The economic value of volatility forecasts:
A conditional approach*

Nicholas Taylor†

Cardiff University

December 31, 2012

Abstract

This paper provides an investigation of the economic value of multivariate volatility forecasting ability using a testing framework in which the quality of competing methods is assessed from a conditional investment perspective. In doing this we are able to provide a novel means by which the benefits of using a particular set of volatility forecasts can be assessed. The results associated with an application to US bond and stock futures markets indicate that investors are willing to pay a significant premium for knowledge of the dynamics of volatility, though the magnitude of this premium varies over time, and depends on risk preferences and economic conditions. The latter variation implies that selection of appropriate dynamics (this is, forecasting method choice) should be a conditional exercise.

Key Words: Economic value, forecasting, volatility.

JEL Classification Code: C53, G11, G17.

1 Introduction

The importance of forecasting risk in financial markets has motivated a vast literature on the evaluation of volatility forecast quality (see Poon and Granger, 2003, for a comprehensive review). The value of any new approach is often assessed by establishing (or refuting) both the statistical *and* economic significance of the proposed forecasts. Assessment of the former is characterised by well-defined metrics (e.g., the mean squared error), while the latter is often based on economic criteria that provide a translation of forecasting ability into meaningful units (e.g., a performance fee measured in basis points). Despite their usefulness, improvements in existing measures are possible. The purpose of the current paper is to introduce a measure of the economic value of volatility forecasting ability such that it varies over time. In doing this, we are able to conduct a more thorough examination of volatility forecast quality by considering unconditional *and* conditional aspects of this ability.

A variety of economic measures of forecasting performance have been considered in the extant literature. Commonly-employed examples include the Sharpe ratio and the performance fee associated with conditionally-optimised portfolios constructed using volatility forecasts; see, e.g., West et al (1993), Fleming et al (2001, 2003), Marquering and Verbeek (2004), Şerban et al (2007), Clements et al (2009), Liu (2009), and Chiriac and Voev (2011).[1] Such measures are unconditional in nature and deliver a single performance measure over the entire sample.[2] For instance, the unconditional performance fee represents the amount an investor is willing to pay to switch from one investment strategy to another in order to equate the unconditional expected utility levels over the two strategies. By contrast, we introduce a measure based on *conditional* expected utility levels. Use of this measure drives the primary contribution of this paper: an assessment of multivariate volatility forecast quality using the Giacomini and White (2006) conditional testing framework under economically meaningful loss function assumptions.

The proposed approach is applied to data pertaining to US bond and stock futures markets. In particular, we consider investors (with various utility functions) who take positions in the 30-year Treasury bond and S&P 500 index futures market at the beginning of each day in order to maximise their conditional expected utility over the day. The investors in our study differentiate themselves in terms of their volatility forecasting method. At the limit we consider investors who select portfolio weights based on a simple random walk model, and those who

---

[1]A similar approach has been adopted with respect to the evaluation of asset return forecast quality; see, e.g., Rapach et al (2010).

[2]Voev (2009) demonstrates that in the presence of noise in the realised volatility measure, use of unconditional economic criteria can lead to sizeable distortion in the ranking of competing forecasts.

select portfolio weights based on advanced volatility methods. By considering the economic cost of competing forecasting methods, we are explicitly examining the benefits of using one volatility method with respect to a competing method within a meaningful investment context. Moreover, we are able to measure how this economic value is distributed over time, and to test whether this value covaries with market conditions (including volatility, correlation, and economic activity levels). Few, if any, studies have conducted such a rich assessment of the economic value of multivariate volatility forecast quality.

To anticipate some of the results, we find that methods based on parameters estimated by optimising investors' utility perform well. Moreover, the conditional testing framework reveals that the performance of each method (relative to a competing method) varies in a predicable fashion over time, with volatility levels contributing to performance differences. For instance, GARCH-based methods typically exhibit below average performance during low volatility periods, but perform very well during high volatility periods. Such predictability means that methods based on functions of forecast sets (including those based on switching to the best method each period) exhibit excellent performance under all conditions. The results also highlight the virtues of using simple methods. The moving average method consistently beats the majority of competing models and enjoys relatively low levels of portfolio turnover (and thus transaction costs). Moreover, a method based on the random walk assumption performs very well when investors face portfolio weight restrictions. Such results have two practical implications: researchers should be mindful of the nature of the data sample used (e.g., a method that is successful during high volatility periods may not be so successful during other periods), and investors should consider switching between a set of forecasting methods that includes those based on economic criteria optimisation and those based on simple dynamics.

2   Multivariate volatility forecasting methods

A wide variety of multivariate volatility forecasting methods have been proposed in the literature, each of which differentiates itself in terms of model specification and/or the estimation technique used. While we cannot include all methods, we do consider a representative subset of this universe. Specifically, we consider methods based on five popular realised volatility (RV) models, four popular conditional volatility (CV) models, combinations thereof, and on a number of different estimation procedures.

2.1 RV-based methods

The first method we consider is the simplest of all methods whereby forecasts of realised covariance are given by realised covariance observed in the preceding period; formally,

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \text{vech}(\mathbf{Y}_{t-1}), \quad t \in \{S+1, \ldots, T\}, \tag{1}$$

where $\text{vech}(\mathbf{A})$ is the half-vectorisation operator such that it stacks the nonredundant elements of the matrix $\mathbf{A}$, and $\mathbf{Y}_t$ is an $(N \times N)$ daily frequency realised covariance matrix measure.[3] As this method assumes that covariances follow a random walk (RW), this method is henceforth referred to as the *RW method.*

The next two methods assume that forecasts of realised covariance are given by the mean of previously observed realised covariances; specifically,

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \widehat{\boldsymbol{\phi}}_{0,t-1}, \tag{2}$$

where $\widehat{\boldsymbol{\phi}}_{0,t-1}$ is the $(N(N+1)/2 \times 1)$ out-of-sample maximum likelihood (ML) coefficient vector. This vector is given by the solution to the following residual sum of squares (RSS) constrained optimisation problem:

$$\widehat{\boldsymbol{\phi}}_{0,t-1} = \operatorname*{argmin}_{\boldsymbol{\phi}_{0,t-1}} \sum_{r=t_0}^{t-1} |(\text{vech}(\mathbf{Y}_r) - \boldsymbol{\phi}_{0,r-1})(\text{vech}(\mathbf{Y}_r) - \boldsymbol{\phi}_{0,r-1})'|, \tag{3a}$$

$$\text{s.t. } \lambda_i(\boldsymbol{\phi}_{0,r-1}) > 0, \ \forall\, r \in \{t_0, \ldots, t-1\}, \text{ and } i \in \{1, \ldots, N\}, \tag{3b}$$

where $\lambda_i(\mathbf{A})$ is the ith eigenvalue of $\mathbf{A}$, and $t_0 = t - t_{\text{win}}$ dictates the amount of past data to use in the estimation process. We use an expanding window of past observations such that $t_{\text{win}} = t - 1$ (henceforth the *MEAN method*), and a 20-day moving average (MA) such that $t_{\text{win}} = 20$ (henceforth the *MA method*).[4]

Given the documented benefits of vector autoregressive (VAR) models of realised volatility (see, e.g., Andersen et al, 2003), the next method we consider assumes that forecasts are given

---

[3]The covariance matrix measure used is based on the subsampling methodology of Zhang et al (2005) within a multivariate setting (see, e.g., Chiriac and Voev, 2011, for use of this methodology in a similar context). Further details are provided in Appendix A.

[4]The 20-day fixed window size is selected because of its use in previous studies within the volatility forecasting literature; see, e.g., Jorion (1995).

4

by the following first-order vector autoregressive moving average (VARMA) structure:

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \widehat{\boldsymbol{\phi}}_{0,t-1} + \widehat{\boldsymbol{\phi}}_{1,t-1}\,\text{vech}(\mathbf{Y}_{t-1}) + \widehat{\boldsymbol{\theta}}_{1,t-1}\widetilde{\boldsymbol{\xi}}_{t-1}, \qquad (4)$$

where $\widetilde{\boldsymbol{\xi}}_{t-1}$ is a $(N(N+1)/2 \times 1)$ vector of residuals, and $\widehat{\boldsymbol{\phi}}_{1,t-1}$ and $\widehat{\boldsymbol{\theta}}_{1,t-1}$ are $(N(N+1)/2 \times N(N+1)/2)$ ML-estimated coefficient matrices based on an expanding window of data (that is, data observed up to time $t-1$). This method is henceforth referred to as the *VARMA method*.

The above approach is augmented in two ways. First, Engle and Rangel (2007) find that macroeconomic variables such as GDP growth have an impact on low frequency volatility. To investigate whether macroeconomic effects are present in high frequency volatility measures, we consider forecasts generated as follows:

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \widehat{\boldsymbol{\phi}}_{0,t-1} + \widehat{\boldsymbol{\phi}}_{1,t-1}\,\text{vech}(\mathbf{Y}_{t-1}) + \widehat{\boldsymbol{\Gamma}}_{t-1}X_{t-1} + \widehat{\boldsymbol{\theta}}_{1,t-1}\widetilde{\boldsymbol{\xi}}_{t-1}, \qquad (5)$$

where $\widehat{\boldsymbol{\Gamma}}_{t-1}$ is an $(N(N+1)/2 \times 1)$ ML-estimated coefficient vector based on an expanding window of data, and X_t is the daily frequency real-time measure of economic activity introduced by Aruoba, Diebold, and Scotti (2009). This method is henceforth referred to as the *VARMAX method*.

The second augmentation adopts the heterogeneous autoregressive (HAR) model of realised variance introduced by Corsi (2009).[5] This provides a convenient way of allowing for long-memory dynamics and is achieved via an additive cascade model of volatility components defined over different time periods. In the spirit of this model, we consider the following HAR-based forecasts:

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \widehat{\boldsymbol{\phi}}_{0,t-1} + \widehat{\boldsymbol{\phi}}_{1,t-1}\,\text{vech}(\mathbf{Y}_{t-1}) + \widehat{\boldsymbol{\phi}}_{2,t-1}\sum_{k=1}^{5}\text{vech}(\mathbf{Y}_{t-k}) + \widehat{\boldsymbol{\phi}}_{3,t-1}\sum_{k=1}^{20}\text{vech}(\mathbf{Y}_{t-k}), \quad (6)$$

where $\widehat{\boldsymbol{\phi}}_{2,t-1}$ and $\widehat{\boldsymbol{\phi}}_{3,t-1}$ are $(N(N+1)/2 \times N(N+1)/2)$ ML-estimated coefficient matrices based on an expanding window of data. This method is henceforth referred to as the *vector-HAR method*.[6]

As an alternative to using ML-estimated parameters based on the RSS function, Kirby and Ostdiek (2011) demonstrate that improvements in performance are available to investors (with

---

[5]Corsi concludes that the performance of the HAR model '*is comparable to the much more complicated and tedious to estimate long-memory ARFIMA model*'.

[6]The subsequent analysis finds that the performance of the VAR-based methods is improved by assuming that all coefficient matrices are diagonal.

unconditional mean-variance risk preferences) if parameters are estimated using an objective function that matches the portfolio management strategy rather than a conventional RSS-type objective function. Within the current context, we examine this potential by considering new versions of the VAR-based methods in which parameters are estimated such that the economic criterion (utility-based loss function) in Corollary 1 in Appendix B is minimised using an expanding window of data; henceforth these methods are referred to as the *VARMA (econ.)*, *VARMAX (econ.)*, and *vector-HAR (econ.) methods*. In contrast, the counterpart methods based on the RSS statistical criterion are henceforth referred to as the *VARMA (stat.)*, *VARMAX (stat.)*, and *vector-HAR (stat.) methods.*[7]

2.2 CV-based methods

In addition to the above RV-based methods, we also consider four CV-based methods that are commonly used in practise (each of which is based on low frequency measures of past volatility). First, motivated by the current RiskMetrics methodology introduced in 2006, we assume that forecasts are generated via the weighted average of several different exponential weighted moving average (EWMA) forecasts:

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \sum_{k=1}^{K} w_k \, \text{vech}(\widetilde{\mathbf{Y}}_{k,t}). \tag{7a}$$

In turn, each EWMA forecast is given by

$$\text{vech}(\widetilde{\mathbf{Y}}_{k,t}) = \mathbf{\Lambda}_k \, \text{vech}(\widetilde{\mathbf{Y}}_{k,t-1}) + (\mathbf{I}_N - \mathbf{\Lambda}_k) \, \text{vech}(\mathbf{R}_{t-1}\mathbf{R}'_{t-1}), \tag{7b}$$

with coefficients,

$$\mathbf{\Lambda}_k = \exp(-1/\tau_k)\mathbf{I}_N, \tag{7c}$$

and weights,

$$w_k = \begin{cases} \frac{1}{C}\left(1 - \frac{\ln(\tau_k)}{\ln(\tau_0)}\right), & \text{if } \tau_k \leq \tau_{max}, \\ 0, & \text{otherwise}, \end{cases} \tag{7d}$$

---

[7]For investors who use the MEAN (or MA) method, it is possible to demonstrate that changing the objective function to the economic criterion does not change the parameter values. Consequently, we make no distinction between these methods with respective to the objection function used.

where $\mathbf{R}_{t-1}$ is an $(N \times 1)$ vector of lagged open-to-close daily returns, and C is a normalising constant designed to ensure that the weights sum to unity.[8] We follow RiskMetrics advice and assume that $\tau_k = \tau_1 \rho^{k-1}, \rho = \sqrt{2}, \tau_0 = 1560, \tau_1 = 4$, and $\tau_{max} = 256$; see Zumbach (2007) for further details. The above method encompasses the original RiskMetrics methodology introduced in 1994. Specifically, if one assumes that $K = 1$ and $\mathbf{\Lambda}_1 = 0.94\mathbf{I}_N$, such that

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = 0.94 \, \text{vech}(\widetilde{\mathbf{Y}}_{t-1}) + 0.06 \, \text{vech}(\mathbf{R}_{t-1}\mathbf{R}'_{t-1}), \tag{8}$$

then the old methodology is replicated. This restricted approach, and its unrestricted counterpart, are henceforth referred to as the *RM1994* and *RM2006 methods*, respectively.

The third CV-based method is similar in spirit to the RM1994 method in that past volatility is exponentially-weighted. However, in contrast, the decay parameter is estimated. Specifically, we follow Fleming et al (2001, 2003) and assume that covariance forecasts are given by

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \exp(-\widehat{\lambda}_{t-1}) \, \text{vech}(\widetilde{\mathbf{Y}}_{t-1}) + \widehat{\lambda}_{t-1} \exp(-\widehat{\lambda}_{t-1}) \, \text{vech}(\mathbf{R}_{t-1}\mathbf{R}'_{t-1}), \tag{9}$$

where $\widehat{\lambda}_{t-1}$ is the ML-estimated decay parameter based an expanding window of data and a Gaussian likelihood function; see Fleming et al (2003) for further details of the estimation methodology. This method is henceforth referred to as the *EXP method*.

While the above method is widely employed it is limited in its ability to capture complex conditional volatility dynamics via a single parameter. To address this issue we consider a method based on the dynamic conditional correlation (DCC) multivariate generalised autoregressive conditional heteroskedasticity (MGARCH) model introduced by Engle (2002). This model implies decomposed forecasts given by

$$\text{vech}(\widetilde{\mathbf{Y}}_t) = \text{vech}(\text{dg}(\widetilde{\mathbf{Y}}_t)^{1/2}\widetilde{\boldsymbol{\rho}}_t \, \text{dg}(\widetilde{\mathbf{Y}}_t)^{1/2}), \tag{10a}$$

where $\widetilde{\boldsymbol{\rho}}_t$ is the value of the conditional correlation matrix. The diagonal elements in $\widetilde{\mathbf{Y}}_t$ each follow a separate univariate GARCH(1,1) process,

$$\text{vech}(\text{dg}(\widetilde{\mathbf{Y}}_t)) = \widehat{\boldsymbol{\alpha}}_{0,t-1} + \widehat{\boldsymbol{\alpha}}_{1,t-1} \odot \text{vech}(\text{dg}(\mathbf{R}_{t-1}\mathbf{R}'_{t-1})) + \widehat{\boldsymbol{\beta}}_{1,t-1} \odot \text{vech}(\text{dg}(\widetilde{\mathbf{Y}}_{t-1})), \tag{10b}$$

where $\widehat{\boldsymbol{\alpha}}_{0,t-1}$, $\widehat{\boldsymbol{\alpha}}_{1,t-1}$, and $\widehat{\boldsymbol{\beta}}_{1,t-1}$ are $(N \times 1)$ ML-estimated parameter vectors for each separate

---

[8]The weighting of different EWMA forecasts is designed to capture long memory in the volatility process (Zumbach, 2007).

