CENTRE FOR
FINANCIAL RISK
Faculty of Business and Economics

MACQUARIE
University

Understanding the Interplay between Covariance Forecasting Factor Models and Risk Based Portfolio Allocations in Currency Carry Trades

WORKING PAPER 17-09

Matthew Ames, Guillaume Bagnarosa, Gareth W. Peters, Pavel V. Shevchenko

Understanding the Interplay Between Covariance Forecasting Factor Models and Risk Based Portfolio Allocations in Currency Carry Trades †

Matthew Ames[a,*], Guillaume Bagnarosa[b,a], Gareth W. Peters[a,c,d], Pavel Shevchenko[e]

[a]*Department of Statistical Science, University College London, UK*
[b]*ESC Rennes School of Business, France*
[c]*Oxford-Man Institute, Oxford University, UK*
[d]*Systemic Risk Center, London School of Economics, UK*
[e]*Department of Applied Finance and Actuarial Studies, Macquarie University, Australia*

Abstract

With the exception of naive methods for portfolio selection, such as the equal weighted approaches, all other methods of portfolio allocation are more or less sensitive to the quality of the inputs considered in constructing the models and risk measures utilised in the allocation framework. The extensively used factor model proposed initially by Sharpe (Sharpe (1963)) has provided a robust backdrop for development of relevant, micro, macro and context specific or asset specific explanatory variables to be incorporated in a statistical manner as inputs to forecasting models that can then be used to obtain risk measures upon which portfolio allocations are based. However, like all statistical models a set of statistical assumptions accompany this factor model regression framework, one of which has recently been highlighted as seemingly non-validated in financial data. This is of course the assumption such factor models make on homoskedasticity or weak sense covariance stationarity of the returns processes being modelled. Such factor models, therefore have typically failed to cope with an important and ubiquitous feature of financial assets data which often demonstrates heteroskedasticity of the returns variances and covariances.

We propose a novel generalised multi-factor forecasting structure utilizing a covariance regression model which allows us to incorporate the required heteroskedasticity effects whilst also admitting potential dependence in the idiosyncratic error terms. We argue that such a modelling approach allows for more explicit relationships to be interpreted between the driving factors and the conditional responses of the portfolio returns. We then compare the forecasting performances of our model with the multi-factor model and the time series DCC (Engle (2002)) model through a currency portfolio application.

Keywords: Covariance Forecasting, Currency Carry Trade, Covariance Regression, Generalised Multi-Factor Model, Portfolio Optimisation

---

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 certainly 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 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. Before proceeding we will split the problem of portfolio allocation into four stylized non-independent stages: the first stage typically involves statistical model estimation and model selection of the portfolio constituent multivariate return processes historically; the second stage typically involves some form of forecasting under the estimated model selected; the third stage involves selection and estimation of a risk measure on which to measure performance of the portfolio; and the third stage involves an optimization criterion upon which one performs portfolio allocation based on the portfolio forecasted risk measure.

With these four stages in mind and returning to the considerations of portfolio allocation under the classical mean-variance based models we note that several challenging model prerequisites for such a framework arise. Most notably these include the estimation of essential but unknown parameters such as each portfolio components 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,

†We would like to thank Eckhard Platen and Stefan Trueck as well as seminar participants at the University of Sydney, University of New South Wales, University of Technology Sydney and Macquarie University for their insightful comments and suggestions.

*Corresponding author
Email address: m.ames.12@ucl.ac.uk (Matthew Ames)

it is often 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 behaviours, examples can be found in both the academic and practitioners literature (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 which 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 and 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 definitely considered to be more reliably predictable and amenable to model developments. Not only have these features have 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 effect directly 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 in two categories which depend 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. We refer to stages one and two as "Upstream" approaches and stages three and four as "Downstream" approaches. In upstream approaches the natural solution would be to improve 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. Improving the modelling at this stage is a statistical pursuit and if achieved has the effect of reduction of the variability and noise on the input sources. An alternative set of solutions instead considers the input noise as an inexorable feature of financial markets data and accordingly focuses on downstream components i.e. stages three and four in the risk measure and optimization program objective function and the various constraints generally going with it. Such methods amount to 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 paper. Firstly, the factor models (Green and Hollifield (1992); Chan et al. (1999) or more recently Santos and Moura (2014)) where the potential portfolio assets have their conditional covariance matrix and drift modelled based on considerations of a constructed value-weighted market index, akin to the approach adopted in single factor models such as (Sharpe (1963)) or the augmented multi-factors models devised by Fama and French (Fama and French (1993)) or Carhart (Carhart (1997)). In the same vein the latent factor models promote instead a transformation and dimension reduction approach based on constructing factors which are orthogonal and typically obtained based on principle component analysis (PCA) based decompositions. This is achieved obviously then at the expense of the economic interpretation that would have been offered by the non-transformed factors (Han (2006), Zhang and Chan (2009)).

The second approach in the literature to tackle issues with the upstream modelling involve 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 models such as the widely considered class of Dynamic Conditional Correlation (DCC) models (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 correlation between the asset returns as a function of their past volatility and the correlations among 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 vector error correction (VEC) formulations developed in (Bollerslev (1990)) and the diagonal VEC (DVEC) and restricted VEC (often called BEKK) models (Engle and Kroner (1995)) models 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 in the model as linear combinations of uncorrelated factors. In this manner it is akin to the approach of principal component analysis dimension reductions. 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 model TVS models, see Christodoulakis and Satchell (2002).

Finally, the third approach has involved Bayesian statistics approaches which have been proposed to reduce the variance of the input estimator among which the technique of shrinkage which was originally applied to the mean parameter estimation by Jorion (1985) and Jorion (1986), subsequently completed by qualitative inputs with the so-called Black-Litterman model (Black and Litterman (1991)) and eventually 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 consists in optimally combining two existing estimators such as the sample estimator (respectively for the expected value or the variance covariance matrix) and for instance a factor model based estimator. More recently (Garlappi et al. (2006); Boyle et al. (2012); Branger et al. (2013)) utilize 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 proposes to focus on the optimization programs objective function and constraint specifications. Often termed robust portfolio theory, it proposes to deal with the upstream input estimator 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 (Scherer (2007); DeMiguel et al. (2009a)).

We clearly see that whatever the angle considered, both noise-reduction alternatives are closely related. In this context, our contribution we present in this paper to modern portfolio theory literature contains multiple aspects. We indeed first propose a new "upstream" model for the portfolio optimization inputs at the crossroad of the time series and multi-factors models. Considering a conditional mean and covariance regression model we accordingly manage to express the heteroskedastic component of the variance and covariance as a function of a set of relevant and known factors which are also intervening in the drift dynamic. We term such a model the Generalised multi-factor model framework (GFM). Furthermore, we demonstrate the influence of heteroskedasticity within the variance and covariance matrix upon the efficient frontier and the optimal mean-variance portfolio weights. We study the sensitivity of the portfolio allocation as well as developing and performing a stress testing framework based on our GFM model to assess the most influential factors in the portfolio allocation and resulting performance.

Finally, we apply our model on currency market data which are rarely considered in portfolio allocation literature notably for factor models empirical applications. In this regard, according to Lustig et al. (2011), we retain the dollar factor and an interest rates differential factor, embodied by the high minus low index proposed in the same article which affects differently the high and the low interest rates baskets of currencies and thus required a different modelling of conditional dependencies. We further complete Lustig (Lustig et al. (2011)) approach as we not only consider price factors as commonly proposed in factor models literature, but also volumes information and more specifically speculative volumes which following Ames et al. (2017) are definitely influencing the covariances between international exchange rates as a function of the interest rates differential prevailing for each country.

The rest of the paper is organised as follows: in section two we provide a brief review of the single factor and DCC models frequently used in the financial literature, we also precisely described the generalised heteroskedastic factor model that we propose to cope with the limitations associated to the standard multi factor homoskedastic version. The third section is devoted to the procedure we utilise to propagate the model covariates and thus forecast the future conditional and unconditional variance covariance matrix for our portfolio optimisation. Section four conducts a detailed review of the existing portfolio selection optimisation approaches, amongst which the risk based models such as the global minimum variance framework is discussed. We utilise this global minimum variance framework to compare the forecasting performance of our covariance model with others such as the standard multi-factor model and the widely considered Dynamic Conditional Correlation (DCC) family of models. In addition, under the Markowitz framework we also proposed in section five the derivation of the optimal portfolio weights sensitivity to the covariates variations. Finally, section six is dedicated to an empirical application to the currency market. Once we have described the currency carry trade strategy and its impact on the dependence structure of the currencies as a function of their relative level of interest rates, we describe the set of factors we retained for our model. Furthermore, in this section we also described precisely how the variance covariance models, being based on different filtrations can be fairly compared and we finally discuss our results.

2. Modelling Covariance

As mentioned previously the broad financial literature about asset price dynamic has already proposed various solutions to model the expected returns and the variance covariance between financial assets. Among them the factor model and its augmented version, the multi-factor model, makes the economic interpretation easier even though the orthogonality is not always guaranteed Christoffersen and Langlois (2013). Besides, while the multivariate GARCH models such as the DCC cope with a salient feature of financial asset prices, namely the returns variance and covariance heteroskedasticity, they do not provide a direct interpretation of the factors intervening in the dynamic of the drift and the conditional variance or covariance, instead they manifest in the form of transformations which obscure the ability to easily interpret and study the direct impact of such factors on resolving issues identified with capturing weak sense covariance stationarity in the portfolio returns. 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.

We begin by first presenting the standard multi-factor model and then we introduce our class of generalised factor models (GFM) based around an explicit covariance regression model, first devised in the statistics literature under a random effects formulation and subsequent EM algorithm estimation framework in Hoff and Niu (2012). We also give a detailed presentation of the estimation procedure we consider for this model, that is via a random-effects representation and expectation maximization (EM) algorithm that is numerically robust and efficient to implement in this context. We then conclude this section with a brief description of the DCC models (Engle (2002)) that we compare to our proposed GFM class of models.

We first define two sets of information filtration in our modelling frameworks. 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 sigma-algebra generated by $\mathcal{F}_t = \sigma\left(\boldsymbol{R}_t, \boldsymbol{R}_{t-1}, \ldots, \boldsymbol{R}_{t-T}\right)$ of length $T+1$, where $\boldsymbol{R}_t$ is the currency return vector at time t. The second filtration we will define is based on exogenous independent explanatory variables or factors $\boldsymbol{X}_t$ and will be denoted by $\mathcal{G}_t$. 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 sigma-algebra generated by $\mathcal{G}_t = \sigma\left(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-T}\right)$. 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 $\widetilde{\mathcal{G}}_t = \cup_{i=1}^{t} \mathcal{G}_i$. This extended filtration $\widetilde{\mathcal{G}}_t$ includes all of the covariate values in the t sliding windows, each of length $T+1$, and thus $\widetilde{\mathcal{G}}_t$ is of length $t \times (T+1)$.

Remark 2.1. *The filtrations above are introduced to address two key challenges:*

- *Firstly, in order to allow for the fact that the covariates $\boldsymbol{X}_t$ are sampled at different rates. For example, DOL and HML_{FX} can be calculated on a daily basis, whilst the speculative open interest rate covariate is only available on a weekly basis. Furthermore, a range of covariates can potentially be considered with different sampling rates.*
- *Secondly, the covariates $\boldsymbol{X}_t$ were found to be only locally stationary in $\mathcal{G}_t$, i.e. in the window $\left(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-T}\right)$, but not globally stationary in $\widetilde{\mathcal{G}}_t$. Therefore it is important to utilise $\mathcal{G}_t$ when considering the conditional covariance.*

2.1. Standard Factor Model

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

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{X}_t + \tilde{\boldsymbol{\epsilon}}_t \; , \tag{2.1}$$

where $N :=$ number of currencies and $K :=$ number of covariates,
$\boldsymbol{R}_t :=$ N-dimensional vector of returns of currencies in basket 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 also include incorporating lagged dependent variables such as in a multivariate ARDL models, cointegration ECM models and more recently (translation invariant copula models) can be written as such model structures with the vector valued innovation error distribution specified through marginals and a copula structure which may also be dynamic itself Salvatierra and Patton (2015) and Hafner and Manner (2012). Under the simple version of this multi-factor model in Equation 2.1 the unconditional covariance matrix (population estimator) is easily obtained according to the following terms:

$$Cov(\boldsymbol{R}_t | \mathcal{F}_{t-1} \cup \widetilde{\mathcal{G}}_{t-1}) = \boldsymbol{\beta}\mathbb{E}[\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}} | \widetilde{\mathcal{G}}_{t-1}]\boldsymbol{\beta}^{\mathrm{T}} + diag(\sigma_1^2, \ldots, \sigma_N^2) \; . \tag{2.2}$$

One can think of this unconditional, in the sense of filtration $\mathcal{G}_{t-1}$, as being a population based realisation. Then there is the conditional covariance matrix given the covariates values $\boldsymbol{X}_t$ according to the expression:

$$Cov(\boldsymbol{R}_t | \mathcal{F}_{t-1} \cup \mathcal{G}_t) = diag(\sigma_1^2, \ldots, \sigma_N^2) \; . \tag{2.3}$$

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. We see 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 $\boldsymbol{X}_t$.

2.2. Generalised Multi-Factor Model Specification

The new covariance regression model developed in this paper, we term 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, we will demonstrate how this model can be estimated 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 2.3.

Weekly carry returns are defined as $\boldsymbol{R}_t = \frac{F_{t,T}}{F_{t-1,T}} - 1$, where $F_{t,T}$ is the price of the future contract at time t with maturity T, i.e. we use a weekly mark-to-market procedure to calculate the relative return on each currency position (see Ames et al. (2017)). In order to capture the heteroskedastic effects of the covariates on the covariance of the currency carry returns, $\boldsymbol{R}_t$, we use the following model:

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{X}_t + \boldsymbol{e}_t \ , \tag{2.4}$$

where $N :=$ is the number of currencies,
$K :=$ is the number of covariates,
$\boldsymbol{R}_t :=$ is the N-dimensional carry returns in basket,
$\boldsymbol{\alpha} :=$ 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,
$\boldsymbol{e}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})$ 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 $\boldsymbol{e}_t$.

Under the multi-factor model we develop we begin by considering the unconditional and conditional covariance matrices of this simple multi-factor regression model presented in Equation 2.4. It is trivial to derive the unconditional covariance matrix as follows:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_{t-1} \cup \widetilde{\mathcal{G}}_{t-1}) = \boldsymbol{\beta}\mathbb{E}[\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}|\widetilde{\mathcal{G}}_{t-1}]\boldsymbol{\beta}^{\mathrm{T}} + \boldsymbol{C}\mathbb{E}(\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}|\widetilde{\mathcal{G}}_{t-1})\boldsymbol{C}^{\mathbf{T}} + \boldsymbol{\Psi} \ , \tag{2.5}$$

where in this case we assume the observed factors $\boldsymbol{X}_t$ are also random vectors and therefore admit a covariance structure that is locally stationary. We can also specify the conditional covariance matrix of this multi-factor model in Equation 2.4, given the factors, as follows:

$$Cov(\boldsymbol{R}_t|\mathcal{F}_{t-1} \cup \mathcal{G}_t) = \boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi} \ , \tag{2.6}$$

where the conditional covariance will be specified according to two terms, the baseline covariance structure 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 portfolio returns over time. From this we first observe that the heteroskedasticity in the conditional covariance is given by the covariance of the error terms $\boldsymbol{e}_t$.

