than the price index factors, DOL and HML_{FX}, previously suggested in the literature. Furthermore, it is demonstrated that the time series for the DOL and HML_{FX} factors are very close to white noise and as such are essentially unforecastable. It is important to note that the DOL and HML_{FX} covariates are risk premia and therefore shouldn't be expected to be forecastable, since otherwise there is no risk to be compensated for. The suggested speculative open interest factors are shown to be amenable to ARIMA model fits and so produce reasonable forecast accuracy.

A sensitivity analysis of the covariance to the factors is also presented allowing the estimation of a confidence interval of the covariance matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. In addition, a regression of the tail dependence measures, obtained from the mixture copula modelling approach, on the $SPEC$ factors illustrates the influence of carry trade speculative behaviour on the extremal joint currency returns. The DOL and HML_{FX} are shown to hold little explanatory power in the joint tails.

Commodity Price Modelling

In addition, this thesis employs a state-space modelling approach to understand the joint dynamic of the commodity spot price and the related futures prices along the curve. This framework is extended to allow for an investigation of the influence of observed macroeconomical covariates on the commodity term structure and in particular whether these covariates affect the short or long end of the curve. This modelling can be used for risk management, derivatives pricing, real options analysis and (carry) strategy development, e.g. backwardation/contango plays.

In particular, in this thesis the focus is on the behaviour of oil prices. Oil has historically been one of the most closely scrutinized commodities in the market. First and foremost, this is because of the important role this commodity plays in the worldwide economy and international relations, which gives it a prominent role, when compared to other energy, agricultural and metals commodities, in many aspects of the global economy and each country's specific macro, micro and monetary economic policy decisions.

Historically, one has seen the importance that economies have placed on the price variation of oil and understanding the factors that affect such a dynamic in order to better understand the determinants of shocks and volatility regimes in the spot price, demand and supply.

Another determining reason for the continued interest lies in the frequent shocks affecting the supply and demand of the so called "black gold" giving birth to sudden and dramatic price movements, such as during the 1973/74 oil crisis. The price of this exhaustible commodity has indeed been in the past heavily impacted by the discovery of new fields or the conflicts in oil-producing countries. On the other hand, the demand behaviour has generally been more influenced by the business cycles or even the evolution of the extracted oil inventories. That being said, according to the US Department of the Interior (DOI) as well as the US Energy Information Administration (EIA), the technology used for its extraction has recently been the main factor influencing the market supply. Over the last decade, advances in the application of horizontal drilling and hydraulic fracturing in shale have indeed drastically modified the international supply and demand equilibrium as well as the existing international relations by allowing the biggest oil consumer, namely the United States, to become over the same time period less and less dependent on its energy imports. According to the EIA, in 2015, 24% of the petroleum consumed in this country was imported which corresponds to the lowest level since 1970.

From a modelling perspective, such changes in the physical market conditions are significantly impacting the commodity price dynamic and need to be incorporated into any interpretable and realistic commodity futures stochastic model. In addition, if the model is developed, as is the case with the class of Hybrid Multi-Factor Models (HMF) introduced in this thesis, to allow for clear closed form representations of structural features such as sensitivity, shock transient response and perturbation influence on the model parameters and the driving exogenous covariates characterizing the features just discussed, then such a class of models has the potential to significantly aid in the study of stochastic variation in oil futures prices and to aid in forecasting and policy decision. The main aim of this research is to provide such a class of models and demonstrate their utility in incorporating a range of exogenous covariates into different structural components that will clearly explain short term and long term speculator and hedger positions in oil futures and their influences.

Finally, the results presented in this thesis shed light upon several topical challenges raised in the literature about the relation between crude oil term structure behaviour and financial or physical information available in the market. One can conclude that the recent increase of the US oil production over the last decade has significantly influenced the behaviour of the crude oil long term equilibrium price and also the dynamics of the futures term structure.

16.4 Future Research Directions

The novel modelling developments proposed and the practical applications considered in this research naturally suggest many interesting questions to investigate. Furthermore, we are currently completing research papers to extend the work performed in this thesis.

From a copula modelling perspective, further investigation into dynamic copulae and a comparison to the sliding window approach adopted in this thesis would be interesting. In particular, whether it is possible to identify periods of currency carry trade construction and unwinding through the change in dynamic copula parameter. In addition, investigating larger baskets of currencies (incorporating a wide range of developing countries) utilising vine copula models would be valuable.

Extending the covariance regression modelling framework to incorporate robust covariate selection and thus optimising portfolio covariance forecasting is a key research direction. This would allow the framework proposed in this thesis to be practically implemented and an optimal currency carry trade strategy to be performed.

The Hybrid Multi-Factor (HMF) modelling framework developed in this thesis has many possible extensions: firstly, the implementation of multiple covariates into each parameter link function; secondly, the implementation of multiple parameter linkings within the same model; thirdly, allowing for a more flexible dependence structure to enter into the model residuals. It would also be of much interest to apply this framework to the investigation of other commodities, for example grains, metals and other energy commodities.

Exploring the relationship between commodity prices and currency dependence dynamics in a causal framework would be particularly revealing in the context of the research presented in this thesis.

Appendices

Appendix A

Archimedean Copula Derivatives

A.1 Multivariate Clayton Copula

A.1.1 $C_\rho^C(\mathbf{u})$

$$C_\rho^C(\mathbf{u}) = \left(\sum_{i=1}^{d} u_i^{-\rho} - d + 1 \right)^{-\frac{1}{\rho}} \quad , \quad \rho > 0 \tag{A–1}$$

A.1.2 $\psi_\rho^{(d)}$: d-th derivative of the Clayton generator

$$(-1)^d \psi_\rho^{(d)}(t) = \frac{\Gamma\left(d + \frac{1}{\rho}\right)}{\Gamma\left(\frac{1}{\rho}\right)} (1+t)^{-\left(d+\frac{1}{\rho}\right)} \tag{A–2}$$

A.1.3 Clayton Copula Density $\left(\frac{\partial^d C}{\partial u_1 \ldots \partial u_d}\right)$

$$c_\rho^C(\mathbf{u}) = \prod_{k=0}^{d-1} (\rho k + 1) \left(\prod_{i=1}^{d} u_i \right)^{-(1+\rho)} \left(1 + t_\rho^C\left(\mathbf{u}\right)\right)^{\left(-d+\frac{1}{\rho}\right)} \tag{A–3}$$

where
$t_\rho^C(\mathbf{u}) = \sum_{i=1}^{d} \psi_C^{-1}(u_i)$

$\psi_C^{-1}(u_i) = (u_i^{-\rho} - 1)$

A.2 Multivariate Frank Copula

A.2.1 $C_\rho^F(\mathbf{u})$

$$C_\rho^F(\mathbf{u}) = -\frac{1}{\rho}\ln\left(1 + \frac{\prod_{i=1}^{d}(e^{-\rho u_i} - 1)}{(e^{-\rho} - 1)^{d-1}}\right) \quad , \quad \rho > 0 \tag{A–4}$$

A.2.2 $\psi_\rho^{(d)}$: d-th derivative of the Frank generator

$$(-1)^d \psi_\rho^{(d)}(t) = \frac{1}{\rho} Li_{-(d-1)}\left\{(1 - e^{-\rho})e^{-t}\right\} \quad , \quad t \in (0, \infty), \;\; d \in \mathbb{N}_0 \tag{A–5}$$

where $Li_s(z) = \sum\limits_{k=1}^{\infty} \frac{z^k}{k^s}$

A.2.3 Frank Copula Density $\left(\frac{\partial^d C}{\partial u_1 \ldots \partial u_d}\right)$

$$c_\rho^F(\mathbf{u}) = \left(\frac{\rho}{1 - e^{-\rho}}\right)^{d-1} Li_{-(d-1)}\left\{h_\rho^F(\mathbf{u})\right\} \frac{e^{\left(-\rho \sum\limits_{j=1}^{d} u_j\right)}}{h_\rho^F(\mathbf{u})} \tag{A–6}$$

where
$h_\rho^F(\mathbf{u}) = (1 - e^{-\rho})^{1-d} \prod\limits_{j=1}^{d} \left\{1 - e^{-\rho u_j}\right\}$

A.3 Multivariate Gumbel Copula

A.3.1 $C_\rho^G(\mathbf{u})$

$$C_\rho^G(\mathbf{u}) = e^{-\left(\sum\limits_{i=1}^{d}(-log\ u_i)^\rho\right)^{\frac{1}{\rho}}} \ ,\ \ \rho \geq 1 \tag{A–7}$$

A.3.2 $\psi_\rho^{(d)}$: d-th derivative of the Gumbel generator

$$(-1)^d \psi_\rho^{(d)}(t) = \frac{\psi_\rho(t)}{t^d} P_{d,\frac{1}{\rho}}^G \left(t^{\frac{1}{\rho}}\right) \ \ ,\ \ t \in (0,\infty),\ d \in \mathbb{N} \tag{A–8}$$

where

$P_{d,\frac{1}{\rho}}^G \left(t^{\frac{1}{\rho}}\right) = \sum\limits_{k=1}^{d} a_{dk}^G \left(\frac{1}{\rho}\right) (t^{\frac{1}{\rho}})^k$

$a_{dk}^G(\frac{1}{\rho}) = \dfrac{d!}{k!} \displaystyle\sum_{i=1}^{k} \binom{k}{i} \binom{\frac{i}{\rho}}{d} (-1)^{d-i}\ ,\ \ k \in 1,...,d$

A.3.3 Gumbel Copula Density $\left(\frac{\partial^d C}{\partial u_1 ... \partial u_d}\right)$

$$c_\rho^G(\mathbf{u}) = \rho^d e^{\left(-t_\rho(\mathbf{u})^{\frac{1}{\rho}}\right)} \frac{\prod\limits_{i=i}^{d}(-log\ u_i)^{\rho-1}}{t_\rho(\mathbf{u})^d \prod\limits_{i=1}^{d} u_i} P_{d,\frac{1}{\rho}}^G(t_\rho^G(\mathbf{u})^{\frac{1}{\rho}}) \tag{A–9}$$

where

$P_{d,\frac{1}{\rho}}^G(t^{\frac{1}{\rho}}) = \sum\limits_{k=1}^{d} a_{dk}^G(\frac{1}{\rho})(t^{\frac{1}{\rho}})^k$

$a_{dk}^G(\frac{1}{\rho}) = \dfrac{d!}{k!} \displaystyle\sum_{i=1}^{k} \binom{k}{i} \binom{\frac{i}{\rho}}{d} (-1)^{d-i}\ ,\ \ k \in 1,...,d$

$t_\rho^G(\mathbf{u}) = \sum\limits_{i=1}^{d} \psi_G^{-1}(u_i)$

$\psi_G^{-1}(u_i) = (-log\ u_i)^\rho$

A.4 Multivariate Clayton-Frank-Gumbel Mixture Copula

A.4.1 $C^{CFG}_{\rho_1,\rho_2,\rho_3}(\mathbf{u})$

$$\begin{aligned}
C^{CFG}_{\rho_C,\rho_F,\rho_G}(\mathbf{u}) = \;& \lambda_C(C^C_{\rho_C}(\mathbf{u})) + \lambda_F(C^F_{\rho_F}(\mathbf{u})) + \lambda_G(C^G_{\rho_G}(\mathbf{u})) \\
= \;& \lambda_C \times \left(\sum_{i=1}^{d} u_i^{-\rho} - d + 1\right)^{-\frac{1}{\rho}} \\
& + \lambda_F \times -\frac{1}{\rho} \ln\left(1 + \frac{\prod_{i=1}^{d}\left(e^{-\rho u_i} - 1\right)}{\left(e^{-\rho} - 1\right)^{d-1}}\right) \\
& + \lambda_G \times e^{-\left(\sum_{i=1}^{d}(-log\ u_i)^{\rho}\right)^{\frac{1}{\rho}}}
\end{aligned} \tag{A–10}$$

A.4.2 Clayton-Frank-Gumbel Mixture Copula Density

$$\begin{aligned}
c^{CFG}_{\rho_C,\rho_F,\rho_G}(\mathbf{u}) =& \lambda_C(c^C_{\rho_C}(\mathbf{u})) + \lambda_F(c^F_{\rho_F}(\mathbf{u})) + \lambda_G(c^G_{\rho_G}(\mathbf{u})) \\
=& \lambda_C \times \prod_{k=0}^{d-1}(\rho k + 1)\left(\prod_{i=1}^{d} u_i\right)^{-(1+\rho)} \left(1 + t^C_{\rho}(\mathbf{u})\right)^{\left(-d+\frac{1}{\rho}\right)} \\
& + \lambda_F \times \left(\frac{\rho}{1 - e^{-\rho}}\right)^{d-1} Li_{-(d-1)}\left\{h^F_{\rho}(\mathbf{u})\right\} \frac{e^{\left(-\rho \sum_{j=1}^{d} u_j\right)}}{h^F_{\rho}(\mathbf{u})} \\
& + \lambda_G \times \rho^d e^{\left(-t_{\rho}(\mathbf{u})^{\frac{1}{\rho}}\right)} \frac{\prod_{i=i}^{d}(-log\ u_i)^{\rho-1}}{t_{\rho}(\mathbf{u})^d \prod_{i=1}^{d} u_i} P^G_{d,\frac{1}{\rho}}(t^G_{\rho}(\mathbf{u})^{\frac{1}{\rho}})
\end{aligned} \tag{A–11}$$

where

$$t_\rho^C(\mathbf{u}) = \sum_{i=1}^{d} \psi_C^{-1}(u_i)$$
$$\psi_C^{-1}(u_i) = (u_i^{-\rho} - 1)$$

$$h_\rho^F(\mathbf{u}) = (1 - e^{-\rho})^{1-d} \prod_{j=1}^{d} \{1 - e^{-\rho u_j}\}$$

$$P_{d,\frac{1}{\rho}}^G(t^{\frac{1}{\rho}}) = \sum_{k=1}^{d} a_{dk}^G(\tfrac{1}{\rho})(t^{\frac{1}{\rho}})^k$$
$$a_{dk}^G(\tfrac{1}{\rho}) = \frac{d!}{k!} \sum_{i=1}^{k} \binom{k}{i} \binom{\frac{i}{\rho}}{d} (-1)^{d-i} \ , \quad k \in 1, ..., d$$
$$t_\rho^G(\mathbf{u}) = \sum_{i=1}^{d} \psi_G^{-1}(u_i)$$
$$\psi_G^{-1}(u_i) = (-log\ u_i)^\rho$$