GARCH equation based on an expanding window of data and a Gaussian likelihood function (as suggested in Engle, 2002), and $\odot$ denotes the Hadamard product. The value of the conditional correlation matrix is obtained by assuming that

$$\widetilde{\boldsymbol{\rho}}_t = \mathrm{dg}(\widetilde{\mathbf{Q}}_t)^{-1/2}\widetilde{\mathbf{Q}}_t\,\mathrm{dg}(\widetilde{\mathbf{Q}}_t)^{-1/2}, \tag{10c}$$

and modelling $\widetilde{\mathbf{Q}}_t$ as follows:

$$\mathrm{vech}(\widetilde{\mathbf{Q}}_t) = (1 - \widehat{\psi}_{1,t-1} - \widehat{\psi}_{2,t-1})\,\mathrm{vech}(\overline{\mathbf{Q}}_{t-1}) + \widehat{\psi}_{1,t-1}\,\mathrm{vech}(\boldsymbol{v}_{t-1}\boldsymbol{v}'_{t-1}) + \widehat{\psi}_{2,t-1}\,\mathrm{vech}(\widetilde{\mathbf{Q}}_{t-1}), \tag{10d}$$

where $\widehat{\psi}_{1,t-1}$ and $\widehat{\psi}_{2,t-1}$ are ML-estimated parameters based on an expanding window of data and the Gaussian likelihood function specified in Engle (2002), and $\boldsymbol{v}_{t-1}$ is the standardised return vector with unconditional covariance $\overline{\mathbf{Q}}_{t-1}$. This method is henceforth referred to as the *DCC-MGARCH method*.[9]

2.3 Combination methods

Given the documented superior performance of combination forecasts in other literatures (see, e.g., Stock and Watson, 2003), we combine the forecasts associated with the above methods (excluding those based on economic criterion minimisation). Specifically, we examine the performance of the median of the forecasts associated with all methods and subsets thereof (viz. RV-based and CV-based methods). These are henceforth referred to as the *combination (RV)*, *combination (CV)*, and *combination methods*.[10]

In addition, we also consider a portfolio weight shrinkage approach in which an investor's weights are given by the median of the weights implied by the forecasting methods described above and subsets thereof. These combined weights are then used to generate the values of the economic loss functions derived in Appendix B.[11] These methods are henceforth referred to as the *shrinkage (RV)*, *shrinkage (CV)*, and *shrinkage methods*.

---

[9]It is possible to allow for leverage effects in the DCC-MGARCH model; see, e.g., the asymmetric DCC-MGARCH model of Cappiello et al (2006). Indeed, we considered such a model in the current analysis. However, as the results were very similar to those of the DCC-MGARCH model, we refrained from presenting these results – though they are available on request.

[10]The mean of the forecasts is also considered. Indeed, Timmermann (2006) demonstrates that mean combining is optimal in the presence of squared error loss. However, we find that median forecasts deliver slightly superior results in the current application. See Andreou et al (2012), for further details of volatility forecast combination performance under a variety of loss functions.

[11]Tu and Zhou (2011) demonstrate that optimal combinations of various sophisticated portfolio rules with the $1/N$ rule are superior to competing strategies from an unconditional perspective. However, in the absence of an optimal conditional combining scheme, we focus on the simplest of all combination strategies in which the median of all portfolio weights is used.

3 Assessing forecast quality

Volatility forecasts are assessed using investor utility-based loss functions and a conditional testing framework. These are described in the following subsections.

3.1 Utility-based loss functions

Under the assumption that investors allocate their wealth each period by selecting their optimal portfolios based on competing covariance matrix forecasts (generated by the above forecasting methods), the economic cost of forecasting method B with respect to forecasting method A is defined by the following conditional utility difference:

$$\Delta L_t = \mathrm{E}(f(R_t^p)|\mathcal{G}_{a,t}) - \mathrm{E}(f(R_t^p)|\mathcal{G}_{b,t}), \tag{11}$$

where $f(.)$ is the utility function, R_t^p is the portfolio return, and $\mathcal{G}_{a,t}$ and $\mathcal{G}_{b,t}$ depend on forecasting methods A and B, respectively. If we replace forecasting method A with a method that generates perfect covariance matrix forecasts (that is, realised values of the covariance matrix), then the expression in (11) can be interpreted as a loss function that measures the cost of forecast imperfection associated with method B. Explicit expressions for the above measure are provided in Appendix B under HARA and quadratic utility.

In addition, the results in Appendix B demonstrate that these differences can be interpreted as the performance fee an investor is willing to pay each period to avoid using a competing investor's forecasting method; see Proposition 3. Specifically, we demonstrate that under quadratic utility, the (maximum) performance fee δ_t^* given by the root of the following equation:

$$\mathrm{E}(f(R_t^p - \delta_t)|\mathcal{G}_{a,t}) - \mathrm{E}(f(R_t^p)|\mathcal{G}_{b,t}) = 0 \tag{12}$$

is equal to ΔL_t. Consequently, the economic cost of forecast method B with respect to method A, is equal to the performance fee an investor who employs forecasting method A is willing to pay each period to avoid using forecasting method B.[12]

---

[12]We also demonstrate that, under certain assumptions, the proposed measure is a member of Patton's (2011) Homogeneous Robust (HR) loss function family; see Proposition 4 in Appendix B. Consequently, we demonstrate that it is robust to the presence of noise in the realised volatility measure; see, e.g., Hansen and Lunde (2006), Laurent et al (2009), and Patton (2011) for details of robust loss functions.

3.2 The testing framework

We wish to test whether the relative costs associated with competing volatility forecasting methods are orthogonal to a particular information set. This can be achieved via the following ordinary least squares (OLS) regression:

$$\Delta L_t = \boldsymbol{\varphi}' \mathbf{Z}_{t-1} + \epsilon_t, \quad t \in \{S+1, \ldots, T\}, \tag{13}$$

where ΔL_t is the out-of-sample relative economic cost of competing forecasting methods, $\boldsymbol{\varphi}$ is a vector of coefficients, $\mathbf{Z}_{t-1}$ is a vector of explanatory variables, and ϵ_t is a suitably defined error term. We express each hypothesis as q zero restrictions $\boldsymbol{\Theta}\boldsymbol{\varphi} = \mathbf{0}_q$ imposed upon the above regression and calculate the following robust Wald test statistic:

$$\mathcal{W} = (T - S)(\boldsymbol{\Theta}\widehat{\boldsymbol{\varphi}})'(\boldsymbol{\Theta}\mathbf{M}^{-1}\widehat{\mathbf{V}}\mathbf{M}^{-1}\boldsymbol{\Theta}')^{-1}(\boldsymbol{\Theta}\widehat{\boldsymbol{\varphi}}) \xrightarrow{d} \chi^2(q). \tag{14}$$

where $\mathbf{M} = (T - S)^{-1}\sum_{t=S+1}^{T} \mathbf{Z}_{t-1}\mathbf{Z}'_{t-1}$ and $\widehat{\mathbf{V}}$ is a heteroscedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance matrix of $\widehat{\boldsymbol{\varphi}}$.

A number of existing tests can be considered within the above framework. First, a test similar to the conditional predictive ability test of Giacomini and White (2006), henceforth the *GW conditional test*, can be performed. Specifically, we test the null hypothesis

$$\text{H}_0: \text{E}(\Delta L_t | \mathcal{I}_{t-1}) = 0, \tag{15}$$

by assuming that $\boldsymbol{\Theta} = \mathbf{I}_q$ (with q set equal to the number of elements in $\mathbf{Z}_{t-1}$), where $\mathcal{I}_{t-1}$ is an information set of interest, and $\mathbf{Z}_{t-1}$ is $\mathcal{I}_{t-1}$-measurable. Note that the test statistic implied by these assumptions is slightly different from the simplified one suggested by Giacomini and White (2006); see equation (4) in their paper. Use of their formulation in the current application is not recommended as expanding window forecasting schemes are used.

Second, it is also possible to examine the impact of more specific information sets on costs. In particular, the unconditional predictive ability test of Giacomini and White (2006) can be conducted, henceforth the *GW unconditional test* (see Giacomini and White, 2006, for a comparison to the related approaches of Diebold and Mariano, 1995, and West, 1996). This test examines the hypothesis

$$\text{H}_0: \text{E}(\Delta L_t) = 0, \tag{16}$$

by assuming that $\boldsymbol{\Theta} = 1$, and $\mathbf{Z}_{t-1} = 1 \, \forall \, t$. Here $\mathcal{I}_{t-1}$ has been replaced with an empty set such that only a constant term is included in the test regression.

Third, the above framework also enables consideration of relative costs within different regimes. In particular, we are interested in testing the null hypotheses that

$$\text{H}_0\text{: } \text{E}(\Delta L_t | \mathcal{R}_t = j) = 0, \quad j \in \{1, \dots, J\}, \tag{17a}$$

where $\mathcal{R}_t$ is the regime at time t. We can go further and test the null hypothesis that costs are equal across all regimes:

$$\text{H}_0\text{: } \text{E}(\Delta L_t) | \mathcal{R}_t = 1) = \dots = \text{E}(\Delta L_t | \mathcal{R}_t = J). \tag{17b}$$

These hypotheses can be tested by respectively assuming that $\boldsymbol{\Theta} = \mathbf{e}'_{j,J}$ (henceforth the *regime test (individual)*), and $\boldsymbol{\Theta} = (\mathbf{1}_{J-1}, -\mathbf{I}_{J-1})$ (henceforth the *regime test (difference)*), where $\mathbf{Z}'_{t-1} = (1_{X_t \leq c_1}, \dots, 1_{X_t > c_{J-1}})$, X_t is the variable of interest underlying the regimes, $\mathbf{e}_{j,J}$ is the jth column of $\mathbf{I}_J$, $1_{\mathcal{K}}$ equals one if condition $\mathcal{K}$ is satisfied (and zero otherwise), and c_j are pre-selected cutoff points.

4   Results

We proceed with an empirical analysis of daily return volatility in the 30-year Treasury bond and S&P 500 index futures markets. This begins with a description of the data used.

4.1   Data

We consider trades in the 30-year Treasury bond and S&P 500 index futures markets over the period, January 1, 1983, to December 31, 2011. In particular, one-second frequency transaction prices were obtained for each market (with ticker symbols US and SP, respectively) from *TickData, Inc.*, with the grid of prices filled with the most recently available price.[13] To obtain a single continuous series for each futures contract, we adopt the common practise that futures contracts with the nearest maturity are replaced (through trading) by contracts with the next nearest maturity when the next contract's daily tick count exceeds the current contract tick

---

[13] All prices pertain to pit transactions, except for the 30-year Treasury bond futures prices over the period July 1, 2003, to December 31, 2011 (reflecting the dominance of the electronic trading mechanism for this market over this period).

count using TickData's automatic roll dates. These data were collected over the daytime trading period in which both markets are open.[14]

The subsequent analysis also makes use of various macroeconomic data. These data were obtained from the following sources. We use the latest vintage of the daily Aruoba-Diebold-Scotti (ADS) index obtained from the Federal Reserve Bank of Philadelphia website: `http://www.philadelphiafed.org/research-and-data/`. The dates of the economic expansion and contraction periods were obtained from the National Bureau of Economic Research (NBER) website: `http://www.nber.org/cycles.html`. Inflation expectations are given by the Thomson Reuters/University of Michigan survey-based inflation expectation series and consist of monthly values of the median expected price change over the next 12 months. This series was obtained from the Federal Reserve Economic Data (FRED) website: `http://research.stlouisfed.org/fred2/`.

4.2 Estimation details

Before proceeding to a description of the results, further details of the assumptions used in the estimation process are required. First, all models are estimated using a minimum of five years of data such that the sample period, $t \in \{5 \times 251, \dots, T\}$ is used, with out-of-sample 1-step ahead forecasts generated at each point in time. These forecasts are then compared to realised variance and covariance (given by the averaged subsample realised covariance matrix measure). Second, the models (where applicable) are estimated using the BFGS optimisation method with Brent line search programmed via use of the Constrained Optimization application in GAUSS 11.0 (64-bit version).[15]

4.3 Summary statistics

A selection of summary statistics associated with the averaged subsample realised covariance matrix measure given by (A.5) are given in Table 1. The results highlight three main characteristics of the data. First, the S&P 500 index futures market is more volatile than the 30-year Treasury bond futures market, reflecting the well established result that stocks are more risky than bonds. Second, all volatility measures are highly time-dependent (as evinced by the

---

[14]The intersection point of trading hours varies over the sample. At the start of the sample in 1983 the trading hours of the 30-year Treasury bond and S&P 500 index futures markets were 8.00am to 2.05pm (CT) and 9.00am to 3.15pm (CT), respectively. However, the current (circa February 2012) trading hours of these markets are 7.20am to 2.00pm (CT) and 8.30am to 3.15pm (CT), respectively.

[15]All code used in this paper is available on request.

first and fifth-order autocorrelations), with fifth-order autocorrelation only slightly lower than first-order autocorrelation. Thus, it is important to allow for strong levels of persistence when modelling these volatility measures. Third, realised correlation is noticeably less than unity and exhibits nontrivial variation over time. For instance, during the sub-periods 1983 to 1997 and 1998 to 2011, average realised correlation was 0.347 and -0.302, respectively; see Yang et al (2009) for an economic rationale for such variation. This has two implications: there would appear to be diversification benefits to those investors who simultaneously trade both of these markets; and one must adequately account for the persistence and time-variation in correlation.

Insert Table 1 here

4.4 Statistical measures of forecast accuracy

The results associated with selective members of Patton's (2011) HR loss function family given by (B.14) are provided in Table 2. Specifically, we consider $b = 0$ (MSE), $b = -1$, $b = -2$ (QLIKE), and $b = -3$. The last of these loss functions is particularly relevant in the current application as it coincides with the loss function derived in the current paper under, inter alia, the single-asset portfolio and quadratic utility assumption; see Proposition 4 in Appendix B. The null hypothesis that the MA method has equal forecast quality to each of the other methods is tested using a Wald test based on Newey-West HAC standard errors.

Insert Table 2 here

The results can be summarised as follows. First, irrespective of the shape of the loss function, only a small number of methods produce variance forecasts that are significantly more accurate than the forecasts associated with the MA method, particularly when considering the utility-based loss function considered in this paper (this is, $b = -3$). Second, the forecasts associated with the utility-optimised and combination methods appear to be highly accurate. Third, the relative performance of the other methods depends upon the measure and/or loss function used. For instance, the RM2006 method produces highly accurate realised variance forecasts when stock data and $b = -3$ are used. However, its relative performance deteriorates when applied to bond data (under the same loss function).[16]

---

[16]The RM2006 method provides uniformly more accurate forecasts than the RM1994 method.

4.5 Economic measures of forecast accuracy

Having established an initial ranking of various forecasting methods in terms of their statistical accuracy, we proceed with an examination of the economic value of this accuracy.