2.3. Generalised Factor Model: Covariance Regression Model Estimation via Random-Effects Representation

To perform estimation it is convenient to formulate our covariance regression model as a special type of random-effects model, see Hoff and Niu (2012), for observed data $\boldsymbol{R}_1, \ldots \boldsymbol{R}_T$ (N-dimensional basket weekly carry returns of length T weeks).

$$\begin{aligned} &\boldsymbol{R_t} = \alpha + \boldsymbol{\beta}\boldsymbol{X}_t + \gamma_t \times \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{\epsilon}_t \ , \\ &\mathbb{E}[\boldsymbol{\epsilon}_t] = 0 \quad , \quad Cov(\boldsymbol{\epsilon}_t) = \boldsymbol{\Psi} \ , \\ &\mathbb{E}[\gamma_t] = 0 \quad , \quad Var[\gamma_t] = 1 \quad , \quad \mathbb{E}[\gamma_t \times \boldsymbol{\epsilon}_t] = 0. \end{aligned} \tag{2.7}$$

Step 1: Mean De-trending of Returns.

The first step is to perform the mean-regression, via in our case a standard linear regression model. This will allow us to obtain zero-mean residuals $\hat{\boldsymbol{e}}_t$, given by $\hat{\boldsymbol{e}}_t = \boldsymbol{R}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}\boldsymbol{X}_t$, where $\hat{\boldsymbol{\beta}}$ 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:

$$\begin{aligned}
&\hat{\boldsymbol{e}}_t = \gamma_t \times \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{\epsilon}_t \ , \\
&\mathbb{E}[\boldsymbol{\epsilon}_t] = 0 \quad , \quad Cov(\boldsymbol{\epsilon}_t) = \boldsymbol{\Psi} \ , \\
&\mathbb{E}[\gamma_t] = 0 \quad , \quad Var[\gamma_t] = 1 \quad , \quad \mathbb{E}[\gamma_t \times \boldsymbol{\epsilon}_t] = 0.
\end{aligned} \tag{2.8}$$

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

$$\begin{aligned}
\Sigma_{\boldsymbol{X}_t} &:= \mathbb{E}[\hat{\boldsymbol{e}}_t \hat{\boldsymbol{e}}_t^{\mathrm{T}} | \boldsymbol{X}_{\boldsymbol{t}}] \\
&= \mathbb{E}[\gamma_t^2 \boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \gamma_t(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{\epsilon}_t^{\mathrm{T}} + \boldsymbol{\epsilon}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}) + \boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^{\mathrm{T}} | \boldsymbol{X}_t] \\
&= \boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi}.
\end{aligned}$$

This random-effects model allows us to perform the maximum likelihood parameter estimation of the coefficients, $\boldsymbol{C}$ and $\boldsymbol{\Psi}$, via the EM algorithm. We proceed by iteratively maximising the complete data log-likelihood of $\hat{\boldsymbol{E}} = \hat{\boldsymbol{e}}_1, \ldots, \hat{\boldsymbol{e}}_T$ 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:

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

We 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 $m_t = v_t(\hat{\boldsymbol{e}}_t^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{C}\boldsymbol{X}_{\boldsymbol{t}})$ and variance $v_t = (1 + \boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{C}\boldsymbol{X}_t)^{-1}$. 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).

2.4. DCC Model

We start by mentioning the Constant Conditional Correlation (CCC) model which would consider the covariance matrix of the returns in the portfolio to be specified according the a structure given by

$$\boldsymbol{H}_t = \boldsymbol{D}_t\boldsymbol{R}\boldsymbol{D}_t \tag{2.9}$$

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 of each asset is assumed constant over time and given by matrix R. Under the DCC model proposed by Engle (2002) this correlation is specified to be dynamically evolving in time. The DCC model has been extensively studied in the literature (Engle and Colacito (2006); Aielli (2013)) and proposed to model the conditional variance as univariate GARCH while the conditional correlations are peculiar functions of the past GARCH standardized returns and the past values of the conditional correlations. More formally, according to this model we can write the conditional covariance matrix according to the following specification:

$$\boldsymbol{H}_t = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t \tag{2.10}$$

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:

$$\mathbf{R_t} = diag(q_{11,t}^{-1/2}, \ldots, q_{NN,t}^{-1/2})\mathbf{Q_t}diag(q_{11,t}^{-1/2}, \ldots, q_{NN,t}^{-1/2}) \tag{2.11}$$

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

$$\mathbf{Q_t} = (1 - \alpha - \beta)\bar{\mathbf{Q}}) + \alpha u_{t-1}u_{t-1}^{\mathrm{T}} + \beta\mathbf{Q_{t-1}} \tag{2.12}$$

with $u_t = \epsilon_{it}/\sqrt{h_{ii,t}}$. $\bar{\mathbf{Q}}$ 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 $\bar{\mathbf{Q}}$ can be estimated jointly with the other model parameters or can be set to the sample estimate to simplify the procedure and reduce the number of parameters, see the papers cited above for estimation procedures for the DCC model family.

We note that in this class of standard DCC models we will consider the models specified with respect to the filtration $\mathcal{F}_t$. As discussed above, there are extensions of such models such as under the GO-GARCH variations of the DCC model that would allow for joint conditioning on filtrations such as $\mathcal{F}_t \cup \mathcal{G}_t$.

3. Covariates and Covariances Forecasting

In this section, we present the method utilised to obtain forecasts of the returns covariance matrix under the Generalised multi-factor model (GFM) framework we have developed in the previous section. In order to obtain forecasts of the covariance of the carry returns we need to first forecast the covariates vector $\boldsymbol{X}$. To do so, we use the Hyndman-Khandakar algorithm for automatic SARIMA modelling as implemented in the *auto.arima* function in the R forecast package Hyndman (2015), see Hyndman and Khandakar (2008) for details.

3.1. Forecasting Covariates

In order to obtain forecasts of the covariance of the returns we need to forecast the covariates $\boldsymbol{X}$. We use the Hyndman-Khandakar algorithm for automatic SARIMA modelling as implemented in the auto.arima function in the R forecast package Hyndman (2015), see Hyndman and Khandakar (2008) for details.

For each of the sliding windows we have a vector valued time series of the covariate values given by $\boldsymbol{X}_t, \boldsymbol{X}_{t+1}, \ldots, \boldsymbol{X}_{t+T}$ where each vector $\boldsymbol{X}_t \in \mathbb{R}^K$. For the forecasting of the covariance at time $t+T+\tau$ on forecast horizon τ relative to the current sliding window, we must forecast the covariates. To achieve this, for each of the sliding windows we first construct a model for each covariate time series marginally and then forecast each from its corresponding model. The outline of the algorithm to fit the ARIMA model to each covariate time series is as follows:

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.

Such an automated procedure for model selection is particularly relevant in the context of the modelling in this paper as we 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 which is equivalent to $270 \times (15 + 11)$ total number of models to be fit. With this many models we need an automatic and efficient procedure.

To make sure the fitted ARIMA models were suitable at each fit we then assess the accuracy of the forecasts using the Mean absolute scaled error (MASE) as given in Definition 3.1 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.

Definition 3.1. *Mean Absolute Scaled Error (MASE) For each time series ARIMA model fit we consider the following metric to assess performance:*

$$MASE_\tau = \frac{1}{\tau}\sum_{t=1}^{\tau}\left(\frac{\tilde{e}_t}{\frac{1}{n-1}\sum_{i=2}^{n}|X_i - X_{i-1}|}\right) \tag{3.1}$$

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 of the one-step naïve (random walk) forecast method, which uses the actual value from the prior period as the forecast, i.e. $\hat{X}_t = X_{t-1}$.

In addition to the MASE criterion we also consider the Mean Absolute Percentage Error (MAPE) criterion to assess the forecast performance.

Definition 3.2 (Mean Absolute Percentage Error (MAPE))**.**

$$MAPE_\tau = 100 \times \frac{1}{\tau}\sum_{t=1}^{\tau}\left(\frac{|X_t - \hat{X}_t|}{|X_t|}\right) \tag{3.2}$$

where the numerator is the forecast error at time t.

Given forecasts of the covariate time series, we forecast the τ-step ahead unconditional covariance matrix:

1. Fit Generalised multi-factor model to the data period $[t-L:t]$ via the method in Section 2.3 to obtain parameter estimates $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\Psi}}$ and $\hat{\boldsymbol{C}}$. L is the lookback period: here $L = 125$ 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: $\widehat{Cov}(\boldsymbol{R}_{t+\tau|t}|\mathcal{F}_t \cup \mathcal{G}_t) = \hat{\boldsymbol{\beta}} Cov(\hat{\boldsymbol{X}}_{t+\tau|t}|\mathcal{G}_t)\hat{\boldsymbol{\beta}}^{\mathrm{T}} + \hat{\boldsymbol{C}}\mathbb{E}(\hat{\boldsymbol{X}}_{t+\tau|t}\hat{\boldsymbol{X}}^{\mathrm{T}}_{t+\tau|t}|\mathcal{G}_t)\hat{\boldsymbol{C}}^{\mathrm{T}} + \hat{\boldsymbol{\Psi}}$.

and the conditional covariance matrix forecast:

$$\widehat{Cov}(\boldsymbol{R}_{t+\tau|t}|\hat{\boldsymbol{X}}_{t+\tau|t}, \mathcal{F}_t \cup \mathcal{G}_t) = \hat{\boldsymbol{C}}\hat{\boldsymbol{X}}_{t+\tau|t}\hat{\boldsymbol{X}}^{\mathrm{T}}_{t+\tau|t}\hat{\boldsymbol{C}}^{\mathrm{T}} + \hat{\boldsymbol{\Psi}} .$$

4. Portfolio Optimisation

In order to compare the accuracy of the variances and covariances forecasts following from our heteroskedastic Generalised multi-factor model (GFM) with those generated through the Standard multi-factor model (SFM) and the DCC model, we discuss a set of allocation approaches which are unevenly impacted by the input sources of uncertainty which 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. In the empirical analysis that follows this section, we illustrate this using one popular approach to portfolio optimisation, namely Global Minimum Variance (GMV).

Typically considered naive methods are the equal weighted method studied recently by DeMiguel et al. (2009b) as well as the 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 assets volatilities.

In addition to these we also discuss the classical approach of the mean-variance optimization program proposed by Markowitz (1952). The prerequisites for this method are expected value and variances covariances assessments. This method is as a consequence highly sensitive to the measurement error affecting the expected value and to a lower extent the variances and the covariances forecasts errors (Jobson and Korkie (1981); Frost and Savarino (1988); Michaud (1989); Chopra and Ziemba (1993); Broadie (1993); Nawrocki (1996)). We then complete our discussion with several other risk based approaches to portfolio allocation that have been more recently proposed, in this regard we focus on the portfolio allocation methods displaying more or less sensitivity to the variances and covariances measurement error. In particular, the equal risk contribution (ERC hereafter) proposed by Maillard et al. (2008), the minimum variance portfolio proposed by Haugen and Baker (1991) as well as the maximum diversification devised by Choueifaty and Coignard (2008) are presented. Interestingly, it has been recently pointed out by Jurczenko et al. (2015) that the measurement error on the variances and the covariances used as input is particularly influencing the optimal weights worked out through this techniques even though 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[2].

4.1. Markowitz Mean-Variance Optimal Portfolio

To begin, we briefly recall 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. The following derivation focuses on incorporating the conditional covariance information as in Equation 2.6, i.e. $\boldsymbol{\Sigma} = \boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi}$. However, if we are interested in the unconditional covariance (as in standard Markowitz optimisation) then we use $\boldsymbol{\Sigma}$ as in Equation 2.5, i.e. $\boldsymbol{\Sigma} = \boldsymbol{\beta}\mathbb{E}[\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}|\widetilde{\mathcal{G}}_{t-1}]\boldsymbol{\beta}^{\mathrm{T}} + \boldsymbol{C}\mathbb{E}(\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}|\widetilde{\mathcal{G}}_{t-1})\boldsymbol{C}^{\mathbf{T}} + \boldsymbol{\Psi}$.

We seek to solve the following unconstrained optimisation problem:

$$\min_{\boldsymbol{w}} \sigma^2_{p,\boldsymbol{w}} = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{w} = \boldsymbol{w}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w} \;\; s.t. \tag{4.1}$$

$$\mu_p = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\mu} = \mu_{p,0} \text{ , and } \boldsymbol{w}^{\mathrm{T}}\mathbf{1} = 1 \; .$$

To solve the unconstrained minimization problem 4.1, first we form the Lagrangian function

$$L(w, \lambda_1, \lambda_2) = \boldsymbol{w}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w} + \lambda_1(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\mu} - \mu_{p,0}) + \lambda_2(\boldsymbol{w}^{\mathrm{T}}\mathbf{1} - 1) \; . \tag{4.2}$$

The first order conditions (FOCs) for a minimum are thus the linear equations

$$\frac{\partial L(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \boldsymbol{w}} = 2(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w} + \lambda_1\boldsymbol{\mu} + \lambda_2\mathbf{1} = \mathbf{0} \; , \tag{4.3}$$

[2] Please refer to Jurczenko et al. (2015) for a more detailed review.

$$\frac{\partial L(\boldsymbol{w},\lambda_1,\lambda_2)}{\partial \boldsymbol{\lambda_1}} = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{\mu} - \mu_{p,0} = 0 \,, \tag{4.4}$$

$$\frac{\partial L(\boldsymbol{w},\lambda_1,\lambda_2)}{\partial \boldsymbol{\lambda_2}} = \boldsymbol{w}^{\mathrm{T}}\mathbf{1} - 1 = 0 \,. \tag{4.5}$$

We can represent the system of linear equations using matrix algebra as

$$\begin{pmatrix} 2(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi}) & \boldsymbol{\mu} & \mathbf{1} \\ \boldsymbol{\mu}^{\mathrm{T}} & 0 & 0 \\ \mathbf{1}^{\mathrm{T}} & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mu_{p,0} \\ 1 \end{pmatrix} \,,$$

or

$$\boldsymbol{A}\boldsymbol{z_w} = \boldsymbol{b_0} \,, \tag{4.6}$$

where

$$\boldsymbol{A} = \begin{pmatrix} 2(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi}) & \boldsymbol{\mu} & \mathbf{1} \\ \boldsymbol{\mu}^{\mathrm{T}} & 0 & 0 \\ \mathbf{1}^{\mathrm{T}} & 0 & 0 \end{pmatrix} \,, \quad \boldsymbol{z_w} = \begin{pmatrix} \boldsymbol{w} \\ \lambda_1 \\ \lambda_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b_0} = \begin{pmatrix} \mathbf{0} \\ \mu_{p,0} \\ 1 \end{pmatrix} \,. \tag{4.7}$$

The solution for $\boldsymbol{z_w}$ is then

$$\boldsymbol{z_w} = \boldsymbol{A}^{-1}\boldsymbol{b_0} \,. \tag{4.8}$$

We note that the first d elements of $\boldsymbol{z_w}$ are the optimal portfolio weights $\boldsymbol{w} = (w_1, \ldots, w_d)$ 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.