Appendix B

Calculating Confidence Intervals for Covariance Regression

Approximate confidence intervals for model parameters can be provided by Wald intervals, i.e. the MLEs plus or minus a multiple of the standard errors, as described in Hoff and Niu [2012]. Standard errors can be obtained from the inverse of the expected information matrix evaluated at the MLEs. The log-likelihood given an observation $\boldsymbol{e}$ is $l(\boldsymbol{B}, \boldsymbol{\Psi} : \boldsymbol{e}) = \log p(\boldsymbol{e}|\boldsymbol{\Sigma}) = -(p \log 2\pi + \log|\boldsymbol{\Sigma}| + \boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e})/2$, where $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{\beta}\boldsymbol{x}$ and $\boldsymbol{\Sigma} = \boldsymbol{\Psi} + \boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{B}^{\boldsymbol{T}}$. The likelihood derivative with respect to $\boldsymbol{B}$ can be obtained as follows:

$$\begin{aligned} \dot{l}_{\mathrm{B}} = \partial l(\boldsymbol{B}, \boldsymbol{\Psi} : \boldsymbol{e})/\partial \boldsymbol{B} &= -(\partial \log|\boldsymbol{\Sigma}|/\partial \boldsymbol{B} + \partial \boldsymbol{e}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{e}/\partial\boldsymbol{B})/2 && \text{(B–1)} \\ &= -\Sigma^{-1}\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}} + \Sigma^{-1}\boldsymbol{e}\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{\boldsymbol{-1}}\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}} && \text{(B–2)} \\ &= H_{\mathrm{z}}\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}, && \text{(B–3)} \end{aligned}$$

where $\boldsymbol{H_z} = \boldsymbol{\Sigma}^{\boldsymbol{-1/2}}(\boldsymbol{z}\boldsymbol{z}^{\boldsymbol{T}} - \boldsymbol{I})\boldsymbol{\Sigma}^{\boldsymbol{-1/2}}$ and $\boldsymbol{z} = \boldsymbol{\Sigma}^{\boldsymbol{-1/2}}\boldsymbol{e}$. The derivative with respect to $\boldsymbol{\Psi}$ is more complicated, as the $p \times p$ matrix $\boldsymbol{\Psi}$ has only $p(p+1)/2$ free parameters. Following McCulloch [1982], we let $\boldsymbol{\psi} = \mathrm{vech}\boldsymbol{\Psi}$ be the $p(p+1)/2$ vector of unique elements of $\boldsymbol{\Psi}$. As described in that article, derivatives of functions with respect to $\boldsymbol{\psi}$ can be obtained as a linear transformation of derivatives with respect to $\boldsymbol{\Psi}$, obtained by ignoring the symmetry in $\boldsymbol{\Psi}$:

$$\begin{aligned} \dot{l}_{\Psi} = \partial l(\boldsymbol{B}, \boldsymbol{\Psi} : \boldsymbol{e})/\partial \boldsymbol{\Psi} &= -(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\boldsymbol{e}\boldsymbol{e}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-1})/2 && \text{(B–4)} \\ &= \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{z}\boldsymbol{z}^T - \boldsymbol{I})\boldsymbol{\Sigma}^{-1/2}/2 = \boldsymbol{H}_z/2, && \text{(B–5)} \\ \dot{l}_{\psi} = \partial l(\boldsymbol{B}, \boldsymbol{\psi} : \boldsymbol{e})/\partial \boldsymbol{\psi} &= \boldsymbol{G}^T \mathrm{vec}\, \dot{l}_{\boldsymbol{\Psi}} = \boldsymbol{G}^T \mathrm{vec}\, \boldsymbol{H}_z/2, && \text{(B–6)} \end{aligned}$$

where $\boldsymbol{G}$ is the matrix such that $\mathrm{vec}\, \boldsymbol{X} = \boldsymbol{G}\mathrm{vech}\, \boldsymbol{X}$, as defined in Henderson and Searle [1979]. Letting $\boldsymbol{b} = \mathrm{vec}\, \boldsymbol{B}$ and $\dot{l}_b = \mathrm{vec}\, \dot{l}_B$ the expected information is

$$\mathcal{I}(\boldsymbol{b}, \boldsymbol{\psi} : \boldsymbol{x}) = E_{b,\psi} \begin{bmatrix} \dot{l}_{\mathrm{b}}\dot{l}_{\mathrm{b}}^T & \dot{l}_{\mathrm{b}}\dot{l}_{\psi}^T \\ \dot{l}_{\psi}\dot{l}_{\mathrm{b}}^T & \dot{l}_{\psi}\dot{l}_{\psi}^T \end{bmatrix} \equiv \begin{pmatrix} \mathcal{I}_{\mathrm{bb}} & \mathcal{I}_{\mathrm{b}\psi} \\ \mathcal{I}_{\mathrm{b}\psi}^T & \mathcal{I}_{\psi\psi} \end{pmatrix}.$$

Calculation of $\mathcal{I}_{bb}$ $\mathcal{I}_{b\psi}$ and $\mathcal{I}_{\psi\psi}$ involves expectations of $(\mathrm{vec}\, \boldsymbol{H}_z)(\mathrm{vec}\, \boldsymbol{H}_z)^T$, which has expected value $(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p})$, where $\boldsymbol{K}_{p,p}$ is the commutation matrix described in Magnus and Neudecker [1979]. Straightforward calculations show that

$$\begin{aligned} \mathcal{I}_{bb} &= (\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{B}^T \otimes \boldsymbol{I}_p)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})(\boldsymbol{I}_{p^2} + \boldsymbol{K}_{p,p})(\boldsymbol{B}\boldsymbol{x}\boldsymbol{x}^T \otimes \boldsymbol{I}_p), && \text{(B–7)} \\ \mathcal{I}_{b\psi} &= (\boldsymbol{x}\boldsymbol{x}^T\boldsymbol{B}^T \otimes \boldsymbol{I}_p)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\boldsymbol{G}, && \text{(B–8)} \\ \mathcal{I}_{\psi\psi} &= \boldsymbol{G}^T(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\boldsymbol{G}/2. && \text{(B–9)} \end{aligned}$$

The expected information contained in observations to be made at $\boldsymbol{x}$-values $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is then $\mathcal{I}(\boldsymbol{b}, \boldsymbol{\psi} : \boldsymbol{X}) = \sum_{i=1}^{n} \mathcal{I}(\boldsymbol{b}, \boldsymbol{\psi} : \boldsymbol{x}_i)$. Plugging the MLEs into the inverse of this matrix gives an estimate of their variance, $\hat{\mathrm{Var}}[\hat{\boldsymbol{b}}^T, \hat{\boldsymbol{\psi}}^T)^T] = \mathcal{I}^{-1}(\hat{\boldsymbol{b}}, \hat{\boldsymbol{\psi}} : \boldsymbol{X})$.

Appendix C

Forward Price Curve Interpolation

In order to calculate daily values of the monthly DOL and HML_{FX} factors of Lustig and Verdelhan [2007], it is necessary to mark to market the one month forward contracts on a daily basis, and hence interpolate the forward curve between the overnight, one week, two week, three week and one month forward contracts. This interpolation of the forward curve was achieved via a linear interpolation on the implied interest rates from the forward contract prices, thus resulting in a monotonic interpolation between available forward price data points along the curve. Equation C–1 shows the interpolation formula used,

$$\log F_{t,t^\star} = \left[\frac{\log F_{t,t_1}}{t_1 - t} \times \frac{t_2 - t^\star}{t_2 - t_1} + \frac{\log F_{t,t_2}}{t_2 - t} \times \frac{t^\star - t_1}{t_2 - t_1}\right] \times (t^\star - t) \tag{C–1}$$

where $t < t_1 < t^\star < t_2$

For each monthly segment m (starting at the beginning of each month), I calculate the mark to market daily carry returns by first calculating the interpolated forward curve each day, then taking the difference between the log forward prices of the appropriate contracts, as shown in equation C–2,

$$\log R_i^m = \log F_{i,T-i} - \log F_{i+1,T-i-1} \qquad i \in 1, 2, \ldots, T \tag{C–2}$$

where $logF_{T,0} = S_T$.

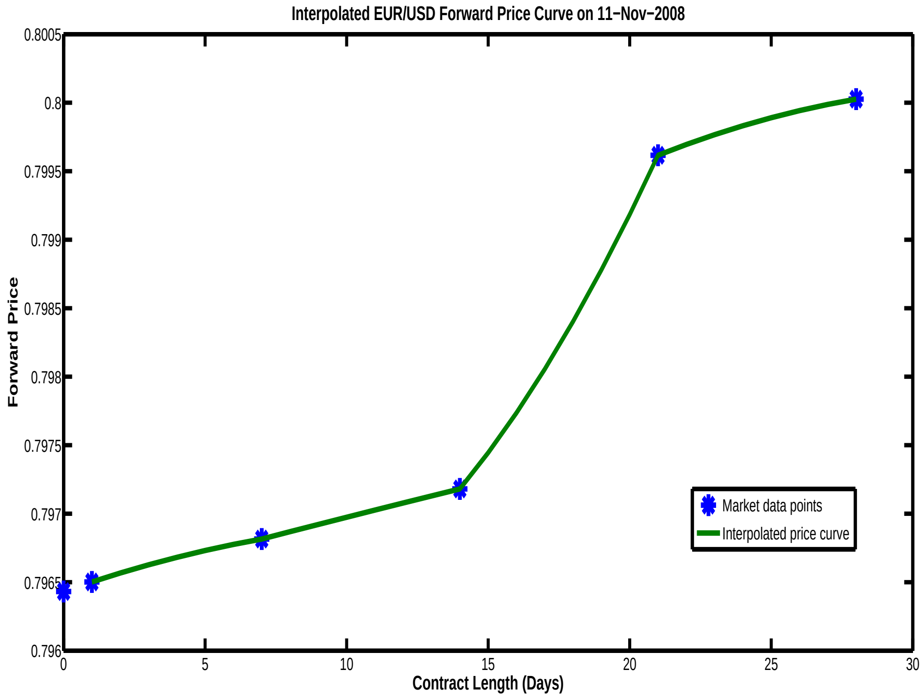


Figure C.1: Forward Price Curve Interpolation. Blue markers show the set of market data points, i.e. spot rate, overnight rate, 1 week rate, 2 week rate, 3 week rate and 1 month rate. The green line shows the interpolated values between these market data points.

Appendix D

Kalman Filter Estimation via Gradient Descent

The score function is given by:

$$S(\theta) = \frac{\partial l(\theta)}{\partial \theta} = \sum_{t=1}^{T} \frac{\partial l_t(\theta)}{\partial \theta} \tag{D–1}$$

where

$$l_t(\Theta) = -\frac{N}{2} log(2\pi) - \frac{1}{2} log|W_t| - \frac{1}{2} v_t^T W_t^{-1} v_t \quad , \ t = 1, 2, \ldots, T. \tag{D–2}$$

Using the following matrix identities from the Matrix Cookbook (Petersen and Pedersen):

$$\frac{\partial |M|}{\partial x} = |M| tr \left[M^{-1} \frac{\partial M}{\partial x} \right] \tag{D–3}$$

$$\frac{\partial M^{-1}}{\partial x} = -M^{-1} \frac{\partial M}{\partial x} M^{-1} \tag{D–4}$$

One obtains:

$$\frac{\partial l_t(\theta)}{\partial \theta_i} = -\frac{1}{2} tr \left[[W_t^{-1} \frac{\partial W_t}{\partial \theta_i}][I - W_t^{-1} v_t v_t^T] \right] - \left(\frac{\partial v_t}{\partial \theta_i} \right)^T W_t^{-1} v_t \tag{D–5}$$

Hence, one requires the derivatives of v_t and $W_t$with respect to θ_i.

$$\frac{\partial W_t}{\partial \theta_i} = \frac{\partial \Lambda_t}{\partial \theta_i} P_{t|t-1} \Lambda_t^T + \Lambda_t \frac{\partial P_{t|t-1}}{\partial \theta_i} \Lambda_t^T + \Lambda_t P_{t|t-1} \frac{\partial \Lambda_t^T}{\partial \theta_i} + \frac{\partial H}{\partial \theta_i} \tag{D–6}$$

$$\frac{\partial v_t}{\partial \theta_i} = -\Lambda_t \frac{\partial \hat{f}_{t|t-1}}{\partial \theta_i} - \frac{\partial \Lambda_t}{\partial \theta_i} \hat{f}_{t|t-1} \tag{D–7}$$

Now one needs derivatives of $\hat{f}_{t|t-1}$ and $P_{t|t-1}$ w.r.t. θ_i:

$$\frac{\partial \hat{f}_{t|t-1}}{\partial \theta_i} = \frac{\partial A}{\partial \theta_i} \hat{f}_{t-1|t-1} + A \frac{\partial \hat{f}_{t-1|t-1}}{\partial \theta_i} + \frac{\partial c}{\partial \theta_i} - \frac{\partial A}{\partial \theta_i} c - A \frac{\partial c}{\partial \theta_i} \tag{D–8}$$

$$\frac{\partial P_{t|t-1}}{\partial \theta_i} = \frac{\partial A}{\partial \theta_i} P_{t-1|t-1} A^T + A \frac{\partial P_{t-1|t-1}}{\partial \theta_i} A^T + A P_{t-1|t-1} \frac{\partial A^T}{\partial \theta_i} + \frac{\partial Q}{\partial \theta_i} \tag{D–9}$$

where the updated derivatives of $\hat{f}_{t|t}$ and $P_{t|t}$ w.r.t. θ_i are given by:

$$\begin{aligned} \frac{\partial \hat{f}_{t|t}}{\partial \theta_i} &= \frac{\partial \hat{f}_{t|t-1}}{\partial \theta_i} + \frac{\partial P_{t|t-1}}{\partial \theta_i} \Lambda_t^T W_t^{-1} v_t + P_{t|t-1} \frac{\partial \Lambda_t^T}{\partial \theta_i} W_t^{-1} v_t \\ & P_{t|t-1} \Lambda_t^T W_t^{-1} \frac{\partial W_t}{\partial \theta_i} W_t^{-1} v_t + P_{t|t-1} \Lambda_t^T W_t^{-1} \frac{\partial v_t}{\partial \theta_i} \end{aligned} \tag{D–10}$$

$$\begin{aligned} \frac{\partial P_{t|t}}{\partial \theta_i} &= \frac{\partial P_{t|t-1}}{\partial \theta_i} - \frac{\partial P_{t|t-1}}{\partial \theta_i} \Lambda_t^T W_t^{-1} \Lambda_t P_{t|t-1} - P_{t|t-1} \frac{\partial \Lambda_t^T}{\partial \theta_i} W_t^{-1} \Lambda_t P_{t|t-1} \\ & + P_{t|t-1} \Lambda_t^T W_t^{-1} \frac{\partial W_t}{\partial \theta_i} W_t^{-1} \Lambda_t P_{t|t-1} - P_{t|t-1} \Lambda_t^T W_t^{-1} \frac{\partial \Lambda_t}{\partial \theta_i} P_{t|t-1} \\ & - P_{t|t-1} \Lambda_t^T W_t^{-1} \Lambda_t \frac{\partial P_{t|t-1}}{\partial \theta_i} \end{aligned} \tag{D–11}$$

One also needs the derivatives of the parameter matrices A, c, Λ_t, H and Q w.r.t θ_i:

$$\frac{\partial c}{\partial \theta_i} = \begin{cases} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{if } \theta_i = \mu \\ \\ \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} & \text{if } \theta_i \neq \mu \end{cases} \tag{D–12}$$

$$\frac{\partial A}{\partial \theta_i} = \begin{cases} \begin{bmatrix} -\Delta t e^{-\beta \Delta t} & 0 \\ 0 & 0 \end{bmatrix} & \text{if } \theta_i = \beta \\ \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } \theta_i \neq \beta \text{ or } \gamma \\ \\ \begin{bmatrix} 0 & 0 \\ 0 & -\Delta t e^{-\gamma \Delta t} \end{bmatrix} & \text{if } \theta_i{=}\gamma \end{cases} \tag{D–13}$$

$$\frac{\partial \Lambda_t}{\partial \theta_i} = \begin{cases} \begin{bmatrix} -\tau_1 e^{-\beta \tau_1} & 0 \\ \vdots & \vdots \\ -\tau_N e^{-\beta \tau_N} & 0 \end{bmatrix} & \text{if } \theta_i = \beta \\ \\ \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} & \text{if } \theta_i \neq \beta \text{ or } \gamma \\ \\ \begin{bmatrix} 0 & -\tau_1 e^{-\beta \tau_1} \\ \vdots & \vdots \\ 0 & -\tau_N e^{-\beta \tau_N} \end{bmatrix} & \text{if } \theta_i{=}\gamma \end{cases} \tag{D–14}$$

$$\frac{\partial H}{\partial \theta_i} = \begin{cases} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 0 \end{bmatrix} & \text{if } \theta_i = s_1 \\ \vdots & \vdots \\ \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix} & \text{if } \theta_i = s_N \end{cases} \tag{D–15}$$

$$
\frac{\partial Q}{\partial \theta_i} = \begin{cases}
\begin{bmatrix} -\frac{1}{2}\sigma_\chi^2(1-e^{-2\beta\Delta t})\beta^{-2} + \sigma_\chi^2 \Delta t e^{-2\beta\Delta t}\beta^{-1} & (\rho_{\chi\xi}\sigma_\chi\sigma_\xi)[(e^{-(\beta+\gamma)\Delta t}-1)(\beta+\gamma)^{-2} + \Delta t e^{-(\beta+\gamma)\Delta t}(\beta+\gamma)^{-1}] \\ (\rho_{\chi\xi}\sigma_\chi\sigma_\xi)[(e^{-(\beta+\gamma)\Delta t}-1)(\beta+\gamma)^{-2} + \Delta t e^{-(\beta+\gamma)\Delta t}(\beta+\gamma)^{-1}] & 0 \end{bmatrix} & \text{if } \theta_i = \beta; \\
\begin{bmatrix} (1-e^{-2\beta\Delta t})\frac{\sigma_\chi}{\beta} & \frac{(1-e^{-(\beta+\gamma)\Delta t})\rho_{\chi\xi}\sigma_\xi}{\beta+\gamma} \\ \frac{(1-e^{-(\beta+\gamma)\Delta t})\rho_{\chi\xi}\sigma_\xi}{\beta+\gamma} & 0 \end{bmatrix} & \text{if } \theta_i = \sigma_\chi; \\
\begin{bmatrix} 0 & \frac{(1-e^{-(\beta+\gamma)\Delta t})\rho_{\chi\xi}\sigma_\chi}{\beta+\gamma} \\ \frac{(1-e^{-(\beta+\gamma)\Delta t})\rho_{\chi\xi}\sigma_\chi}{\beta+\gamma} & (1-e^{-2\gamma\Delta t})\frac{\sigma_\xi}{\gamma} \end{bmatrix} & \text{if } \theta_i = \sigma_\xi; \\
\begin{bmatrix} 0 & (\rho_{\chi\xi}\sigma_\chi\sigma_\xi)[(e^{-(\beta+\gamma)\Delta t}-1)(\beta+\gamma)^{-2} + \Delta t e^{-(\beta+\gamma)\Delta t}(\beta+\gamma)^{-1}]; \\ (\rho_{\chi\xi}\sigma_\chi\sigma_\xi)[(e^{-(\beta+\gamma)\Delta t}-1)(\beta+\gamma)^{-2} + \Delta t e^{-(\beta+\gamma)\Delta t}(\beta+\gamma)^{-1}] & -\frac{1}{2}\sigma_\chi^2(1-e^{-2\gamma\Delta t})\gamma^{-2} + \sigma_\chi^2 \Delta t e^{-2\gamma\Delta t}\gamma^{-1} \end{bmatrix} & \text{if } \theta_i = \gamma; \\
\begin{bmatrix} 0 & \frac{(1-e^{-(\beta+\gamma)\Delta t})\sigma_\chi\sigma_\xi}{\beta+\gamma} \\ \frac{(1-e^{-(\beta+\gamma)\Delta t})\sigma_\chi\sigma_\xi}{\beta+\gamma} & 0 \end{bmatrix} & \text{if } \theta_i = \rho_{\chi\xi}; \\
\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } \theta_i \neq \beta \text{ or } \sigma_\chi \text{ or } \sigma_\xi \text{ or } \gamma \text{ or } \rho_{\chi\xi}.
\end{cases}
\tag{D–16}
$$

Therefore, the score function is:

$$S(\theta) = \frac{\partial l(\theta)}{\partial \theta} \tag{D–17}$$

where

$$\frac{\partial l(\theta)}{\partial \theta_i} = \sum_{t=1}^{T} \left\{ -\frac{1}{2} tr \left[[W_t^{-1} \frac{\partial W_t}{\partial \theta_i}][I - W_t^{-1} v_t v_t^T] \right] - \left(\frac{\partial v_t}{\partial \theta_i} \right)^T W_t^{-1} v_t \right\} \tag{D–18}$$

for , $i = 1, 2, \ldots, p$.