4.5.1 Mean imperfection costs

The economic costs of forecast imperfection for each forecasting method are calculated under the assumption that investors hold unbounded positions in either bonds, stocks, or a combination of the two, have log or quadratic utility, $\gamma \in \{1, 2, 4, 8\}$, and $\boldsymbol{\mu}' - R^f = (0.017, 0.055)$ (annualised), where γ is the coefficient of relative risk aversion, $\boldsymbol{\mu}$ is the expected return vector, and R^f is the risk-free rate.[17] The means of these costs are provided in Table 3, with differences between these costs representing the mean performance fees associated with competing methods.[18]

Insert Table 3 here

A number of features are apparent amongst the considerable variation in the mean costs over the portfolios, utility functions, and methods considered. First, as γ is proportional to costs (see (B.11a) in Corollary 1), it follows that these fall as γ increases. Second, the ranking of the methods is consistent over the utility function assumptions, with log utility results similar to the quadratic utility ($\gamma = 1$) results. Finally, combination-type methods perform particularly well, with those based on all methods outperforming those based on RV or CV methods. For instance, when investors are permitted to simultaneously hold bonds and stocks, the mean cost of forecast imperfection to an investor who uses the shrinkage method under log utility is 8.0% per annum. This compares to 8.6% for the MA method, and 16.3% for the MEAN method (the worst method).[19]

The mean costs of forecast imperfection in Table 3 also highlight the virtues of using forecasting methods with parameters optimised according the objective function used. In particular, the VARMA (econ.), VARMAX (econ.), and vector-HAR (econ.) methods deliver mean costs of

---

[17]The assumption that $\boldsymbol{\mu}' - R^f = (0.017, 0.055)$ (annualised) is based on the average mean returns to 30-year Treasury bond and S&P 500 index futures over the 1983 to 2011 period. Unless otherwise stated, this assumption is maintained in the subsequent analysis.

[18]All costs are given in annual percentage terms.

[19]Despite their simplicity, the quality of MA-type methods has been documented previously. For instance, González-Rivera et al (2004) demonstrate that an MA method based on 20 days of past squared returns is no worse than a variety of GARCH-based methods when forecasting S&P 500 index volatility under an option-based loss function assumption. The excellent performance of such methods may explain their popularity within the practitioner community.

forecast imperfection that are uniformly lower than the competing methods. For instance, for investors with log utility who trade both futures markets, the mean cost of forecast imperfection equals 7.1% (vector-HAR (econ.) method), compared to 10.5% (vector-HAR (stat.) method) – implying a benefit to using parameters based on the utility-based loss function of 3.4%. It is also noticeable that the utility-optimised methods are the only ones (apart from the shrinkage method) to provide significantly lower costs than the MA method. This result is consistent with the findings of Kirby and Ostdiek (2011) regarding the superior performance of methods based on parameters that are estimated using an objective function that matches the portfolio management strategy.[20]

4.5.2 Portfolio turnover

To give an indication of the magnitude of transaction costs that are likely to be incurred by investors in the current analysis, we report mean portfolio turnover for each asset (including the risk-free asset). This is defined as the mean of the absolute change in portfolio weights associated with each asset and forecasting method. Results provided in Table 4 are calculated under the assumption that investors trade both futures markets.

Insert Table 4 here

The results indicate that for low risk aversion levels the majority of methods exhibit a fairly large amount of portfolio turnover, with only the MEAN and MA methods providing low levels of portfolio turnover. Indeed, only the portfolio turnover associated with the MEAN method is significantly lower than the turnover associated with the MA method. Thus, apart from these methods, the performance of the forecasting methods documented elsewhere in this paper should be tempered by the fact that they are likely to deliver significant transaction costs.

4.5.3 The distribution of performance

To gain an appreciation for the distribution of relative costs over time, plots of these costs (and associated portfolio weights) are provided in Figure 1 for a selection of methods under the assumption that investors trade both futures markets and have quadratic utility (with $\gamma = 4$).

---

[20]These methods perform well in the current context as the loss function upon which they are based (and assessed) is highly asymmetric and more heavily penalises volatility forecasts that are below realised volatility (cf. the symmetric RSS loss function). This will manifest itself in generally higher volatility forecasts associated. Consequently, when volatility increases unexpectedly, costs are more likely to be lower for these methods in comparison to methods based on the RSS loss function. See subsection B.4.2 in Appendix B for details of the utility-based loss function used in this paper.

The use of this value of γ is motivated by the fact that the portfolio weights take reasonable values under this assumption. For instance, the mean portfolio weights associated with the MEAN method are 0.473 (bond futures) and 0.755 (stock futures).

Insert Figure 1 here

It is apparent that imperfection costs associated with the MEAN method are highly variable. This variation is due to the static nature of the forecasts (and hence portfolio weights). The relative costs associated with this method (with respect to the MA method) appear to show two peaks: around 1994/95 and 2008/09 (see panel (h)). The first of these is due to MEAN method users not taking offsetting positions in bond and stock futures when their correlation was high (cf. weights in panels (d) and (e)). The second is due to these users taking too high a position in bond and stock futures during the recent financial crisis period. In both cases it is the failure of the MEAN method to adjust to current conditions that causes the high relative costs. By contrast, the MA and vector-HAR (econ.) both adjust to market conditions, such that the relative costs of these methods is relatively small over the sample (with the latter method almost always less costly).

To explore unconditional distributional features of the performance of all competing methods, we present histograms of the annualised relative costs between the MA method and all the other methods under the assumption that investors trade both futures markets and have quadratic utility (with $\gamma = 4$). These are provided in Figure 2. Under the performance fee interpretation, these costs represent the amount an investor is willing to pay to avoid switching from using each method to the MA method during each period.

Insert Figure 2 here

The plots highlight the merits of using the MA, VARMA (econ.), VARMAX (econ.), and vector-HAR (econ.) methods. For the majority of methods, relative costs appear to be more often negative than positive – implying that investors are less often willing to pay a fee to avoid switching to the MA method. However, there are some exceptions. Most notably, the VARMA (econ.), VARMAX (econ.), and vector-HAR (econ.) methods appear to be the only methods with positive modal performance fees.[21]

---

[21]The RW method delivers a small number of highly negative relative costs not captured by these plots. Hence, the apparent contradiction of the symmetry of the RW plot and the large mean RW imperfection costs found in Table 3.

4.5.4 Predicting future performance

To gain further insight into the conditional nature of performance, we present the results of the test regression given by (13) estimated over the out-of-sample period. Specifically, coefficients (and adjusted R^2 values) associated with regressions of relative costs (for investors who trade both futures markets and have quadratic utility (with $\gamma = 4$)) upon a constant, lagged relative costs, lagged $1/N$ volatility (that is, volatility associated with a portfolio containing equal weights in bond and stock futures), and lagged values of the ADS index of economic activity are presented in Table 5. In addition to relative costs based on unrestricted method comparisons, we also consider relative costs with respect to restricted versions of each method in which all correlation forecasts are set to zero.[22] For all comparisons, we indicate whether the null hypothesis that each coefficient equals zero can be rejected using a HAC-based Wald test.

Insert Table 5 here

The results indicate that relative costs are dependent on each of the independent variables considered (though there is considerable variation across method comparisons). For instance, when comparing the RM1994 and RM2006 methods, 19.4% of the variation in relative costs is due to these variables, with the future relative performance of RM2006 over RM1994 improving as past relative performance increases, and volatility decreases. By contrast, when comparing the MA and vector-HAR (econ.) methods, only 0.6% of the variation in relative costs is explained, with future relative performance unrelated to any of the independent variables. However, it is noticeable that in this instance the intercept is positive and significant implying that the vector-HAR (econ.) method generally exhibits superior performance. The results also indicate that for the majority of methods it is better to avoid imposing the zero-correlation forecast restriction (the exception applies to the MEAN method). However, there is considerable predictability in this relative cost – suggesting potential benefits if one could switch to the restricted method when the predicted relative cost of the restriction is negative in the next period.

The above regression results suggest that for some methods it may be possible to predict relative performance in the next period. This motivates a new method in which one switches

---

[22]As the non-intercept independent variables are demeaned, the intercept term can be interpreted as the mean relative cost net of independent variable effects. Furthermore, the coefficients associated with non-intercept independent variables are standardised such that they can be interpreted as the standard deviation change in the dependent variable when the independent variables change by one standard deviation, holding all other variables constant.

from a poor forecasting method to a good forecasting method each period. To this end (and in the spirit of Giacomini and White, 2006), we consider a method based on the following decision rule:

Step 1: Perform the regression of ΔL_t on $\mathbf{Z}_{t-1}$ over the out-of-sample period, where ΔL_t is the relative cost of a forecasting method with respect to a benchmark method, and $\mathbf{Z}_{t-1}$ contains the independent variables described above. Let $\widehat{\varphi}$ denote the estimated coefficient vector from this regression.

Step 2: Repeat Step 1 until all forecasting methods have been examined. Switch to the forecasting method that delivers the lowest predicted relative cost at each point in time (that is, the lowest value of $\widehat{\varphi}'\mathbf{Z}_{t-1}$).

Given the excellent performance of the vector-HAR (econ.) method in the previous analysis, this method is selected as the benchmark method. Two versions of the above decision rule are considered. This first compares the benchmark method and unrestricted versions of each method, henceforth referred to as the *hybrid method*.[23] The second version attempts to exploit the predictability in switching to zero-correlation forecasts documented in Table 5. To this end, the benchmark method is compared to unrestricted *and* restricted (zero-correlation forecast) versions of each method. This method is henceforth referred to as the *hybrid (zero-corr.) method*. The conditional performance of these and competing methods is formally tested in the next subsection.

4.5.5 Testing conditional performance

The GW conditional and unconditional tests described in subsection 3.2 are performed. Regarding the former test, we also adopt the test rule proposed by Giacomini and White (2006). Specifically, we define method B as superior to method A if the proportion of time the expected (predicted) relative costs (the fitted value from (13)) are positive is below 50% *and* the GW conditional test rejects the null at the 5% level. These proportions (and an indication of which method is superior according to the above rule) are presented in Table 6 for various method comparisons under the assumption that investors trade both futures markets and have quadratic utility (with $\gamma = 4$). Regarding the GW unconditional test (and associated test rule), we report the ratio of the costs of forecast imperfection and provide an indication as to which

---

[23]The performance of this method appears best when the vector-HAR (econ.) method is only compared to the MA and DCC-MGARCH methods. This assumption is maintained in the subsequent analysis.

method is superior (that is, a ratio below unity and a significant GW unconditional test result at the 5% level).

Insert Table 6 here

A clear ranking emerges regarding the quality of each method. In particular, the hybrid method is the best, following closely by the VARMA (econ.), VARMAX (econ.), Vector-HAR (econ.) methods, and then the combination-type and MA methods.[24] Of these, the first four methods exhibit very similar performance and are the clear winners in that they are almost always selected as the superior methods irrespective of the method it is compared to or the test used. While the quality of these methods is unequivocal, for other comparisons the conditional and unconditional tests deliver different method rankings. For instance, the DCC-MGARCH method beats five other methods (that is, it has five successes) when the GW conditional test rule is adopted, but only four when the GW unconditional test rule is used. This difference is ultimately due to the superior power of the GW conditional test derived from the dependence of relative costs on the independent variables.

To gain further insight into the conditional nature of method performance, we report the results of the regime tests described in subsection 3.2 for a selection of the best forecasting methods. Specifically, the results in Table 7 provide the results of the regime (individual) and regime (difference) tests under the assumption that investors trade both futures markets and have quadratic utility (with $\gamma = 4$). The following regimes are considered: low and high volatility regimes (based on $1/N$ volatility), low and high (absolute) bond-stock correlation periods (based on the correlation between bond and stock futures returns), economic contraction and expansion periods (based on NBER dates), and low and high inflation expectation periods (based on Thomson Reuters/University of Michigan inflation expectations).[25] With the exception of the economic contraction/expansion regime, the median value is used to determine the distinction between the low and high regimes.

Insert Table 7 here

The results indicate that the superiority of the hybrid and vector-HAR (econ.) methods is somewhat dependent on the level of volatility. In the low volatility regime, these methods

---

[24]The hybrid and hybrid (zero-corr.) methods exhibit very similar performance. This suggests that the predictability of the zero-correlation restriction documented in Table 5 does not have value in the current context. For this reason the subsequent analysis focuses exclusively on the hybrid method.

[25]The inclusion of the economic activity and inflation expectation regimes is motivated by their documented relationship with asset return correlations (see, e.g., Yang et al, 2009).

dominate the other methods considered. However, in the high volatility regime, they are significantly better than the MA method, but not significantly better than the shrinkage method. For instance, the ratio of the hybrid mean forecast imperfection cost to the shrinkage mean forecast imperfection cost is 0.81 (and significantly different from unity) in the former regime, but 1.04 (and insignificantly different from unity) in the latter regime. Other changes in performance are noteworthy: the MA method appears useful during low volatility regimes (it is superior to the shrinkage method), but less so during high volatility regimes (it is not superior to any of the featured methods), and the superiority of the hybrid and vector-HAR (econ.) methods is maintained, with only minor variation in their dominance over the non-volatility regimes.

4.6 Summarising method performance

To summarise the above results we provide the ranks of the methods for all time periods (based on the GW conditional and unconditional test rules), and within the four types of regimes used previously, under the assumption that investors trade both futures markets and have quadratic utility (with $\gamma = 4$). The ranks are based on the number of successes achieved, where successes are defined as occasions when the method is selected by the test rule. Results are given in Table 8.

Insert Table 8 here

A number of features are apparent from this table. The hybrid and vector-HAR (econ.) methods are dominant though they weaken slightly during high volatility periods. Here combination-type methods are the best – a result consistent with the literature on optimal weight combining in the presence of parameter uncertainty (see, e.g., Tu and Zhou, 2011). For the other methods there is clear variation in the ranks across the regimes. For instance, the MA method is particularly strong (3rd best) during low volatility periods, but only 7th best during high volatility periods. By contrast, the DCC-MGARCH method is the 9th and 3th best method during these regimes, respectively. Results of this type imply that variation in the conclusions of previous studies may be due to use of different sample periods in which certain events (such as high volatility episodes) are more prominent. Finally, there appears to be less variation in the ranks when other regimes are considered.

4.7 Robustness checks

The robustness of the above results is examined with respect to variation in the assumptions of the investment framework: viz. the mean return and portfolio weight assumptions.

4.7.1 Asset return assumptions

To examine the robustness of the superiority of the vector-HAR (econ.) method to the investment framework assumptions, the mean relative costs (and associated HAC standard errors) are calculated for investors who trade both futures markets, have quadratic utility (with $\gamma = 4$), and $\boldsymbol{\mu} - R^f \in (-0.08, 0.08)$ (annualised). The contour plots in Figure 3 present the relative costs of the MA and combination methods with respect to the vector-HAR (econ.) method.

Insert Figure 3 here

The contour plots exhibit fairly regular non-linearities, with relative costs increasing as the absolute value of the risk premia increases (unsurprising given the presence of $(\boldsymbol{\mu}' - R^f)^2$ in (B.11a) in Corollary 1). The plots also indicate that mean relative costs from 20 to 40 basis points exist for realistic risk premia levels. Moreover, these costs are likely to be significantly different from zero (in a statistical sense), as the mean relative costs are typically five to ten times larger than the associated standard errors.

4.7.2 Portfolio weight assumptions

As a second check of the robustness of the results to variation in the investment environment, we consider the ranks of the methods in the absence of a risk-free asset (with and without short-sale restrictions). The results in Table 8 indicate that the hybrid and vector-HAR (econ.) methods remain dominant. However, the ranking of the other methods does change. Most notably, there is a clear improvement in the relative performance of the RW method. For instance, under the GW conditional test rule with no portfolio restrictions, it is the worst method. However, it is the 7th best method (out of 12 methods) under the no risk-free asset restriction (without short-sale restrictions), and 2nd best (jointly) if we go further and forbid short-selling.