Having obtained an efficient frontier we may select a criterion to consider the portfolio of interest on this mean-variance efficient frontier. For example, the Markowitz portfolio with the maximum Sharpe ratio, where the Sharpe ratio is defined as expected return divided by expected volatility. Also one can consider the largest diversification portfolio on the efficient frontier, where diversification is related to volatility based risk measures.

In the empirical investigation performed in this paper we utilise the global minimum variance portfolio (GMV) in order to focus on the covariance modelling techniques introduced. The GMV portfolio can be found as the solution to the Markowitz problem under the assumption that all assets have equal returns.

Definition 4.1. ***Global Minimum Variance Portfolio***
The unconstrained Global Minimum Variance (GMV) portfolio satisfies the following minimisation problem:

$$\min_{\boldsymbol{w}} \sigma^2_{p,\boldsymbol{w}} = \boldsymbol{w^T\Sigma w} = \boldsymbol{w^T(CX_tX_t^TC^T + \Psi)w} \;\; s.t. \;\; \boldsymbol{w^T1} = 1 \,. \tag{4.9}$$

which has the solution

$$\boldsymbol{w}^{\star}_{GMV} = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1}} = \frac{(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})^{-1}\mathbf{1}}{\mathbf{1}^T(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})^{-1}\mathbf{1}} \tag{4.10}$$

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 (for more detail, please refer to Haugen and Baker (1991); Clarke et al. (2011)).

Remark 4.2. *In the currency portfolio carry trade examples we study below we will require constraints on the weights of the currencies to be positive in the long basket and negative in the short basket. For this constrained case, there is no closed form solution available and so we utilise a quadratic programming approach. See Boyd and Vandenberghe (2004); Palomar and Eldar (2010) for detailed references.*

4.2. Risk Based and Naive Optimisation 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. Therefore, risk-based techniques have arisen as an alternative, the example considered here is Equal Risk Contribution (ERC), see Maillard et al. (2008). In the version of ERC explained below the aim is to find the portfolio in which each asset contributes equally to the total portfolio variance. A more general approach to risk-based portfolio construction can be seen in Jurczenko et al. (2015); Stefanovits et al. (2014). The optimal weights vector $\boldsymbol{w}^{\star}$ associated to this model can be expressed as:

$$\boldsymbol{w}^{\star} = \{\boldsymbol{w} \in [0,1]^N : \sum w_i = 1, w_i \times \partial_{w_i}\sigma(\boldsymbol{w}) = w_j \times \partial_{w_j}\sigma(\boldsymbol{w}) \; \forall i,j\} \tag{4.11}$$

where $\partial_{w_i}\sigma(\boldsymbol{w}) = \frac{\partial\sigma(\boldsymbol{w})}{\partial w_i}$ and $\sigma(\boldsymbol{w}) = \sqrt{\boldsymbol{w}^{\mathrm{T}}\Sigma\boldsymbol{w}} = \sqrt{\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w}}$. Noting 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 $\boldsymbol{w}$, we have the following optimisation problem:

$$
\begin{aligned}
\boldsymbol{w}^{\star} &= \{\boldsymbol{w} \in [0,1]^N : \sum w_i = 1, w_i \times (\Sigma\boldsymbol{w})_i = w_j \times (\Sigma\boldsymbol{w})_j \ \forall i,j\} && (4.12)\\
&= \{\boldsymbol{w} \in [0,1]^N : \sum w_i = 1, w_i \times ((\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w})_i = w_j \times ((\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w})_j \ \forall i,j\} && (4.13)
\end{aligned}
$$

This problem can be solved using a Sequential Quadratic Programming (SQP) algorithm, see details in Maillard et al. (2008).

Choueifaty and Coignard (2008) proposed another risk based method in their article, which involved maximising a measurement of the portfolio diversification which was expressed according to the following ratio:

$$D = \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\sigma}}{\boldsymbol{w}^{\mathrm{T}}\Sigma\boldsymbol{w}} = \frac{\boldsymbol{w}^{\mathrm{T}}\sqrt{diag(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})}}{\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi})\boldsymbol{w}}. \tag{4.14}$$

If we consider synthetic portfolios P_i combining $1/\sigma_i$ of the i-th asset A_i and the rest in cash such that:

$$P_i = \frac{A_i}{\sigma_i} + (1 - 1/\sigma_i)\,A_f = \frac{A_i}{\sqrt{(\boldsymbol{c}_i^T\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{c}_i^{\mathrm{T}} + \boldsymbol{\psi}_{i,i})}} + \left(1 - \frac{1}{\sqrt{(\boldsymbol{c}_i^T\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{c}_i^{\mathrm{T}} + \boldsymbol{\psi}_{i,i})}}\right)A_f \tag{4.15}$$

where $\boldsymbol{c}_i$ is the i-th column of the matrix C and A_f embodies the risk free asset while we assume that we can borrow and lend cash at the same rate. Then, the optimisation (4.14) can be rewritten such as a minimum variance optimisation program[3]. The optimal weights $w_i^{\star}$ we thus obtain have to be adjusted after computation to take into consideration the initial normalisation. Eventually, the optimal weights vector invested in the risky assets available can thus be written such as:

$$\boldsymbol{w}^{\star} = \left(\frac{w_1^{\star}}{\sigma_1}, \ldots, \frac{w_N^{\star}}{\sigma_N}\right) = \left(\frac{w_1^{\star}}{\sqrt{(\boldsymbol{c}_1^T\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{c}_1^{\mathrm{T}} + \boldsymbol{\psi}_{1,1})}}, \ldots, \frac{w_N^{\star}}{\sqrt{(\boldsymbol{c}_N^T\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{c}_N^{\mathrm{T}} + \boldsymbol{\psi}_{N,N})}}\right), \tag{4.16}$$

while the remaining money $\left(1 - \sum_{i=1}^{N}\frac{w_i^{\star}}{\sigma_i}\right) = \left(1 - \sum_{i=1}^{N}\frac{w_i^{\star}}{\sqrt{(\boldsymbol{c}_i^T\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{c}_i^{\mathrm{T}} + \boldsymbol{\psi}_{i,i})}}\right)$ is invested in the risk free asset A_f.

Finally, naive methods typically considered are the equal weighted approach where each asset receives a $1/N$ allocation and the risk parity approach described in Anderson et al. (2012); Asness et al. (2012) where an equal correlation among each pair of assets is assumed which leads to a $1/\sigma_i$ allocation for each asset i.

5. Conditional Covariance and Optimal Markowitz Weight Sensitivity to Factors

In this section, we provide expressions for the sensitivity of the conditional covariance, the optimal Markowitz portfolio weights and the GMV weights to the explanatory exogenous factors that make up the filtration $\mathcal{G}_t$.

5.1. Conditional Covariance Sensitivity to Covariates

We begin by expressing the sensitivity of the conditional covariance of the portfolio under GFM model to each factor in the model.

$$
\begin{aligned}
\Sigma_{X_t} &= \mathbb{E}[e_t e_t^{\mathrm{T}}|\mathcal{F}_{t-1} \cup \mathcal{G}_t]\\
&= \Psi + CX_tX_t^{\mathrm{T}}C^{\mathrm{T}}\,.
\end{aligned} \tag{5.1}
$$

$$Cov(e_m, e_n|\mathcal{F}_{t-1} \cup \mathcal{G}_t) = \Sigma_{X_t}^{m,n} = \Psi^{m,n} + (C_{m,:}X_t) \times (C_{n,:}X_t)\,, \tag{5.2}$$

where $m = 1, \ldots, d$ and $n = 1, \ldots, d$ and d is the number of assets in the portfolio.
Differentiating Σ_{X_t} w.r.t covariate $X_{k,t}$ gives:

$$\frac{\partial\Sigma_{X_t}^{m,n}}{\partial X_{k,t}} = (C_{m,k} \times (C_{n,:}X_t)) + (C_{n,k} \times (C_{m,:}X_t))\,. \tag{5.3}$$

In the results section we will then be able to utilise these results to study the influence that each factor has on the portfolio allocation and performance in the SFM and GFM frameworks. We can also utilise these results to study the effect of the factors on the forecast covariance performance.

[3]In other words maximising diversification is similar to minimising $(\boldsymbol{w}^{\mathrm{T}}C\boldsymbol{w})$ where C corresponds to the correlation matrix of the original assets.

5.2. Optimal Markowitz Weights Sensitivity to Covariates (Weights can be negative here)

Having obtained the sensitivity of the conditional covariance of the portfolio under GFM model to each factor in the model we can now extend this to study the sensitivity of the allocation weights selected for the portfolio to the factors.

$$\begin{aligned}
\frac{\partial \boldsymbol{z_w}}{\partial X_{k,t}} &= \left(\frac{\partial \boldsymbol{A}^{-1}}{\partial X_{k,t}} \times \boldsymbol{b_0}\right) + \underbrace{\left(\frac{\partial \boldsymbol{b_0}}{\partial X_{k,t}} \times \boldsymbol{A}^{-1}\right)}_{=0,\text{ since } \boldsymbol{b_0} \text{ doesn't depend on } X_{k,t}} && (5.4)\\
&= \left(-\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial X_{k,t}}\boldsymbol{A}^{-1}\right)\boldsymbol{b_0} && (5.5)\\
&= \left(-\boldsymbol{A}^{-1}\begin{bmatrix} 2\left((\boldsymbol{C}_{n,:}X_t)\boldsymbol{C}_{m,k} + (\boldsymbol{C}_{m,:}X_t)\boldsymbol{C}_{n,k}\right) & \boldsymbol{\beta}_{:,k} & 0\\ \boldsymbol{\beta}^{\mathrm{T}}_{:,k} & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\boldsymbol{A}^{-1}\right)\boldsymbol{b_0}\,. && (5.6)
\end{aligned}$$

where $\boldsymbol{\beta}_{:,k}$ is the mean regression loading for the k-th covariate, as in equation 2.3.

5.3. Global Minimum Variance Weights Sensitivity to Covariates (Weights can be negative here)

Similarly, the sensitivity of the GMV portfolio allocation weights to the covariates can also be obtained as follows:

$$\begin{aligned}
\frac{\partial \boldsymbol{z}^{GMV}_{\boldsymbol{w}}}{\partial X_{k,t}} &= \left(\frac{\partial \boldsymbol{A}^{-1}_{GMV}}{\partial X_{k,t}} \times \boldsymbol{b}^{GMV}_{\boldsymbol{0}}\right) + \underbrace{\left(\frac{\partial \boldsymbol{b}^{GMV}_{\boldsymbol{0}}}{\partial X_{k,t}} \times \boldsymbol{A}^{-1}_{GMV}\right)}_{=0,\text{ since } \boldsymbol{b}^{GMV}_{\boldsymbol{0}} \text{ doesn't depend on } X_{k,t}} && (5.7)\\
&= \left(-\boldsymbol{A}^{-1}_{GMV}\frac{\partial \boldsymbol{A}_{GMV}}{\partial X_{k,t}}\boldsymbol{A}^{-1}_{GMV}\right)\boldsymbol{b}^{GMV}_{\boldsymbol{0}} && (5.8)\\
&= \left(-\boldsymbol{A}^{-1}_{GMV}\begin{bmatrix} 2\left((\boldsymbol{C}_{n,:}X_t)\boldsymbol{C}_{m,k} + (\boldsymbol{C}_{m,:}X_t)\boldsymbol{C}_{n,k}\right) & 0\\ 0 & 0\end{bmatrix}\boldsymbol{A}^{-1}_{GMV}\right)\boldsymbol{b}^{GMV}_{\boldsymbol{0}}\,. && (5.9)
\end{aligned}$$

where

$$\boldsymbol{A}_{GMV} = \begin{pmatrix} 2(\boldsymbol{C}\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{\Psi}) & \boldsymbol{1}\\ \boldsymbol{1}^{\mathrm{T}} & 0\end{pmatrix}\,,\quad \boldsymbol{z}^{GMV}_{\boldsymbol{w}} = \begin{pmatrix}\boldsymbol{w}\\ \lambda\end{pmatrix} \text{ and } \boldsymbol{b}^{GMV}_{\boldsymbol{0}} = \begin{pmatrix}\boldsymbol{0}\\ 1\end{pmatrix}\,. \tag{5.10}$$

Remark 5.1. *These sensitivity results can now be utilised in future sections 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.*

6. Empirical Application to the Currency Portfolio Analysis

This section being devoted to our empirical application to the currency market we first describe some of its characteristics such as the currency carry trade strategy. Subsequently we enumerate and analyse the set of factors that we retain for our generalised factor model while we explain in detail the procedure we followed for the comparison of variance covariance forecasting models and finally discuss the resulting performances of our approach.

6.1. The Currency Market and The Carry Trade

Understanding the dynamic of exchange rates has always been a central question in economics and finance. The information carried by such exchange rate signals or variables in our globalised economies is indeed at the crossroad of several other dynamic variables such as the balance of payments between the country considered and the rest of the world, the international interest rates and yield curves joint dynamics. In addition, it is known that such signals can be strongly influenced by speculators[4] interventions on financial markets (Brunnermeier et al. (2008); Anzuini and Fornari (2012); Hutchison and Sushko (2013); Fong (2013)).