Appendix E

Sensitivity of Average Backwardation to Parameter Shocks

Here, I give the expressions for the derivatives of the average backwardation with respect to the three parameters to which the observable covariates are linked in the HMF models.

The sensitivity of the Average Backwardation of Log Prices$_t$ to a shock on μ is given as:

$$\frac{\partial AB_t}{\partial \mu} = -\frac{e^{-\gamma_t t}}{\gamma_t}(e^{-\gamma_t \tau_1} - e^{-\gamma_t \tau_5}) \quad \text{(E–1)}$$

The sensitivity of the Average Backwardation of Log Prices$_t$ to a shock on β_t

is given as:

$$
\begin{aligned}
\frac{\partial AB_t}{\partial \beta_t} &= -t(e^{-\beta_t\tau_1} - e^{-\beta_t\tau_5})\chi_0 e^{-\beta_t t} - \tau_1 e^{-\beta_t\tau_1}\chi_0 e^{-\beta_t t} \\
&\quad + \tau_5 e^{-\beta_t\tau_5}\chi_0 e^{-\beta_t t} + 0.25\sigma_\chi^2\beta_t^{-2}(e^{-2\beta_t\tau_1} - 1) \\
&\quad + 0.5\tau_1\sigma_\chi^2\beta_t^{-1}e^{-2\beta_t\tau_1} - \lambda_\chi\beta_t^{-2}(e^{-\beta_t\tau_1} - 1) \\
&\quad - \tau_1\lambda_\chi\beta_t^{-1}e^{-\beta_t\tau_1} + \rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-2}(e^{-(\beta_t+\gamma_t)\tau_1} - 1) \\
&\quad + \tau_1\rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-1}e^{-(\beta_t+\gamma_t)\tau_1} - 0.25\sigma_\chi^2\beta_t^{-2}(e^{-2\beta_t\tau_5} - 1) \\
&\quad - 0.5\tau_5\sigma_\chi^2\beta_t^{-1}e^{-2\beta_t\tau_5} + \lambda_\chi\beta_t^{-2}(e^{-\beta_t\tau_5} - 1) \\
&\quad + \tau_5\lambda_\chi\beta_t^{-1}e^{-\beta_t\tau_5} - \rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-2}(e^{-(\beta_t+\gamma_t)\tau_5} - 1) \\
&\quad - \tau_5\rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-1}e^{-(\beta_t+\gamma_t)\tau_5}
\end{aligned}
\tag{E–2}
$$

The sensitivity of the Average Backwardation of Log Prices$_t$ to a shock on γ_t is given as:

$$
\begin{aligned}
\frac{\partial AB_t}{\partial \gamma_t} &= -t\xi_0 e^{-\gamma_t t}(e^{-\gamma_t\tau_1} - e^{-\gamma_t\tau_5}) - \tau_1 e^{-\gamma_t\tau_1}\xi_0 e^{-\gamma_t t} \\
&\quad + \tau_5 e^{-\gamma_t\tau_5}\xi_0 e^{-\gamma_t t} + \frac{\mu}{\gamma_t}(-\tau_1 e^{-\gamma_t\tau_1} + \tau_5 e^{-\gamma_t\tau_5})(1 - e^{-\gamma_t t}) \\
&\quad - \mu\gamma_t^{-2}(1 - e^{-\gamma_t t})(e^{-\gamma_t\tau_1} - e^{-\gamma_t\tau_5}) + \frac{\mu}{\gamma_t}te^{-\gamma_t t}(e^{-\gamma_t\tau_1} - e^{-\gamma_t\tau_5}) \\
&\quad + 0.25\sigma_\xi^2\gamma_t^{-2}(e^{-2\gamma_t\tau_1} - 1) + 0.5\tau_1\sigma_\xi^2\gamma_t^{-1}e^{-2\gamma_t\tau_1} \\
&\quad + \gamma_t^{-2}(\mu - \lambda_\xi)(e^{-\gamma_t\tau_1} - 1) + \tau_1\gamma_t^{-1}(\mu - \lambda_\xi)e^{-\gamma_t\tau_1} \\
&\quad + \rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-2}(e^{-(\beta_t+\gamma_t)\tau_1} - 1) + \tau_1\rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-1}e^{-(\beta_t+\gamma_t)\tau_1} \\
&\quad - 0.25\sigma_\xi^2\gamma_t^{-2}(e^{-2\gamma_t\tau_5} - 1) - 0.5\tau_5\sigma_\xi^2\gamma_t^{-1}e^{-2\gamma_t\tau_5} \\
&\quad - \gamma_t^{-2}(\mu - \lambda_\xi)(e^{-\gamma_t\tau_5} - 1) - \tau_5\gamma_t^{-1}(\mu - \lambda_\xi)e^{-\gamma_t\tau_5} \\
&\quad - \rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-2}(e^{-(\beta_t+\gamma_t)\tau_5} - 1) - \tau_5\rho_{\chi\xi}\sigma_\chi\sigma_\xi(\beta_t + \gamma_t)^{-1}e^{-(\beta_t+\gamma_t)\tau_5}
\end{aligned}
\tag{E–3}
$$

Appendix F

HMF SSX Results Tables

F.1 2011 - 2016 Results

Table F.1: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into μ parameter. Data period 2011 - 2016.

Covariate	β	σ_χ	λ_χ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.091 (0.007)	0.235 (0.011)	0.008 (0.004)	0.193 (0.011)	0.961 (0.036)	-0.061 (0.083)	0.111 (0.074)	3.673 (0.160)	-	-4501
BDI	0.093 (0.007)	0.237 (0.011)	0.007 (0.004)	0.195 (0.010)	0.964 (0.035)	-0.064 (0.085)	0.096 (0.072)	3.677 (0.160)	1.02E-2 (8.33E-3)	-4502
DXY	0.091 (0.008)	0.235 (0.011)	0.008 (0.005)	0.193 (0.011)	0.961 (0.036)	-0.061 (0.084)	0.111 (0.074)	3.673 (0.162)	-7.15E-5 (1.59E-2)	-4501
End Stocks	0.100 (0.010)	0.241 (0.012)	0.005 (0.005)	0.195 (0.010)	0.968 (0.035)	-0.075 (0.085)	0.069 (0.076)	3.689 (0.161)	-3.60E-2 (2.68E-2)	-4502
GSCI	0.139 (0.019)	0.206 (0.010)	-0.037 (0.014)	0.244 (0.019)	0.930 (0.039)	-0.114 (0.107)	0.026 (0.090)	3.395 (0.199)	1.68E-1 (3.37E-2)	-4512***
Lev Rat	0.101 (0.008)	0.243 (0.011)	0.003 (0.005)	0.193 (0.010)	0.974 (0.035)	-0.080 (0.084)	0.063 (0.071)	3.697 (0.159)	-3.60E-2 (1.38E-2)	-4504***
Ref Util	0.091 (0.007)	0.236 (0.011)	0.009 (0.004)	0.193 (0.011)	0.963 (0.036)	-0.059 (0.084)	0.108 (0.073)	3.692 (0.161)	7.20E-3 (4.31E-3)	-4502
SP500	0.082 (0.007)	0.224 (0.011)	0.014 (0.004)	0.198 (0.011)	0.947 (0.034)	-0.031 (0.086)	0.163 (0.076)	3.690 (0.157)	7.03E-2 (1.77E-2)	-4509***
SPEC	0.091 (0.007)	0.235 (0.011)	0.008 (0.004)	0.195 (0.011)	0.958 (0.035)	-0.060 (0.084)	0.107 (0.073)	3.662 (0.161)	6.50E-3 (7.27E-3)	-4501
US Prod	0.090 (0.007)	0.235 (0.011)	0.010 (0.005)	0.193 (0.011)	0.959 (0.036)	-0.058 (0.084)	0.111 (0.074)	3.677 (0.161)	1.04E-2 (2.14E-2)	-4501

Table F.2: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into β parameter. Data period 2011 - 2016.

Covariate	σ_χ	λ_χ	μ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.235 (0.011)	0.008 (0.004)	3.673 (0.161)	0.193 (0.011)	0.961 (0.036)	-0.061 (0.084)	0.111 (0.073)	0.091 (0.007)	-	-4501
BDI	0.241 (0.011)	0.007 (0.032)	3.603 (0.134)	0.195 (0.012)	0.949 (0.008)	-0.068 (0.126)	0.072 (0.081)	0.096 (0.030)	-0.003 (0.002)	-4509 ***
DXY	0.227 (0.011)	0.013 (0.052)	3.697 (0.207)	0.195 (0.014)	0.960 (0.015)	-0.051 (0.173)	0.155 (0.109)	0.077 (0.044)	-0.003 (0.004)	-4504 **
End Stocks	0.233 (0.011)	0.008 (0.051)	3.676 (0.202)	0.193 (0.014)	0.962 (0.015)	-0.060 (0.169)	0.121 (0.106)	0.088 (0.043)	-0.001 (0.004)	-4501
GSCI	0.216 (0.011)	0.014 (0.041)	3.894 (0.159)	0.200 (0.014)	1.002 (0.017)	-0.036 (0.128)	0.229 (0.096)	0.052 (0.037)	0.011 (0.004)	-4531 ***
Lev Rat	0.222 (0.013)	0.007 (0.051)	3.818 (0.365)	0.200 (0.021)	0.972 (0.021)	-0.044 (0.309)	0.115 (0.165)	0.088 (0.066)	-0.010 (0.004)	-4503 **
Ref Util	0.230 (0.012)	0.016 (0.034)	3.579 (0.117)	0.193 (0.011)	0.927 (0.009)	-0.048 (0.109)	0.153 (0.085)	0.076 (0.031)	-0.005 (0.001)	-4522 ***
SP500	0.213 (0.011)	0.023 (0.037)	3.261 (0.126)	0.197 (0.011)	0.843 (0.009)	-0.033 (0.117)	0.227 (0.078)	0.056 (0.030)	-0.005 (0.003)	-4512 ***
SPEC	0.237 (0.012)	0.006 (0.032)	3.776 (0.116)	0.192 (0.011)	0.989 (0.009)	-0.064 (0.107)	0.107 (0.086)	0.094 (0.031)	0.001 (0.002)	-4501
US Prod	0.214 (0.034)	0.023 (0.042)	3.287 (0.113)	0.197 (0.031)	0.847 (0.010)	-0.032 (0.106)	0.223 (0.149)	0.055 (0.024)	-0.005 (0.004)	-4513 ***

Table F.3: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into γ parameter. Data period 2011 - 2016.

Covariate	β	σ_χ	λ_χ	μ	σ_ξ	λ_ξ	$\rho_{\chi\xi}$	$\psi_{constant}$	ψ_1	NLL
None	0.091 (0.008)	0.235 (0.011)	0.008 (0.005)	3.673 (0.164)	0.193 (0.011)	-0.061 (0.085)	0.111 (0.075)	0.961 (0.036)	-	-4501
BDI	0.094 (0.007)	0.237 (0.011)	0.007 (0.004)	3.680 (0.159)	0.195 (0.010)	-0.065 (0.085)	0.088 (0.070)	0.965 (0.035)	-0.004 (0.002)	-4502
DXY	0.093 (0.008)	0.236 (0.011)	0.007 (0.005)	3.662 (0.163)	0.193 (0.011)	-0.064 (0.084)	0.104 (0.075)	0.961 (0.036)	0.002 (0.004)	-4501
End Stocks	0.105 (0.010)	0.244 (0.012)	0.003 (0.005)	3.695 (0.159)	0.196 (0.009)	-0.081 (0.086)	0.047 (0.073)	0.970 (0.035)	0.014 (0.007)	-4503*
GSCI	0.138 (0.016)	0.209 (0.010)	-0.036 (0.012)	3.411 (0.188)	0.238 (0.016)	-0.109 (0.104)	0.035 (0.081)	0.933 (0.038)	-0.042 (0.008)	-4514***
Lev Rat	0.103 (0.008)	0.245 (0.012)	0.002 (0.005)	3.701 (0.159)	0.193 (0.009)	-0.083 (0.084)	0.052 (0.070)	0.976 (0.035)	0.011 (0.004)	-4506***
Ref Util	0.091 (0.007)	0.236 (0.011)	0.009 (0.004)	3.693 (0.161)	0.193 (0.011)	-0.059 (0.084)	0.107 (0.073)	0.964 (0.036)	-0.002 (0.001)	-4502*
SP500	0.082 (0.007)	0.224 (0.011)	0.014 (0.004)	3.692 (0.158)	0.197 (0.011)	-0.033 (0.086)	0.164 (0.076)	0.949 (0.034)	-0.017 (0.004)	-4508***
SPEC	0.091 (0.007)	0.235 (0.011)	0.008 (0.004)	3.661 (0.160)	0.195 (0.011)	-0.060 (0.085)	0.107 (0.073)	0.958 (0.035)	-0.002 (0.002)	-4501
US Prod	0.090 (0.008)	0.235 (0.011)	0.009 (0.005)	3.675 (0.161)	0.193 (0.011)	-0.060 (0.084)	0.111 (0.074)	0.960 (0.036)	-0.001 (0.005)	-4501

F.2 2006 - 2011 Results

Table F.4: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into μ parameter. Data period 2006 - 2011.

Covariate	β	σ_χ	λ_χ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.000 (0.014)	0.266 (0.013)	0.065 (0.011)	0.265 (0.014)	0.675 (0.027)	-0.074 (0.118)	0.122 (0.066)	2.374 (0.147)	-	-4443
BDI	0.000 (0.014)	0.264 (0.013)	0.065 (0.011)	0.267 (0.014)	0.676 (0.027)	-0.075 (0.118)	0.127 (0.066)	2.382 (0.147)	2.28E-2 (1.59E-2)	-4444
DXY	0.000 (0.013)	0.260 (0.013)	0.064 (0.010)	0.268 (0.014)	0.676 (0.025)	-0.059 (0.118)	0.140 (0.067)	2.419 (0.145)	-2.79E-2 (9.04E-3)	-4448***
End Stocks	0.000 (0.013)	0.266 (0.013)	0.065 (0.009)	0.265 (0.014)	0.675 (0.025)	-0.070 (0.119)	0.123 (0.066)	2.384 (0.144)	1.04E-2 (1.63E-2)	-4443
GSCI	0.000 (0.010)	0.185 (0.010)	0.038 (0.009)	0.367 (0.023)	0.683 (0.025)	-0.048 (0.160)	-0.019 (0.100)	2.371 (0.182)	1.91E-1 (3.22E-2)	-4462***
Lev Rat	0.000 (0.013)	0.265 (0.013)	0.065 (0.009)	0.265 (0.014)	0.675 (0.025)	-0.074 (0.118)	0.124 (0.067)	2.375 (0.143)	7.47E-3 (7.10E-3)	-4444
Ref Util	0.000 (0.013)	0.265 (0.013)	0.065 (0.009)	0.266 (0.014)	0.675 (0.025)	-0.073 (0.118)	0.123 (0.066)	2.375 (0.143)	-3.33E-3 (4.23E-3)	-4443
SP500	0.000 (0.011)	0.259 (0.012)	0.063 (0.008)	0.268 (0.014)	0.676 (0.023)	-0.072 (0.119)	0.147 (0.065)	2.372 (0.142)	4.55E-2 (1.26E-2)	-4449***
SPEC	0.000 (0.013)	0.266 (0.013)	0.065 (0.009)	0.265 (0.014)	0.675 (0.024)	-0.073 (0.118)	0.124 (0.066)	2.377 (0.143)	6.10E-3 (5.28E-3)	-4444
US Prod	0.000 (0.013)	0.266 (0.013)	0.065 (0.009)	0.265 (0.014)	0.675 (0.025)	-0.076 (0.118)	0.124 (0.067)	2.370 (0.143)	-4.95E-3 (4.95E-3)	-4444

Table F.5: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into β parameter. Data period 2006 - 2011.