The success of the RW method under portfolio restrictions can be explained as follows. The nature of the RW method means that the forecasts produced are highly variable. In

the absence of portfolio weight restrictions this means that extreme positions in the assets are required (as evinced by the portfolio turnover in Table 4) leading, in turn, to potentially outlying performance. By contrast, when restrictions are imposed the effect of these positions is mitigated. At the limit, when there is no risk-free asset and no short-selling, portfolio weights must lie between zero and unity (and must sum to unity); consequently, the extreme RW method positions are truncated and there adverse effects limited to reveal the quality of this method.

5 Concluding remarks

A conditional assessment of the economic value of volatility prediction is provided within a dynamic investment environment. The evidence presented suggests that it is worth investing in knowledge of volatility dynamics – though this result is highly temporal and depends on investor risk preferences and economic conditions (including volatility levels); see Henkel et al (2011) for similar conditional evidence within the asset return predictability literature.[26] Of the methods considered there is overwhelming support for use of forecasts based on hybrid-type methods and those in which the underling model parameters are estimated by minimising the economic criterion considered in this paper; however, simple moving average and random walk-type methods also have value depending on the nature of the sample of data used and the investment environment adopted.

It is possible to augment the investment environment assumed in this paper in a number of ways. First, investors with Epstein-Zin preferences could be considered. However, the recursive nature of the utility function would mean that deriving expressions for the associated forecasting performance loss function would present a significant challenge that would undoubtedly require use of simplifying assumptions within a numerical optimisation framework. Second, transaction costs could be incorporated by changing the objective functions given by (B.7a) and (B.7d), such that portfolio returns that appear in these equations are adjusted to reflect the cost of changing portfolio weights. Finally, the analysis assumes that only second moments are predictable. This could be relaxed by dropping the Gaussian assumption and allowing for time variation in higher moments (see, e.g., Brooks et al, 2005, for evidence of such time variation). This would then result in a new set of approximations in Proposition 1; specifically, (B.8) in the proof of Proposition 1 would replace (B.7b), such that higher conditional moments are considered. Implementation of these improvements in left for future work.

---

[26]Henkel et al (2011) do not consider the economic value of asset return predictability. Rather, they examine the adjusted R^2 statistics associated with predictor equations within economic contraction and expansion periods.

References

Andersen T, Bollerslev T, Diebold F, Labys P. 2003. Modeling and forecasting realized volatility. *Econometrica* **71**: 579–625.

Andreou E, Kourouyiannis C, Kourtellos A. 2012. Volatility forecast combinations using asymmetric loss functions. University of Cyprus Discussion Paper 07-2012.

Aruoba B, Diebold F, Scotti C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* **27**: 417–427.

Barndorff-Nielsen O, Shephard N. 2004. Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica* **72**: 885–925.

Brooks C, Burke S, Heravi S, Persand G. 2005. Autoregressive conditional kurtosis. *Journal of Financial Econometrics* **3**: 399–421.

Cappiello L, Engle R, Sheppard K. 2006. Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics* **4**: 537–572.

Chiriac R, Voev V. 2011. Modelling and forecasting multivariate realised volatility. *Journal of Applied Econometrics* **26**: 922–947.

Clements A, Doolan M, Hurn S, and Becker R. 2009. Evaluating multivariate volatility forecasts. NCER Working Paper.

Corsi F. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* **7**: 174–196.

Diebold F, Mariano R. 1995. Comparing predictive ability. *Journal of Business and Economic Statistics* **13**: 253–265.

Engle R. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* **20**: 339-350.

Engle R, Rangel J. 2008. The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies* **21**: 1187–1222.

Fleming J, Kirby C, Ostdiek B. 2001. The economic value of volatility timing. *Journal of Finance* **56**: 329–352.

———. 2003. The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics* **67**: 473–509.

Giacomini R, White H. 2006. Tests of conditional predictive ability. *Econometrica* **74**: 1545–1578.

González-Rivera G, Lee T-H, Mishra S. 2004. Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting* **20**: 629–645.

Hansen P, Lunde A. 2006. Consistent ranking of volatility models. *Journal of Econometrics* **131**: 97–121.

Henkel S, Martin S, Nardari F. 2011. Time-varying short-horizon predictability, *Journal of Financial Economics*, **99**: 560–580.

Hlawitschka W. 1994. The empirical nature of Taylor-series approximations to expected utility. *American Economic Review* **84**: 713–719.

Jorion P. 1995. Predicting volatility in the foreign exchange market. *Journal of Finance* **50**: 507–528.

Kirby C, Ostdiek B. 2011. Optimal active portfolio management with unconditional mean-variance risk preferences. Available at SSRN: http://ssrn.com/abstract=1821284.

Laurent S, Rombouts J, Violante F. 2009. Consistent ranking of multivariate volatility models. CORE Working Paper.

Levy H, Markowitz H. 1979. Approximating expected utility by a function of mean and variance. *American Economic Review* **69**: 308–317.

Liu Q. 2009. On portfolio optimization: How and when do we benefit from high-frequency data? *Journal of Applied Econometrics* **24**: 560–582.

Marquering W, Verbeek M. 2004. The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis* **39**: 407–429.

Merton R. 1980. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* **8**: 323–361.

Patton A. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* **160**: 246–256.

Poon S, Granger C. 2003. Forecasting volatility in financial markets. *Journal of Economics Literature* **41**: 478–539.

Rapach D, Strauss J, Zhou G. 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* **23**: 821–862.

Şerban M, Brockwell A, Lehoczky J, Srivastava S. 2007. Modelling the dynamic dependence structure in multivariate financial time series. *Journal of Time Series Analysis* **28**: 763–782.

Stock J, Watson M. 2003. How did leading indicator forecasts perform during the 2001 recession? *Federal Reserve Bank of Richmond, Economic Quarterly* **89**: 71–90.

Timmermann A. 2006. Forecast Combinations, in *Handbook of Economic Forecasting*, in Elliott G, Granger C, and Timmermann A (eds.). Elsevier.

Tu J, Zhou G. 2011. Markowitz meets Talmud: A combination of sophisticated and naïve diversification strategies. *Journal of Financial Economics* **99**: 204–215.

Voev V. 2009. On the economic evaluation of volatility forecasts, CREATES Working Paper.

West K. 1996. Asymptotic inference about predictive ability, *Econometrica* **64**: 1067–1084.

West K, Edison H, Cho D. 1993. A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics* **35**: 23–45.

Yang J, Zhou Y, Wang Z. 2009. The stock-bond correlation and macroeconomic conditions: One and a half centuries of evidence. *Journal of Banking and Finance* **33**: 670–680.

Zhang L, Mykland P, and Aït-Sahalia Y. 2005. A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of The American Statistical Association* **100**: 1394–1411.

Zumbach G. 2007. The RiskMetrics methodology. Risk Technical Documents, RiskMetrics Group.

Appendices

A Realised covariance matrix measures

As we allow for simultaneous positions in more than one market in the analysis, we require realised covariance matrix measures. Prior to defining these measures, some preliminary notation is required: let day t have unit length, and let the full grid of all observation points be given by $G = \{t_0, \ldots, t_m\}$. Given this notation, the following realised covariance matrix estimator can be defined:

$$\mathbf{Y}_t^{(all)} = \sum_{t_i \in G} (\mathbf{P}_{t_{i,+}} - \mathbf{P}_{t_i})(\mathbf{P}_{t_{i,+}} - \mathbf{P}_{t_i})', \qquad (A.1)$$

where $\mathbf{P}_{t_i}$ is an $(N \times 1)$ vector of log prices observed during the ith intraday period of day t, and $t_{i,+}$ represents the next observation after t_i on G. We use 1-second frequency data to construct this particular realised covariance measure. While the asymptotic distribution theory for the realised covariance estimator in (A.1) has been established (Barndorff-Nielsen and Shephard, 2004), it is known to be adversely affected by the microstructure noise observed at high frequencies.

The non-robustness of the above measure has spawned a number of competing measures. We follow Zhang et al (2005) and apply their subsampling methodology within a multivariate context. In particular, we first partition the grid G into J nonoverlapping subgrids $G^{(j)}$, such that,

$$G = \bigcup_{j=1}^{J} G^{(j)}, \qquad (A.2)$$

where $G^{(j)} \cap G^{(k)} = \emptyset$ when $j \neq k$, and the jth subgrid $G^{(j)}$ contains s_j equally spaced observations such that $G^{(j)} = \{t_{j-1}, t_{j-1+J}, t_{j-1+2J}, \ldots, t_{j-1+s_j J}\}$. Based on these subgrids we can construct the following (sparse) realised covariance matrix estimator for each grid:

$$\mathbf{Y}_t^{(j)} = \sum_{t_i \in G^{(j)}} (\mathbf{P}_{t_{i,+}} - \mathbf{P}_{t_i})(\mathbf{P}_{t_{i,+}} - \mathbf{P}_{t_i})', \qquad (A.3)$$

where if $t_i \in G^{(j)}$, then $t_{i,+}$ is the following element in $G^{(j)}$. We follow the extant literature and construct each subgrid realised covariance measure using five-minute frequency data ($K = 300$).

To construct a single realised covariance measure based on the J measures in (A.3), the average is taken over all subgrid measures to give the following realised covariance matrix measure:

$$\mathbf{Y}_t^{(ave)} = \frac{1}{J} \sum_{j=1}^{J} \mathbf{Y}_t^{(j)}. \qquad (A.4)$$

25

Within a univariate context, Zhang et al (2005) demonstrate that the asymptotic bias and variance associated with the estimator given by (A.4) are both smaller than those associated with the measure given by (A.1). However, it remains biased. To correct for this, we consider a multivariate version of their bias-corrected measure:

$$\mathbf{Y}_t = \mathbf{Y}_t^{(ave)} - \frac{\overline{m}}{m}\mathbf{Y}_t^{(all)}, \tag{A.5}$$

where $\overline{m} = (m - J + 1)/J$ is the (average) number of $G^{(j)}$ subgrid elements. This measure is henceforth referred to as the *averaged subsample realised covariance matrix measure*. Given its demonstrated superiority over competing measures (see the Monte Carlo simulation evidence in Table 2 of Zhang et al, 2005), only this measure is considered in the analysis.

B   Economic measures of forecasting method quality

Methods for calculating the economic value of competing forecasting methods are established in two stages: the assumptions of the investors are stated, and measures of the comparative performance of two competing investor strategies are derived.

B.1   The investment framework

Assumption 1. A myopic investor selects an optimal portfolio in order to maximise the conditional expectation of next period utility given by

$$\mathrm{E}(U_t|\mathcal{M}(\mathcal{F}_{t-1})) = \mathrm{E}(f(R_t^p)|\mathcal{M}(\mathcal{F}_{t-1})), \tag{B.1}$$

where $\mathcal{M}(\mathcal{F}_{t-1})$ is the method used to generate forecasts of utility, $\mathcal{F}_{t-1}$ is chosen as the σ-field generated by a set of conditioning instruments available at time $t-1$, R_t^p is the gross return to the investor's risky asset portfolio, and the function f is a monotonic, quasi-concave, and globally non-satiated utility function dependent on R_t^p such that f: $[0, +\infty) \to \mathbb{R}$.

Assumption 2. The risky asset portfolio consists of N risky assets, with returns to this portfolio given by

$$R_t^p = R^f + \mathbf{w}_t'(\mathbf{R}_t - R^f), \tag{B.2}$$

where $\mathbf{R}_t$ is an $(N \times 1)$ vector of gross returns to the risky assets, R^f is the risk-free rate, and $\mathbf{w}_t$ is an $(N \times 1)$ vector containing the fractions of wealth allocated to the risky assets constructed using the set of conditional moments of these returns based on a forecasting method (with $1 - \mathbf{1}'\mathbf{w}_t$ allocated to the risk-free asset).

Assumption 3. Risky asset returns evolve as follows:

$$\mathbf{R}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t, \tag{B.3}$$

where $\boldsymbol{\mu}$ is an $(N \times 1)$ vector of risky asset return means, $\boldsymbol{\epsilon}_t = \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{\nu}_t$ is an $(N \times 1)$ vector of errors, $\boldsymbol{\Sigma}_t^{1/2}$ is the (unique) square root of an $(N \times N)$ covariance matrix $\boldsymbol{\Sigma}_t$, and $\boldsymbol{\nu}_t \sim \mathrm{IN}(0, 1)$.

Remark. The covariance matrix $\boldsymbol{\Sigma}_t$ is given by the averaged subsample realised covariance matrix measure $\mathbf{Y}_t$ described in Appendix A.

Assumption 4. Two types of (imperfect) investor exist (referred to as *investor A* and *investor*

B), each with a different forecast of the next period return covariance matrix.[27] These investors can be represented in terms of the following conditional distributions:

$$\mathbf{R}_t | \mathcal{M}(\mathcal{F}_{i,t-1}) \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{i,t|t-1}), \tag{B.4}$$

where $\boldsymbol{\mu}$ is the expected return vector, $\boldsymbol{\Sigma}_{i,t|t-1}$ is the expected covariance matrix of the i-type investor generated by their forecasting method ($\mathcal{M}(\mathcal{F}_{i,t-1})$).

Given the above assumptions, we wish to assign an interpretable economic value to the expectation of the difference in utility levels associated with the two investor strategies. The extant literature often considers the performance fee required to force the unconditional expectation of utility differences to zero (see, e.g., Fleming et al, 2001, 2003). By contrast, we consider the conditional expectation of these differences in order to provide a time-specific measure of economic value. We penalise use of incorrect conditional expectations by conditioning on the current value of the realised covariance matrix (in addition to the portfolio weights based on the expected covariance matrix). Thus, we have the following definitions:

Definition 1. The economic cost of a volatility forecasting method is given by the following conditional expectation:

$$L(\boldsymbol{\Sigma}_t, \boldsymbol{\Sigma}_{i,t|t-1}) = \mathrm{E}(f(R_t^p)|\mathbf{w}_{i,t}^*(\boldsymbol{\Sigma}_t), \boldsymbol{\Sigma}_t) - \mathrm{E}(f(R_t^p)|\mathbf{w}_{i,t}^*(\boldsymbol{\Sigma}_{i,t|t-1}), \boldsymbol{\Sigma}_t), \tag{B.5a}$$

where $\mathbf{w}_{i,t}^*(\boldsymbol{\Sigma}_t)$ is the perfect foresight vector of optimal fractions of wealth allocated to the risky assets, and $\mathbf{w}_{i,t}^*(\boldsymbol{\Sigma}_{i,t|t-1})$ is the i-type investor allocation constructed using $\boldsymbol{\Sigma}_{i,t|t-1}$ (based on $\mathcal{M}(\mathcal{F}_{i,t-1})$).

Remark. The first term on the RHS of (B.5a) measures the highest available conditional expectation, while the second expression gives the conditional expectation achieved using $\boldsymbol{\Sigma}_{i,t|t-1}$. Conditioning on realised covariance ensures that the above cost is non-negative, is negatively related to volatility forecast accuracy, and equals zero for perfect-foresight volatility forecasts.

Definition 2. The economic cost of volatility forecasting method B with respect to forecasting method A is given by the difference in conditional expectations:

$$\begin{aligned} \Delta L_t &\equiv L(\boldsymbol{\Sigma}_t, \boldsymbol{\Sigma}_{b,t|t-1}) - L(\boldsymbol{\Sigma}_t, \boldsymbol{\Sigma}_{a,t|t-1}) \\ &= \mathrm{E}(f(R_t^p)|\underbrace{\mathbf{w}_{a,t}^*(\boldsymbol{\Sigma}_{a,t|t-1}), \boldsymbol{\Sigma}_t}_{\mathcal{G}_{a,t}}) - \mathrm{E}(f(R_t^p)|\underbrace{\mathbf{w}_{b,t}^*(\boldsymbol{\Sigma}_{b,t|t-1}), \boldsymbol{\Sigma}_t}_{\mathcal{G}_{b,t}}), \end{aligned} \tag{B.5b}$$

where $\mathcal{F}_{i,t-1} \subset \mathcal{G}_{i,t}$.