In response to the behaviour of the exchange rate signals and knowledge of their influence and reactivity to different macroeconomic environments, there has developed a popular currency trading strategy named the currency carry trade. This has been commonly implemented over the last decades by investors looking for supposedly low risk strategies to leverage. This strategy consists in constructing portfolios by selling low interest rate currencies in order to buy

[4]Following the definition provided by the CFTC, futures market positions are identified as non-speculative when “their purpose is to offset price risks incidental to commercial cash or spot operations and such positions are established and liquidated in an orderly manner in accordance with sound commercial practices”. CFTC Regulation 1.3, 17 CFR 1.3(z).

high interest rate currencies and thus profit from the interest rate differentials between the two countries involved in the cross rate. Under the uncovered interest rates parity hypothesis (UIP thereafter), such profit opportunities should not occur recurrently or at least should not be profitable on average. However, several empirical works already demonstrated the existence of such UIP violations and the resulting sizable profits (Hansen and Hodrick (1980); Fama (1984); 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. (2012); Ames et al. (2015)). Before introducing UIP we first have to define a no-arbitrage relationship well known in finance which establishes the relation between spot and forward exchange rates as a function of the two nominal interest rates prevailing in the respective countries. Under the efficient market hypothesis, borrowing in the low interest rates countries to invest on a certain time horizon in the high interest rates countries and covering the currency risk through a position in the same maturity forward contract should not be profitable since the forward premium or discount should offset the differential of interest rates existing between the two countries. More formally we can write the CIP relation as follows:

Definition 6.1. *Covered Interest Rate Parity (CIP)*
This relation states that the forward price at time t of one unit of foreign currency against the base currency (which here is the US dollar) with maturity T can be expressed as:

$$F_t^T = e^{(r_{t,T} - r_{t,T}^f)(T-t)} S_t, \tag{6.1}$$

where S_t denotes the price of one unit of foreign currency at time t (spot price). While $r_{t,T}$ and $r_{t,T}^f$ represent the domestic[5] *and foreign risk free interest rate yields for maturity T. The CIP condition states that one should not be able to make a risk free profit by selling a forward contract and replicating its payoff through the spot market.*

Furthermore, the validity of this arbitrage relation on a daily basis, which has been demonstrated empirically in the currency market by Juhl et al. (2006); Akram et al. (2008), also permits to use the forward contracts in order to simply replicate the involved financial operation described earlier which consists in borrowing in one country at the risk free rates and investing it right away in the other countries rate. This statement also justifies that we use forward contracts to build carry trade strategies which benefits from the differential of interest rates between the low and high interest rates countries. Moreover, this relation also intervenes in the UIP validity conditions where the CIP is assumed to hold. As a matter of fact, the UIP affirms that under the historical probability distribution the expected change in the currency spot rates equals the differential of interest rates as follows:

Definition 6.2. *Uncovered Interest Rate Parity (UIP)*
Considering Equation 6.1 the UIP condition can be defined as:

$$E\left[\frac{S_T}{S_t}\middle|\mathcal{F}_t\right] = \frac{F_t^T}{S_t} = e^{(r_{t,T} - r_{t,T}^f)(T-t)}, \tag{6.2}$$

where $\mathcal{F}_t$ is the filtration associated to the stochastic process S_t. The UIP equation indeed states that the expected variation of the exchange rate S_t should equal the differential of interest rate between the two countries.

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

6.2. Currency Data and Currency Factors Description

We consider for our empirical analysis 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). We 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.

For the explanatory factors in our currency analysis we consider a range of different factors that we 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 piled up within assets returns would thus be received by investors accepting to bear the associated sources of risk. Lustig et al. (2011)

[5] Domestic risk free yield means the interest rate yield in the reference country, which would be for instance the dollar for an American investor.

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[6]. 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}. Lustig et al. (2011) prove 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 we retain for our covariance regression model. Moreover, we also take into consideration the respective factors volatility σ_{DOL}, σ_{HML} as well as the covariance between the factors $\sigma_{DOL,HML}$. Besides these factors, in Ames et al. (2017) they recently demonstrate 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 future market. Leaning on a very rich academic literature, the relation between assets variance or covariance and trading volume has been recurrently demonstrated and utterly admitted 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). We should draw a parallel between this branch of the literature and the empirical works about 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. (2017) which entails us to augment our price-based covariance regression model with speculative volumes information provided in a weekly report published by the CFTC, see CFTC (2015). Doing so, we assume the financial inflows and outflows resulting from the adjustments of the speculative long or short positions generate and thus could help explain dependencies between international exchange rates as demonstrated in Ames et al. (2017).

6.3. Covariate SARIMA Forecast Results

In order to assess the accuracy of our SARIMA forecasting models for the individual covariates we utilise the MASE and the MAPE measures, as discussed in Section 3.1. The MASE forecast accuracy results, shown as a time series in Figure 1 and as a boxplot summary in Figure 2, suggest that on average all of the models constructed from the ARIMA automated fitting procedure described in Section 3.1 behave as expected with the covariates with non-trivial ARIMA model structure producing reasonably accurate forecasts performance over the one month forecast horizon, which is required for the applications to carry trade strategies considered in this paper.

We note that in the case of the DOL and HML covariates the ARIMA model arrived at seems 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, if we also look at the boxplot summary of the Mean Absolute Percentage Error (MAPE) forecast accuracy results in Figure 3 we see that the DOL and HML have median MAPEs of 100%, i.e. often we forecast these covariates as zero.

If instead of just considering DOL and HML, we instead consider additional covariates based on the volatility in these factors and there covariance, then the fitted models for these factors demonstrate that for the σ_{DOL}, σ_{HML} and $\sigma_{DOL,HML}$ a much more accurate forecast performance for the σ_{DOL}, σ_{HML} and $\sigma_{DOL,HML}$ 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 6.2 the factors known as DOL and HML were shown to have strong explanatory power of the carry trade portfolio returns when studied from an in-sample analysis via PCA. However, as we have demonstrated in this section, this has not carried forward to good forecast performance for the models fitted for these DOL and HML factors. The 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 MAPE and MASE criterion.

In fact these findings further strengthens our arguments 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 mean and importantly in the covariance regression structures as well as good out-of-sample forecast performance under the models selected for these factors from our ARIMA studies. This will mean that such factors can be both significant in interpreting inter-temporal variation

[6] When we consider an American investor. 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.

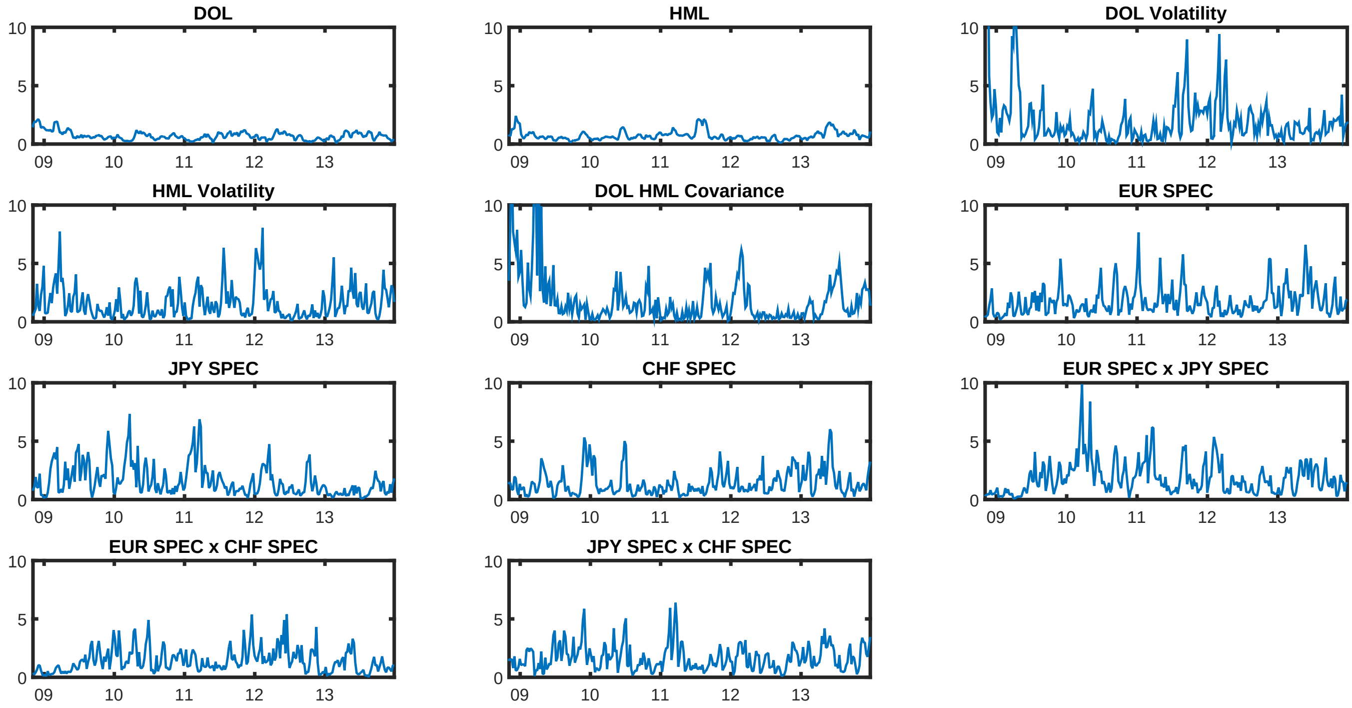

Figure 1: Mean Absolute Scaled Errors (MASE) for Low Interest Rate Basket Covariate Forecasts.

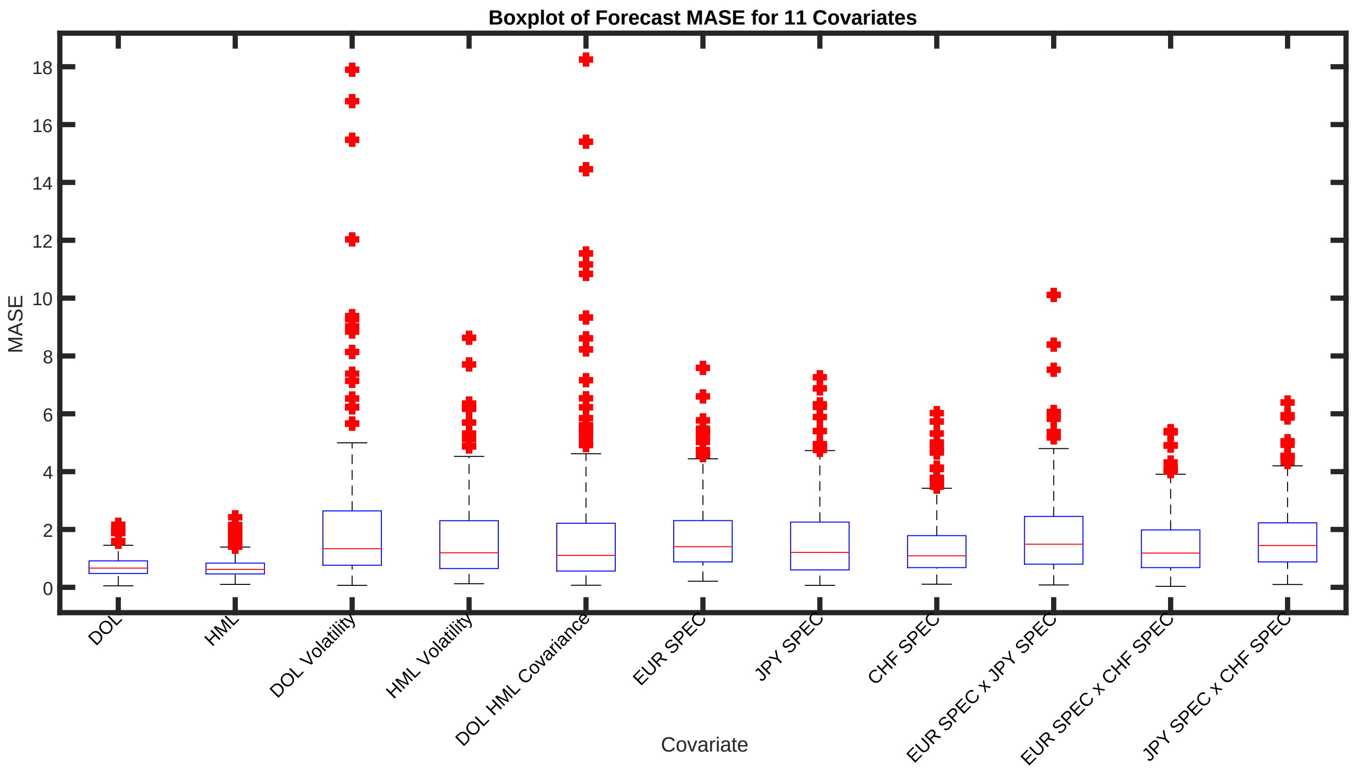

Figure 2: Boxplots of Mean Absolute Scaled Errors (MASE) for Low Interest Rate Basket Covariate Forecasts.

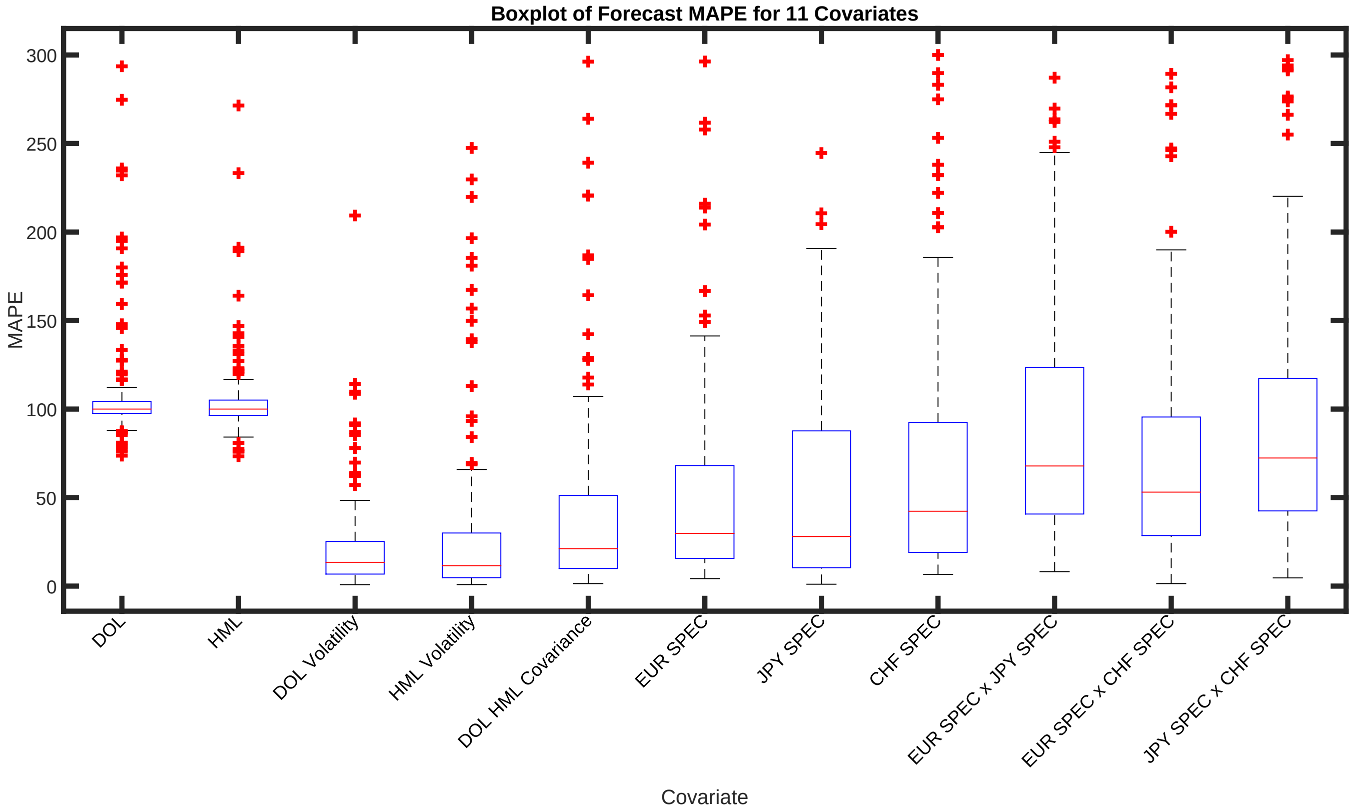

Figure 3: Mean Absolute Percentage Errors (MAPE) for Low Interest Rate Basket Covariate Forecasts.

in carry returns as well as instrumental in improving covariance model forecasts in the GFM model we propose and therefore may contribute to improving the portfolio performance that results from such a model. We will investigate this second aspect further in the studies contained in the remaining sections.

