FORECASTING COVARIANCE FOR OPTIMAL CARRY TRADE PORTFOLIO ALLOCATIONS

Matthew Ames⋆ Guillaume Bagnarosa†⋆ Gareth W. Peters⋆†×, Pavel Shevchenko§⋆ Tomoko Matsui††

⋆ Department of Statistical Science, University College London, UK
† Oxford-Man Institute, Oxford University, UK
× Systemic Risk Center, London School of Economics, UK
† ESC Rennes School of Business, France
§ Macquarie University, Australia
†† Institute of Statistical Mathematics, Japan

ABSTRACT
Modelling and forecasting of asset volatility and covariance is of prime importance in the construction of portfolios. In this paper, we present a generalised multi-factor model that incorporates heteroskedasticity and dependence in the idiosyncratic error terms. We apply this model to forecasting the time-varying covariances in a basket of high interest rate and a basket of low interest rate carry trade currencies and then utilise these forecasts for portfolio optimisation. We compare traditional Markowitz portfolio optimisation to the more recently popular risk-based portfolio optimisation. Our model is shown to provide superior risk-adjusted returns for a currency carry trade strategy over the period 1999 - 2014.

Index Terms— Covariance Forecasting, Currency Carry Trade, Covariance Regression, Markowitz Portfolio, Equal Risk Contribution

1. INTRODUCTION
Understanding the behaviour of currency markets has been an active area of research for the past few decades. Much of the literature has focused on the marginal behaviours of exchange rates and carry trade portfolio returns resulting from the established violations of uncovered interest rate parity. Such investment strategies are popular approaches which involve constructing portfolios by selling low interest rate currencies and buying high interest rate currencies, thus profiting from the interest rate differentials. When such portfolios are highly leveraged this can result in sizeable profits. The presence of such opportunities, pointed out by [1, 2, 3] and more recently by [4, 5, 6, 7, 8, 9, 10], violates the fundamental relation of uncovered interest rate parity (UIP). The UIP refers to the parity condition in which exposure to foreign exchange risk, with unanticipated changes in exchange rates, is uninhibited and therefore changes in the exchange rates should offset the potential to profit from the interest rate differentials between high interest rate (investment) currencies and low interest rate (funding) currencies. We briefly introduce the definition of covered and uncovered interest rate parity. The UIP condition is directly linked to the arbitrage relation existing between the spot and forward prices of a given currency, namely the covered interest rate parity (CIP) condition.

Definition 1.1. 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:

\[ F^T_t = e^{(r_{t,T} - r_{t,T}^f)(T-t)} S_t, \]  

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 domestic¹ 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.

It is worth emphasizing that under the hypothesis of an absence of arbitrage opportunities, the CIP relation is directly resulting from the replication of the forward contract payoff using a self financed strategy. Moreover, the validity of this arbitrage relation has been demonstrated empirically in the currency market by [11, 12] when one considers daily data. From this arbitrage relation follows the UIP condition.

Definition 1.2. Uncovered Interest Rate Parity (UIP)
Considering Equation 1.1 the UIP condition can be defined as:

\[ E \left[ \frac{S^T_t}{S_t} \mid \mathcal{F}_t \right] = \frac{F^T_t}{S_t} = e^{(r_{t,T} - r_{t,T}^f)(T-t)}, \]

¹Domestic risk free yield means the interest rate yield in the reference country, which would be for instance the dollar for an American investor.
where \( F_t \) 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.

The existence of the UIP relationship is related to two primary assumptions, which are capital mobility and perfect substitutability of domestic and foreign assets. When UIP holds, then given foreign exchange market equilibrium, the interest rate parity condition implies that the expected return on domestic assets will equal the exchange rate-adjusted expected return on foreign currency assets. Therefore, no arbitrage opportunities should arise in practice; however such opportunities are routinely observed and exploited by large volume trading strategies. Numerous empirical studies [2, 1, 13, 4] have previously demonstrated the failure of UIP, i.e. that investors can actually earn profits by borrowing in a country with a lower interest rate, exchanging for foreign currency, and investing in a foreign country with a higher interest rate. Therefore, it is believed that trading strategies that aim to exploit the interest rate differentials may be profitable on average. In this paper we study how to identify trading strategies and portfolio selection to exploit violation of UIP in carry trades.