Definition 3. The (maximum) performance fee δ_t^* investor A is willing to pay each period to avoid use of investor B's volatility forecasts is given by the root of the following equation:

$$\mathrm{E}(f(R_t^p - \delta_t)|\mathcal{G}_{a,t}) - \mathrm{E}(f(R_t^p)|\mathcal{G}_{b,t}) = 0, \tag{B.5c}$$

where $\delta_t \in \mathbb{R}$.

B.2 A solution under HARA utility

Expressions for the above loss functions are available under a general class of utility functions.

---

[27] As it is widely accepted that accurate forecasts of expected returns are particularly difficult to obtain (Merton, 1980), we surrender all knowledge of expected returns to the most basic level. In particular, we assume that expected (and mean) returns to the risky assets are constant over time.

Assumption 5. The i-type investor employs the hyperbolic absolute risk aversion (HARA) class of utility functions given by

$$f_h(R_t^p) = \frac{\gamma_1}{1 - \gamma_1} \left(\frac{\gamma_2}{\gamma_1} R_t^p + \gamma_3 \right)^{1-\gamma_1}, \tag{B.6}$$

where $\gamma_2 > 0$ and $\gamma_2 R_t^p / \gamma_1 + \gamma_3 > 0$.

Remark. This class of utility functions contains a number of important special cases: isoelastic (power) utility ($\gamma_1 > 0$, $\gamma_2 = \gamma_1$, $\gamma_3 = 0$, with the Arrow-Pratt coefficient of relative risk aversion given by γ_1), logarithmic utility ($\gamma_1 = 1$, $\gamma_2 = 1$, $\gamma_3 = 0$), and negative exponential utility ($\gamma_1 = \infty$, $\gamma_2 > 0$, $\gamma_3 = 1$, with the Arrow-Pratt coefficient of constant risk aversion given by γ_2).

Closed-form expressions are unlikely to be available given the complex nature of the conditional expectations involved. Therefore, we propose the following numerical procedure.

Proposition 1. *Under assumptions 1 to 5, and Definition 2, the economic cost of forecast differences to investor B (with respect to investor A) is given by the following non-linear equation:*

$$\Delta L_{h,t} = \mathrm{E}(f_h(R_t^p)|\mathcal{G}_{a,t}) - \mathrm{E}(f_h(R_t^p)|\mathcal{G}_{b,t}), \tag{B.7a}$$

where

$$\mathrm{E}(f_h(x)|\mathcal{G}_{i,t}) \approx f_h(\mathrm{E}(x|\mathcal{G}_{i,t})) + \sum_{k=1}^{K} \frac{1}{2k!} f_h^{(2k)}(\mathrm{E}(x|\mathcal{G}_{i,t})) \, \mathrm{var}(x|\mathcal{G}_{i,t})^k \prod_{j=1}^{k} (2j-1), \tag{B.7b}$$

and $f_h^{(2k)}(\mathrm{E}(x|\mathcal{G}_{i,t}))$ is the $2k$th derivative of $f_h(x)$ with respect to x evaluated at $\mathrm{E}(x|\mathcal{G}_{i,t})$ such that

$$f_h^{(2k)}(\mathrm{E}(x|\mathcal{G}_{i,t})) = \left(\frac{\gamma_2}{\gamma_1} \right)^{2k} \left(\frac{\gamma_1}{1 - \gamma_1} \right) \left(\frac{\gamma_2}{\gamma_1} \mathrm{E}(x|\mathcal{G}_{i,t}) + \gamma_3 \right)^{1-\gamma_1-2k} \prod_{j=0}^{2k-1} (1 - \gamma_1 - j), \tag{B.7c}$$

with x denoting either R_t^p or $R_t^p - \delta_t$. In turn, the optimal HARA portfolio weights embedded in $\mathcal{G}_{i,t}$ are given by

$$\mathbf{w}_{i,t|t-1}^* = \operatorname*{argmin}_{\mathbf{w}_{i,t|t-1} \in (w_l, w_u)} -\mathrm{E}(f_h(R_t^p)|\mathcal{F}_{i,t-1}), \tag{B.7d}$$

and w_l and w_u are the user-specified lower and upper bounds on portfolio weights, respectively. The solutions to (B.7d) is obtained by numerical methods.

Proof. Taking the conditional expectation of a Taylor series expansion of the HARA utility function in (B.6) about the conditional mean of x we obtain

$$\mathrm{E}(f_h(x)|\mathcal{G}_{i,t}) = \sum_{k=0}^{\infty} \frac{1}{k!} f_h^{(k)}(\mathrm{E}(x|\mathcal{G}_{i,t})) \, \mathrm{E}((x - \mathrm{E}(x|\mathcal{G}_{i,t}))^k|\mathcal{G}_{i,t}). \tag{B.8}$$

Truncating to a $2K$th-order series expansion, and invoking the conditional Gaussian assumption (see Assumption 4), leads to (B.7b) such that it involves nonlinear transformations of the first two conditional moments only.[28] □

---

[28] We use the BFGS optimisation method with STEPBT line search programmed via the Constrained Optimization application in GAUSS 11.0 (64-bit version) as our numerical method for portfolio weights under HARA

B.3 A closed-form expression under quadratic utility

If we make the following quadratic (mean-variance) utility assumption, then a closed-form expression is available.

Assumption 6. The *i*-type investor maximises the conditional expectation of quadratic utility given by

$$\mathrm{E}(f_q(R_t^p)|\mathcal{G}_{i,t}) = \mathrm{E}(R_t^p|\mathcal{G}_{i,t}) - \frac{\gamma}{2}\mathrm{var}(R_t^p|\mathcal{G}_{i,t}), \tag{B.9}$$

where γ is the coefficient of relative risk aversion.

Proposition 2. *Under Assumptions 1 to 4, and 6, and Definition 2, the economic cost of forecast differences to investor B (with respect to investor A) is given by*

$$\gamma\Delta L_{q,t} = (\boldsymbol{\omega}_{a,t|t-1} - \boldsymbol{\omega}_{b,t|t-1})'(\boldsymbol{\mu} - R^f) - \frac{1}{2}\mathrm{tr}\left((\boldsymbol{\omega}_{a,t|t-1} - \boldsymbol{\omega}_{b,t|t-1})(\boldsymbol{\omega}_{a,t|t-1} + \boldsymbol{\omega}_{b,t|t-1})'\boldsymbol{\Sigma}_t\right), \tag{B.10a}$$

where

$$\boldsymbol{\omega}_{i,t|t-1} = \boldsymbol{\Sigma}_{i,t|t-1}^{-1}(\boldsymbol{\mu} - R^f), \tag{B.10b}$$

and $\mathrm{tr}(\mathbf{X})$ *is the trace of the* $\mathbf{X}$.

Proof. Replacing the conditional expectations in (B.5b) with those in (B.9), using the expressions in (B.2), (B.3), and (B.4), and rearranging, eventually leads to (B.10a) and (B.10b). □

It is possible to go further and express the above values in terms of a familiar statistical loss function via the following additional assumption.

Assumption 7. Investor A possesses perfect-foresight covariance expectations (referred to as the *perfect investor*), while investor B remains imperfect in this regard (referred to as the *imperfect investor*).

Corollary 1. *Under Assumptions 1 to 4, 6 and 7, and Definition 1, the cost of forecast imperfection* $(L_q(\boldsymbol{\Sigma}_t, \boldsymbol{\Sigma}_{t|t-1}) \in \mathbb{R}_+)$ *can be decomposed as follows:*

$$\gamma L_q(\boldsymbol{\Sigma}_t, \boldsymbol{\Sigma}_{t|t-1}) = \frac{1}{2}\left(\mathbf{1}'((\boldsymbol{\Sigma}_t^d)^{1/2}(\boldsymbol{\Sigma}_{t|t-1}^d)^{-1} - (\boldsymbol{\Sigma}_t^d)^{-1})^2(\boldsymbol{\mu} - R^f)^2\right) + C_t, \tag{B.11a}$$

where $\boldsymbol{\Sigma}_t^d = \mathrm{dg}(\boldsymbol{\Sigma}_t)$, $\boldsymbol{\Sigma}_{t|t-1}^d = \mathrm{dg}(\boldsymbol{\Sigma}_{t|t-1})$, $\boldsymbol{\Sigma}_t$ $(\boldsymbol{\Sigma}_{t|t-1})$ *is the expected covariance matrix of the perfect (imperfect) investor, and* $\mathrm{dg}(\mathbf{X})$ *is a diagonal matrix containing the diagonal elements of* $\mathbf{X}$. *The term* $C_t \in \mathbb{R}$ *represents the sum of all covariance effects such that it equals zero in the presence of zero-correlated asset returns (and zero-covariance forecasts), and is given by*

$$C_t = \mathbf{A}_t(\boldsymbol{\mu} - R^f) - \frac{1}{2}\left(\sum_{j=1}^{4}\mathrm{tr}(\mathbf{B}_{j,t}\boldsymbol{\Sigma}_t^c) + \sum_{j=2}^{4}\mathrm{tr}(\mathbf{B}_{j,t}\boldsymbol{\Sigma}_t^d)\right), \tag{B.11b}$$

where $\mathbf{A}_t = (\boldsymbol{\omega}_t^c - \boldsymbol{\omega}_{t|t-1}^c)'$, $\mathbf{B}_{1,t} = (\boldsymbol{\omega}_t^d - \boldsymbol{\omega}_{t|t-1}^d)(\boldsymbol{\omega}_t^d + \boldsymbol{\omega}_{t|t-1}^d)'$, $\mathbf{B}_{2,t} = (\boldsymbol{\omega}_t^c - \boldsymbol{\omega}_{t|t-1}^c)(\boldsymbol{\omega}_t^d + \boldsymbol{\omega}_{t|t-1}^d)'$, $\mathbf{B}_{3,t} = (\boldsymbol{\omega}_t^d - \boldsymbol{\omega}_{t|t-1}^d)(\boldsymbol{\omega}_t^c + \boldsymbol{\omega}_{t|t-1}^c)'$, $\mathbf{B}_{4,t} = (\boldsymbol{\omega}_t^c - \boldsymbol{\omega}_{t|t-1}^c)(\boldsymbol{\omega}_t^c + \boldsymbol{\omega}_{t|t-1}^c)'$, $\boldsymbol{\omega}_t^c = -(\boldsymbol{\Sigma}_t^d)^{-1}(\mathbf{I}_N +$

---

utility with a 10th-order Taylor approximation imposed and no bounds placed on the admissible weight values. The choice of K has been investigated within an unconditional context. Levy and Markowitz (1979) demonstrate the quality of second-order Taylor series approximations of standard utility functions. However, higher order approximations may be more appropriate particularly when the series converge, though this choice is ultimately an empirical issue (Hlawitschka, 1994).

$(\boldsymbol{\Sigma}_t^c)(\boldsymbol{\Sigma}_t^d)^{-1})^{-1}(\boldsymbol{\Sigma}_t^c)(\boldsymbol{\Sigma}_t^d)^{-1}(\boldsymbol{\mu}-R^f)$, $\boldsymbol{\omega}_t^d = (\boldsymbol{\Sigma}_t^d)^{-1}(\boldsymbol{\mu}-R^f)$, $\boldsymbol{\omega}_{t|t-1}^c = -(\boldsymbol{\Sigma}_{t|t-1}^d)^{-1}(\mathbf{I}_N + (\boldsymbol{\Sigma}_{t|t-1}^c)(\boldsymbol{\Sigma}_{t|t-1}^d)^{-1})^{-1}$ $(\boldsymbol{\Sigma}_{t|t-1}^c)(\boldsymbol{\Sigma}_{t|t-1}^d)^{-1}(\boldsymbol{\mu}-R^f)$, $\boldsymbol{\omega}_{t|t-1}^d = (\boldsymbol{\Sigma}_{t|t-1}^d)^{-1}(\boldsymbol{\mu}-R^f)$, $\boldsymbol{\Sigma}_t^c = \boldsymbol{\Sigma}_t - \boldsymbol{\Sigma}_t^d$, $\boldsymbol{\Sigma}_{t|t-1}^c = \boldsymbol{\Sigma}_{t|t-1} - \boldsymbol{\Sigma}_{t|t-1}^d$, and $\mathbf{I}_N$ is an N-dimension identity matrix.

Proof. The result follows by substituting (B.10b) into (B.10a), decomposing $\boldsymbol{\omega}_t$ and $\boldsymbol{\omega}_{i,t|t-1}$ into variance and covariance effects, and rearranging. The statement that C_t equals zero for zero-correlated asset returns follows from the fact that $\mathbf{A}_t$, $\mathbf{B}_{2,t}$, $\mathbf{B}_{3,t}$, $\mathbf{B}_{4,t}$, and $\mathrm{tr}(\mathbf{B}_{j,t}\boldsymbol{\Sigma}_t^c)$ in (B.11b) all equal zero under such conditions. $\square$

B.4 Relations to existing measures of performance

The results established in this appendix are related to previously considered measures of performance as described below.

B.4.1 Performance fees

The relationship between relative costs and performance fees is formalised in the following proposition.

Proposition 3. *Under Assumptions 1 to 4, and 6, and Definition 3, the economic cost of forecast differences to investor B (with respect to investor A) is equal to the (maximum) performance fee investor A is willing to pay each period to avoid use of investor B's volatility forecasts, that is, $\Delta L_{q,t} = \delta_{q,t}^*$.*

Proof. Under (B.9), we have

$$\mathrm{E}(R_t^p - \delta_{q,t}|\mathcal{G}_{a,t}) - \frac{\gamma}{2}\mathrm{var}(R_t^p - \delta_{q,t}|\mathcal{G}_{a,t}) - \left(\mathrm{E}(R_t^p|\mathcal{G}_{b,t}) - \frac{\gamma}{2}\mathrm{var}(R_t^p|\mathcal{G}_{b,t})\right) = 0. \qquad \text{(B.12a)}$$

Solving for $\delta_{q,t}$ gives

$$\delta_{q,t}^* = \underbrace{\mathrm{E}(R_t^p|\mathcal{G}_{a,t}) - \frac{\gamma}{2}\mathrm{var}(R_t^p|\mathcal{G}_{a,t})}_{E(f_q(R_t^p)|\mathcal{G}_{a,t})} - \underbrace{\left(\mathrm{E}(R_t^p|\mathcal{G}_{b,t}) - \frac{\gamma}{2}\mathrm{var}(R_t^p|\mathcal{G}_{b,t})\right)}_{E(f_q(R_t^p)|\mathcal{G}_{b,t})}, \qquad \text{(B.12b)}$$

which, in turn, equals $\Delta L_{q,t}$. $\square$

Performance fees have been considered previously; see, e.g., Fleming et al (2001, 2003). These are traditionally defined in terms of the unconditional expectation of utility differences, and deliver a single measure for the entire sample period. By contrast, we consider the conditional expectation of these differences such that a time-specific performance measure can be obtained.

B.4.2 Loss functions

Imperfection costs (albeit under a restricted investment strategy) can be considered in terms of an existing loss function as formalised in the following proposition.