6.4. Variance and covariance dynamic 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 2 for the SFM, GFM and DCC models behaves 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, we demonstrate that each model's forecast covariance produces significantly different behaviours over time in both the information content captured by each model and more importantly in the reactivity of the covariance model forecasts to inter-temporal variation in the information content contained in 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 we used it in portfolio allocation, as described in Section 6.2. Performing these studies can be achieved in a number of different ways, the approach we have selected to present below is based on a similar type of analysis performed in Engle and Colacito (2006).

Our first study consist in pointing out the distinctive features and benefits of using the GFM model we proposed in this article versus the other SFM and DCC models also discussed. Demonstrating the differences in the second order modelled information content is achieved through analysis of the forecast covariance matrix: we use the trace, to study the variation and reactivity of each model forecast to marginal volatility fluctuations; and we use 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, we have considered the global minimum variance portfolio allocation framework 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 Section 4.

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 types of analysis performed in Engle and Colacito (2006), modified for the context of the models in this paper. This required us to consider 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). We will denote this time series of estimators as the “ex-post” portfolio volatility that our 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. We draw 1000 bootstraps replicate samples and then calculate the covariance of these bootstrapped weekly portfolio volatilities. We compare the ex-post portfolio volatility obtained 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 we note in the model description section 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 $\mathcal{F}_{t-1}$, $\mathcal{G}_t$ and $\widetilde{\mathcal{G}}_{t-1}$, depending on whether they contain factors and whether they are population based estimations such as for the SFM and GFM models in Equations 2.2 and 2.5 respectively, or locally adapted conditional estimations as in the SFM and GFM models in Equations 2.3 and 2.6 respectively. The utility of these three different filtrations is due to the very practical consideration that the covariates have different sampling rates, as discussed in Remark 2.1. Hence, different short term and long term covariance properties can be captured by considering these filtrations.

To interpret the comparison between the "ex-post" portfolio volatility and each of the SFM, GFM and DCC different forecast results one must do so with care. We shall undertake this comparison under the following statistical assumptions. We assume that the population based covariance estimate for the model factors, that are constructed from filtration $\widetilde{\mathcal{G}}_{t-1}$, form an unbiased and consistent estimator of a stationary population based covariance. Furthermore, since the filtration $\widetilde{\mathcal{G}}_{t-1}$ is comprised of a time series of length $(t-1)\times(T+1)$ whereas the filtration $\mathcal{F}_{t-1}$ is of length $T+1$ for each sliding window, we will assume that for comparison purposes the contribution to the unconditional covariance, for the SFM and GFM models in Equations 2.2 and 2.5 respectively is approximately "exact". To be more precise, we assume the convergence rate of the second order moments of $\boldsymbol{X}_{t-1}$ which are constructed based on $\widetilde{\mathcal{G}}_{t-1}$ are a function of $(t-1)\times(T+1)$ and as such, we will assume that as T and t go to infinity, asymptotically we are only seeing the leading contribution to the portfolio volatility from the SFM and GFM unconditional covariance models which is arising from local (in the current sliding window) variability due to the filtration $\mathcal{F}_{t-1}$. In this sense we are then able to compare the models for the SFM, GFM which are based on $\mathcal{F}_{t-1}\cup\widetilde{\mathcal{G}}_{t-1}$ with the version of the DCC model which is based only on $\mathcal{F}_{t-1}$. If this were not the case, the results are still valid but direct comparison between model performance would be less obvious.

We note that an alternative approach would be to extend the bootstrap procedure to also sampling multiple realizations of the factors $\boldsymbol{X}_{t-1}$ that make up the filtration $\widetilde{\mathcal{G}}_{t-1}$. These sampled bootstrap replicates could then be used to numerically average out the variability due to the realization of the factors attributed to the terms such as $Cov(\boldsymbol{X}_t|\widetilde{\mathcal{G}}_{t-1})$ and $\mathbb{E}(\boldsymbol{X}_t\boldsymbol{X}_t^{\mathrm{T}}|\widetilde{\mathcal{G}}_{t-1})$ in the SFM and GFM models when considering the unconditional covariance, in order to again isolate the influence on portfolio volatility attributed to $\mathcal{F}_{t-1}$.

We demonstrate that a key difference among the set of models described earlier stands indeed in the conditioning filtrations considered for each of them. Furthermore, another distinguishing feature involves the choice of conditional variance and covariance dynamic considered, which means in our case either heteroskedastic or on the contrary homoskedastic models in the SFM, GFM and DCC models.

In the following, we emphasize the differences of reactivity among the set of variance covariance models under scrutiny and demonstrate that not only does the conditional dynamic of the dependence structure have a role to play, but in addition the filtration utilized in constructing the portfolio variance also has an important role to play in determining how fast each estimator can adapt to abrupt change of environment. Therefore, it is interesting to then study if a particular model is found to be more reactive to the local environment, as we will show with versions of our GFM model, does this necessarily translate into better portfolio performance and in what sense?

To this end we 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. We can notice in the upper panels of Figures 4 and 5 that the variance covariance matrix traces resulting from the GFM model are more reactive than those generated by the SFM model or the historical variance covariance matrix model even though the amplitude of the adjustment stayed restrained with respect to the DCC. While the trace embodies the average variability of the matrix diagonal elements, in other words 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 4 and 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 assets dependence level. Said differently the heteroskedasticity seems to be more influential at the covariance level of the GFM model generated variance and covariance matrices while the DCC generated matrices react more significantly to the variance heteroskedasticity component.

As discussed, to further our comparison between the GFM and the DCC we propose to compare the forecasting accuracy of the two models 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 6 and 7 the accuracy of the two methods is quite similar and remains within the $+/-15\%$

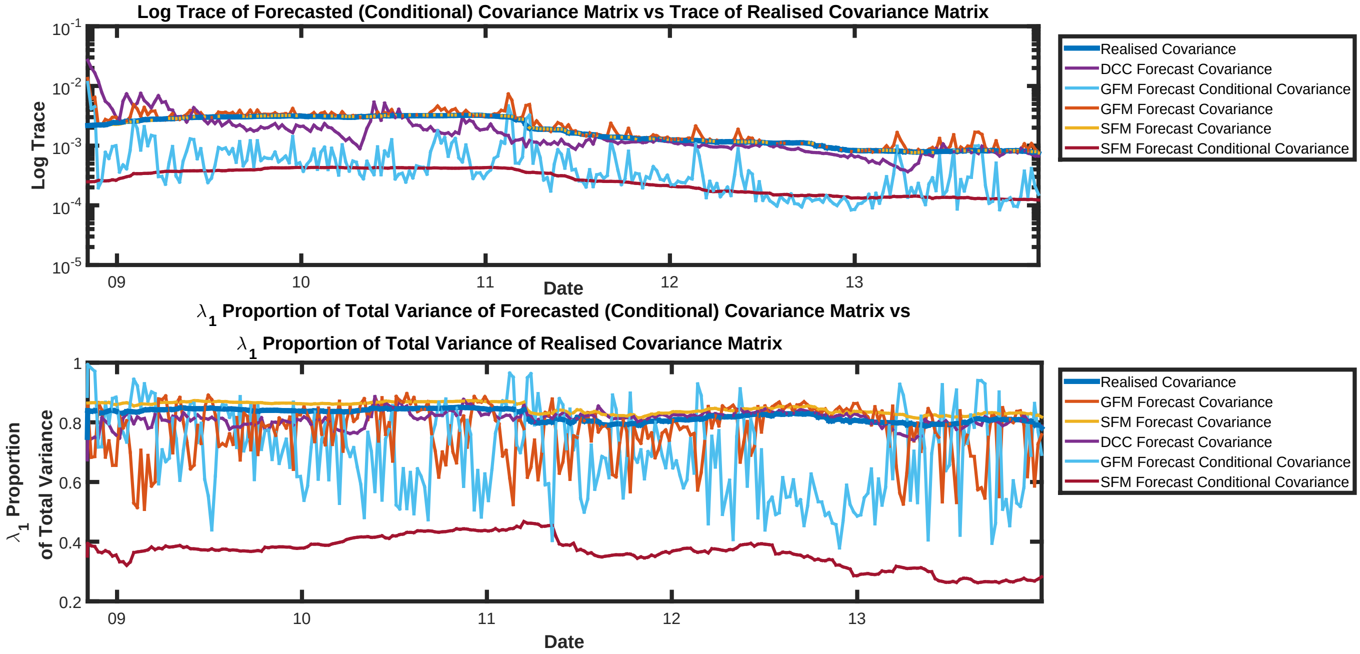

Figure 4: High Basket. Upper panel: Trace of covariance matrix.
Lower panel: Proportion of variance explained by first principal component.

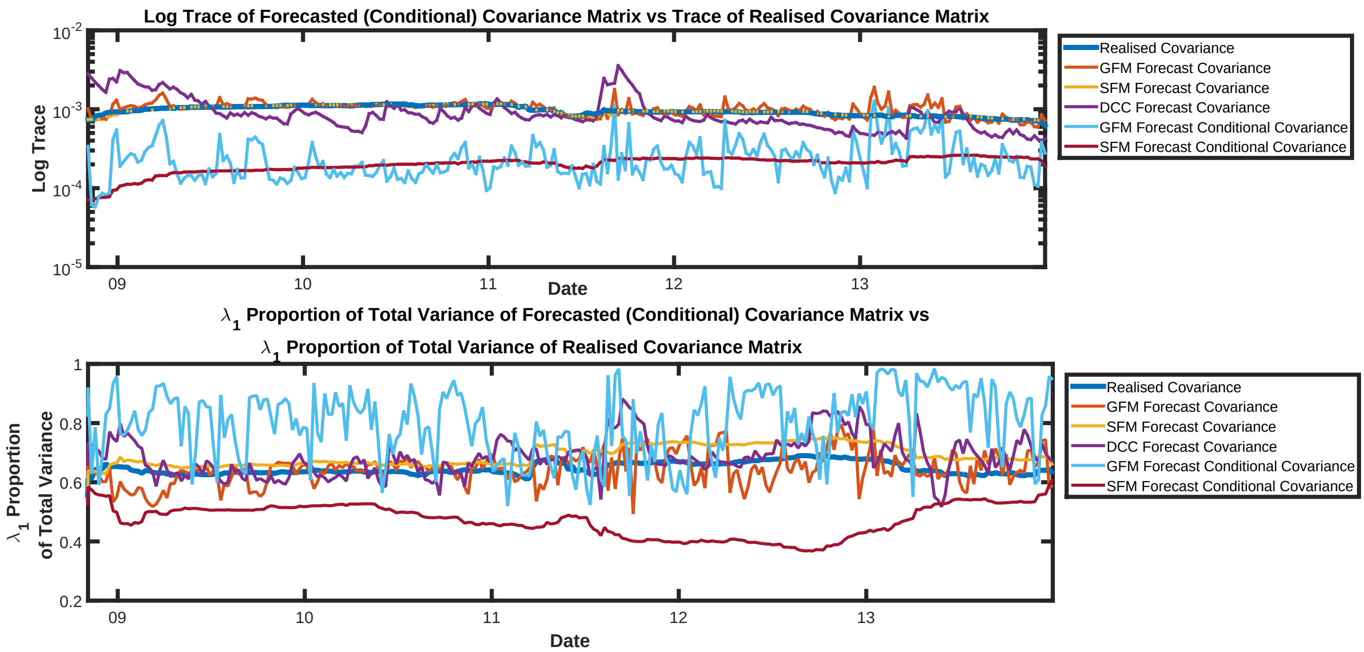

Figure 5: Low Basket. Upper panel: Trace of covariance matrix.
Lower panel: Proportion of variance explained by first principal component.

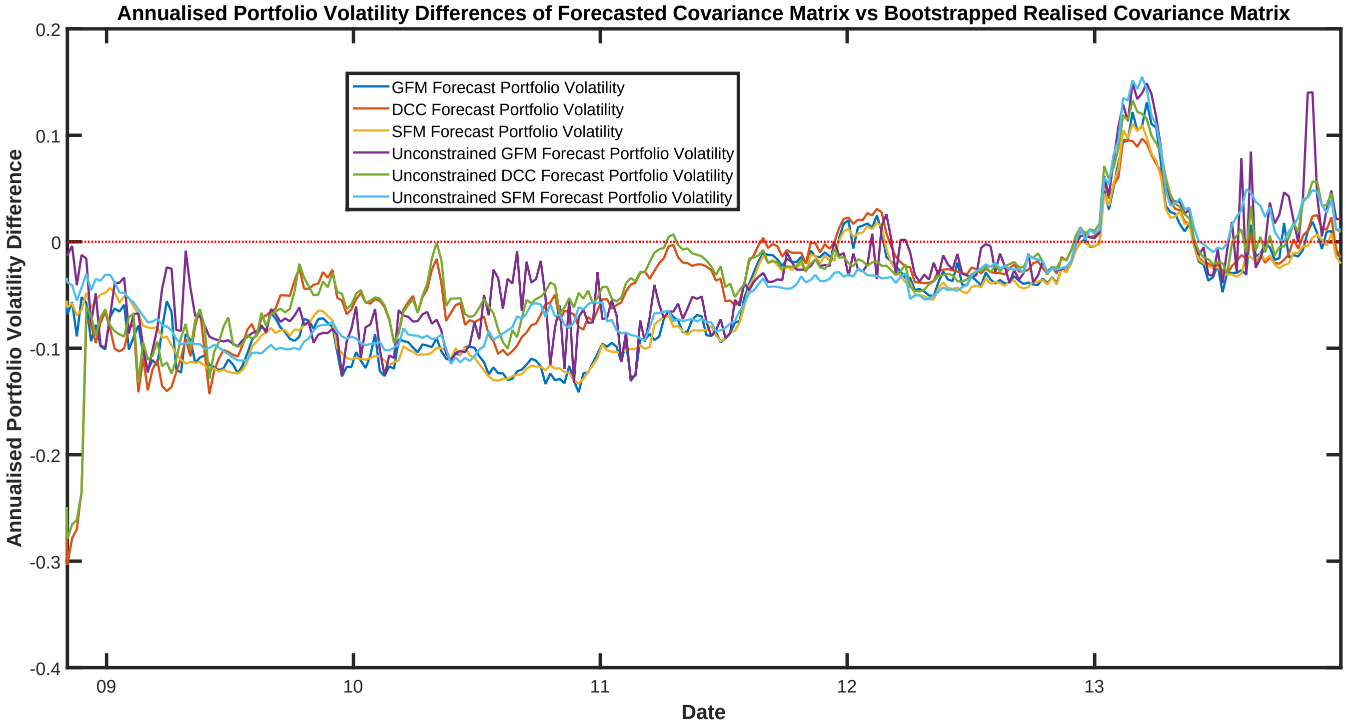

Figure 6: High basket. Annualised portfolio volatility differences between forecast covariance matrix and realised bootstrapped covariance matrix for different covariance forecasting models.

annualise 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 variance and covariance matrices.