2. COVARIANCE REGRESSION MODEL

We introduce a new covariance regression model, termed the Generalised Multi-Factor model, which extends the traditional Multi-Factor model by allowing the factors to appear in the covariance of the idiosyncratic error terms and hence in both the unconditional and conditional covariance matrices. Weekly carry returns are defined as \( R_t = (F_t^T / F_t^T) - 1 \), i.e. we use a weekly mark-to-market procedure to calculate the relative return on each currency position (see [14]). In order to capture the heteroskedasticity of the covariates on the covariance of the carry returns, \( R_t \), we use the following model:

\[
R_t = \alpha + \beta X_t + \epsilon_t ,
\]

where \( N \) := number of currencies, \( K \) := number of covariates, \( R_t \) := N-dimensional carry returns in basket, \( \alpha \) := N-dimensional constant, \( \beta \) := N-by-K-dimensional matrix of mean covariate loadings, \( X_t \) := K-dimensional vector of covariate values, \( \epsilon_t \sim N(0, C X_t X_t^T C^T + \Psi) \) are the N-dimensional errors, \( C \) := N-by-K matrix of covariate loadings, \( \Psi \) := N-by-N baseline covariance of the errors \( \epsilon_t \).

Then unconditional covariance matrix is given by:

\[
Cov(R_t) = \beta Cov(X_t) \beta^T + C X_t X_t^T C^T + \Psi ,
\]

and the conditional covariance matrix is given by:

\[
Cov(R_t | X_t) = C X_t X_t^T C^T + \Psi .
\]

We observe that the heteroskedasticity in the conditional covariance is given by the covariance of the error terms \( \epsilon_t \).

Remark 2.1. In this paper, the following covariates are used: DOL, HML,FX, \( \sigma_{DOL} \), \( \sigma_{HML} \), and \( \sigma_{DOL,HML} \). DOL is the return an investor would earn by being long all currencies against the USD dollar. HML,FX is the return an investor would earn by being long the high interest rate basket and short the low interest rate basket. \( \sigma_{DOL} \) and \( \sigma_{HML} \) are the volatilities of the DOL and HML,FX respectively, \( \sigma_{DOL,HML} \) is the covariance of the DOL and HML,FX.

3. 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 [15], for observed data \( R_1, \ldots R_T \) (\( N \)-dimensional basket weekly carry returns of length \( T \) weeks).

\[
R_t = \alpha + \beta X_t + \gamma_t \times C X_t + \epsilon_t ,
\]

\[
\mathbb{E}[\epsilon_t] = 0 , \quad Cov(\epsilon_t) = \Psi ,
\]

\[
\mathbb{E}[\gamma_t] = 0 , \quad Var[\gamma_t] = 1 , \quad \mathbb{E}[\gamma_t \times \epsilon_t] = 0 .
\]  

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 \( \hat{\epsilon}_t \), given by \( \hat{\epsilon}_t = R_t - \hat{\alpha} - \hat{\beta} X_t \), where \( \hat{\beta} \) is the vector of mean regression loading estimates.

Step 2: Covariance Regression of Mean-Detrended Returns. Next, perform the covariance regression of these residuals on the factors, using the random-effects representation:

\[
\hat{\epsilon}_t = \gamma_t \times C X_t + \epsilon_t ,
\]

The conditional covariance matrix for \( \hat{\epsilon}_t \) is then given by,

\[
\Sigma_{\hat{X}_t} := \mathbb{E}[\hat{\epsilon}_t \hat{\epsilon}_t^T | X_t] = \mathbb{E}[\gamma_t^2 C X_t \hat{X}_t^T C^T + \gamma_t (C X_t \epsilon_t^T + \epsilon_t X_t^T C^T) + \epsilon_t \epsilon_t^T | X_t] = C X_t \hat{X}_t^T C^T + \Psi .
\]

This random-effects model allows us to perform the maximum likelihood parameter estimation of the coefficients, \( C \) and \( \Psi \), via the Expectation Maximization (EM) algorithm. We proceed by iteratively maximising the complete data log-likelihood of \( \hat{\epsilon} = \hat{\epsilon}_1, \ldots, \hat{\epsilon}_T \) denoted \( l(C, \Psi) \), which is obtained from the multivariate normal density given by:

\[
-2l(C, \Psi) = T N \log(2 \pi) + T \log |\Psi| + \sum_{t=1}^{T} (\hat{\epsilon}_t - \gamma_t C X_t)^T \Psi^{-1} (\hat{\epsilon}_t - \gamma_t C X_t).
\]

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 \( \gamma_t | \hat{\epsilon}, X, \Psi, C \) := \( N(m_t, v_t) \) with mean given by \( m_t = \gamma_t (\hat{\epsilon}_t^T \Psi^{-1} C X_t) \) and variance \( v_t = (1 + X_t^T C^T \Psi^{-1} C X_t)^{-1} \).
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 $\gamma_i$, 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 [15].