Proposition 4. *For a single asset portfolio (or for each risky asset within a zero-correlated portfolio), the loss function in Corollary 1 is robust to noise in the realised variance measure and is a member of the Homogeneous Robust (HR) loss function family proposed by Patton (2011) with a degree of homogeneity equal to -1.*

Proof. For a single asset portfolio (or for each risky asset within a zero-correlated portfolio), the loss function in Corollary 1 simplifies to

$$L_q(\sigma_t, h_t) = \left(\frac{\sigma_t^2 - h_t}{\sigma_t h_t} \right)^2, \tag{B.13}$$

where σ_t^2 (h_t) is the volatility forecast made by the perfect (imperfect) investor, and L_q: $\mathbb{R}_+^2 \to \mathbb{R}_+$. The HR loss function proposed by Patton (2011) is given by

$$L_r(\sigma_t, h_t; b) = \begin{cases} \frac{1}{(b+1)(b+2)}(\sigma_t^{2b+4} - h_t^{b+2}) - \frac{1}{b+1}h_t^{b+1}(\sigma_t^2 - h_t), & \text{if } b \notin \{-1, -2\}, \\ h_t - \sigma_t^2 + \sigma_t^2 \ln \frac{\sigma_t^2}{h_t}, & \text{if } b = -1, \\ \frac{\sigma_t^2}{h_t} - \ln \frac{\sigma_t^2}{h_t} - 1, & \text{if } b = -2, \end{cases} \tag{B.14}$$

where b governs the shape of the loss function ($b < 0$ indicates heavier penalisation of under-prediction, $b = 0$ corresponds to the MSE loss function, and $b = -2$ corresponds to the QLIKE loss function, with degree of homogeneity given by $b + 2$), and L_r: $\mathbb{R}_+^2 \to \mathbb{R}_+$. If we set $b = -3$ in (B.14) we obtain

$$L_r(\sigma_t, h_t; -3) = \frac{1}{2} \left(\frac{\sigma_t^2 - h_t}{\sigma_t h_t} \right)^2. \tag{B.15}$$

Thus, as $L_q(\sigma_t, h_t) \propto L_r(\sigma_t, h_t; -3)$, it follows that the loss function in Corollary 1 is robust with degree of homogeneity equal to -1. □

The above loss function is asymmetric with respect to forecast errors ($\sigma_t^2 - h_t$). Under-prediction of volatility ($\sigma_t^2 > h_t$) is more heavily penalised than over-prediction ($\sigma_t^2 < h_t$); cf. the MSE function – a result also theoretically established by West et al (1993) within a utility framework. The reasons for this asymmetry stem from the fact that it is based on investors maximising per period conditional utility in which under-prediction (hence an excessively high risky asset portfolio weight) is more heavily penalised in the utility function. In this sense, the proposed loss function provides a formal rationale for the commonly advised use of asymmetric loss functions when using financial data.

Table 1 – Summary statistics

Volatility Measure	Statistic			
	Mean	SD	$\widehat{\rho}_1$	$\widehat{\rho}_5$
Panel A: Full Sample (1983-2011)				
Realised Variance (Bond)	7.137	2.823	0.578	0.493
Realised Variance (Stock)	11.861	7.497	0.741	0.571
Realised Correlation	0.034	0.405	0.894	0.846
Panel B: Sub-sample 1 (1983-1997)				
Realised Variance (Bond)	7.277	2.832	0.623	0.506
Realised Variance (Stock)	10.279	6.614	0.634	0.379
Realised Correlation	0.347	0.211	0.646	0.495
Panel C: Sub-sample 2 (1998-2011)				
Realised Variance (Bond)	6.985	2.807	0.527	0.476
Realised Variance (Stock)	13.567	8.003	0.798	0.674
Realised Correlation	-0.302	0.273	0.743	0.618

Notes: This table contains summary statistics (including first and fifth-order autocorrelations, $\widehat{\rho}_1$ and $\widehat{\rho}_5$, respectively) associated with realised variance in the 30-year Treasury bond futures market, realised variance in the S&P 500 index futures market, and realised correlation associated with these markets. The realised variance and covariance measures are constructed using the averaged subsample approach of Zhang et al (2005). The variance measures are given in square root form and are in annualised percentage terms. The sample period used is January 1, 1983, to December 31, 2011.

Table 2 – Statistical measures of forecasting method quality

Method	Measure	$b = 0$	$b = -1$	$b = -2$	$b = -3$
		Loss Function Parameter			
RW	Realised Variance (Bond)	0.071	0.056	0.238	1.947
MEAN		0.046	0.071	0.281	1.515
MA		0.040	0.042	0.153	0.919
VARMA (stat.)		0.041	0.042	0.149	0.894*
VARMAX (stat.)		0.041	0.042	0.149*	0.895*
Vector-HAR (stat.)		0.042	0.041	0.146**	0.876**
VARMA (econ.)		0.041	0.040*	0.143**	0.865**
VARMAX (econ.)		0.041	0.040*	0.144**	0.866**
Vector-HAR (econ.)		0.041	0.040*	0.144**	0.868**
RM1994		0.040	0.045	0.172	1.100
RM2006		0.040	0.043	0.162	0.992
EXP		0.040	0.044	0.172	1.116
DCC-MGARCH		0.039	0.043	0.162	0.976
Combination (RV)		0.040	0.041	0.146**	0.882**
Combination (CV)		0.039	0.043	0.164	1.024
Combination		0.038	0.039**	0.145**	0.875**
RW	Realised Variance (Stock)	0.947	0.186*	0.212	0.815
MEAN		1.151	0.516	0.543	0.994
MA		0.739	0.211	0.209	0.505
VARMA (stat.)		0.675	0.194*	0.225	0.591
VARMAX (stat.)		0.668	0.200	0.239	0.630
Vector-HAR (stat.)		0.680*	0.200	0.231	0.603
VARMA (econ.)		0.646*	0.156**	0.154**	0.397**
VARMAX (econ.)		0.645*	0.156**	0.154**	0.397**
Vector-HAR (econ.)		0.645*	0.156**	0.153**	0.394**
RM1994		0.718	0.215	0.217	0.520
RM2006		0.717	0.214	0.212	0.499
EXP		0.748	0.226	0.228	0.543
DCC-MGARCH		0.684	0.204	0.210	0.513
Combination (RV)		0.670*	0.191**	0.220	0.581
Combination (CV)		0.711	0.210	0.209	0.499
Combination		0.669*	0.184**	0.182**	0.456**
RW	Realised Covariance	0.027			
MEAN		0.043			
MA		0.022			
VARMA (stat.)		0.021			
VARMAX (stat.)		0.021			
Vector-HAR (stat.)		0.019**			
VARMA (stat.)		0.019**			
VARMAX (stat.)		0.019**			
Vector-HAR (stat.)		0.019**			
RM1994		0.023			
RM2006		0.022			
EXP		0.022			
DCC-MGARCH		0.024			
Combination (RV)		0.020**			
Combination (CV)		0.022			
Combination		0.020**			

Notes: This table contains the mean values of the Patton (2011) HR loss function: b governs the shape of the loss function, with $b < 0$ indicating heavier penalisation of under-prediction, $b = 0$ corresponds to the MSE loss function, $b = -2$ corresponds to QLIKE loss function, and $b = -3$ corresponds to a scaled version of the utility-based loss function derived in Appendix B. The null hypothesis that the evaluation measure associated with the MA method equals the evaluation measure associated with each other method is tested using a Wald test based on Newey-West HAC standard errors. Rejections at the 5% and 1% level are indicated by superscripts $*$ and $**$, respectively.

Table 3 – Economic measures of forecasting method quality

Method	Portfolio	Utility Function				
		LN	Q(1)	Q(2)	Q(4)	Q(8)
RW	Bond	2.263	2.250	1.125	0.563	0.281
MEAN		1.752	1.752	0.876	0.438	0.219
MA		1.063	1.062	0.531	0.266	0.133
VARMA (stat.)		1.034*	1.034*	0.517*	0.258*	0.129*
VARMAX (stat.)		1.035*	1.035*	0.517*	0.259*	0.129*
Vector-HAR (stat.)		1.012**	1.012**	0.506**	0.253**	0.127**
VARMA (econ.)		1.000**	1.000**	0.500**	0.250**	0.125**
VARMAX (econ.)		1.001**	1.001**	0.501**	0.250**	0.125**
Vector-HAR (econ.)		1.003**	1.003**	0.502**	0.251**	0.125**
RM1994		1.273	1.272	0.636	0.318	0.159
RM2006		1.147	1.146	0.573	0.287	0.143
EXP		1.292	1.290	0.645	0.322	0.161
DCC-MGARCH		1.129	1.129	0.564	0.282	0.141
Combination (RV)		1.019**	1.019**	0.510**	0.255**	0.127**
Combination (CV)		1.185	1.184	0.592	0.296	0.148
Combination		1.012**	1.012**	0.506**	0.253**	0.126**
Shrinkage (RV)		1.019**	1.019**	0.509**	0.255**	0.127**
Shrinkage (CV)		1.190	1.189	0.595	0.297	0.149
Shrinkage		1.012**	1.012**	0.506**	0.253**	0.126**
RW	Stock	10.140	9.514	4.931	2.466	1.233
MEAN		12.036	12.018	6.014	3.007	1.503
MA		6.155	6.095	3.052	1.526	0.763
VARMA (stat.)		7.127	7.141	3.575	1.788	0.894
VARMAX (stat.)		7.623	7.618	3.814	1.907	0.953
Vector-HAR (stat.)		7.268	7.283	3.646	1.823	0.912
VARMA (econ.)		4.804**	4.799**	2.404**	1.202**	0.601**
VARMAX (econ.)		4.795**	4.790**	2.400**	1.200**	0.600**
Vector-HAR (econ.)		4.766**	4.761**	2.385**	1.193**	0.596**
RM1994		6.322	6.288	3.149	1.574	0.787
RM2006		6.043	6.025	3.017	1.509	0.754
EXP		6.606	6.564	3.287	1.643	0.822
DCC-MGARCH		6.195	6.198	3.104	1.552	0.776
Combination (RV)		7.003	7.017	3.513	1.757	0.878
Combination (CV)		6.052	6.029	3.019	1.510	0.755
Combination		5.502**	5.506**	2.758**	1.379**	0.689**
Shrinkage (RV)		6.984	6.998	3.504	1.752	0.876
Shrinkage (CV)		6.065	6.041	3.025	1.513	0.756
Shrinkage		5.441**	5.445**	2.727**	1.364**	0.682**
RW	Bond/Stock	15.627	14.727	7.540	3.770	1.885
MEAN		16.257	16.265	8.137	4.069	2.034
MA		8.608	8.550	4.280	2.140	1.070
VARMA (stat.)		10.203	10.223	5.116	2.558	1.279
VARMAX (stat.)		12.780	12.662	6.336	3.168	1.584
Vector-HAR (stat.)		10.518	10.526	5.268	2.634	1.317
VARMA (econ.)		7.196**	7.192**	3.601**	1.800**	0.900**
VARMAX (econ.)		7.180**	7.176**	3.593**	1.797**	0.898**
Vector-HAR (econ.)		7.088**	7.084**	3.547**	1.774**	0.887**
RM1994		11.204	11.115	5.562	2.781	1.391
RM2006		9.985	9.942	4.976	2.488	1.244
EXP		11.469	11.365	5.687	2.844	1.422
DCC-MGARCH		9.370	9.383	4.696	2.348	1.174
Combination (RV)		9.982	10.002	5.006	2.503	1.251
Combination (CV)		9.710	9.668	4.839	2.420	1.210
Combination		8.225	8.235	4.122	2.061	1.031
Shrinkage (RV)		9.857	9.876	4.943	2.472	1.236
Shrinkage (CV)		10.306	10.247	5.129	2.564	1.282
Shrinkage		7.975**	7.982**	3.996**	1.998**	0.999**

Notes: This table contains the mean annualised costs of forecast imperfection. These costs assume that investors hold one or two risky asset portfolios and have either logarithmic (LN) or quadratic utility (Q(γ)). The null hypothesis that the mean costs associated with the MA method equals the mean costs associated with each other method is tested using a Wald test based on Newey-West HAC standard errors. Rejections at the 5% and 1% level are indicated by superscripts $*$ and $**$, respectively.

Table 4 – Portfolio turnover

Model	Position	Utility Function				
		LN	Q(1)	Q(2)	Q(4)	Q(8)
RW	Bond	3.574	3.588	1.794	0.897	0.448
MEAN		0.001**	0.001**	0.000**	0.000**	0.000**
MA		0.208	0.208	0.104	0.052	0.026
VARMA (stat.)		0.676	0.677	0.339	0.169	0.085
VARMAX (stat.)		0.822	0.824	0.412	0.206	0.103
Vector-HAR (stat.)		0.846	0.848	0.424	0.212	0.106
VARMA (econ.)		0.771	0.773	0.386	0.193	0.097
VARMAX (econ.)		0.762	0.764	0.382	0.191	0.096
Vector-HAR (econ.)		0.772	0.774	0.387	0.193	0.097
RM1994		0.565	0.567	0.284	0.142	0.071
RM2006		0.645	0.647	0.323	0.162	0.081
EXP		0.503	0.504	0.252	0.126	0.063
DCC-MGARCH		0.255	0.255	0.128	0.064	0.032
Combination (RV)		0.692	0.693	0.347	0.173	0.087
Combination (CV)		0.515	0.517	0.258	0.129	0.065
Combination		0.584	0.585	0.293	0.146	0.073
Shrinkage (RV)		0.672	0.673	0.336	0.168	0.084
Shrinkage (CV)		0.544	0.545	0.273	0.136	0.068
Shrinkage		0.575	0.576	0.288	0.144	0.072
RW	Stock	3.542	3.567	1.793	0.896	0.448
MEAN		0.001**	0.001**	0.000**	0.000**	0.000**
MA		0.205	0.206	0.103	0.051	0.026
VARMA (stat.)		0.278	0.279	0.139	0.070	0.035
VARMAX (stat.)		0.438	0.440	0.220	0.110	0.055
Vector-HAR (stat.)		0.311	0.312	0.156	0.078	0.039
VARMA (econ.)		0.868	0.872	0.436	0.218	0.109
VARMAX (econ.)		0.859	0.862	0.431	0.216	0.108
Vector-HAR (econ.)		0.965	0.969	0.484	0.242	0.121
RM1994		0.483	0.486	0.243	0.121	0.061
RM2006		0.535	0.538	0.269	0.134	0.067
EXP		0.440	0.442	0.221	0.111	0.055
DCC-MGARCH		0.497	0.498	0.249	0.125	0.062
Combination (RV)		0.292	0.292	0.146	0.073	0.037
Combination (CV)		0.500	0.502	0.251	0.125	0.063
Combination		0.525	0.526	0.263	0.132	0.066
Shrinkage (RV)		0.293	0.294	0.147	0.074	0.037
Shrinkage (CV)		0.494	0.496	0.248	0.124	0.062
Shrinkage		0.507	0.509	0.254	0.127	0.064
RW	Risk-free	5.729	5.762	2.891	1.445	0.723
MEAN		0.001**	0.001**	0.000**	0.000**	0.000**
MA		0.314	0.315	0.158	0.079	0.039
VARMA		0.784	0.786	0.393	0.196	0.098
VARMAX		0.834	0.836	0.418	0.209	0.105
Vector-HAR		1.045	1.048	0.524	0.262	0.131
VARMA (econ.)		1.217	1.222	0.611	0.305	0.153
VARMAX (econ.)		1.205	1.209	0.605	0.302	0.151
Vector-HAR (econ.)		1.305	1.310	0.655	0.328	0.164
RM1994		0.738	0.741	0.370	0.185	0.093
RM2006		0.837	0.840	0.420	0.210	0.105
EXP		0.650	0.652	0.326	0.163	0.082
DCC-MGARCH		0.511	0.513	0.256	0.128	0.064
Combination (RV)		0.829	0.831	0.415	0.208	0.104
Combination (CV)		0.702	0.705	0.353	0.176	0.088
Combination		0.826	0.828	0.414	0.207	0.104
Shrinkage (RV)		0.802	0.804	0.402	0.201	0.100
Shrinkage (CV)		0.722	0.725	0.363	0.181	0.091
Shrinkage		0.840	0.842	0.421	0.210	0.105

Notes: This table contains portfolio turnover defined as the mean absolute change in portfolio weights associated with each market. We assume that investors trade both futures markets and have either logarithmic (LN) or quadratic utility $(Q(\gamma))$. The null hypothesis that the turnover associated with the MA method equals the turnover associated with each other method is tested using a Wald test based on Newey-West HAC standard errors. Rejections at the 5% and 1% level are indicated by superscripts $*$ and $**$, respectively.