6.5. Portfolio performance and conditioning of the variance covariance matrix

In this section, we consider 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 effects implicitly induced on the portfolio variance through the optimization routine and not directly through the stage one or two statistical model estimations. This is particularly interesting to consider in the context of the models studied in this paper for the SFM, GFM and DCC covariance forecast models.

Therefore, we 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 we are long the high interest rate currencies while financing this position through short positions on the low interest rate currencies. This implies obviously that we constrain the weights to be positive in the high interest rate currencies basket optimization program, whereas we enforce that the optimal positions in the low interest rates currencies basket to be negative. In the aftermath of this supposedly slight modification of the global minimum variance optimization program the variance and the covariance entries used as input are accordingly affected and even more precisely an implicit form of shrinkage occurs on these matrices. For instance, one sees that the objective function for the global minimum variance will contain, in the resulting constrained Lagrangian, a form of 'penalty' term given by $(\lambda 1^{\mathrm{T}} - 1\lambda^{\mathrm{T}})$, 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 variance 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 the stages one and two of 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 variance 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 paper goes in the same direction. We indeed notice that the risk return profiles associated to the global minimum variance portfolios built on various estimators are barely distinguishable when we append the positivity constraint as shown in Figure 8.

Contrary to this, the unconstrained results, plotted in Figure 9, we see that the differences among estimators chosen are clearly noticeable on a rolling window basis when these constraints are not imposed. We emphasize that through

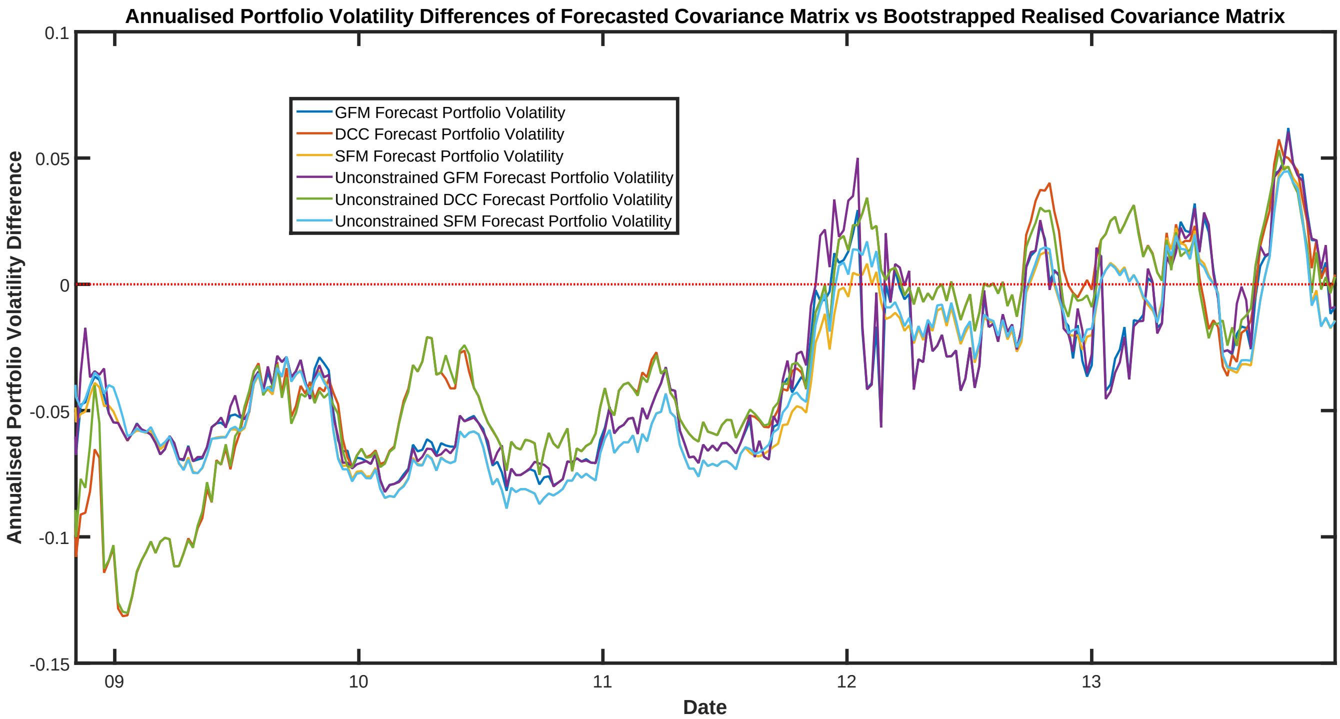

Figure 7: Low basket. Annualised portfolio volatility differences between forecast covariance matrix and realised bootstrapped covariance matrix for different covariance forecasting models.

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 variance covariance estimator choice.

This is indisputably true as far as we content ourselves with the filtrations $\widetilde{\mathcal{G}}_t$ or $\mathcal{F}_t$, however, our method in the GFM family of models is interestingly also enabling us to condition our variance 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 expression (2.6) derived earlier. Hence, 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 $(\mathcal{F}_t \cup \mathcal{G}_t)$. As such this conditional covariance model thus allows us to build different high and low interest rates currency optimal portfolios which will represent in a sense the non-diagonal heteroskedastic component of the diffusion and the dependence structure characterizing separately form the related high and low interest rates sets of currencies. Then we find that if we if we compare the 12-month rolling Sharpe ratio of the constrained minimum variance portfolio based on the GFM conditional variance covariance matrix with the unconditional GFM and the conditional DCC models, all of them being conditioned on different filtration or combination of filtration, the former strikingly stands out from the two others.

Figures 10 and 11 indeed display a significantly different behaviour of the rolling Sharpe ratio for the portfolios based on the GFM conditional covariance matrix even though we appended the non-negativity constraint to the global minimum variance optimization program. As we focus on the carry trade portfolio we also notice 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 6.7.

6.6. Sensitivity Analysis

In this section we 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 we systematically vary each individual covariate/factor, one-by-one, from the set of currency factors we considered in Section 6.2. 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. We then summarise the behaviour through portfolio based metrics of the perturbation effect of each covariate. This allows us to study which covariates are most influential in driving the portfolio performance and which covariates are likely to result in largest sensitivity of results. We note that such an analysis is easily undertaken due to the specific model structure we have developed for our GFM structure where the covariates enter explicitly into the covariance matrix.

This enlightening robustness analysis allows us to estimate a confidence interval or our variance covariances matrix entries as a function of the marginal distribution of each covariate used for the covariance regression. The formula

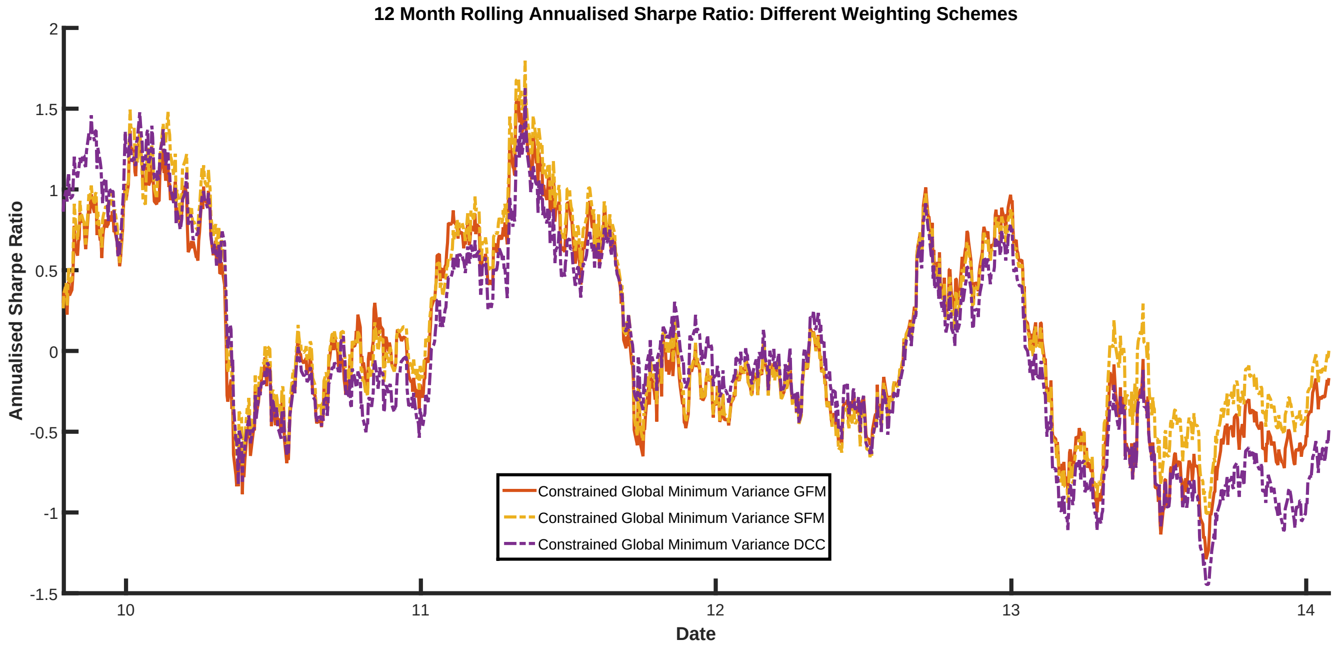

Figure 8: High basket. Constrained GMV 12 month rolling Sharpe ratio comparison.

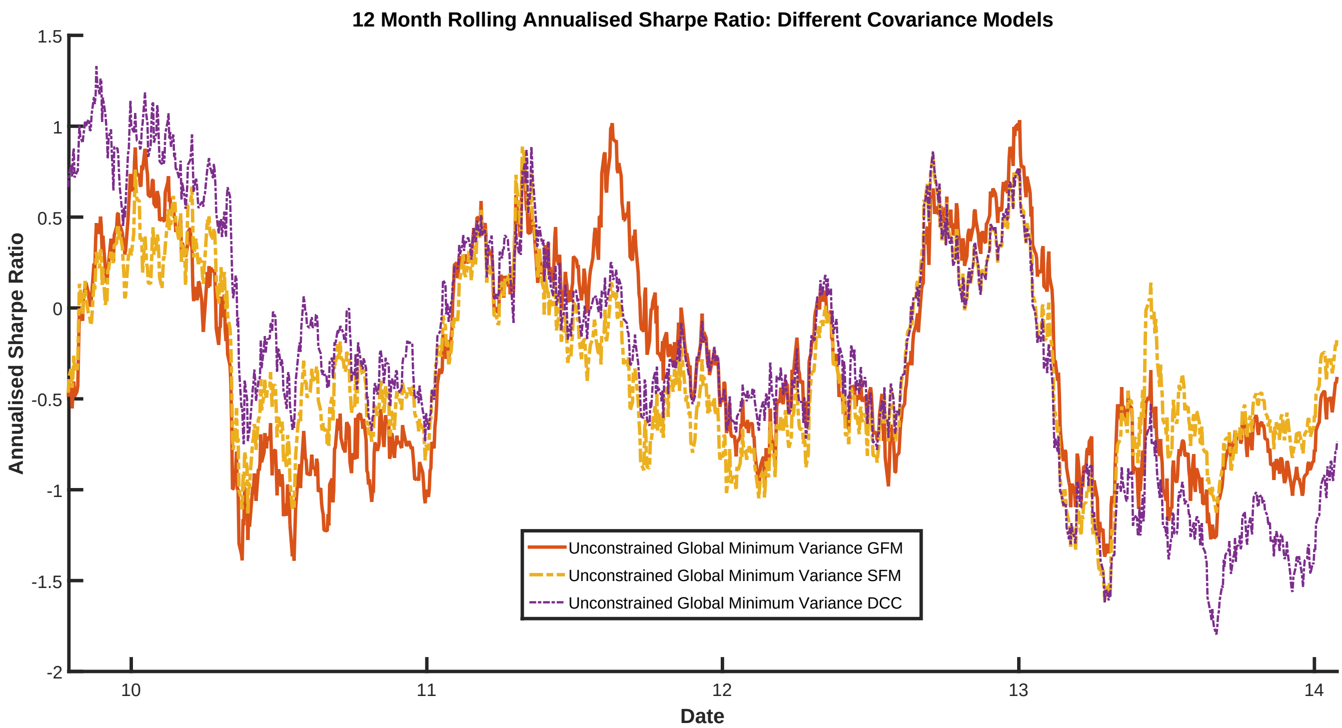

Figure 9: High basket. Unconstrained GMV 12 month rolling Sharpe ratio comparison.

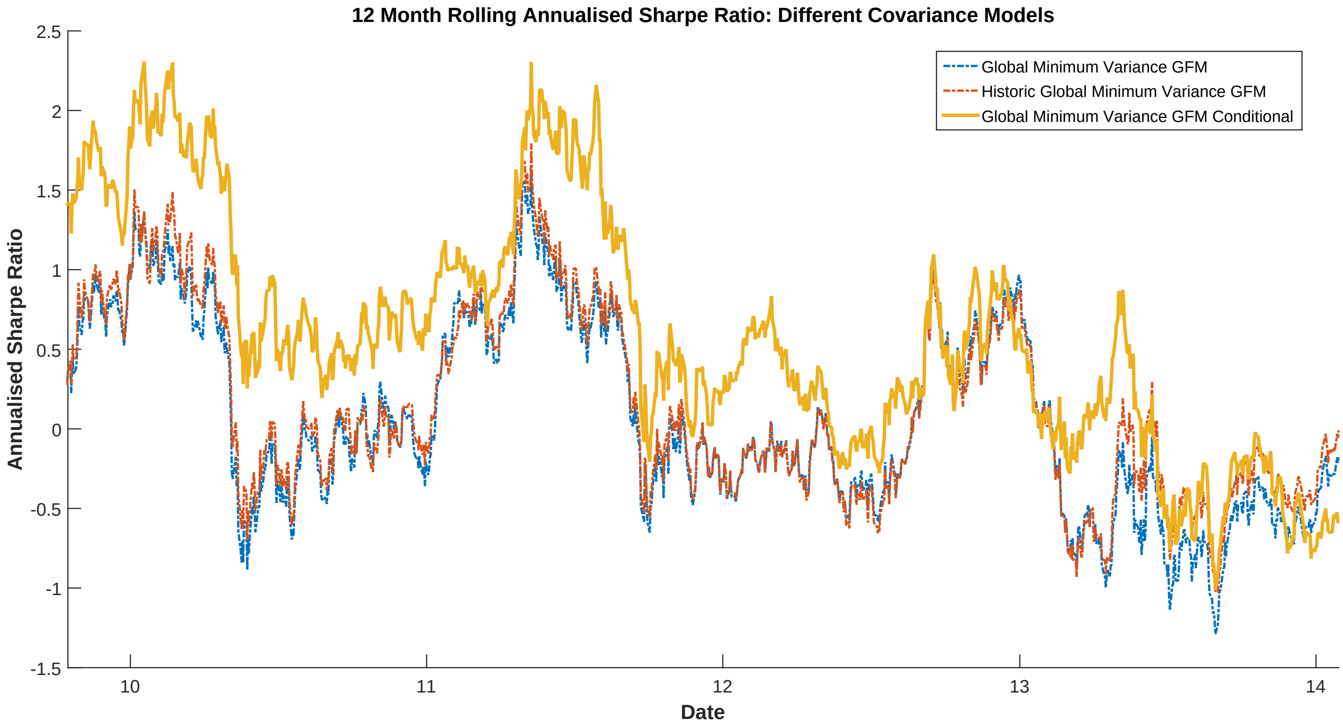

Figure 10: High basket. 12 month annualised rolling Sharpe ratio. Comparison of Conditional GFM and Unconditional GFM.