4. FORECASTING COVARIANCE USING COVARIATE TIME SERIES FORECASTS

In this section, we present the method utilised to obtain forecasts of the returns covariance matrix. In order to obtain forecasts of the covariance of the carry returns we need to forecast the covariates vector $X$. 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 [16], see [17] for details. We then assess the accuracy of the forecasts using the Mean absolute scaled error (MASE) as given in Definition 4.1 and introduced in [18]. The MASE measure scales the error based on the in-sample MAE from the naive (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 4.1.** *Mean Absolute Scaled Error (MASE)*

$$MASE_\tau = \frac{1}{\tau} \sum_{t=1}^\tau \left( \frac{\bar{e}_t}{\frac{1}{n-1} \sum_{i=2}^n |X_i - X_{i-1}|} \right)$$ (4.1)

where the numerator $\bar{e}_t$ is the forecast error at time $t$, defined as the actual value ($X_i$) minus the forecast value ($\hat{X}_i$) for that period, i.e. $\bar{e}_t = X_t - \hat{X}_t$, and the denominator is the average in-sample forecast error of the one-step naive (random walk) forecast method, which uses the actual value from the prior period as the forecast, i.e. $\hat{X}_1 = X_{t-1}$.

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

1. Fit Generalised Multi-Factor Model to the data period $[t - L : t]$ via the method in Section 3 to obtain parameter estimates $\hat{\beta}$, $\hat{\Psi}$ and $\hat{C}$. $L = 125$ data points.

2. Forecast $\tau$-step ahead covariate values, $\hat{X}_{t+\tau}$ for each covariate individually, as described by the SARIMA forecasting method in [17].

3. The $\tau$-step ahead covariance matrix is calculated as:

$$Cov(R_{t+\tau}|X_{t+\tau}) = \hat{\beta}Cov(\hat{X}_{t+\tau})\hat{\beta}^T + \hat{C}\hat{X}_{t+\tau}\hat{X}_{t+\tau}^T\hat{C}^T + \hat{\Psi}$$

and the conditional covariance matrix forecast:

$$Cov(R_{t+\tau}|\hat{X}_{t+\tau}|X_{t+\tau}) = \hat{C}\hat{X}_{t+\tau}\hat{X}_{t+\tau}^T\hat{C}^T + \hat{\Psi}$$

5. PORTFOLIO OPTIMISATION

We compare the traditional Markowitz mean-variance framework [19] to the recent Equal Risk Contribution approach introduced in [20]. There are a number of related papers from within the signal processing community, such as [21] in which a novel portfolio mixing method via weight shrinkage is proposed. Here, we consider the long and short baskets separately in the portfolio optimisation procedures below. The long or short basket returns are the weighted sums of the individual currency returns within each basket, i.e. $R_{t,basket} = \sum_{i=1}^d w_{t,i}R_{t,i}$, for currencies $i = 1, \ldots, d$ in the basket.

5.1. Markowitz Mean-Variance Optimal Portfolio

We briefly recall the general closed-form Markowitz framework for calculating the optimal portfolio weights in the unconstrained case, i.e. when weights $w$ are allowed to be negative. The following derivation is for a generic covariance matrix $\Sigma$. Thus, if we want to use the conditional covariance we use $\Sigma$ as in Equation 2.3. However, if we are interested in the unconditional covariance (as in standard Markowitz optimisation) then we use $\Sigma$ as in Equation 2.2. We seek to solve the following unconstrained optimisation problem:

$$\min_w \sigma^2_{p,w} = w^T\Sigma w \quad s.t. \quad \mu_p = w^T\mu = \mu_{p,0} \quad \text{and} \quad w^T1 = 1.$$ (5.1)

This is achieved via a Lagrangian producing first order conditions for a minimum which are linear equations with solution:

$$z_w = A^{-1}b_0, \quad \text{with}$$ (5.2)

$$A = \begin{pmatrix} \Sigma & \mu^T \\ \mu^T & 0 \end{pmatrix} \quad \text{and} \quad b_0 = \begin{pmatrix} 0 \\ \mu_{p,0} \end{pmatrix}.$$ (5.2)

Note that the first $d$ elements of $z_w$ are the optimal portfolio weights $w = (w_1, \ldots, w_d)$ 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 $w$ is an efficient portfolio. In our case we need to constrain 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 [22, 23] for detailed references. We select the Markowitz portfolio with the maximum Sharpe ratio, i.e. return/volatility ratio.