Table 5 – The determinants of forecasting method quality

Method		Standardised Regression Coefficient/Fit				
A	B	$\widehat{\psi}_0$	$\widehat{\psi}_1$	$\widehat{\psi}_2$	$\widehat{\psi}_3$	$\overline{R}^2$
Panel A: Unrestricted Method Comparisons						
MEAN	RW	−0.299	−0.054**	−0.090**	0.031*	1.212
MA		1.630**	−0.041**	−0.045*	0.025**	0.431
VARMA (econ.)		1.970**	−0.050**	−0.043*	0.026**	0.517
VARMAX (econ.)		1.973**	−0.049**	−0.043*	0.026**	0.509
Vector-HAR (econ.)		1.997**	−0.047**	−0.043*	0.027**	0.488
RM1994		0.989**	0.001	−0.061*	0.004	0.339
RM2006		1.282**	−0.025	−0.057*	0.014	0.392
EXP		0.926**	0.009	−0.064*	0.001	0.384
DCC-MGARCH		1.422**	−0.055**	−0.036*	0.023**	0.483
Combination		1.708**	−0.055**	−0.041*	0.026**	0.548
Shrinkage		1.773**	−0.054**	−0.041*	0.027**	0.534
MA	MEAN	1.929**	0.236**	0.077*	−0.019	6.958
VARMA (econ.)		2.268**	0.362**	0.025	−0.042	14.260
VARMAX (econ.)		2.272**	0.357**	0.020	−0.045	13.821
Vector-HAR (econ.)		2.295**	0.363**	0.025	−0.041	14.343
RM1994		1.288**	0.361**	0.025	−0.046	13.866
RM2006		1.581**	0.351**	0.023	−0.044	13.114
EXP		1.225**	0.377**	0.024	−0.045	14.923
DCC-MGARCH		1.721**	0.425**	0.000	−0.062	18.975
Combination		2.007**	0.454**	−0.007	−0.053	21.004
Shrinkage		2.072**	0.430**	0.013	−0.045	19.492
VARMA (econ.)	MA	0.340**	0.074*	−0.006	−0.007	0.492
VARMAX (econ.)		0.343**	0.072*	−0.007	−0.010	0.470
Vector-HAR (econ.)		0.367**	0.079	−0.007	−0.006	0.560
RM1994		−0.641**	0.303*	−0.038	−0.047*	9.658
RM2006		−0.348**	0.160*	−0.048*	−0.037*	2.796
EXP		−0.704**	0.337*	−0.033	−0.046*	11.934
DCC-MGARCH		−0.208**	0.108**	0.002	−0.015	1.155
Combination		0.078	0.087*	−0.006	−0.006	0.694
Shrinkage		0.143**	0.086*	−0.002	−0.001	0.688
VARMAX (econ.)	VARMA (econ.)	0.004	−0.005	0.001	−0.030	0.042
Vector-HAR (econ.)		0.027**	−0.030	0.004	0.014	0.055
RM1994		−0.981**	0.407**	0.021	−0.020	16.424
RM2006		−0.687**	0.327**	−0.005	−0.018	10.748
EXP		−1.043**	0.428**	0.028	−0.018	18.013
DCC-MGARCH		−0.548**	0.140**	0.020	−0.014	2.032
Combination		−0.261**	0.136**	0.004	0.004	1.792
Shrinkage		−0.197**	0.091*	0.014	0.014	0.801
Vector-HAR (econ.)	VARMAX (econ.)	0.023*	−0.032	0.000	0.024	0.105
RM1994		−0.985**	0.402**	0.027	−0.017	15.890
RM2006		−0.691**	0.319**	0.001	−0.014	10.157
EXP		−1.047**	0.423**	0.033	−0.015	17.460
DCC-MGARCH		−0.552**	0.138**	0.025*	−0.009	1.969
Combination		−0.265**	0.121**	0.008	0.009	1.408
Shrinkage		−0.201**	0.081*	0.016	0.020	0.636
RM1994	Vector-HAR (econ.)	−1.008**	0.428**	0.026	−0.019	18.117
RM2006		−0.714**	0.340**	−0.003	−0.019	11.640
EXP		−1.070**	0.450**	0.033	−0.017	19.818
DCC-MGARCH		−0.575**	0.149**	0.019	−0.018	2.327
Combination		−0.288**	0.147**	0.004	−0.001	2.098
Shrinkage		−0.224**	0.103**	0.015	0.010	1.023
RM2006	RM1994	0.293**	0.443**	−0.041*	0.022	19.430
EXP		−0.063**	0.331**	0.001	−0.019	10.976
DCC-MGARCH		0.433**	0.398**	−0.018	0.014	15.611
Combination		0.720**	0.394**	−0.014	0.024	15.448
Shrinkage		0.784**	0.415**	−0.012	0.026	17.230
EXP	RM2006	−0.356**	0.435**	0.042	−0.021	18.562
DCC-MGARCH		0.140*	0.335**	0.008	0.009	11.301
Combination		0.426**	0.312**	0.014	0.023	9.889
Shrinkage		0.491**	0.333**	0.019	0.027	11.362

Table 5 – The determinants of forecasting method quality (cont.)

Method		Standardised Regression Coefficient/Fit				
A	B	$\widehat{\psi}_0$	$\widehat{\psi}_1$	$\widehat{\psi}_2$	$\widehat{\psi}_3$	$\overline{R}^2$
DCC-MGARCH	EXP	0.496**	0.423**	−0.026	0.012	17.508
Combination		0.782**	0.417**	−0.021	0.022	17.206
Shrinkage		0.847**	0.440**	−0.020	0.023	19.188
Combination	DCC-MGARCH	0.287**	0.269**	−0.008	0.033	7.519
Shrinkage		0.351**	0.242**	0.006	0.046*	6.077
Shrinkage	Combination	0.064**	0.224**	0.036	0.027	5.035
Panel B: Restricted Method Comparisons (zero-correlation forecasts imposed)						
RW	zero-corr. imposed	0.603**	−0.011	0.002	0.060**	0.309
MEAN		−0.149**	0.327**	−0.021	−0.023	10.245
MA		0.650**	0.164**	−0.022	0.035	2.999
VARMA (econ.)		0.638**	0.105**	−0.030*	0.037*	1.428
VARMAX (econ.)		0.636**	0.117**	−0.032*	0.033*	1.671
Vector-HAR (econ.)		0.658**	0.101**	−0.031*	0.034	1.330
RM1994		0.141*	0.388**	−0.033	−0.026	15.495
RM2006		0.323**	0.284**	−0.053	−0.009	8.681
EXP		0.184**	0.419**	−0.034	−0.026	18.100
DCC-MGARCH		0.348**	0.231**	−0.032	0.020	5.487
Combination		0.469**	0.206**	−0.046	0.030	4.896
Shrinkage		0.523**	0.169**	−0.037	0.040	3.473

Notes: This table contains the estimated standardised coefficients (and adjusted R^2 statistics) associated with the regression given by (13). Specifically, regressions of relative cost (B with respect to A in annualised percentage terms) upon a constant (with coefficient ψ_0), lagged relative costs (ψ_1), lagged $1/N$ volatility (ψ_2), and lagged ADS economic activity (ψ_3). The relative costs assume that investors trade both futures markets and have quadratic utility (Q(4)). The null hypothesis that each coefficient equals zero is tested using a HAC-based Wald test. Rejections at the 5% and 1% level are indicated by superscripts $*$ and $**$, respectively.

Table 6 – Conditional and unconditional tests of forecasting method quality

							Method							
Method	A	B	C	D	E	F	G	H	I	J	K	L	M	N
							Panel A: Conditional Tests							
A. RW		0.460^+	0.042^+	0.036^+	0.035^+	0.035^+	0.053^+	0.047^+	0.060^+	0.050^+	0.044^+	0.042^+	0.034^+	0.034^+
B. MEAN			0.019^+	0.020^+	0.019^+	0.019^+	0.096^+	0.063^+	0.107^+	0.030^+	0.022^+	0.021^+	0.018^+	0.018^+
C. MA				0.012^+	0.010^+	0.007^+	0.888^-	0.888^-	0.881^-	0.894^-	0.205^+	0.087^+	0.003^+	0.004^+
D. VARMA (econ.)					0.400	0.053^+	0.918^-	0.929^-	0.915^-	0.980^-	0.965^-	0.973^-	0.025^+	0.013^+
E. VARMAX (econ.)						0.137	0.923^-	0.936^-	0.919^-	0.980^-	0.971^-	0.975^-	0.085^+	0.066
F. Vector-HAR (econ.)							0.923^-	0.934^-	0.919^-	0.977^-	0.966^-	0.974^-	0.038^+	0.004
G. RM1994								0.189^+	0.757^-	0.311^+	0.147^+	0.126^+	0.072^+	0.072^+
H. RM2006									0.834^-	0.444^+	0.145^+	0.116^+	0.060^+	0.060^+
I. EXP										0.306^+	0.151^+	0.129^+	0.075^+	0.076^+
J. DCC-MGARCH											0.063^+	0.040^+	0.017^+	0.019^+
K. Combination												0.106^+	0.027^+	0.030^+
L. Shrinkage													0.019^+	0.021^+
M. Hybrid														0.606^-
N. Hybrid (zero-corr.)														
# Successes (out of thirteen)	0	1	6	9	9	10	3	4	2	5	7	8	13	10
							Panel B: Unconditional Tests							
A. RW		1.079	0.568^+	0.478^+	0.477^+	0.470^+	0.738^+	0.660^+	0.754^+	0.623^+	0.547^+	0.530^+	0.467^+	0.468^+
B. MEAN			0.526^+	0.443^+	0.442^+	0.436^+	0.684^+	0.611^+	0.699^+	0.577^+	0.507^+	0.491^+	0.433^+	0.433^+
C. MA				0.841^+	0.840^+	0.829^+	1.300^-	1.163^-	1.329^-	1.097^-	0.963	0.933^+	0.823^+	0.824^+
D. VARMA (econ.)					0.998	0.985^+	1.545^-	1.382^-	1.579^-	1.304^-	1.145^-	1.109^-	0.979^+	0.979^+
E. VARMAX (econ.)						0.987^+	1.548^-	1.385^-	1.583^-	1.307^-	1.148^-	1.112^-	0.981^+	0.982^+
F. Vector-HAR (econ.)							1.568^-	1.403^-	1.603^-	1.324^-	1.162^-	1.126^-	0.994	0.994
G. RM1994								0.895^+	1.022^-	0.844^+	0.741^+	0.718^+	0.634^+	0.634^+
H. RM2006									1.143^-	0.944	0.829^+	0.803^+	0.708^+	0.709^+
I. EXP										0.826^+	0.725^+	0.702^+	0.620^+	0.620^+
J. DCC-MGARCH											0.878^+	0.851^+	0.750^+	0.751^+
K. Combination												0.969^+	0.855^+	0.855^+
L. Shrinkage													0.882^+	0.883^+
M. Hybrid														1.001
N. Hybrid (zero-corr.)														
# Successes (out of thirteen)	0	0	6	9	9	11	3	4	2	4	6	8	11	11

Notes: This table contains the results of conditional and unconditional tests for method quality using the GW tests described in the main text, respectively. The relative costs assume that investors trade both futures markets and have quadratic utility (Q(4)). The entries in Panel A show the proportion of time the expected value of relative costs (as given by the fitted values from (13)) associated with the method in the column with respect to the method in the row are positive. In this panel, rejections of the GW conditional test at the 5% level are indicated by superscripts $-$ and $+$, with the former (latter) denoting the case where the GW conditional test is rejected at the 5% level *and* the entries are above (below) 0.5. The entries in Panel B show the ratio of the mean forecast imperfection costs associated with the method in the column to the mean forecast imperfection costs associated with the method in the row. In this panel, rejections of the GW unconditional test at the 5% level are indicated by superscripts $-$ and $+$, with the former (latter) denoting the case where the GW unconditional test is rejected at the 5% level *and* the entries are above (below) unity. Successes are defined as occasions when the method in the column is selected by the GW conditional and unconditional test rules.

Table 7 – Regime-based conditional tests of forecasting method quality

| Method	Regime 1				Regime 2				Absolute Difference			
	A	B	C	D	A	B	C	D	A	B	C	D
Panel A: Volatility Regimes												
A. MA		0.957+	1.186−	0.958+		0.680+	0.641+	0.668+		0.277−	0.546−	0.290−
B. Vector-HAR (econ.)			1.240−	1.001			0.942	0.982			0.298−	0.020
C. Shrinkage				0.808+				1.042				0.235+
D. Hybrid												
# Successes (out of three)	1	2	0	2	0	1	1	1				
Panel B: Correlation Regimes												
A. MA		0.902+	1.024	0.902+		0.751+	0.837+	0.740+		0.151−	0.187−	0.161−
B. Vector-HAR (econ.)			1.135−	1.000			1.114−	0.986			0.021	0.013
C. Shrinkage				0.880+				0.885+				0.005
D. Hybrid												
# Successes (out of three)	0	2	0	2	0	2	1	2				
Panel C: Macroeconomic Regimes (real activity)												
A. MA		0.706+	0.936	0.707+		0.839+	0.933	0.833+		0.133	0.003	0.125
B. Vector-HAR (econ.)			1.326−	1.002			1.112−	0.993			0.213	0.009
C. Shrinkage				0.756+				0.893+				0.137
D. Hybrid												
# Successes (out of three)	0	2	0	2	0	2	0	2				
Panel D: Macroeconomic Regimes (inflation expectations)												
A. MA		0.831+	0.917	0.828+		0.827+	0.943	0.821+		0.003	0.026	0.007
B. Vector-HAR (econ.)			1.104−	0.996			1.140−	0.992			0.036	0.004
C. Shrinkage				0.902+				0.870+				0.032
D. Hybrid												
# Successes (out of three)	0	2	0	2	0	2	0	2				

Notes: This table contains the results of conditional tests for forecasting method quality using the regime (individual) and regime (difference) tests described in the main text, respectively. The relative costs assume that investors trade both futures markets and have quadratic utility (Q(4)). The entries in Panel A show the ratio of the mean forecast imperfection costs associated with the method in the column to the mean forecast imperfection costs associated with the method in the row, for two regimes: regime 1 is the low volatility regime and regime 2 is the high volatility regime (with the cutoff point given by the median volatility level). In this panel, rejections of the regime (individual) test at the 5% level are indicated by superscripts − and +, with the former (latter) denoting the case where the regime (individual) test is rejected at the 5% level *and* the entries are above (below) unity. The entries in Panels B, C, and D provide similar information except here a correlation and two macroeconomic regimes are considered: regime 1 is the low correlation (in absolute terms), NBER-defined recession, or the Thomson Reuters/University of Michigan low inflation expectation regime, and regime 2 is the high correlation (in absolute terms), NBER-defined expansion, or the Thomson Reuters/University of Michigan high inflation expectation regime. For all panels we include the absolute difference in the forecast imperfection cost ratios. Here rejections of the regime (difference) test at the 5% level are indicated by superscripts − and +, with the former (latter) denoting the case where the regime (difference) test is rejected at the 5% level *and* the forecast imperfection differences in regime 1 are higher (lower) than in regime 2. Successes are defined as occasions when the method in the column is selected by the GW unconditional test rule.