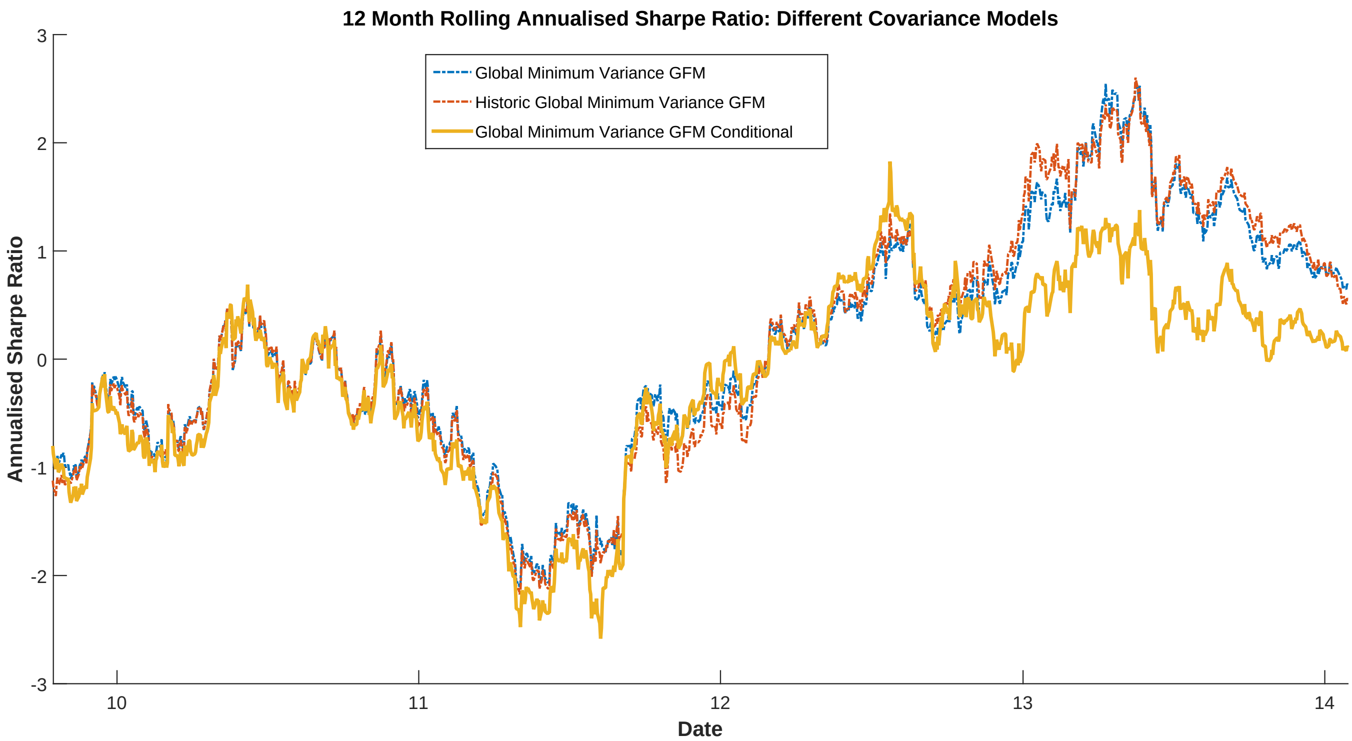

Figure 11: Low basket. 12 month annualised rolling Sharpe ratio. Comparison of Conditional GFM and Unconditional GFM.

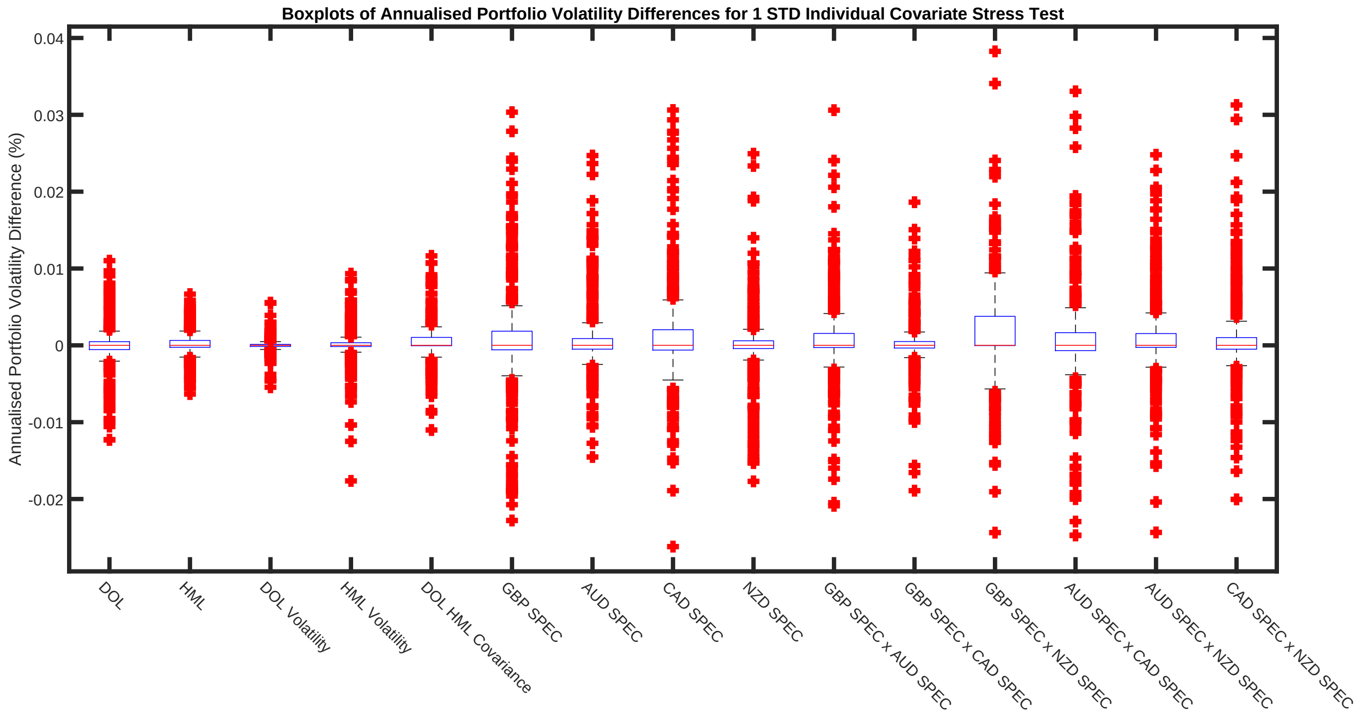

Figure 12: High Basket. Boxplot of annualised portfolio volatility differences resulting from one standard deviation individual perturbation of each covariate for GFM model with GMV weights.

5.3, we 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 and 13 show that some covariates changes can lead to larger effect on the structure of dependence among assets and their respective marginal features leading accordingly to large modification of the optimal portfolio volatility. As an example, we see in Figure 12 that the GBP speculative interest have 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 factors should be added to the limited forecasting quality highlighted earlier. Furthermore, we can globally notice 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 factors as well as their respective volatility and the covariance between them. Contrarily 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 variance covariance matrix which will give different results for a positive or a negative modification of a given variance covariance matrix entry.

6.7. The Carry Trade Portfolio

This last 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 variance covariance estimators under the SFM, GFM and DCC models. We stress the fact that this could not to be considered as the optimal carry trade portfolio as we 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. (2017, 2015). Considering the carry trade portfolio configuration, we 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) boils down to very similar 12-month rolling Sharpe for the DCC, the GFM unconditional and the SFM unconditional estimator. Nevertheless, in Figures 14 and 15 we easily notice the difference of behaviour of the carry trade portfolio optimised using the GFM conditional estimator. Apart from the second half of 2013 this portfolio has always shown significantly higher Sharpe ratio on a 12-month rolling basis which demonstrates the robustness of the improvement. Furthermore, Table 1 underpins 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. The downside volatility is not penalizing either

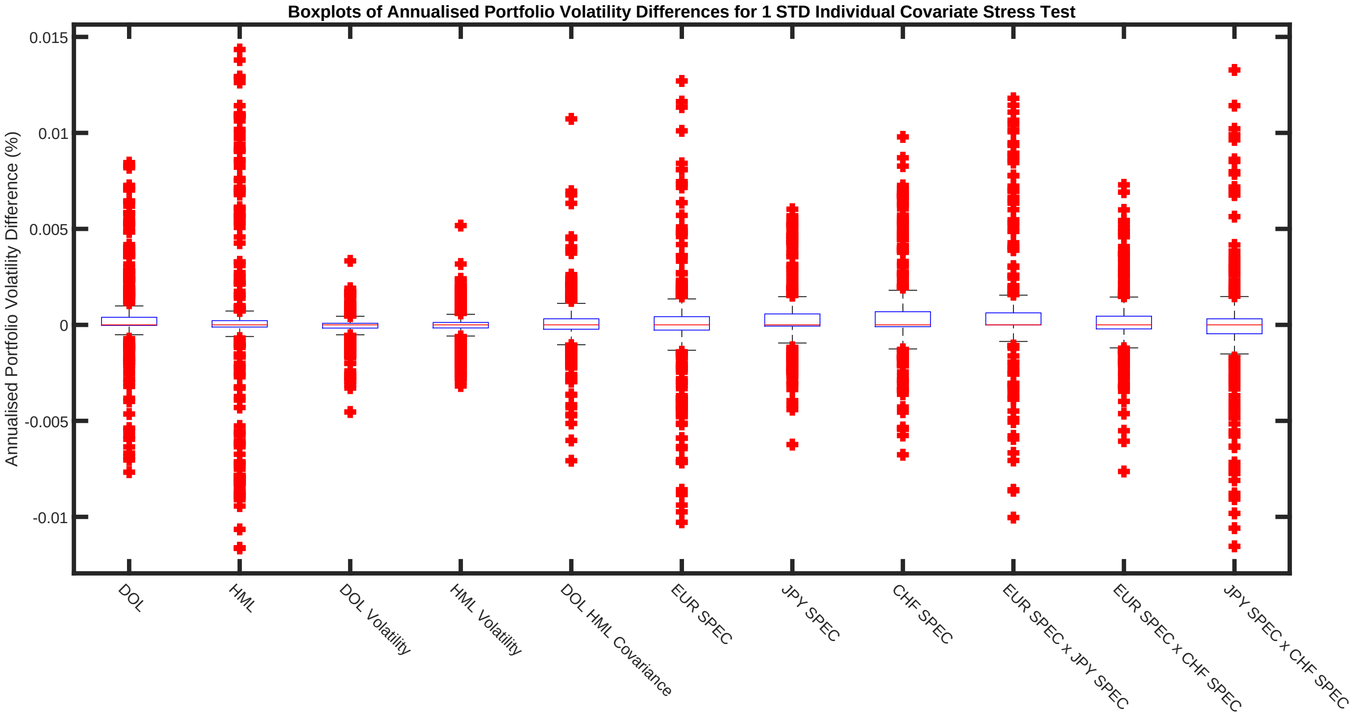

Figure 13: Low Basket. Boxplot of annualised portfolio volatility differences resulting from one standard deviation individual perturbation of each covariate for GFM model with GMV weights.

Table 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 Conditional	GFM Conditional (No SPEC)	GFM	SFM	DCC
Sharpe	**0.31**	0.27	0.1	0.1	0.21
Sortino	**0.43**	0.38	0.14	0.14	0.3
Omega	**1.87**	1.78	1.27	1.28	1.58
Max DD	31.7	**28.8**	30.3	30	31.5
Calmar	**0.15**	0.15	0.05	0.05	0.1

the GFM conditional estimator as the Sortino ratio demonstrates this feature. Succinctly, we may conclude from this analysis that the conditional variance covariance estimator we developed in this paper under the GFM model family is clearly displaying interesting properties such as its lower sensitivity to the shrinkage effect resulting from the weight constraints but also on the resulting improvements of the global minimum variance portfolio risk and return profiles.

7. Conclusions

The standard multi-factor model (SFM) family discussed in this paper 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 is known to fail to account for an important feature displayed by financial assets real data returns, that is the heteroskedastic nature of the assets covariance structure over time. In this paper we develop a generalised version of the multi-factor model, the GFM family of models. 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 we demonstrated in this paper that we are able 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. We show that the GFM model is directly interpretable in terms of how it depends on the assets return filtration but also on the sigma-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 variance covariance matrix assuming the stationarity of the trend and the variance covariance regression parameters. The GFM model we developed allows us to devise and estimate robust forecasting models for the set of independent variables we picked and we demonstrate this 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 we

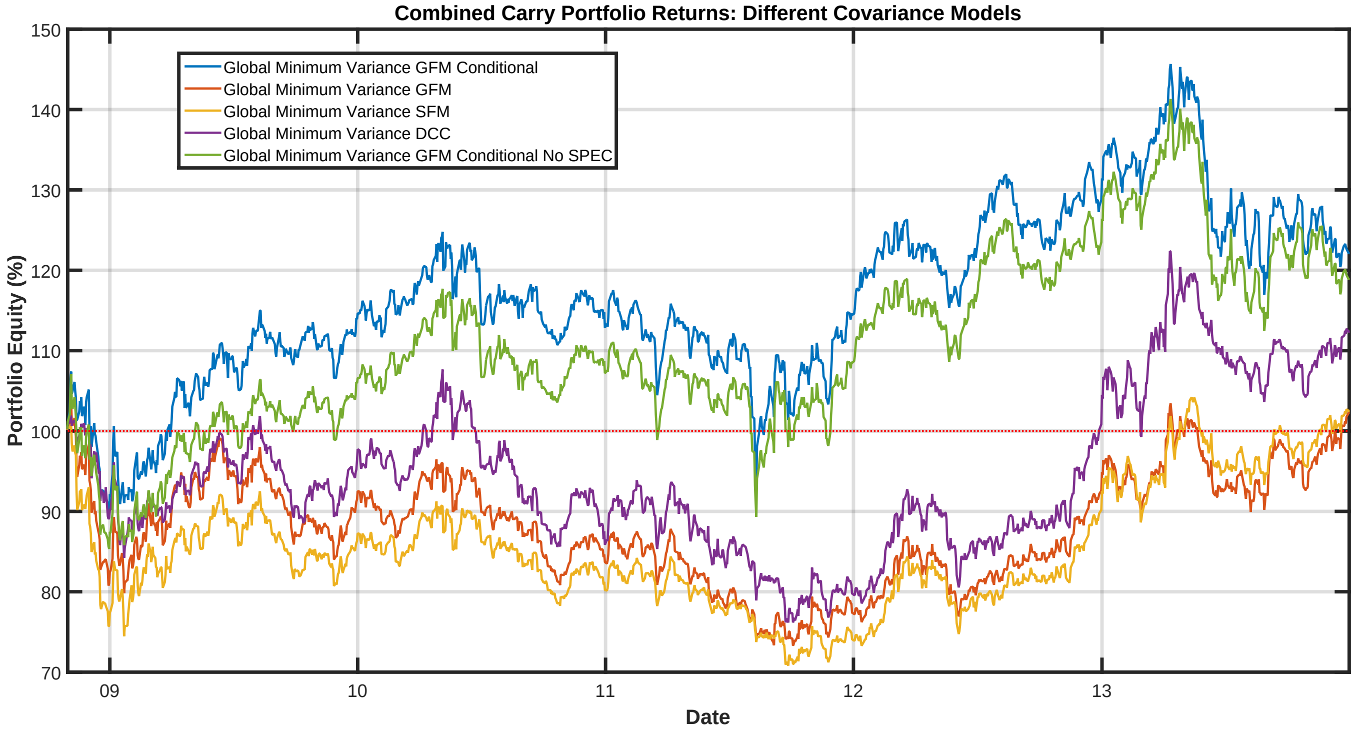

Figure 14: Carry trade portfolio performance. We re-optimise the portfolios on a monthly basis using an annual portfolio volatility target of 15% and hence scale the monthly returns according to the expected portfolio volatility for each method. We assume that we initially capitalise the strategy to the value of the unleveraged baskets.