5.2. Equal Risk Contribution Portfolio

Due to the high sensitivity of mean-variance portfolio optimisation approaches to the input parameters, and in particular to the expected returns, risk-based techniques have arisen as an alternative, the example we consider is Equal Risk Contribution (ERC), see [20]. Each asset contributes equally to...
the total portfolio variance. A more general approach to risk-based portfolio construction can be seen in [24, 25].

**Definition 5.1. Equal Risk Contribution Portfolio (ERC)**

\[
\mathbf{w}^* = \{ \mathbf{w} \in [0,1]^N : \sum w_i = 1, \mathbf{w} \times \partial w_j \sigma(\mathbf{w}) = w_j \times \partial w_j \sigma(\mathbf{w}) \forall i, j \} \\
\text{where } \partial w_j \sigma(\mathbf{w}) = \frac{\partial \sigma(\mathbf{w})}{\partial w_j}, \text{ and } \sigma(\mathbf{w}) = \sqrt{\mathbf{w}^T \Sigma \mathbf{w}}.
\]  

(5.3)

Noting that \( \partial w_j \sigma(\mathbf{w}) \propto (\Sigma \mathbf{w})_j \), where \((\Sigma \mathbf{w})_j\) denotes the \(i\)-th row of the vector issued from the product of \(\Sigma\) with \(\mathbf{w}\), we have the following optimisation problem:

\[
\mathbf{w}^* = \{ \mathbf{w} \in [0,1]^N : \sum w_i = 1, \mathbf{w} \times (\Sigma \mathbf{w})_i = w_j \times (\Sigma \mathbf{w})_j \forall i, j \} \\
\]  

(5.4)

This problem can be solved using a Sequential Quadratic Programming (SQP) algorithm, see details in [20].

**6. CURRENCY PORTFOLIO ANALYSIS**

We consider for our empirical analysis a carry trade portfolio consisting of a long basket and 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 [5], 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, \(R_i\). The daily time series analysed were obtained from Bloomberg and range from 04/01/1999 to 29/01/2014.

**6.1. Covariate SARIMA Forecast Results**

The MASE forecast accuracy results suggest that on average all five of the covariates have reasonable forecast performance. However, we note that the \(DOL\) and \(HML\) covariates are very close to white noise and hence the naïve in-sample forecasting method is very poor. Thus, if we also look at the Mean Absolute Percentage Error (MAPE) we see that the \(DOL\) and \(HML\) have median MAPEs of 99% whereas the \(\sigma_{DOL}, \sigma_{HML}\) and \(\sigma_{DOL,HML}\) have median MAPEs of 11%, 13% and 21% respectively. However, it is noticeable that for the \(\sigma_{DOL}, \sigma_{HML}\) and \(\sigma_{DOL,HML}\) there are periods of poor forecast performance, e.g. the 2008 Financial Crisis.

**6.2. Portfolio Risk and Performance**

The portfolio performance and risk of the four different approaches are presented in Figure 1 and in Tables 1 and 2. Here, 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. 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. Max DD is the maximum decline from historical peak. It can be seen that the traditional Markowitz approach is much improved when using our forecasting approach. The ERC approach is also slightly improved by incorporating the heteroskedastic covariance forecast.

In this paper, we present the Generalised Multi-Factor model, which incorporates the factors into the covariance of the idiosyncratic error term and hence allows for heteroskedastic covariance. We utilise this model in the setting of forecasting currency carry trade returns and covariances. Two portfolio optimisation approaches are considered: Markowitz and Equal-Risk Contribution. Equal-Risk Contribution is shown to outperform the Markowitz approach on all measures except for the maximum drawdown. We note that when some or all of the expected returns are negative then the Markowitz weights are concentrated to the point of excluding the currencies with the most negative expected returns. Our Generalised Multi-Factor model is shown to provide an improvement to the performance and risk characteristics of the portfolios.
8. REFERENCES