Table 8 – Forecasting method ranks

|	All		Vol.		Corr.		Act.		Inf.	
Method	T1	T2	R1	R2	R1	R2	R1	R2	R1	R2
Panel A: Unrestricted portfolio										
Hybrid	1	1	1	3	1	1	1	1	1	1
Vector-HAR (econ.)	2	1	1	3	1	1	1	1	1	1
Shrinkage	3	3	7	1	3	3	3	3	3	3
Combination	4	4	8	1	5	4	4	4	4	4
MA	5	4	3	7	3	5	4	4	5	5
DCC-MGARCH	6	6	9	3	7	5	7	6	7	6
RM2006	7	7	5	8	6	7	7	7	6	7
Vector-HAR (stat.)	8	9	10	3	10	7	4	9	7	9
RM1994	9	8	5	9	7	9	7	8	7	8
EXP	10	9	4	11	7	10	7	9	7	9
MEAN	11	11	12	9	11	10	7	11	11	11
RW	12	11	10	11	11	10	7	11	11	11
Panel B: Restricted portfolio (no risk-free asset)										
Hybrid	1	1	1	1	1	1	1	1	1	1
Vector-HAR (econ.)	1	1	1	1	1	1	1	1	1	1
Combination	3	3	7	1	3	3	4	3	3	3
Shrinkage	3	3	7	1	3	3	4	3	3	3
MA	5	5	3	7	5	5	6	5	5	5
DCC-MGARCH	6	6	9	6	7	6	6	6	8	5
RM2006	7	7	3	8	5	6	6	7	6	5
RW	7	9	9	9	9	6	6	9	8	8
Vector-HAR (stat.)	9	9	11	1	9	6	3	9	8	8
RM1994	10	8	5	9	8	6	6	8	6	8
EXP	11	9	5	11	9	11	6	11	8	11
MEAN	12	12	12	11	12	11	6	11	12	11
Panel C: Restricted portfolio (no risk-free asset and no short-sales)										
Hybrid	1	1	1	1	1	1	1	1	1	1
Vector-HAR (econ.)	2	1	2	2	2	1	1	1	2	1
RW	2	3	3	3	3	3	3	3	5	3
Combination	4	3	3	3	3	3	5	3	3	3
Shrinkage	4	3	3	3	3	3	5	3	4	3
Vector-HAR (stat.)	6	6	11	3	6	6	3	7	6	6
MA	7	6	9	7	6	8	7	7	9	6
RM2006	8	9	9	7	6	8	7	7	7	6
RM1994	9	6	7	7	6	7	7	6	7	6
EXP	9	9	7	10	6	8	7	7	9	6
DCC-MGARCH	11	9	6	10	6	8	7	7	9	6
MEAN	12	12	12	12	12	12	7	12	12	12

Notes: This table contains the ranks of the forecasting methods for all time periods (based on the GW conditional and unconditional test rules, denoted T1 and T2, respectively), and within various regimes, under the assumption that investors trade both futures markets (with and without portfolio weight restrictions), and have quadratic utility (with $\gamma = 4$). Volatility, correlation and two macroeconomic regimes are considered: regime 1 (R1) is the low volatility, low correlation (in absolute terms), NBER-defined recession, or the Thomson Reuters/University of Michigan low inflation expectation regime, and regime 2 (R2) is the high volatility, high correlation (in absolute terms), NBER-defined expansion, or the Thomson Reuters/University of Michigan high inflation expectation regime. The ranks are based on the number of successes achieved, where successes are defined as occasions when the method is selected by GW conditional or unconditional test rule.
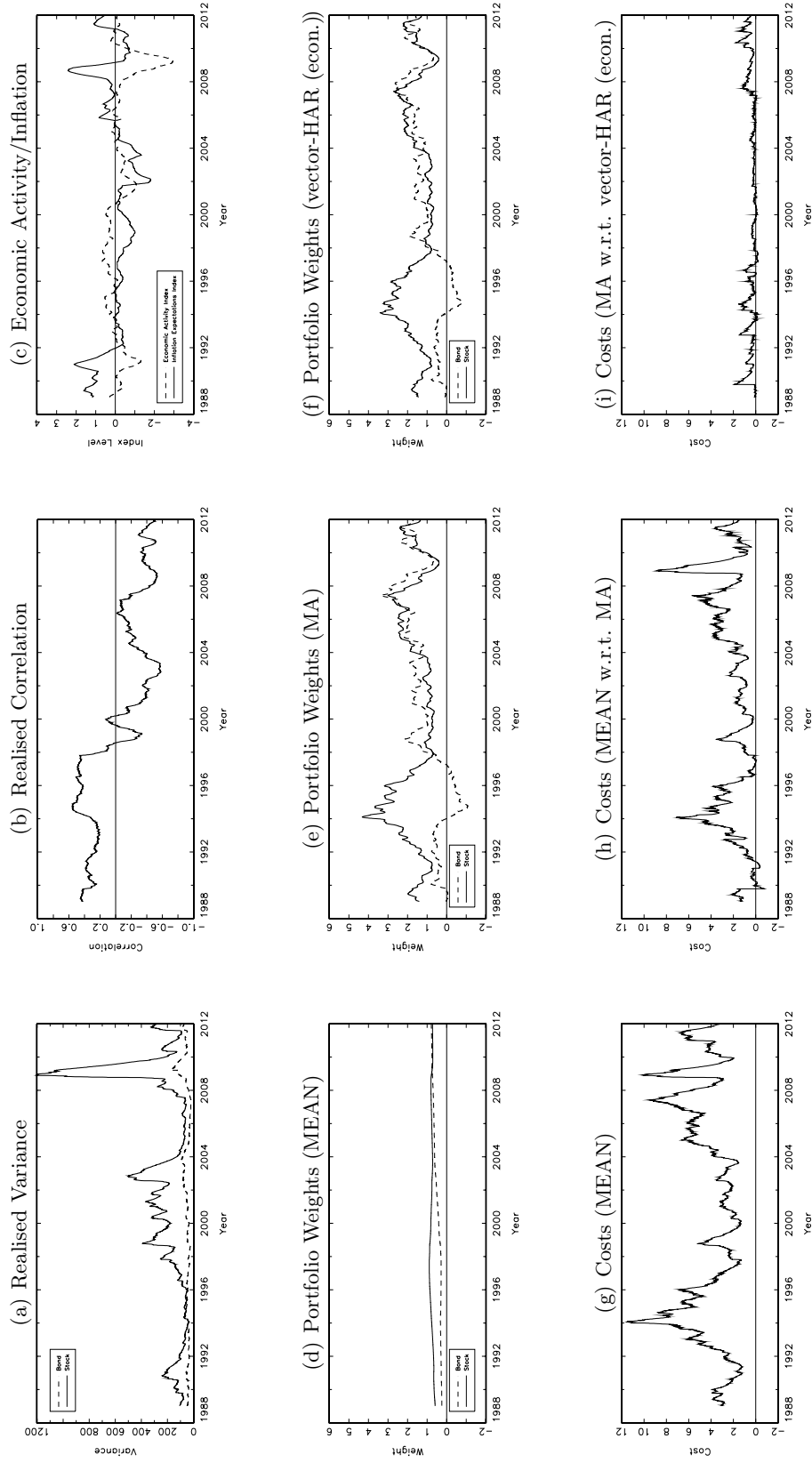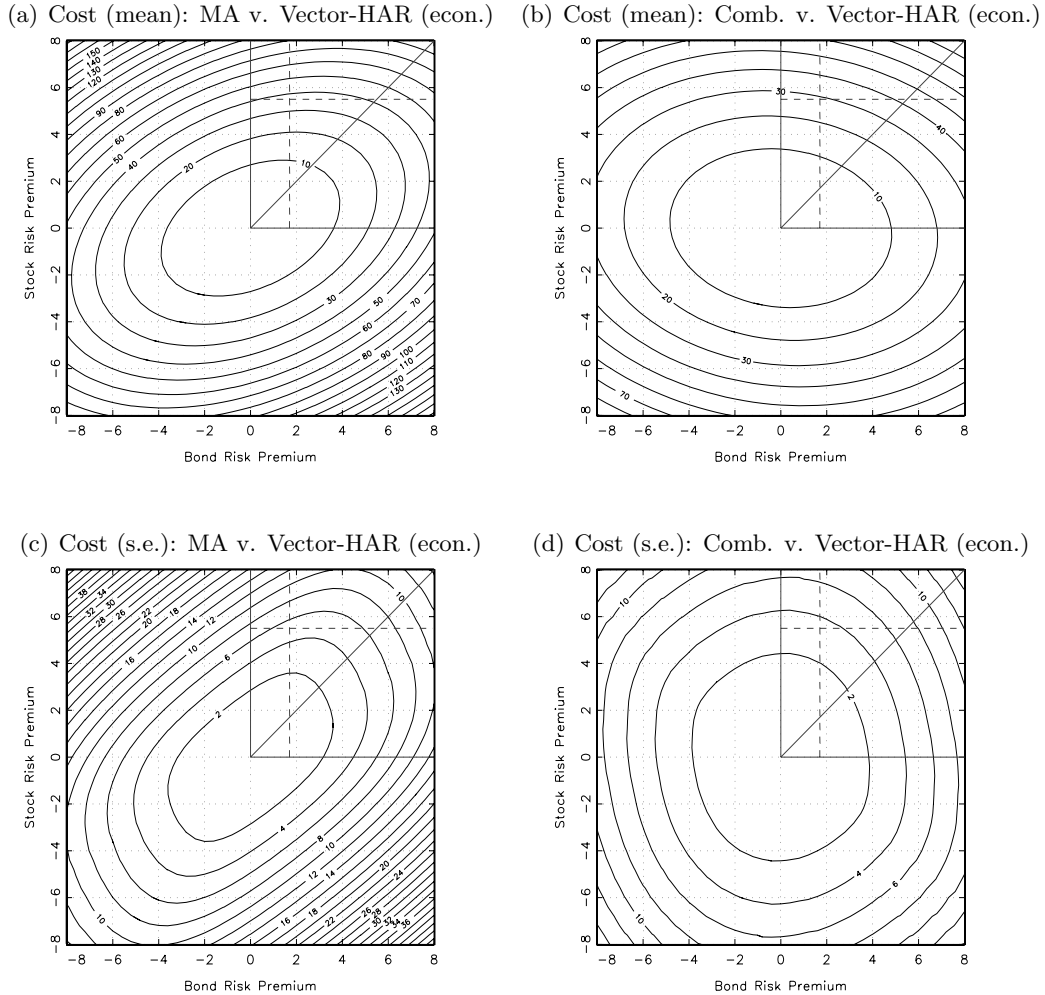
Figure 1 – Economic measures of forecast accuracy

This figure contains time series plots of: *i.* annualised averaged subsample realised variance and correlation measures, the ADS economic activity index, and the Thomson Reuters/University of Michigan inflation expectations index (both standardised to have a mean of zero and a standard deviation of one); *ii.* the optimal portfolio weights associated with each method; *iii.* the forecast imperfection costs (in annualised percentage terms) associated with each method (or with respect to each other). The weights and costs assume that investors trade both futures markets, have quadratic utility (with $\gamma = 4$), and $\boldsymbol{\mu}' - R^f = (0.017, 0.055)$ (annualised). For presentation purposes, an EWMA model (with a 250-day window and a smoothing coefficient of 0.99) is used to smooth all of the plots.

41
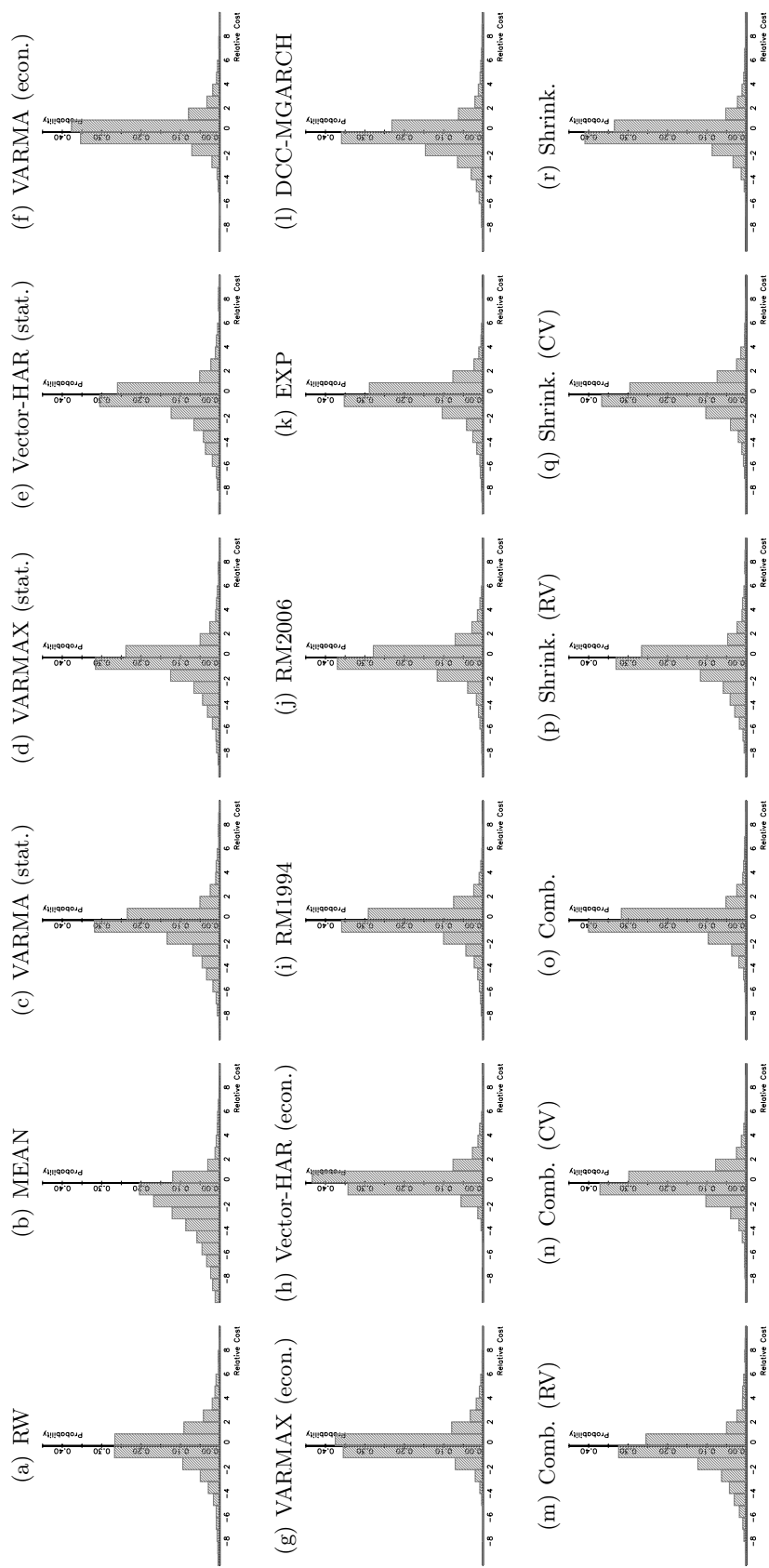
Figure 2 – Histograms of relative costs

This figure contains distributional plots of the mean relative costs associated with the MA and various competing forecasting methods. The costs assume that investors trade both futures markets, have quadratic utility (with $\gamma = 4$), and $\boldsymbol{\mu}' - R^f = (0.017, 0.055)$ (annualised). The relative cost can be interpreted as the performance fees an investor is willing to pay to avoid switching from using the MA method to the competing method.

(a) Cost (mean): MA v. Vector-HAR (econ.)   (b) Cost (mean): Comb. v. Vector-HAR (econ.)

(c) Cost (s.e.): MA v. Vector-HAR (econ.)   (d) Cost (s.e.): Comb. v. Vector-HAR (econ.)

Figure 3 – Contour plots of relative costs

This figure contains contour plots of the relative costs (measured in basis points) between the MA (or combination) method and the vector-HAR (econ.) method. In addition, we provide contour plots of the HAC standard errors associated with these costs. The costs assume that investors trade both futures markets, have quadratic utility (with $\gamma = 4$), and $\boldsymbol{\mu} - R^f \in (-0.08, 0.08)$ (annualised). The intersection of the dashed lines represent the values of $\boldsymbol{\mu}$ used in the paper, viz. $\boldsymbol{\mu}' - R^f = (0.017, 0.055)$. The relative cost can be interpreted as the performance fee an investor is willing to pay to avoid switching from using the vector-HAR (econ.) method to the competing method.