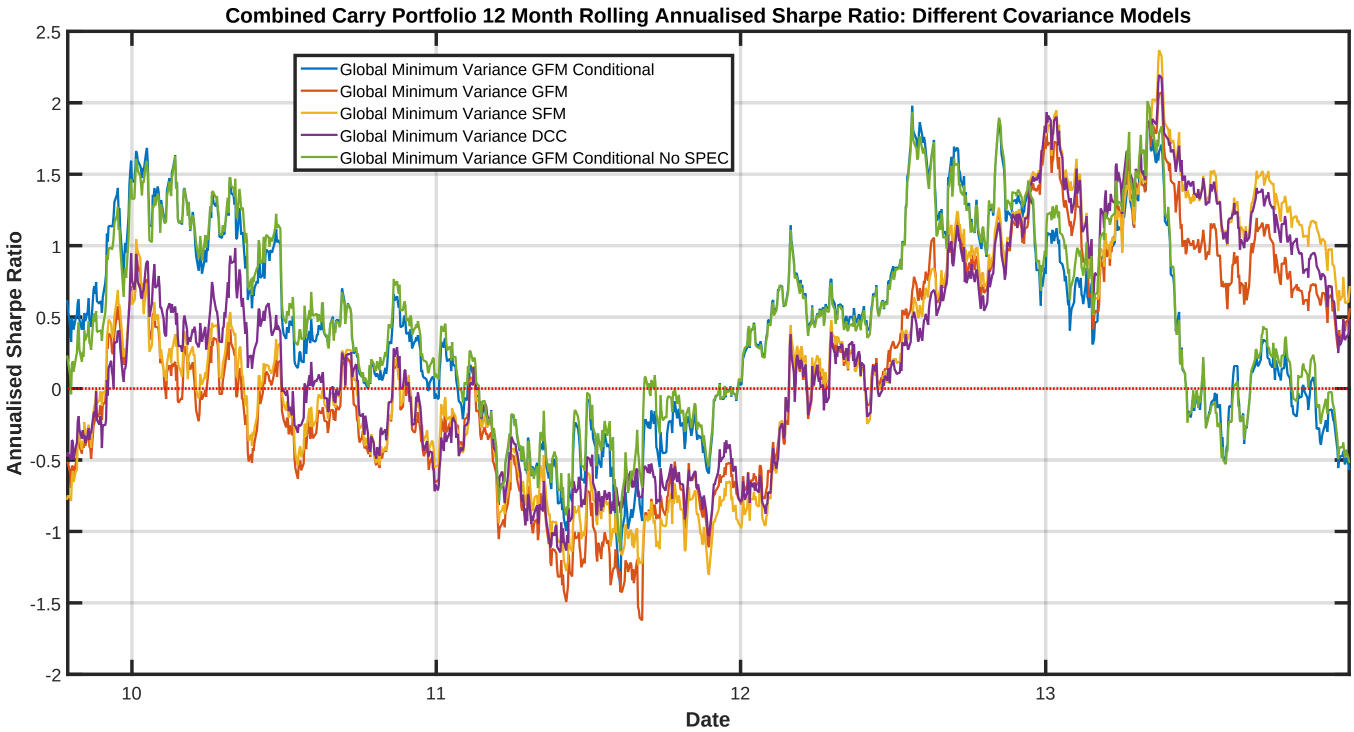

Figure 15: Carry trade portfolio 12 month annualised rolling Sharpe ratio. The Sharpe ratio is defined as return divided by volatility.

develop in this paper 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. We demonstrate that in the currency portfolio studies we perform, whilst two well known factors studied in the literature, the DOL and HML, are providing strong in-sample explanatory power, their out-of-sample forecast performance is very poor. This makes them difficult to utilise in portfolio selection frameworks which require the forecast portfolio trend and covariance. In this paper we 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 we consider are directly interpretable and we may relate there attributes to an established literature in economics, relating returns to volume and liquidity of an asset Ames et al. (2017). This established relation between the speculative positions and the asset returns dependence structure allows us to better capture the heteroskedasticity prevailing in the asset returns notably in high volatility environment. Through our empirical application to the currency market we demonstrated that the conditional formulation of the variance covariance we proposed in this paper definitely outperforms on a risk-return basis several widely implemented models such as the DCC or the single factor models.

This established relation between the speculative positions and the asset returns dependence structure combined with other currency market admitted market factors allows us to better capture the heteroskedasticity prevailing in the asset returns notably in high volatility environment. Through our empirical application to the currency market we demonstrated that the conditional formulation of the variance covariance we proposed in this paper definitely outperforms on a risk-return basis several widely implemented models such as the DCC or the standard factor models even though non-negativity and non-positivity constraints are necessarily appended to our high and low interest rates baskets optimization in order to build a self-financing portfolios.

References

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299. 2, 6

Akram, Q. F., Rime, D., and Sarno, L. (2008). Arbitrage in the foreign exchange market: Turning on the microscope. *Journal of International Economics*, 76(2):237–253. 12

Alexander, C. (2000). A primer on the orthogonal garch model. *manuscript ISMA Centre, University of Reading, UK*. 2

Ames, M., Bagnarosa, G., and Peters, G. W. (2017). Violations of uncovered interest rate parity and international exchange rate dependences. *Journal of International Money and Finance*, 73, Part A:162 – 187. 3, 5, 13, 22, 25

Ames, M., Peters, G. W., Bagnarosa, G., and Kosmidis, I. (2015). Upside and Downside Risk Exposures of Currency Carry Trades via Tail Dependence. In *Innovations in Quantitative Risk Management*, pages 163–181. Springer. 12, 22

Anderson, R. M., Bianchi, S. W., and Goldberg, L. R. (2012). Will my risk parity strategy outperform? *Financial Analysts Journal*, 68(6):75–93. 8, 10

Anzuini, A. and Fornari, F. (2012). Macroeconomic determinants of carry trade activity. *Review of International Economics*, 20(3):468–488. 11, 13

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59. 8, 10

Backus, D. K., Foresi, S., and Telmer, C. I. (2001). Affine term structure models and the forward premium anomaly. *The Journal of Finance*, 56(1):279–304. 12

Black, F. and Litterman, R. B. (1991). Asset allocation: combining investor views with market equilibrium. *The Journal of Fixed Income*, 1(2):7–18. 3

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505. 2

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press. 9

Boyle, P., Garlappi, L., Uppal, R., and Wang, T. (2012). Keynes meets markowitz: The trade-off between familiarity and diversification. *Management Science*, 58(2):253–272. 3

Branger, N., Larsen, L. S., and Munk, C. (2013). Robust portfolio choice with ambiguity and learning about return predictability. *Journal of Banking & Finance*, 37(5):1397 – 1411. 3

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1):21–58. 2, 8

Brunnermeier, M. K., Nagel, S., and Pedersen, L. H. (2008). Carry trades and currency crashes. Technical Report 23, Nber Macroeconomics Annual 2008. 11, 12, 13

Brunnermeier, M. K. and Pedersen, L. H. (2009). Market liquidity and funding liquidity. *Review of Financial studies*, 22(6):2201–2238. 13

Burnside, C., Eichenbaum, M., Kleshchelski, I., and Rebelo, S. (2011). Do peso problems explain the returns to the carry trade? *Review of Financial Studies*, 24(3):853–891. 12

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82. 2

CFTC (2015). US CFTC Commitments of Traders: Explanatory Notes. [Online; accessed 9-November-2015]. 13

Chan, L. K. C., Karceski, J., and Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies*, 12(5):937–974. 2, 4

Chopra, V. K. and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11. 2, 8

Choueifaty, Y. and Coignard, Y. (2008). Toward maximum diversification. *Journal of Portfolio Management*, 35(1):40. 8, 10

Christiansen, C., Ranaldo, A., and Söderlind, P. (2011). The time-varying systematic risk of carry trade strategies. *Journal of Financial and Quantitative Analysis*, 46(04):1107–1125. 12

Christodoulakis, G. A. and Satchell, S. E. (2002). Correlated arch (corrarch): Modelling the time-varying conditional correlation between financial asset returns. *European Journal of Operational Research*, 139(2):351–370. 3

Christoffersen, P. and Langlois, H. (2013). The joint dynamics of equity market factors. *Journal of Financial and Quantitative Analysis*, 48:1371–1404. 4

Clarke, R., De Silva, H., and Thorley, S. (2011). Minimum-variance portfolio composition. *Journal of Portfolio Management*, 37(2):31. 9

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5):798–812. 3

DeMiguel, V., Garlappi, L., and Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953. 8

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350. 1, 2, 4, 6

Engle, R. and Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2):238–253. 2, 6, 15

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric theory*, 11(01):122–150. 2

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3):319–338. 12

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3 – 56. 2, 12

Fong, W. M. (2013). Footprints in the market: Hedge funds and the carry trade. *Journal of International Money and Finance*, 33:41 – 59. 11, 13

Frost, P. A. and Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *The Journal of Portfolio Management*, 15(1):29–34. 2, 3, 8

Garlappi, L., Uppal, R., and Wang, T. (2006). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20(1):41–81. 3

George E. Tauchen, M. P. (1983). The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505. 13

Goldfarb, D. and Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38. 3

Green, R. C. and Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified? *The Journal of Finance*, 47(5):1785–1809. 2, 4

Hafner, C. M. and Manner, H. (2012). Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, 27(2):269–295. 4

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies*, 19(1):237–271. 2

Hansen, L. P. and Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *The Journal of Political Economy*, pages 829–853. 12

Haugen, R. A. and Baker, N. L. (1991). The efficient market inefficiency of capitalization-weighted stock portfolios. *The Journal of Portfolio Management*, 17(3):35–40. 8, 9

He, X. and Velu, R. (2014). Volume and Volatility in a Common-Factor Mixture of Distributions Model. *Journal of Financial and Quantitative Analysis*, 49(01):33–49. 13

Hoff, P. D. and Niu, X. (2012). A covariance regression model. *Statistica Sinica*, 22(2):729–753. 4, 5, 6

Hutchison, M. and Sushko, V. (2013). Impact of macro-economic surprises on carry trade activity. *Journal of Banking & Finance*, 37(4):1133 – 1147. 11, 13

Hyndman, R. J. (2015). *forecast: Forecasting functions for time series and linear models*. R package version 6.1. 7

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22. 7, 8

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688. 7

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, 58(4):1651–1684. 3, 18

Jobson, J. D. and Korkie, R. M. (1981). Putting markowitz theory to work. *The Journal of Portfolio Management*, 7(4):70–74. 2, 8

Jorion, P. (1985). International portfolio diversification with estimation risk. *The Journal of Business*, 58(3):259–278. 3

Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *The Journal of Financial and Quantitative Analysis*, 21(3):279–292. 3

Juhl, T., Miles, W., and Weidenmier, M. D. (2006). Covered interest arbitrage: then versus now. *Economica*, 73(290):341–352. 12

Jurczenko, E., Michel, T., and Teiletche, J. (2015). A unified framework for risk-based investing. *The Journal of Investment Strategies*, 4(4). 8, 9

Kourtis, A., Dotsis, G., and Markellos, R. N. (2012). Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *Journal of Banking & Finance*, 36(9):2522 – 2531. 3

Lanne, M. and Saikkonen, P. (2007). A multivariate generalized orthogonal factor garch model. *Journal of Business & Economic Statistics*, 25(1):61–75. 4

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603 – 621. 3

Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common Risk Factors in Currency Markets. *Review of Financial Studies*, 24(11):3731–3777. 3, 12, 13

Lustig, H. and Verdelhan, A. (2007). The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk. *American Economic Review*, 97(1):89–117. 12

Maillard, S., Roncalli, T., and Teïletche, J. (2008). On the properties of equally-weighted risk contributions portfolios. *Available at SSRN 1271972*. 8, 9, 10

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91. 1, 8

Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2012). Carry trades and global foreign exchange volatility. *Journal of Finance*, 67(2):681–718. 12

Michaud, R. O. (1989). The markowitz optimization enigma: is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42. 2, 8

Nawrocki, D. (1996). Portfolio analysis with a large universe of assets. *Applied Economics*, 28(9):1191–1198. 2, 8

Palomar, D. P. and Eldar, Y. C. (2010). *Convex optimization in signal processing and communications*. Cambridge university press. 9

Salvatierra, I. D. L. and Patton, A. J. (2015). Dynamic copula models and high frequency data. *Journal of Empirical Finance*, 30:120–135. 4

Santos, A. A. and Moura, G. V. (2014). Dynamic factor multivariate garch model. *Computational Statistics & Data Analysis*, 76:606 – 617. CFEnetwork: The Annals of Computational and Financial Econometrics2nd Issue. 2

Scherer, B. (2007). Can robust portfolio optimisation help to build better portfolios? *Journal of Asset Management*, 7(6):374–387. 3

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2):277–293. 1, 2

Stefanovits, D., Schubiger, U., and Wüthrich, M. V. (2014). Model risk in portfolio optimization. *Risks*, 2(3):315–348. 9

Tutuncu, R. and Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187. 3

Van der Weide, R. (2002). Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564. 3

van der Weide, R. (2002). Go-garch: a multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564. 4

Zhang, K. and Chan, L. (2009). Efficient factor garch models and factor-dcc models. *Quantitative Finance*, 9(1):71–91. 2