Covariate	σ_χ	λ_χ	μ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.235 (0.013)	0.100 (0.009)	2.288 (0.151)	0.277 (0.015)	0.641 (0.024)	-0.051 (0.127)	0.224 (0.072)	-0.051 (0.013)	-	-4452
BDI	0.238 (0.013)	0.093 (0.009)	2.080 (0.150)	0.292 (0.015)	0.582 (0.024)	-0.057 (0.126)	0.137 (0.073)	-0.037 (0.013)	-0.029 (0.005)	-4472***
DXY	0.264 (0.015)	0.063 (0.010)	2.239 (0.147)	0.285 (0.015)	0.620 (0.024)	-0.046 (0.124)	0.066 (0.080)	0.008 (0.015)	0.032 (0.005)	-4474***
End Stocks	0.232 (0.013)	0.104 (0.009)	2.206 (0.147)	0.282 (0.014)	0.620 (0.025)	-0.054 (0.120)	0.210 (0.067)	-0.055 (0.013)	0.012 (0.004)	-4456***
GSCI	0.232 (0.013)	0.098 (0.009)	2.234 (0.154)	0.282 (0.015)	0.626 (0.028)	-0.051 (0.120)	0.208 (0.069)	-0.048 (0.013)	-0.005 (0.005)	-4453***
Lev Rat	0.240 (0.013)	0.094 (0.009)	2.262 (0.144)	0.278 (0.014)	0.635 (0.024)	-0.056 (0.119)	0.196 (0.068)	-0.041 (0.013)	-0.008 (0.003)	-4456***
Ref Util	0.239 (0.013)	0.094 (0.009)	2.336 (0.146)	0.277 (0.014)	0.656 (0.026)	-0.053 (0.118)	0.212 (0.067)	-0.043 (0.013)	0.006 (0.002)	-4456***
SP500	0.259 (0.014)	0.082 (0.009)	2.691 (0.180)	0.263 (0.013)	0.764 (0.042)	-0.073 (0.111)	0.222 (0.065)	-0.030 (0.012)	0.026 (0.005)	-4461***
SPEC	0.236 (0.013)	0.101 (0.009)	2.308 (0.144)	0.277 (0.014)	0.649 (0.025)	-0.057 (0.118)	0.226 (0.066)	-0.052 (0.012)	0.006 (0.003)	-4454***
US Prod	0.236 (0.000)	0.098 (0.000)	2.282 (0.000)	0.280 (0.000)	0.642 (0.000)	-0.057 (0.000)	0.213 (0.000)	-0.049 (0.000)	0.010 (0.000)	-4460***

Table F.6: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into γ parameter. Data period 2006 - 2011.

Covariate	β	σ_χ	λ_χ	μ	σ_ξ	λ_ξ	$\rho_{\chi\xi}$	$\psi_{constant}$	ψ_1	NLL
None	-0.051 (0.000)	0.235 (0.000)	0.100 (0.000)	2.288 (0.000)	0.277 (0.000)	-0.051 (0.000)	0.224 (0.000)	0.641 (0.000)	-	-4452
BDI	0.542 (0.025)	0.310 (0.016)	-0.429 (0.020)	-0.108 (0.130)	0.231 (0.015)	0.120 (0.112)	0.095 (0.077)	-0.045 (0.015)	-0.009 (0.001)	-4473***
DXY	0.598 (0.032)	0.298 (0.021)	-0.483 (0.026)	0.110 (0.289)	0.265 (0.026)	0.131 (0.240)	0.019 (0.135)	0.013 (0.021)	0.009 (0.002)	-4474***
End Stocks	0.609 (0.028)	0.290 (0.015)	-0.446 (0.021)	-0.198 (0.119)	0.231 (0.014)	0.077 (0.100)	0.211 (0.067)	-0.058 (0.017)	0.003 (0.001)	-4449***
GSCI	0.622 (0.030)	0.288 (0.015)	-0.458 (0.022)	-0.180 (0.122)	0.233 (0.014)	0.069 (0.104)	0.213 (0.069)	-0.051 (0.016)	0.000 (0.001)	-4446**
Lev Rat	0.621 (0.028)	0.287 (0.015)	-0.459 (0.021)	-0.132 (0.130)	0.239 (0.015)	0.081 (0.112)	0.184 (0.073)	-0.041 (0.015)	-0.003 (0.001)	-4451***
Ref Util	0.651 (0.034)	0.286 (0.014)	-0.472 (0.022)	-0.141 (0.130)	0.241 (0.015)	0.068 (0.108)	0.203 (0.070)	-0.040 (0.018)	0.002 (0.001)	-4451***
SP500	0.756 (0.042)	0.271 (0.016)	-0.529 (0.025)	-0.106 (0.127)	0.261 (0.016)	0.055 (0.104)	0.209 (0.068)	-0.027 (0.019)	0.005 (0.001)	-4453***
SPEC	0.635 (0.028)	0.286 (0.015)	-0.464 (0.020)	-0.189 (0.122)	0.235 (0.014)	0.066 (0.104)	0.222 (0.068)	-0.053 (0.016)	0.001 (0.001)	-4447***
US Prod	0.631 (0.028)	0.288 (0.014)	-0.458 (0.020)	-0.172 (0.121)	0.235 (0.014)	0.075 (0.104)	0.210 (0.068)	-0.051 (0.015)	0.002 (0.001)	-4454***

F.3 2000 - 2006 Results

Table F.7: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into μ parameter. Data period 2000 - 2006.

Covariate	β	σ_χ	λ_χ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.869 (0.054)	0.281 (0.032)	0.037 (0.050)	0.182 (0.034)	0.042 (0.030)	0.276 (0.162)	-0.101 (0.078)	0.383 (0.238)	-	-4298
BDI	0.808 (0.029)	0.292 (0.017)	0.046 (0.040)	0.196 (0.016)	0.065 (0.018)	0.285 (0.118)	-0.217 (0.057)	0.475 (0.143)	2.42E-2 (3.58E-3)	-4311***
DXY	0.727 (0.055)	0.326 (0.028)	0.020 (0.045)	0.239 (0.034)	0.127 (0.041)	0.297 (0.181)	-0.432 (0.094)	0.698 (0.284)	-7.04E-2 (6.51E-3)	-4357***
End Stocks	0.851 (0.088)	0.304 (0.050)	-0.050 (0.052)	0.259 (0.056)	0.186 (0.049)	0.322 (0.310)	-0.398 (0.117)	0.895 (0.444)	1.17E-1 (6.51E-3)	-4365***
GSCI	0.787 (0.056)	0.322 (0.022)	0.080 (0.043)	0.192 (0.023)	0.091 (0.021)	0.279 (0.153)	-0.318 (0.073)	0.565 (0.199)	3.12E-2 (4.13E-3)	-4300**
Lev Rat	0.787 (0.040)	0.295 (0.019)	0.010 (0.043)	0.202 (0.019)	0.063 (0.018)	0.278 (0.143)	-0.245 (0.060)	0.458 (0.174)	-5.41E-2 (2.43E-3)	-4391***
Ref Util	0.866 (0.032)	0.281 (0.016)	0.038 (0.041)	0.182 (0.016)	0.042 (0.016)	0.276 (0.128)	-0.102 (0.058)	0.384 (0.153)	1.49E-3 (3.00E-3)	-4298
SP500	0.844 (0.045)	0.291 (0.024)	-0.027 (0.044)	0.184 (0.025)	0.022 (0.021)	0.271 (0.145)	-0.155 (0.066)	0.309 (0.183)	-3.72E-2 (2.88E-3)	-4344***
SPEC	0.829 (0.045)	0.289 (0.019)	0.052 (0.042)	0.191 (0.019)	0.049 (0.021)	0.278 (0.145)	-0.178 (0.065)	0.413 (0.187)	1.11E-2 (3.49E-3)	-4305***
US Prod	0.871 (0.000)	0.280 (0.000)	0.036 (0.000)	0.183 (0.000)	0.044 (0.000)	0.276 (0.000)	-0.102 (0.000)	0.389 (0.000)	-1.74E-3 (0.00E0)	-4298

Table F.8: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into β parameter. Data period 2000 - 2006.

Covariate	σ_χ	λ_χ	μ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.281 (0.018)	0.037 (0.028)	0.383 (0.117)	0.182 (0.011)	0.042 (0.011)	0.276 (0.092)	-0.101 (0.106)	0.869 (0.063)	-	-4298
BDI	0.288 (0.017)	0.042 (0.026)	0.430 (0.111)	0.189 (0.011)	0.054 (0.010)	0.284 (0.090)	-0.153 (0.101)	0.894 (0.060)	0.302 (0.048)	-4319***
DXY	0.305 (0.021)	0.088 (0.026)	0.462 (0.130)	0.199 (0.014)	0.063 (0.014)	0.275 (0.097)	-0.242 (0.118)	0.855 (0.070)	-0.422 (0.048)	-4354***
End Stocks	0.306 (0.023)	0.106 (0.032)	0.449 (0.141)	0.195 (0.015)	0.061 (0.017)	0.260 (0.096)	-0.256 (0.133)	0.862 (0.083)	0.430 (0.053)	-4330***
GSCI	0.281 (0.016)	0.040 (0.030)	0.387 (0.103)	0.183 (0.010)	0.043 (0.010)	0.276 (0.086)	-0.106 (0.100)	0.866 (0.057)	0.020 (0.069)	-4298
Lev Rat	0.313 (0.019)	0.114 (0.022)	0.454 (0.113)	0.202 (0.013)	0.060 (0.011)	0.276 (0.091)	-0.273 (0.100)	0.791 (0.054)	-0.344 (0.034)	-4368***
Ref Util	0.281 (0.016)	0.037 (0.027)	0.383 (0.099)	0.182 (0.009)	0.042 (0.008)	0.276 (0.085)	-0.101 (0.095)	0.869 (0.053)	0.000 (0.038)	-4298
SP500	0.298 (0.017)	0.100 (0.017)	0.444 (0.099)	0.186 (0.011)	0.051 (0.009)	0.295 (0.084)	-0.193 (0.096)	0.743 (0.045)	-0.289 (0.028)	-4352***
SPEC	0.291 (0.017)	0.049 (0.016)	0.390 (0.109)	0.186 (0.010)	0.044 (0.009)	0.280 (0.090)	-0.149 (0.096)	0.885 (0.058)	0.327 (0.041)	-4340***
US Prod	0.284 (0.015)	0.041 (0.021)	0.368 (0.099)	0.184 (0.010)	0.039 (0.008)	0.269 (0.086)	-0.110 (0.090)	0.904 (0.049)	-0.247 (0.064)	-4306***

Table F.9: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into γ parameter. Data period 2000 - 2006.

Covariate	β	σ_χ	λ_χ	μ	σ_ξ	λ_ξ	$\rho_{\chi\xi}$	$\psi_{constant}$	ψ_1	NLL
None	0.869 (0.056)	0.281 (0.034)	0.037 (0.049)	0.383 (0.248)	0.182 (0.036)	0.276 (0.167)	-0.101 (0.081)	0.042 (0.032)	-	-4298
BDI	0.816 (0.056)	0.291 (0.034)	0.045 (0.049)	0.464 (0.248)	0.194 (0.036)	0.285 (0.167)	-0.203 (0.081)	0.063 (0.032)	-0.006 (0.002)	-4309***
DXY	0.739 (0.029)	0.323 (0.016)	0.019 (0.040)	0.704 (0.143)	0.239 (0.016)	0.297 (0.118)	-0.422 (0.058)	0.130 (0.018)	0.021 (0.001)	-4354***
End Stocks	0.898 (0.058)	0.298 (0.029)	-0.065 (0.045)	0.942 (0.301)	0.262 (0.036)	0.325 (0.188)	-0.380 (0.099)	0.203 (0.044)	-0.037 (0.002)	-4363***
GSCI	0.780 (0.094)	0.332 (0.053)	0.086 (0.053)	0.617 (0.476)	0.195 (0.059)	0.277 (0.333)	-0.361 (0.124)	0.107 (0.052)	-0.011 (0.002)	-4302***
Lev Rat	0.789 (0.057)	0.295 (0.022)	0.010 (0.043)	0.452 (0.203)	0.203 (0.023)	0.276 (0.154)	-0.243 (0.074)	0.062 (0.022)	0.017 (0.001)	-4390***
Ref Util	0.865 (0.040)	0.281 (0.019)	0.038 (0.043)	0.384 (0.174)	0.181 (0.019)	0.276 (0.143)	-0.102 (0.060)	0.042 (0.018)	-0.001 (0.001)	-4298
SP500	0.838 (0.032)	0.292 (0.016)	-0.024 (0.040)	0.300 (0.153)	0.183 (0.016)	0.272 (0.128)	-0.157 (0.059)	0.018 (0.016)	0.012 (0.001)	-4343***
SPEC	0.828 (0.046)	0.289 (0.024)	0.052 (0.044)	0.413 (0.184)	0.190 (0.025)	0.278 (0.145)	-0.176 (0.066)	0.049 (0.021)	-0.003 (0.001)	-4305***
US Prod	0.869 (0.045)	0.281 (0.019)	0.037 (0.042)	0.384 (0.189)	0.182 (0.019)	0.276 (0.146)	-0.101 (0.065)	0.042 (0.022)	0.000 (0.001)	-4298

F.4 1995 - 2000 Results

Table F.10: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into μ parameter. Data period 1995 - 2000.

Covariate	β	σ_χ	λ_χ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	1.132 (0.054)	0.340 (0.032)	0.171 (0.050)	0.297 (0.034)	0.378 (0.030)	0.122 (0.162)	-0.564 (0.078)	1.256 (0.238)	-	-4326
BDI	1.079 (0.029)	0.367 (0.017)	0.264 (0.040)	0.336 (0.016)	0.430 (0.018)	0.080 (0.118)	-0.648 (0.057)	1.413 (0.143)	1.98E-2 (3.58E-3)	-4341***
DXY	1.119 (0.055)	0.314 (0.028)	0.215 (0.045)	0.254 (0.034)	0.294 (0.041)	0.115 (0.181)	-0.458 (0.094)	1.007 (0.284)	-3.21E-2 (6.51E-3)	-4366***
End Stocks	1.082 (0.088)	0.382 (0.050)	0.174 (0.052)	0.344 (0.056)	0.431 (0.049)	0.116 (0.310)	-0.668 (0.117)	1.415 (0.444)	-1.43E-2 (6.51E-3)	-4332***
GSCI	1.049 (0.056)	0.438 (0.022)	0.192 (0.043)	0.380 (0.023)	0.477 (0.021)	0.100 (0.153)	-0.739 (0.073)	1.552 (0.199)	3.40E-2 (4.13E-3)	-4363***
Lev Rat	1.178 (0.040)	0.332 (0.019)	0.154 (0.043)	0.290 (0.019)	0.378 (0.018)	0.131 (0.143)	-0.537 (0.060)	1.256 (0.174)	9.33E-3 (2.43E-3)	-4331***
Ref Util	1.151 (0.032)	0.338 (0.016)	0.171 (0.041)	0.296 (0.016)	0.378 (0.016)	0.124 (0.128)	-0.558 (0.058)	1.256 (0.153)	4.36E-3 (3.00E-3)	-4328*
SP500	1.158 (0.045)	0.326 (0.024)	0.213 (0.044)	0.276 (0.025)	0.342 (0.021)	0.111 (0.145)	-0.508 (0.066)	1.147 (0.183)	-2.52E-2 (2.88E-3)	-4361***
SPEC	1.189 (0.045)	0.327 (0.019)	0.148 (0.042)	0.287 (0.019)	0.377 (0.021)	0.136 (0.145)	-0.524 (0.065)	1.254 (0.187)	-8.39E-3 (3.49E-3)	-4331***
US Prod	1.186 (0.000)	0.325 (0.000)	0.175 (0.000)	0.278 (0.000)	0.351 (0.000)	0.127 (0.000)	-0.508 (0.000)	1.175 (0.000)	1.50E-2 (0.00E0)	-4341***

Table F.11: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into β parameter. Data period 1995 - 2000.

Covariate	σ_χ	λ_χ	μ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.340 (0.009)	0.171 (0.004)	1.256 (0.198)	0.297 (0.013)	0.378 (0.053)	0.122 (0.124)	-0.564 (0.061)	1.132 (0.013)	-	-4326
BDI	0.385 (0.009)	0.307 (0.004)	1.424 (0.200)	0.351 (0.013)	0.440 (0.054)	0.041 (0.124)	-0.677 (0.061)	1.064 (0.013)	0.136 (0.027)	-4339***
DXY	0.332 (0.009)	0.250 (0.004)	1.124 (0.206)	0.281 (0.013)	0.336 (0.050)	0.097 (0.127)	-0.535 (0.061)	1.129 (0.012)	-0.193 (0.019)	-4350***
End Stocks	0.349 (0.009)	0.185 (0.005)	1.296 (0.183)	0.307 (0.013)	0.392 (0.044)	0.115 (0.126)	-0.590 (0.063)	1.115 (0.014)	-0.036 (0.012)	-4327
GSCI	0.394 (0.009)	0.258 (0.008)	1.414 (0.303)	0.336 (0.018)	0.433 (0.071)	0.075 (0.147)	-0.668 (0.100)	1.100 (0.027)	0.177 (0.027)	-4354***
Lev Rat	0.333 (0.009)	0.115 (0.003)	1.253 (0.194)	0.288 (0.013)	0.376 (0.045)	0.147 (0.129)	-0.538 (0.064)	1.232 (0.015)	0.249 (0.028)	-4345***
Ref Util	0.336 (0.009)	0.092 (0.003)	1.233 (0.189)	0.291 (0.013)	0.369 (0.046)	0.154 (0.130)	-0.547 (0.066)	1.189 (0.015)	0.170 (0.023)	-4336***
SP500	0.334 (0.020)	0.226 (0.011)	1.173 (0.341)	0.286 (0.016)	0.352 (0.081)	0.105 (0.157)	-0.535 (0.120)	1.180 (0.044)	-0.185 (0.034)	-4353***
SPEC	0.329 (0.009)	0.170 (0.004)	1.250 (0.182)	0.290 (0.013)	0.377 (0.045)	0.124 (0.124)	-0.530 (0.061)	1.198 (0.013)	-0.105 (0.018)	-4331***
US Prod	0.336 (0.000)	0.178 (0.000)	1.212 (0.000)	0.291 (0.000)	0.364 (0.000)	0.121 (0.000)	-0.544 (0.000)	1.175 (0.000)	0.139 (0.000)	-4339***

Table F.12: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into γ parameter. Data period 1995 - 2000.

Covariate	β	σ_χ	λ_χ	μ	σ_ξ	λ_ξ	$\rho_{\chi\xi}$	$\psi_{constant}$	ψ_1	NLL
None	1.132 (0.056)	0.340 (0.034)	0.171 (0.049)	1.256 (0.248)	0.297 (0.036)	0.122 (0.167)	-0.564 (0.081)	0.378 (0.032)	-	-4326
BDI	1.038 (0.056)	0.407 (0.034)	0.305 (0.049)	1.476 (0.248)	0.377 (0.036)	0.030 (0.167)	-0.718 (0.081)	0.459 (0.032)	-0.008 (0.002)	-4342***
DXY	1.121 (0.029)	0.314 (0.016)	0.210 (0.040)	1.001 (0.143)	0.253 (0.016)	0.118 (0.118)	-0.455 (0.058)	0.292 (0.018)	0.010 (0.001)	-4364***
End Stocks	1.077 (0.058)	0.386 (0.029)	0.175 (0.045)	1.430 (0.301)	0.350 (0.036)	0.114 (0.188)	-0.677 (0.099)	0.436 (0.044)	0.005 (0.002)	-4332***
GSCI	1.049 (0.094)	0.440 (0.053)	0.191 (0.053)	1.561 (0.476)	0.383 (0.059)	0.100 (0.333)	-0.743 (0.124)	0.480 (0.052)	-0.011 (0.002)	-4365***
Lev Rat	1.176 (0.057)	0.333 (0.022)	0.155 (0.043)	1.254 (0.203)	0.290 (0.023)	0.130 (0.154)	-0.538 (0.074)	0.378 (0.022)	-0.003 (0.001)	-4331***
Ref Util	1.151 (0.040)	0.337 (0.019)	0.171 (0.043)	1.256 (0.174)	0.296 (0.019)	0.124 (0.143)	-0.557 (0.060)	0.378 (0.018)	-0.001 (0.001)	-4328**
SP500	1.160 (0.032)	0.325 (0.016)	0.209 (0.040)	1.138 (0.153)	0.274 (0.016)	0.113 (0.128)	-0.505 (0.059)	0.339 (0.016)	0.008 (0.001)	-4360***
SPEC	1.190 (0.046)	0.327 (0.024)	0.148 (0.044)	1.254 (0.184)	0.287 (0.025)	0.136 (0.145)	-0.525 (0.066)	0.377 (0.021)	0.003 (0.001)	-4331***
US Prod	1.185 (0.045)	0.325 (0.019)	0.174 (0.042)	1.170 (0.189)	0.277 (0.019)	0.127 (0.146)	-0.506 (0.065)	0.349 (0.022)	-0.005 (0.001)	-4340***

F.5 1990 - 1995 Results

Table F.13: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into μ parameter. Data period 1990 - 1995.

Covariate	β	σ_χ	λ_χ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.179 (0.014)	0.187 (0.009)	-0.005 (0.005)	0.293 (0.013)	1.790 (0.047)	0.115 (0.125)	0.201 (0.061)	5.404 (0.186)	-	-4381
BDI	0.186 (0.014)	0.189 (0.009)	-0.005 (0.005)	0.293 (0.013)	1.805 (0.046)	0.110 (0.125)	0.198 (0.061)	5.430 (0.186)	-1.76E-2 (1.64E-2)	-4382
DXY	0.182 (0.013)	0.188 (0.009)	-0.004 (0.005)	0.294 (0.013)	1.807 (0.048)	0.113 (0.125)	0.199 (0.061)	5.453 (0.193)	9.62E-3 (1.17E-2)	-4382
End Stocks	0.188 (0.013)	0.191 (0.009)	-0.009 (0.005)	0.293 (0.013)	1.803 (0.044)	0.103 (0.125)	0.187 (0.062)	5.379 (0.185)	2.18E-2 (9.76E-3)	-4384**
GSCI	0.291 (0.028)	0.167 (0.009)	0.005 (0.009)	0.375 (0.022)	1.766 (0.055)	0.076 (0.176)	-0.169 (0.111)	5.387 (0.267)	1.27E-1 (1.60E-2)	-4404***
Lev Rat	0.166 (0.012)	0.183 (0.009)	-0.002 (0.004)	0.296 (0.013)	1.816 (0.045)	0.120 (0.125)	0.222 (0.060)	5.514 (0.190)	5.31E-2 (1.68E-2)	-4387***
Ref Util	0.186 (0.013)	0.189 (0.009)	-0.005 (0.005)	0.294 (0.013)	1.795 (0.044)	0.111 (0.125)	0.188 (0.063)	5.407 (0.185)	-1.33E-2 (9.11E-3)	-4382*
SP500	0.287 (0.021)	0.217 (0.011)	-0.014 (0.008)	0.301 (0.013)	1.794 (0.049)	0.073 (0.133)	0.015 (0.074)	5.297 (0.216)	-1.29E-1 (1.39E-2)	-4407***
SPEC	0.176 (0.013)	0.186 (0.009)	-0.004 (0.004)	0.294 (0.013)	1.781 (0.043)	0.119 (0.124)	0.202 (0.061)	5.387 (0.183)	1.20E-2 (7.36E-3)	-4383*
US Prod	0.188 (0.000)	0.188 (0.000)	-0.005 (0.000)	0.296 (0.000)	1.799 (0.000)	0.112 (0.000)	0.184 (0.000)	5.425 (0.000)	1.49E-2 (0.00E0)	-4383*

Table F.14: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into β parameter. Data period 1990 - 1995.

Covariate	σ_χ	λ_χ	μ	σ_ξ	γ	λ_ξ	$\rho_{\chi\xi}$	ψ_{const}	ψ_1	NLL
None	0.187 (0.009)	-0.005 (0.004)	5.404 (0.198)	0.293 (0.013)	1.790 (0.053)	0.115 (0.124)	0.201 (0.061)	0.179 (0.013)	-	-4381
BDI	0.194 (0.009)	-0.003 (0.004)	5.691 (0.200)	0.294 (0.013)	1.892 (0.054)	0.101 (0.124)	0.208 (0.061)	0.183 (0.013)	-0.084 (0.027)	-4386***
DXY	0.189 (0.009)	0.002 (0.004)	5.581 (0.206)	0.299 (0.013)	1.832 (0.050)	0.119 (0.127)	0.185 (0.061)	0.199 (0.012)	0.110 (0.019)	-4398***
End Stocks	0.191 (0.009)	-0.013 (0.005)	5.243 (0.183)	0.295 (0.013)	1.764 (0.044)	0.108 (0.126)	0.170 (0.063)	0.204 (0.014)	-0.060 (0.012)	-4392***
GSCI	0.174 (0.009)	-0.019 (0.008)	4.661 (0.303)	0.325 (0.018)	1.567 (0.071)	0.120 (0.147)	0.058 (0.100)	0.217 (0.027)	-0.143 (0.027)	-4412***
Lev Rat	0.182 (0.009)	-0.002 (0.003)	5.292 (0.194)	0.301 (0.013)	1.755 (0.045)	0.124 (0.129)	0.190 (0.064)	0.183 (0.015)	0.210 (0.028)	-4421***
Ref Util	0.182 (0.009)	-0.016 (0.003)	4.899 (0.189)	0.301 (0.013)	1.646 (0.046)	0.122 (0.130)	0.148 (0.066)	0.207 (0.015)	0.172 (0.023)	-4413***
SP500	0.220 (0.020)	-0.048 (0.011)	4.187 (0.341)	0.316 (0.016)	1.435 (0.081)	0.134 (0.157)	-0.130 (0.120)	0.353 (0.044)	0.266 (0.034)	-4467***
SPEC	0.185 (0.009)	-0.002 (0.004)	5.374 (0.182)	0.293 (0.013)	1.770 (0.045)	0.122 (0.124)	0.199 (0.061)	0.178 (0.013)	0.030 (0.018)	-4383*
US Prod	0.180 (0.000)	-0.012 (0.000)	5.032 (0.000)	0.307 (0.000)	1.691 (0.000)	0.114 (0.000)	0.151 (0.000)	0.179 (0.000)	-0.140 (0.000)	-4402***

Table F.15: HMF SSX Model parameter estimates and negative log likelihoods obtained when incorporating covariates into γ parameter. Data period 1990 - 1995.

Covariate	β	σ_χ	λ_χ	μ	σ_ξ	λ_ξ	$\rho_{\chi\xi}$	$\psi_{constant}$	ψ_1	NLL
None	0.179 (0.014)	0.187 (0.009)	-0.005 (0.005)	5.404 (0.187)	0.293 (0.013)	0.115 (0.125)	0.201 (0.061)	1.790 (0.046)	-	-4381
BDI	0.187 (0.014)	0.189 (0.009)	-0.006 (0.005)	5.438 (0.187)	0.293 (0.013)	0.108 (0.125)	0.198 (0.061)	1.810 (0.046)	0.008 (0.006)	-4382
DXY	0.182 (0.013)	0.188 (0.009)	-0.004 (0.005)	5.456 (0.192)	0.294 (0.013)	0.113 (0.125)	0.198 (0.061)	1.808 (0.048)	-0.004 (0.004)	-4382
End Stocks	0.187 (0.013)	0.191 (0.009)	-0.009 (0.005)	5.390 (0.185)	0.293 (0.013)	0.104 (0.125)	0.189 (0.062)	1.805 (0.044)	-0.007 (0.003)	-4383**
GSCI	0.288 (0.027)	0.166 (0.009)	0.006 (0.009)	5.486 (0.255)	0.376 (0.022)	0.067 (0.174)	-0.190 (0.111)	1.801 (0.051)	-0.042 (0.005)	-4404***
Lev Rat	0.165 (0.012)	0.183 (0.008)	-0.002 (0.004)	5.514 (0.189)	0.297 (0.013)	0.120 (0.125)	0.222 (0.060)	1.815 (0.045)	-0.020 (0.006)	-4388***
Ref Util	0.185 (0.013)	0.189 (0.009)	-0.005 (0.005)	5.411 (0.185)	0.294 (0.013)	0.111 (0.125)	0.190 (0.063)	1.796 (0.044)	0.004 (0.003)	-4382
SP500	0.288 (0.021)	0.217 (0.011)	-0.014 (0.008)	5.364 (0.216)	0.301 (0.013)	0.070 (0.133)	0.007 (0.075)	1.819 (0.049)	0.046 (0.005)	-4404***
SPEC	0.176 (0.013)	0.186 (0.009)	-0.004 (0.004)	5.388 (0.183)	0.294 (0.013)	0.118 (0.124)	0.202 (0.061)	1.781 (0.043)	-0.004 (0.002)	-4383*
US Prod	0.187 (0.000)	0.188 (0.000)	-0.005 (0.000)	5.427 (0.000)	0.296 (0.000)	0.112 (0.000)	0.185 (0.000)	1.800 (0.000)	-0.005 (0.000)	-4383*

References

Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198, 2009. 73

Viral V. Acharya, Lars A. Lochstoer, and Tarun Ramadorai. Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics*, 109(2):441 – 465, 2013. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/j.jfineco.2013.03.003. URL http://www.sciencedirect.com/science/article/pii/S0304405X13000780. 43, 305, 309, 331

Fabian Ackermann, Walt Pohl, and Karl Schmedders. Optimal and naive diversification in currency markets. *Management Science*, 2016. 241

Tobias Adrian, Emanuel Moench, and Hyun Song Shin. Financial intermediation, asset prices and macroeconomic dynamics. *FRB of New York Staff Report*, 2010. 310

Tobias Adrian, Erkko Etula, and Tyler Muir. Financial intermediaries and the cross-section of asset returns. *The Journal of Finance*, 69(6):2557–2596, 2014. ISSN 1540-6261. doi: 10.1111/jofi.12189. URL http://dx.doi.org/10.1111/jofi.12189. 43, 305, 310

Giorgio Dall Aglio, Samuel I Kotz, and Gabriella Salinetti. *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, volume 67. Springer, 1991. 58

Roberto F. Aguilera and Marian Radetzki. Introduction and overview. In *The Price of Oil*, pages 1–8. Cambridge University Press, 2015. ISBN 9781316272527. URL http://dx.doi.org/10.1017/CBO9781316272527.001. Cambridge Books Online. 304

Gian Piero Aielli. Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299, 2013. doi: 10.1080/07350015.2013.771027. URL http://dx.doi.org/10.1080/07350015.2013.771027. 177, 238

Q. Farooq Akram. Commodity prices, interest rates and the dollar. *Energy Economics*, 31(6):838 – 851, 2009. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2009.05.016. URL http://www.sciencedirect.com/science/article/pii/S0140988309000917. Energy Sector Pricing and Macroeconomic Dynamics. 327

Qaisar Farooq Akram, Dagfinn Rime, and Lucio Sarno. Arbitrage in the foreign exchange market: Turning on the microscope. *Journal of International Economics*, 76(2):237–253, 2008. 108

Carol Alexander. A primer on the orthogonal garch model. *manuscript ISMA Centre, University of Reading, UK*, 2000. 174, 239

Carol Alexander and A Chibumba. Multivariate orthogonal factor garch. *University of Sussex, Mimeo*, 1997. 173

Matthew Ames, Guillaume Bagnarosa, and Gareth William Peters. Violations of uncovered interest rate parity and international exchange rate dependences. *Available at SSRN*, 2015a. 192, 193, 273, 277

Matthew Ames, Guillaume Bagnarosa, Gareth William Peters, and Pavel V Shevchenko. Understanding the interplay between covariance forecasting factor models and risk based portfolio allocations in currency carry trades. *Available at SSRN 2699020*, 2015b. 180

Matthew Ames, Gareth W Peters, Guillaume Bagnarosa, and Ioannis Kosmidis. Upside and Downside Risk Exposures of Currency Carry Trades via Tail Dependence. In *Innovations in Quantitative Risk Management*, pages 163–181. Springer, 2015c. 81, 109, 273

Matthew Ames, Gareth William Peters, Guillaume Bagnarosa, Pavel V. Shevchenko, and Tomoko Matsui. Consistently combining multi-factor stochastic oil commodity models with observed exogenous explanatory regression factors: Perspectives from speculators and hedgers. *Available at SSRN 2840355*, 2016. 284, 290, 301, 307

Robert M Anderson, Stephen W Bianchi, and Lisa R Goldberg. Will my risk parity strategy outperform? *Financial Analysts Journal*, 68(6):75–93, 2012. doi: http://dx.doi.org/10.2469/faj.v68.n6.7. 248, 254

Thierry Ané and Hélyette Geman. Order flow, transaction clock, and normality of asset returns. *The Journal of Finance*, 55(5):2259–2284, 2000. 196

Alessio Anzuini and Fabio Fornari. Macroeconomic determinants of carry trade activity. *Review of International Economics*, 20(3):468–488, 2012. ISSN 1467-9396. doi: 10.1111/j.1467-9396.2012.01034.x. URL `http://dx.doi.org/10.1111/j.1467-9396.2012.01034.x`. 193

Clifford S Asness, Andrea Frazzini, and Lasse H Pedersen. Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59, 2012. doi: http://dx.doi.org/10.2469/faj.v68.n1.1. 248, 254

T Bacigál and R Mesiar. 3-dimensional archimax copulas and their fitting to real data. COMPSTAT, 2012. 97

David K. Backus and Mario J. Crucini. Oil prices and the terms of trade. *Journal of International Economics*, 50(1):185 – 213, 2000. ISSN 0022-1996. doi: http://dx.doi.org/10.1016/S0022-1996(98)00064-6. URL `http://www.sciencedirect.com/science/article/pii/S0022199698000646`. 304

David K Backus, Silverio Foresi, and Chris I Telmer. Affine term structure models and the forward premium anomaly. *The Journal of Finance*, 56(1):279–304, 2001. 108

Gurdip Bakshi, George Panayotov, and Georgios Skoulakis. The baltic dry index as a predictor of global stock returns, commodity returns, and global economic activity. In *AFA 2012 Chicago Meetings Paper. Available at SSRN:*

http://ssrn.com/abstract=1787757 or http://dx.doi.org/10.2139/ssrn.1787757, 2011. 308

Nick Baltas. Trend-following, risk-parity and the influence of correlations. In *Risk-Based and Factor Investing*, pages 65–96. ISTE Press, Elsevier, 2015. 109

Nick Baltas. Multi-asset seasonality and trend-following strategies. *Bankers, Markets & Investors*, 140:47–62, 2016. 109

Nick Baltas and Robert Kosowski. Momentum strategies in futures markets and trend-following funds. 2013. 109

Nick Baltas and Robert Kosowski. Demystifying time-series momentum strategies: Volatility estimators, trading rules and pairwise correlations. *Available at SSRN 2140091*, 2015. 109, 249

David Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, 2012. 73

Pedro Barroso and Pedro Santa-Clara. Beyond the carry trade: Optimal currency portfolios. *Journal of Financial and Quantitative Analysis*, 50(05):1037–1056, 2015. 240

Devraj Basu and Jolle Miffre. Capturing the risk premium of commodity futures: The role of hedging pressure. *Journal of Banking & Finance*, 37(7): 2652 – 2664, 2013. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2013.02.031. URL `http://www.sciencedirect.com/science/article/pii/S0378426613001283`. 309, 330

Luc Bauwens, Sébastien Laurent, and Jeroen VK Rombouts. Multivariate garch models: a survey. *Journal of Applied Econometrics*, 21(1):79–109, 2006. 169

Tim Bedford and Roger M Cooke. Vines–a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068, 2002. 73

Geert Bekaert and George Panayotov. Good carry, bad carry. 2016. 112, 119

Daniel Berg and Kjersti Aas. Models for construction of multivariate dependence: A comparison study. *The European Journal of Finance*, 15(7):639–659, 2009. 73

Dan Bernhardt and Bart Taub. Cross-Asset Speculation in Stock Markets. *The Journal of Finance*, 63(5):2385–2427, 2008. 196

Hendrik Bessembinder. Systematic risk, hedging pressure, and risk premiums in futures markets. *The Review of Financial Studies*, 5(4):637–667, 1992. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/2962144. 43, 305

Hendrik Bessembinder, Jay f. Coughenour, Paul j. Seguin, and Margaret Monroe Smoller. Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *The Journal of Finance*, 50(1):361–375, 1995. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1995.tb05178.x. URL http://dx.doi.org/10.1111/j.1540-6261.1995.tb05178.x. 317, 318

Karol Binkowski, Pavel Shevchenko, and Nino Kordzakhia. Modelling commodity prices. *CSIRO technical report CMIS 09/43*, 2009. 289

Fischer Black and Robert B Litterman. Asset allocation: combining investor views with market equilibrium. *The Journal of Fixed Income*, 1(2):7–18, 1991. URL http://dx.doi.org/10.3905/jfi.1991.408013. 239

Nils Blomqvist. On a measure of dependence between two random variables. *The Annals of Mathematical Statistics*, pages 593–600, 1950. 67

Zvi Bodie. Common stocks as a hedge against inflation. *The Journal of Finance*, 31(2):459–470, 1976. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1976.tb01899.x. URL http://dx.doi.org/10.1111/j.1540-6261.1976.tb01899.x. 329

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986. 167

Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505, 1990. 176, 239

Tim Bollerslev, Robert F Engle, and Jeffrey M Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131, 1988. 170, 171

Tim Bollerslev, Robert F Engle, and Daniel B Nelson. Arch models. *Handbook of Econometrics*, 4:2959–3038, 1994. 168

E. Bouyé, V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli. Copulas for finance-a reading guide and some applications. *Available at SSRN 1032533*, 2000. 58, 74

George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015. 166, 183, 184

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004. 243

Phelim Boyle, Lorenzo Garlappi, Raman Uppal, and Tan Wang. Keynes meets markowitz: The trade-off between familiarity and diversification. *Management Science*, 58(2):253–272, 2012. doi: 10.1287/mnsc.1110.1349. URL `http://dx.doi.org/10.1287/mnsc.1110.1349`. 239

Nicole Branger, Linda Sandris Larsen, and Claus Munk. Robust portfolio choice with ambiguity and learning about return predictability. *Journal of Banking & Finance*, 37(5):1397 – 1411, 2013. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2012.05.009. URL `http://www.sciencedirect.com/science/article/pii/S0378426612001379`. 239

Mark Broadie. Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1):21–58, 1993. ISSN 0254-5330. doi: 10.1007/BF02282040. URL `http://dx.doi.org/10.1007/BF02282040`. 236, 254

Chris Brooks. *Introductory econometrics for finance*. Cambridge university press, 2014. 167

Markus K Brunnermeier and Lasse Heje Pedersen. Market liquidity and funding liquidity. *Review of Financial studies*, 22(6):2201–2238, 2009. 42, 111, 193

Markus K. Brunnermeier, Stefan Nagel, and Lasse H. Pedersen. Carry Trades and Currency Crashes. Working Paper 14473, National Bureau of Economic Research, November 2008. URL http://www.nber.org/papers/w14473. 42, 108, 111, 115, 146, 191, 193, 195, 210, 211, 216, 220, 255

Craig Burnside, Martin S Eichenbaum, and Sergio Rebelo. Understanding the forward premium puzzle: A microstructure approach. Technical report, National Bureau of Economic Research, 2007. 42, 110

Craig Burnside, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. Do peso problems explain the returns to the carry trade? *Review of Financial Studies*, 24(3):853–891, 2011. 109, 112

Bahattin Büyükşahin and Michel A. Robe. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, 42: 38 – 70, 2014. ISSN 0261-5606. doi: http://dx.doi.org/10.1016/j.jimonfin.2013.08.004. URL http://www.sciencedirect.com/science/article/pii/S026156061300106X. Understanding International Commodity Price Fluctuations. 310, 311, 330

Bahattin Büyükşahin, Michael S. Haigh, and Michel A. Robe. Commodities and equities: Ever a "market of one"? *Journal of Alternative Investments*, 12(3): 76 – 95, 2009. ISSN 15203255. URL http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=47798193&site=bsi-live. 330

Philippe Capéraà, Anne-Laure Fougères, and Christian Genest. Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, 72(1):30–49, 2000. 96

Massimiliano Caporin and Michael McAleer. Ten things you should know about the dynamic conditional correlation representation. *Econometrics*, 1(1):115–126, 2013. 177

Lorenzo Cappiello, Robert F Engle, and Kevin Sheppard. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial econometrics*, 4(4):537–572, 2006. 177

Mark M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82, 1997. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1997.tb03808.x. URL http://dx.doi.org/10.1111/j.1540-6261.1997.tb03808.x. 238

Jaime Casassus and Pierre Collin-Dufresne. Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance*, 60(5):2283–2331, 2005. 44, 284, 306, 328

Stephen G Cecchetti, Ingo Fender, and Patrick McGuire. Toward a global risk map. 2010. 193

CFTC. US CFTC Commitments of Traders: Explanatory Notes, 2015. URL http://www.cftc.gov/MarketReports/CommitmentsofTraders/ExplanatoryNotes/index.htm. [Online; accessed 9-November-2015]. 193

Louis K. C. Chan, Jason Karceski, and Josef Lakonishok. On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies*, 12(5):937–974, 1999. doi: 10.1093/rfs/12.5.937. URL http://rfs.oxfordjournals.org/content/12/5/937.abstract. 179, 238

Arthur Charpentier, Jean-David Fermanian, and Olivier Scaillet. The estimation of copulas: Theory and practice. *Copulas: from theory to applications in finance*, pages 35–62, 2007. 104

Arthur Charpentier, A-L Fougères, Christian Genest, and JG Nešlehová. Multivariate archimax copulas. *Journal of Multivariate Analysis*, 126:118–136, 2014. 96, 97, 98

Denis B Chaves, Jason C Hsu, Feifei Li, and Omid Shakernia. Efficient algorithms for computing risk parity portfolio weights. *Journal of Investing*, 21:150–163, 2012. 249

Vijay K Chopra and William T Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11, 1993. doi: 10.3905/jpm.1993.409440. URL http://dx.doi.org/10.3905/jpm.1993.409440. 236, 254

Yves Choueifaty and Yves Coignard. Toward maximum diversification. *Journal of Portfolio Management*, 35(1):40, 2008. URL http://dx.doi.org/10.3905/JPM.2008.35.1.40. 246, 254

Yves Choueifaty, Tristan Froidure, and Julien Reynier. Properties of the most diversified portfolio. *Journal of Investment Strategies*, 2(2):49–70, 2013. 246

Charlotte Christiansen, Angelo Ranaldo, and Paul Söderlind. The time-varying systematic risk of carry trade strategies. *Journal of Financial and Quantitative Analysis*, 46(04):1107–1125, 2011. 109

Charlotte Christiansen, Maik Schmeling, and Andreas Schrimpf. A comprehensive look at financial volatility prediction by economic variables. *Journal of Applied Econometrics*, 27(6):956–977, 2012. 195

George A Christodoulakis and Stephen E Satchell. Correlated arch (corrarch): Modelling the time-varying conditional correlation between financial asset returns. *European Journal of Operational Research*, 139(2):351–370, 2002. 176, 239

Giulio Cifarelli and Giovanna Paladino. Oil price dynamics and speculation: A multivariate financial approach. *Energy Economics*, 32(2):363 – 372, 2010. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2009.08.014. URL http://www.sciencedirect.com/science/article/pii/S0140988309001480. 331

Richard Clarida, Josh Davis, and Niels Pedersen. Currency carry trade regimes: Beyond the fama regression. *Journal of International Money and Finance*, 28 (8):1375–1389, 2009. 220

John H Cochrane and Monika Piazzesi. Bond risk premia. Technical report, National Bureau of Economic Research, 2002. 195

Niall Coffey, Warren B Hrung, and Asani Sarkar. Capital constraints, counterparty risk, and deviations from covered interest rate parity. Technical report, Staff Report, Federal Reserve Bank of New York, 2009. 108

Stuart Coles, Janet Heffernan, and Jonathan Tawn. Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365, 1999. 72

GM Constantine and TH Savits. A multivariate faa di bruno formula with applications. *Transactions of the American Mathematical Society*, 348(2):503–520, 1996. 99

PC Consul and GC Jain. On the log-gamma distribution and its properties. *Statistical Papers*, 12(2):100–106, 1971. 124

Gonzalo Cortazar, Ivo Kovacevic, and Eduardo S. Schwartz. Expected commodity returns and pricing models. *Energy Economics*, 49:60 – 71, 2015. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2015.01.015. URL `http://www.sciencedirect.com/science/article/pii/S0140988315000316`. 317, 318

Drew Creal, Siem Jan Koopman, and André Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795, 2013. 74

Marcelo G Cruz, Gareth W Peters, and Pavel V Shevchenko. *Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk*. John Wiley & Sons, 2015. 71

Mark Cummins, Michael Dowling, and Fearghal Kearney. Oil market modelling: A comparative analysis of fundamental and latent factor approaches. *International Review of Financial Analysis*, 46:211 – 218, 2016. ISSN 1057-5219. doi: http://dx.doi.org/10.1016/j.irfa.2016.05.010. URL `http://www.sciencedirect.com/science/article/pii/S1057521916300850`. 44, 306

Kent Daniel, Robert J Hodrick, and Zhongjin Lu. The carry trade: Risks and drawdowns. Technical report, National Bureau of Economic Research, 2014. 112, 241

Charoula Daskalaki, Alexandros Kostakis, and George Skiadopoulos. Are there common factors in individual commodity futures returns? *Journal of Banking & Finance*, 40:346–363, 2014. 44, 306, 310

G. De Luca and G. Rivieccio. Multivariate tail dependence coefficients for archimedean copulae. *Advanced Statistical Methods for the Analysis of Large Data-Sets*, page 287, 2012. 69, 80

Casper G de Vries and Chen Zhou. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):23–25, 2006. 62

Petros Dellaportas and Ioannis D Vrontos. Modelling volatility asymmetries: a bayesian analysis of a class of tree structured multivariate garch models. *The Econometrics Journal*, 10(3):503–520, 2007. 175

Victor DeMiguel, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5):798–812, 2009a. doi: 10.1287/mnsc.1080.0986. URL `http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1080.0986`. 240

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009b. doi: 10.1093/rfs/hhm075. URL `http://rfs.oxfordjournals.org/content/22/5/1915.abstract`. 248, 254

Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977. 105

MAH Dempster, Elena Medova, and Ke Tang. Determinants of oil futures prices and convenience yields. *Quantitative Finance*, 12(12):1795–1809, 2012. 44, 283, 284, 285, 299, 306, 307, 309, 327

A. Dias and P. Embrechts. Dynamic copula models for multivariate high-frequency data in finance. *Manuscript, ETH Zurich*, 2004. 73

Nebojsa Dimic, Vitaly Orlov, and Vanja Piljak. The effect of political risk on currency carry trades. *Finance Research Letters*, 19:75–78, 2016. 118

Jadran Dobrić and Friedrich Schmid. Nonparametric estimation of the lower tail dependence λ_l in bivariate copulas. *Journal of Applied Statistics*, 32(4):387–407, 2005. 71

Ali Dolati and Manuel Úbeda-Flores. On measures of multivariate concordance. *Journal of Probability and Statistical Science*, 4(2):147–164, 2006. 67

Catherine Donnelly and Paul Embrechts. The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. *Astin Bulletin*, 40(1):1–33, 2010. 61

Fabrizio Durante and Carlo Sempi. Copula theory: an introduction. In *Copula theory and its applications*, pages 3–31. Springer, 2010. 74

Eyal Dvir and Kenneth Rogoff. Demand effects and speculation in oil markets: Theory and evidence. *Journal of International Money and Finance*, 42: 113 – 128, 2014. ISSN 0261-5606. doi: http://dx.doi.org/10.1016/j.jimonfin.2013.08.007. URL http://www.sciencedirect.com/science/article/pii/S0261560613001095. Understanding International Commodity Price Fluctuations. 309, 328, 329, 337

P. Embrechts. Copulas: A personal view. *Journal of Risk and Insurance*, 76(3): 639–650, 2009. 62

Paul Embrechts. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):45–47, 2006. 62

Paul Embrechts, Alexander McNeil, and Daniel Straumann. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, pages 176–223, 2002. 59

Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas and applications to risk management. *Handbook of heavy tailed distributions in finance*, 8(329-384):1, 2003. 58

Paul Embrechts, Rdiger Frey, and Alexander McNeil. Quantitative risk management. *Princeton Series in Finance, Princeton*, 2005. 59

Charles Engel. The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of empirical finance*, 3(2):123–192, 1996. 41, 108

Charles M. Engel. Testing for the absence of expected real profits from forward market speculation. *Journal of International Economics*, 17(34):299 – 308, 1984. ISSN 0022-1996. doi: 10.1016/0022-1996(84)90025-4. URL http://www.sciencedirect.com/science/article/pii/0022199684900254. 107

Robert Engle. Garch 101: The use of arch/garch models in applied econometrics. *The Journal of Economic Perspectives*, 15(4):157–168, 2001. 167

Robert Engle. Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002. doi: 10.1198/073500102288618487. URL http://dx.doi.org/10.1198/073500102288618487. 176, 238

Robert Engle and Riccardo Colacito. Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2):238–253, 2006. doi: 10.1198/073500106000000017. URL http://dx.doi.org/10.1198/073500106000000017. 177, 238, 260, 261

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982. 166

Robert F Engle and Tim Bollerslev. Modelling the persistence of conditional variances. *Econometric reviews*, 5(1):1–50, 1986. 167

Robert F Engle and Kenneth F Kroner. Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150, 1995. 171, 239

Robert F Engle and Victor K Ng. Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778, 1993. 167

Robert F Engle, David M Lilien, and Russell P Robins. Estimating time varying risk premia in the term structure: the arch-m model. *Econometrica: Journal of the Econometric Society*, pages 391–407, 1987. 168

Robert F Engle, Victor K Ng, and Michael Rothschild. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45(1):213–237, 1990. 172, 178

Erkko Etula. Broker-dealer risk appetite and commodity returns. *Journal of Financial Econometrics*, 2013. doi: 10.1093/jjfinec/nbs024. URL http://jfec.oxfordjournals.org/content/early/2013/01/10/jjfinec.nbs024.abstract. 43, 305

Cavaliere Francesco Faa di Bruno. Note sur une nouvelle formule de calcul différentiel. *Quarterly J. Pure Appl. Math*, 1:359–360, 1857. 99

Eugene F. Fama. Stock returns, real activity, inflation, and money. *The American Economic Review*, 71(4):545–565, 1981. ISSN 00028282. URL `http://www.jstor.org/stable/1806180`. 329

Eugene F Fama. Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3):319–338, 1984. 41, 107, 108, 110

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993. 191, 195, 238, 255

Emmanuel Farhi and Xavier Gabaix. Rare Disasters and Exchange Rates. Working Paper 13805, National Bureau of Economic Research, February 2008. URL `http://www.nber.org/papers/w13805`. 42, 111, 146, 220

William Feller. *An Introduction to Probability Theory and Its Applications. Vol. 2*. New York: Wiley, 1971. 83, 99

NI Fisher. *Copulas. In: Kotz S, Read CB, Banks DL (eds) Encyclopedia of Statistical Sciences, Update Vol 1.* Wiley, New York, 1997. 56

Wai Mun Fong. Footprints in the market: Hedge funds and the carry trade. *Journal of International Money and Finance*, 33:41 – 59, 2013. ISSN 0261-5606. doi: http://dx.doi.org/10.1016/j.jimonfin.2012.09.001. URL `http://www.sciencedirect.com/science/article/pii/S026156061200160X`. 193

Gregory A Fredricks and Roger B Nelsen. On the relationship between spearman's rho and kendall's tau for pairs of continuous random variables. *Journal of Statistical Planning and Inference*, 137(7):2143–2150, 2007. 67

E.W. Frees and E.A. Valdez. Understanding relationships using copulas. *North American actuarial journal*, 2(1), 1998. 58

Peter A Frost and James E Savarino. For better performance: Constrain portfolio weights. *The Journal of Portfolio Management*, 15(1):29–34, 1988. URL http://dx.doi.org/10.3905/jpm.1988.409181. 236, 240, 254

Gabriele Galati, Alexandra Heath, and Patrick McGuire. Evidence of carry trade activity. *BIS Quarterly Review*, 3:27–41, 2007. 193

A Ronald Gallant, Peter Eric Rossi, and George Tauchen. Stock prices and volume. *Review of Financial studies*, 5(2):199–242, 1992. 196

Francis Galton. *Natural inheritance*, volume 42. Macmillan, 1889. 64

Lorenzo Garlappi, Raman Uppal, and Tan Wang. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20(1):41–81, 2006. doi: 10.1093/rfs/hhl003. URL http://rfs.oxfordjournals.org/content/early/2005/12/31/rfs.hhl003.short. 239

Andrew Gelman. Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, 27(15):2865–2873, 2008. 33, 311, 312

Helyette Geman and William O Smith. Shipping markets and freight rates: an analysis of the baltic dry index. *Journal of Alternative Investments*, 15(1): 98–109, 2012. 308, 309

Christian Genest and Anne-Catherine Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12(4):347–368, 2007. 58

Christian Genest and Johanna Neslehova. A primer on copulas for count data. *Astin Bulletin*, 37(2):475, 2007. 74

Christian Genest and Bruno Rémillard. Discussion of copulas: tales and facts, by thomas mikosch. *Extremes*, 9(1):27–36, 2006. 62

Christian Genest and Louis-Paul Rivest. On the multivariate probability integral transformation. *Statistics & Probability Letters*, 53(4):391–399, 2001. 94

Christian Genest, Kilani Ghoudi, and L-P Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995. 125

Christian Genest, Michel Gendron, and Michaël Bourdeau-Brien. The advent of copulas in finance. *The European Journal of Finance*, 15(7-8):609–618, 2009. 55

Christian Genest, Alberto Carabarín-Aguirre, and Fanny Harvey. Copula parameter estimation using blomqvists beta. *Journal de la Société Française de Statistique*, 154(1):5–24, 2013. 68

Mark Pitts George E. Tauchen. The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505, 1983. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1912002. 192

D. Giannikis, I.D. Vrontos, and P. Dellaportas. Modelling nonlinearities and heavy tails via threshold normal mixture {GARCH} models. *Computational Statistics & Data Analysis*, 52(3):1549 – 1571, 2008. ISSN 0167-9473. doi: http://dx.doi.org/10.1016/j.csda.2007.05.007. URL http://www.sciencedirect.com/science/article/pii/S0167947307002010. 175

Rajna Gibson and Eduardo S Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990. 36, 44, 284, 285, 288, 306, 311

Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993. 168

D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, February 2003. ISSN 0364-765X. doi: 10.1287/moor.28.1.1.14260. URL http://dx.doi.org/10.1287/moor.28.1.1.14260. 240

Itay Goldstein and Liyan Yang. Commodity financialization: Risk sharing and price discovery in commodity futures markets. *Available at SSRN 2555996*, 2015. 331

Gary B. Gorton, Fumio Hayashi, and K. Geert Rouwenhorst. The fundamentals of commodity futures returns. *Review of Finance*, 17(1):35 – 105, 2013. ISSN 15723097. URL http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=84556160&site=bsi-live. 309, 328

Christian Gouriéroux. *ARCH models and financial applications*. Springer Science & Business Media, 2012. 170

Pierre-Olivier Gourinchas and Helene Rey. From world banker to world venture capitalist: US external adjustment and the exorbitant privilege. In *G7 Current Account Imbalances: Sustainability and Adjustment*, pages 11–66. University of Chicago Press, 2007. 118

Richard C. Green and Burton Hollifield. When will mean-variance efficient portfolios be well diversified? *The Journal of Finance*, 47(5):1785–1809, 1992. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2328996. 179, 238

Steffen Grønneberg. The copula information criterion and its implications for the maximum pseudo-likelihood estimator. *Dependence Modelling: The Vine Copula Handbook, World Scientific Books*, pages 113–138, 2010. 130

Marc Gronwald, Janina Ketterer, and Stefan Trück. The relationship between carbon, commodity and financial markets: A copula analysis. *Economic Record*, 87(s1):105–124, 2011. 62

Gred M Gupton, Christopher Clemens Finger, and Mickey Bhatia. *Creditmetrics: technical document*. JP Morgan & Co., 1997. 60

Christian M Hafner and Hans Manner. Dynamic stochastic copula models: estimation, inference and applications. *Journal of Applied Econometrics*, 27(2): 269–295, 2010. 125, 180

Yufeng Han. Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies*, 19(1):237–271, 2006. doi: 10.1093/rfs/hhj002. URL http://rfs.oxfordjournals.org/content/19/1/237.abstract. 238

Lars Peter Hansen and Robert J Hodrick. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *The Journal of Political Economy*, pages 829–853, 1980. 41, 107, 108, 110

Lars Peter Hansen and Robert J Hodrick. Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models. In *Exchange rates and international macroeconomics*, pages 113–152. University of Chicago Press, 1983. 107, 110

Joel Hasbrouck and Duane J Seppi. Common factors in prices, order flows, and liquidity. *Journal of financial Economics*, 59(3):383–411, 2001. 196

Robert A Haugen and Nardin L Baker. The efficient market inefficiency of capitalization-weighted stock portfolios. *The Journal of Portfolio Management*, 17(3):35–40, 1991. URL `http://dx.doi.org/10.3905/jpm.1991.409335`. 254

Xiaojun He and Raja Velu. Volume and Volatility in a Common-Factor Mixture of Distributions Model. *Journal of Financial and Quantitative Analysis*, 49(01): 33–49, 2014. 193, 196

Brian J. Henderson, Neil D. Pearson, and Li Wang. New evidence on the financialization of commodity markets. *Review of Financial Studies*, 2014. doi: 10.1093/rfs/hhu091. URL `http://rfs.oxfordjournals.org/content/early/2014/12/25/rfs.hhu091.abstract`. 196, 308, 311

Harold V Henderson and SR Searle. Vec and vech operators for matrices, with some uses in jacobians and multivariate statistics. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, pages 65–81, 1979. 356

Ludger Hentschel. All in the family nesting symmetric and asymmetric garch models. *Journal of Financial Economics*, 39(1):71–104, 1995. 168

David Hirshleifer. Hedging pressure and futures price movements in a general equilibrium model. *Econometrica*, 58(2):411–428, 1990. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/2938209`. 330

Wassily Hoeffding. Scale-invariant correlation theory. In *The collected works of Wassily Hoeffding*, pages 109–133. Springer, 1994a. 55

Wassily Hoeffding. Scale-invariant correlation theory. In *The collected works of Wassily Hoeffding*, pages 57–107. Springer, 1994b. 55

J.M. Hofert. *Sampling Nested Archimedean Copulas: With Applications to CDO Pricing*. Suedwestdeutscher Verlag fuer Hochschulschriften, 2010a. 82

Marius Hofert. Construction and sampling of nested Archimedean copulas. *Copula Theory and Its Applications*, pages 147–160, 2010b. 85

Marius Hofert, Martin Mächler, and Alexander J Mcneil. Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, 2012. 93, 103

Peter D. Hoff and Xiaoyue Niu. A covariance regression model. *Statistica Sinica*, 22(2):729–753, 2012. ISSN 10170405, 19968507. URL `http://www.jstor.org/stable/24310032`. 178, 182, 183, 201, 204, 355

Harrison Hong and Motohiro Yogo. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics*, 105(3):473–490, 2012. 196

Yongmiao Hong, Jun Tu, and Guofu Zhou. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies*, 20(5):1547–1581, 2007. 61

Michael Hutchison and Vladyslav Sushko. Impact of macro-economic surprises on carry trade activity. *Journal of Banking & Finance*, 37(4):1133 – 1147, 2013. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2012.10.022. URL `http://www.sciencedirect.com/science/article/pii/S0378426612003330`. 193

Rob J Hyndman. *forecast: Forecasting functions for time series and linear models*, 2015. URL `http://github.com/robjhyndman/forecast`. R package version 6.1. 185

Rob J Hyndman and Yeasmin Khandakar. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22, 2008. URL `http://ideas.repec.org/a/jss/jstsof/27i03.html`. 185, 187

Rob J Hyndman and Anne B Koehler. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688, 2006. 185

J. P. Morgan. Riskmetrics technical document. 1996. 168

Ravi Jagannathan and Tongshu Ma. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1684, 2003. ISSN 1540-6261. doi: 10.1111/1540-6261.00580. URL http://dx.doi.org/10.1111/1540-6261.00580. 240, 265, 266, 267

Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993. 109

Narasimhan Jegadeesh and Sheridan Titman. Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of finance*, 56(2):699–720, 2001. 109

J David Jobson and Robert M Korkie. Putting markowitz theory to work. *The Journal of Portfolio Management*, 7(4):70–74, 1981. URL http://dx.doi.org/10.3905/jpm.1981.408816. 236, 254

Harry Joe. Multivariate concordance. *Journal of multivariate analysis*, 35(1):12–30, 1990. 67

Harry Joe. *Multivariate models and multivariate dependence concepts*, volume 73. Chapman & Hall/CRC, 1997. 58

Harry Joe. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419, 2005. 125

Harry Joe. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):37–41, 2006. 62

Philippe Jorion. International portfolio diversification with estimation risk. *The Journal of Business*, 58(3):259–278, 1985. ISSN 00219398, 15375374. URL http://www.jstor.org/stable/2352997. 239

Philippe Jorion. Bayes-stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis*, 21(3):279–292, 1986. ISSN 00221090, 17566916. URL http://www.jstor.org/stable/2331042. 239

Ted Juhl, William Miles, and Marc D Weidenmier. Covered interest arbitrage: then versus now. *Economica*, 73(290):341–352, 2006. 108, 118

Emmanuel Jurczenko, Thierry Michel, and Jerome Teiletche. A unified framework for risk-based investing. *The Journal of Investment Strategies*, 4(4), 2015. 250, 254

Nicholas Kaldor. Speculation and economic stability. *The Review of Economic Studies*, 7(1):1–27, 1939. 44, 306

Robert K. Kaufmann. The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy*, 39(1):105 – 115, 2011. ISSN 0301-4215. doi: http://dx.doi.org/10.1016/j.enpol.2010.09.018. URL http://www.sciencedirect.com/science/article/pii/S0301421510007044. 331

Robert K. Kaufmann and Ben Ullman. Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices. *Energy Economics*, 31(4):550 – 558, 2009. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2009.01.013. URL http://www.sciencedirect.com/science/article/pii/S0140988309000243. 43, 305, 331

Robert K. Kaufmann, Stephane Dees, Audrey Gasteuil, and Michael Mann. Oil prices: The role of refinery utilization, futures markets and non-linearities. *Energy Economics*, 30(5):2609 – 2622, 2008. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2008.04.010. URL http://www.sciencedirect.com/science/article/pii/S0140988308000649. 309

Hakan Kaya and Wai Lee. Demystifying risk parity. 2012. 249

Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2): 81–93, 1938. 66

John Maynard Keynes. A treatise on money, 1930. 330

Lutz Kilian and Thomas K. Lee. Quantifying the speculative component in the real price of oil: The role of global oil inventories. *Journal of International Money and Finance*, 42:71 – 87, 2014. ISSN 0261-5606. doi: http://dx.doi.org/10.1016/j.jimonfin.2013.08.005. URL http://www.sciencedirect.com/science/article/pii/S0261560613001071. Understanding International Commodity Price Fluctuations. 331

Lutz Kilian and Daniel P. Murphy. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3):454–478, 2014. ISSN 1099-1255. doi: 10.1002/jae.2322. URL http://dx.doi.org/10.1002/jae.2322. 43, 305

Lutz Kilian and Cheolbeom Park. The impact of oil price shocks on the u.s. stock market*. *International Economic Review*, 50(4):1267–1287, 2009. ISSN 1468-2354. doi: 10.1111/j.1468-2354.2009.00568.x. URL http://dx.doi.org/10.1111/j.1468-2354.2009.00568.x. 330

Clark H Kimberling. A probabilistic interpretation of complete monotonicity. *Aequationes mathematicae*, 10(2):152–164, 1974. 81

Claudia Klüppelberg, Alexander Lindner, and Ross Maller. A continuous-time garch process driven by a lévy process: stationarity and second-order behaviour. *Journal of Applied Probability*, 41(03):601–622, 2004. 168

Ioannis Kosmidis and Dimitris Karlis. Model-based clustering using copulas with applications. *Statistics and Computing*, pages 1–21, 2014. 105

Apostolos Kourtis, George Dotsis, and Raphael N. Markellos. Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking & Finance*, 36(9):2522 – 2531, 2012. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2012.05.005. URL http://www.sciencedirect.com/science/article/pii/S0378426612001331. 239

William H Kruskal. Ordinal measures of association. *Journal of the American Statistical Association*, 53(284):814–861, 1958. 66

Dorota Kurowicka and Harry Joe. *Dependence Modeling: Vine Copula Handbook*. World Scientific, 2011. 73

Markku Lanne and Pentti Saikkonen. A multivariate generalized orthogonal factor garch model. *Journal of Business & Economic Statistics*, 25(1):61–75, 2007. ISSN 07350015. URL http://www.jstor.org/stable/27638907. 174, 178

Jerry F Lawless. Inference in the generalized gamma and log gamma distributions. *Technometrics*, 22(3):409–419, 1980. 124, 126

Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603 – 621, 2003. ISSN 0927-5398. doi: http://dx.doi.org/10.1016/S0927-5398(03)00007-0. URL http://www.sciencedirect.com/science/article/pii/S0927539803000070. 239

Bong Soo Lee. Stock returns and inflation revisited: An evaluation of the inflation illusion hypothesis. *Journal of Banking & Finance*, 34(6):1257 – 1273, 2010. ISSN 0378-4266. doi: http://dx.doi.org/10.1016/j.jbankfin.2009.11.023. URL http://www.sciencedirect.com/science/article/pii/S0378426609003203. 329

Wai Lee. Risk-based asset allocation: A new answer to an old question?, 2011. 249

Yves Lempérière, Cyril Deremble, Trung-Tu Nguyen, Philip Seager, Marc Potters, and Jean-Philippe Bouchaud. Risk premia: Asymmetric tail risks and excess returns. *Quantitative Finance*, 17(1):1–14, 2017. 112

David X Li. On default correlation: A copula function approach. *The Journal of Fixed Income*, 9(4):43–54, 2000. 60

Alexander Lindner. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):43–44, 2006. 62

Cho-hsin Ling. *Representation of associative functions*. 1964. 81

X. Luo and P.V. Shevchenko. The t copula with multiple parameters of degrees of freedom: bivariate characteristics and application to risk management. *Quantitative Finance*, 10(9):1039–1054, 2010. 77

Hanno Lustig and Adrien Verdelhan. The cross section of foreign currency risk premia and consumption growth risk. *American Economic Review*, 97(1):89–117, September 2007. doi: 10.1257/aer.97.1.89. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.97.1.89. 41, 108, 110, 195, 198, 199, 221, 224, 225, 357

Hanno Lustig, Nikolai Roussanov, and Adrien Verdelhan. Common risk factors in currency markets. *Review of Financial Studies*, 24(11):3731–3777, 2011. doi: 10.1093/rfs/hhr068. URL http://rfs.oxfordjournals.org/content/24/11/3731.abstract. 109, 110, 112, 191, 192, 195, 255

Donald MacKenzie and Taylor Spears. the formula that killed wall street? the gaussian copula and the material cultures of modelling. *University of Edinburgh, Mimeo*, 2012. 61

Jan R Magnus and Heinz Neudecker. The commutation matrix: some properties and applications. *The Annals of Statistics*, pages 381–394, 1979. 356

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4):60–70, 2010. 249, 250, 254

Dominique Drouet Mari and Samuel Kotz. *Correlation and dependence*, volume 2. World Scientific, 2001. 64

Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1952.tb01525.x. URL http://dx.doi.org/10.1111/j.1540-6261.1952.tb01525.x. 235, 241, 242, 254

Albert W Marshall and Ingram Olkin. Families of multivariate distributions. *Journal of the American Statistical Association*, 83(403):834–841, 1988. 85

Svetlana Maslyuk and Russell Smyth. Unit root properties of crude oil spot and futures prices. *Energy Policy*, 36(7):2591 – 2600, 2008. ISSN 0301-4215. doi: http://dx.doi.org/10.1016/j.enpol.2008.03.018. URL http://www.sciencedirect.com/science/article/pii/S0301421508001547. 317

Charles E McCulloch. Symmetric matrix derivatives with applications. *Journal of the American Statistical Association*, 77(379):679–682, 1982. 355

Alexander J McNeil and Johanna Nešlehová. Multivariate archimedean copulas, d-monotone functions and l1-norm symmetric distributions. *The Annals of Statistics*, 37(5B):3059–3097, 2009. 56, 58, 79, 82, 83, 84, 86, 87, 88, 89

Rajnish Mehra and Edward C Prescott. The equity premium: A puzzle. *Journal of monetary Economics*, 15(2):145–161, 1985. 45, 113, 114

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Carry trades and global foreign exchange volatility. *Journal of Finance*, 67(2):681–718, April 2012a. doi: 10.1111/j.1540-6261.2012.01728.x. 41, 42, 109, 110, 115, 146, 150, 195, 220

Lukas Menkhoff, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf. Currency momentum strategies. *Journal of Financial Economics*, 106(3):660–684, 2012b. 109

Robert C. Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323 – 361, 1980. ISSN 0304-405X. doi: http://dx.doi.org/10.1016/0304-405X(80)90007-0. URL http://www.sciencedirect.com/science/article/pii/0304405X80900070. 248

Radko Mesiar and Andrea StupňAnová. Decomposition integrals. *International Journal of Approximate Reasoning*, 54(8):1252–1259, 2013. 97

Attilio Meucci. A short, comprehensive, practical guide to copulas. *GARP Risk Professional*, pages 22–27, 2011. 29, 58, 59

Richard O Michaud. The markowitz optimization enigma: is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42, 1989. URL http://dx.doi.org/10.2469/faj.v45.n1.31. 236, 254

Miloud Mihoubi. Bell polynomials and binomial type sequences. *Discrete Mathematics*, 308(12):2450–2459, 2008. 100

Thomas Mikosch. Copulas: Tales and facts. *Extremes*, 9(1):3–20, 2006a. 62

Thomas Mikosch. Copulas: Tales and facts rejoinder. *Extremes*, 9(1):55–62, 2006b. ISSN 1386-1999. doi: 10.1007/s10687-006-0024-9. URL http://dx.doi.org/10.1007/s10687-006-0024-9. 62

Nikolaos T. Milonas and Thomas Henker. Price spread and convenience yield behaviour in the international oil market. *Applied Financial Economics*, 11(1):23–36, 2001. doi: 10.1080/09603100150210237. URL http://dx.doi.org/10.1080/09603100150210237. 309

Patricia Mariela Morillas. A method to obtain new copulas from a given one. *Metrika*, 61(2):169–184, 2005. 94, 95

Tobias J Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2012. 109

David Nawrocki. Portfolio analysis with a large universe of assets. *Applied Economics*, 28(9):1191–1198, 1996. doi: 10.1080/000368496328038. URL http://dx.doi.org/10.1080/000368496328038. 236, 254

Roger B Nelsen. Dependence and order in families of Archimedean copulas. *Journal of Multivariate Analysis*, 60(1):111–122, 1997. 99, 101

Roger B Nelsen. Concordance and copulas: A survey. In *Distributions with given marginals and statistical modelling*, pages 169–177. Springer, 2002. 66, 67, 68

Roger B Nelsen. *An introduction to copulas*. Springer, 2006. 55, 58, 74, 78, 80, 101, 102

Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991. 168

Whitney K Newey and Kenneth D West. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pages 777–787, 1987. 228

Roger Newson. Parameters behind nonparametric statistics: Kendalls tau, somers d and median differences. *Stata Journal*, 2(1):45–64, 2002. 67

Vitaly Orlov. Currency momentum, carry trade, and market illiquidity. *Journal of Banking & Finance*, 67:1–11, 2016. 110

Annual Energy Outlook. Us energy information administration washington, 2013. 328

Alessandro Palandri. Sequential conditional correlations: Inference and evaluation. *Journal of Econometrics*, 153(2):122–132, 2009. 177

Franz C Palm. 7 garch models of volatility. *Handbook of statistics*, 14:209–240, 1996. 168

Daniel P Palomar and Yonina C Eldar. *Convex optimization in signal processing and communications*. Cambridge university press, 2010. 243

A.J. Patton. A review of copula models for economic time series. *Journal of Multivariate Analysis*, 2012. 73

Andrew J Patton. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, 2(1): 130–168, 2004. 61

Karl Pearson. Mathematical contributions to the theory of evolution.–on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, 60(359-367):489–498, 1896. 64

Liang Peng. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):49–50, 2006. 62

Gareth W Peters, Alice XD Dong, and Robert Kohn. A copula based bayesian approach for paid–incurred claims models for non-life insurance reserving. *Insurance: Mathematics and Economics*, 59:258–278, 2014. 102

Gareth William Peters, Mark Briers, Pavel Shevchenko, and Arnaud Doucet. Calibration and filtering for multi factor commodity models with seasonality: incorporating panel data from futures contracts. *Methodology and Computing in Applied Probability*, 15(4):841–874, 2013. 289, 297

Robert S. Pindyck. Inventories and the short-run dynamics of commodity prices. *The RAND Journal of Economics*, 25(1):141–159, 1994. ISSN 07416261. URL `http://www.jstor.org/stable/2555858`. 328

Robert S. Pindyck. The long-run evolution of energy prices. *The Energy Journal*, 20(2):1–27, 1999. ISSN 01956574, 19449089. URL `http://www.jstor.org/stable/41322828`. 317

Harry Pollard et al. The bernstein-widder theorem on completely monotonic functions. *Duke Math. J*, 11:427–430, 1944. 83

Fernando A.S. Postali and Paulo Picchetti. Geometric brownian motion and structural breaks in oil prices: A quantitative analysis. *Energy Economics*, 28 (4):506 – 522, 2006. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2006.02.011. URL `http://www.sciencedirect.com/science/article/pii/S0140988306000338`. 311, 317, 318

Marcel Prokopczuk and Yingying Wu. The determinants of convenience yields. *Available at SSRN 2372838*, 2013. 283, 284, 285, 299, 311

Robert Ready, Nikolai Roussanov, and Colin Ward. After the tide: Commodity currencies and global trade. *Journal of Monetary Economics*, 85:69–86, 2017. 111, 281

Juan C. Reboredo. Modelling oil price and exchange rate co-movements. *Journal of Policy Modeling*, 34(3):419 – 440, 2012. ISSN 0161-8938. doi: http://dx.doi.org/10.1016/j.jpolmod.2011.10.005. URL http://www.sciencedirect.com/science/article/pii/S0161893811001220. 328

Juan Carlos Reboredo, Miguel A. Rivera-Castro, and Gilney F. Zebende. Oil and {US} dollar exchange rate dependence: A detrended cross-correlation approach. *Energy Economics*, 42:132 – 139, 2014. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2013.12.008. URL http://www.sciencedirect.com/science/article/pii/S0140988313002910. 328

Thomas A Rietz. The equity risk premium a solution. *Journal of monetary Economics*, 22(1):117–131, 1988. 45, 114

John Riordan. Derivatives of composite functions. *Bull. Amer. Math. Soc*, 52(19):664–667, 1946. 100

Steven Roman. The formula of Faa di Bruno. *The American Mathematical Monthly*, 87(10):805–809, 1980. 99

Thierry Roncalli. *Introduction to risk parity and budgeting*. CRC Press, 2013. 248

Bryan R Routledge, Duane J Seppi, and Chester S Spatt. Equilibrium forward curves for commodities. *The Journal of Finance*, 55(3):1297–1338, 2000. 328

Felix Salmon. The formula that killed wall street. *Significance*, 9(1):16–20, 2012. 61

Irving De Lira Salvatierra and Andrew J Patton. Dynamic copula models and high frequency data. *Journal of Empirical Finance*, 30:120–135, 2015. 180

Andr A.P. Santos and Guilherme V. Moura. Dynamic factor multivariate garch model. *Computational Statistics & Data Analysis*, 76:606 – 617, 2014. ISSN 0167-9473. doi: http://dx.doi.org/10.1016/j.csda.2012.09.010. URL http://www.sciencedirect.com/science/article/pii/S0167947312003398. CFEnetwork: The Annals of Computational and Financial Econometrics2nd Issue. 238

Marco Scarsini. On measures of concordance. *Stochastica: revista de matemática pura y aplicada*, 8(3):201–218, 1984. 64

Bernd Scherer. Can robust portfolio optimisation help to build better portfolios? *Journal of Asset Management*, 7(6):374–387, 2007. doi: 10.1057/palgrave.jam.2250049. 240

Rafael Schmidt and Ulrich Stadtmüller. Non-parametric Estimation of Tail Dependence. *Scandinavian Journal of Statistics*, 33(2):307–335, 2006. 71

T. Schmidt. Coping with copulas. *Chapter forthcoming in Risk Books: Copulas from theory to applications in finance*, 2006. 58, 74

Philipp Schönbucher and Dirk Schubert. Copula-dependent defaults in intensity models. *Available at SSRN 301968*, 2001. 61

Eduardo Schwartz and James E Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000. 44, 284, 285, 286, 289, 297, 306, 308, 316, 318, 319, 322, 334

Eduardo S Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997. 284, 311, 318

Berthold Schweizer. Thirty years of copulas. In *Advances in probability distributions with given marginals*, pages 13–50. Springer, 1991. 55

Johan Segers. Discussion of copulas: Tales and facts, by thomas mikosch. *Extremes*, 9(1):51–53, 2006. 62

Enrique Sentana. Quadratic arch models. *The Review of Economic Studies*, 62 (4):639–661, 1995. 168

Robert J Serfling. *Approximation theorems of mathematical statistics*, volume 162. John Wiley & Sons, 2009. 105

William F. Sharpe. A simplified model for portfolio analysis. *Management Science*, 9(2):277–293, 1963. doi: 10.1287/mnsc.9.2.277. URL `http://dx.doi.org/10.1287/mnsc.9.2.277`. 238

Ali Shehadeh, Youwei Li, and Michael Moore. The forward premium bias, carry trade return and the risks of volatility and liquidity. 2016. 110

Neil Shephard. Statistical aspects of arch and stochastic volatility. *Monographs on Statistics and Applied Probability*, 65:1–68, 1996. 168

Annastiina Silvennoinen and Timo Teräsvirta. Multivariate garch models. In *Handbook of financial time series*, pages 201–229. Springer, 2009. 169

Renan Silvrio and Alexandre Szklo. The effect of the financial sector on the evolution of oil prices: Analysis of the contribution of the futures market to the price discovery process in the {WTI} spot market. *Energy Economics*, 34 (6):1799 – 1808, 2012. ISSN 0140-9883. doi: http://dx.doi.org/10.1016/j.eneco.2012.07.014. URL http://www.sciencedirect.com/science/article/pii/S0140988312001508. 43, 305

Kenneth J Singleton. Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60(2):300–318, 2013. 196

Kenneth J. Singleton. Investor flows and the 2008 boom/bust in oil prices. *Manage Sci*, 60(2):300–318, 2014. doi: 10.1287/mnsc.2013.1756. URL http://dx.doi.org/10.1287/mnsc.2013.1756. 311

Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8(1):11, 1959. 55, 56

Abe Sklar. Random variables, distribution functions, and copulas: a personal look backward and forward. *Lecture notes-monograph series*, pages 1–14, 1996. 55

Charles Spearman. "general intelligence", objectively determined and measured. *The American Journal of Psychology*, 15(2):201–292, 1904. 65

David Stefanovits, Urs Schubiger, and Mario V Wüthrich. Model risk in portfolio optimization. *Risks*, 2(3):315–348, 2014. 250

Ke Tang and Wei Xiong. Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(5):54–74, 2012. 309

Pravin K Trivedi and David M Zimmer. *Copula modeling: an introduction for practitioners*. Now Publishers Inc, 2007. 74

Yiu K Tse and Albert K C Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics*, 20(3):351–362, 2002. 176

R.H. Tutuncu and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187, 2004. ISSN 0254-5330. doi: 10.1023/B:ANOR.0000045281.41041.ed. URL http://dx.doi.org/10.1023/B%3AANOR.0000045281.41041.ed. 240

Roy Van der Weide. Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564, 2002. 174, 178, 239

Stephen Craig Pirrong Victor K. Ng. Fundamentals and volatility: Storage, spreads, and the dynamics of metals prices. *The Journal of Business*, 67(2):203–230, 1994. ISSN 00219398, 15375374. URL http://www.jstor.org/stable/2353103. 309

Ioannis D Vrontos, Petros Dellaportas, and Dimitris N Politis. A full-factor multivariate garch model. *The Econometrics Journal*, 6(2):312–334, 2003. 174

Martin L Weitzman. Subjective expectations and asset-return puzzles. *The American Economic Review*, 97(4):1102–1130, 2007. 41, 110

Peter Whittle. *Hypothesis testing in time series analysis*, volume 4. Almqvist & Wiksells, 1951. 166

Richard Edmund Williamson et al. Multiply monotone functions and their laplace transforms. *Duke Math. J*, 23(2):189–207, 1956. 87, 88

CF Jeff Wu. On the convergence properties of the em algorithm. *The Annals of statistics*, pages 95–103, 1983. 105

Chih-Chiang Wu, Huimin Chung, and Yu-Hsien Chang. The economic value of co-movement between oil price and exchange rate using copula-based garch models. *Energy Economics*, 34(1):270–282, 2012. 61

Jean-Michel Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, 18(5):931–955, 1994. 168

Kun Zhang and Laiwan Chan. Efficient factor garch models and factor-dcc models. *Quantitative Finance*, 9(1):71–91, 2009. doi: 10.1080/14697680802039840. URL http://dx.doi.org/10.1080/14697680802039840. 175, 238