

NCER Working Paper Series

# Forecast performance of implied volatility and the impact of the volatility risk premium

Ralf Becker  
Adam Clements  
Christopher Coleman-Fenn

Working Paper #45  
July 2009

# Forecast performance of implied volatility and the impact of the volatility risk premium

Ralf Becker\*, Adam E. Clements#, and Christopher A. Coleman-Fenn#

\* Economics, School of Social Sciences, University of Manchester

# School of Economics and Finance, Queensland University of Technology

July 13, 2009

**Keywords:** Implied volatility, volatility forecasts, volatility models, volatility risk premium, model confidence sets.

**JEL Classification:** C12, C22, G00.

## Abstract

Forecasting volatility has received a great deal of research attention, with the relative performance of econometric models based on time-series data and option implied volatility forecasts often being considered. While many studies find that implied volatility is the preferred approach, a number of issues remain unresolved. Implied volatilities are risk-neutral forecasts of spot volatility, whereas time-series models are estimated on risk-adjusted or real world data of the underlying. Recently, an intuitive method has been proposed to adjust these risk-neutral forecasts into their risk-adjusted equivalents, possibly improving on their forecast accuracy. By utilising recent econometric advances, this paper considers whether these risk-adjusted forecasts are statistically superior to the unadjusted forecasts, as well as a wide range of model based forecasts. It is found that an unadjusted risk-neutral implied volatility is an inferior forecast. However, after adjusting for the risk premia it is of equal predictive accuracy relative to a number of model based forecasts.

## Corresponding author

Adam Clements

School of Economics and Finance

Queensland University of Technology

GPO Box 2434, Brisbane, Qld, Australia 4001

email [a.clements@qut.edu.au](mailto:a.clements@qut.edu.au)

Ph +61 7 3864 2525, Fax +61 7 3864 1500.

# 1 Introduction

Estimates of the future volatility of asset returns are of great interest to many financial market participants. Generally, there are two approaches which may be employed to obtain such estimates. *First*, predictions of future volatility may be generated from econometric models of volatility given historical information (model based forecasts, *MBF*). For surveys of common modeling techniques see Campbell, Lo and MacKinlay (1997), Gouriéroux and Jasiak (2001), and Taylor (2005). *Second*, estimates of future volatility may be derived from option prices using implied volatility (*IV*). *IV* should represent a market's best risk-neutral prediction of the future volatility of the underlying (see, amongst others, Jorion, 1995, Poon and Granger, 2003, 2005).

Given the practical importance of volatility forecasting (i.e. portfolio allocation problems, Value-at-Risk, and option valuation), there exists a wide body of literature examining the relative forecast performance of various approaches. While the results of individual studies are mixed, the survey of 93 articles compiled by Poon and Granger (2003, 2005) reports that, overall, *IV* estimates often provide more accurate volatility forecasts than competing *MBF*. In the context of this study, Day and Lewis (1992), Canina and Figlewski (1993), Ederington and Guan (2002), and Koopman, Jungbacker and Hol (2005) report that *MBF* of equity index volatility provide more information relative to *IV*. On the other hand, Fleming, Ostdiek and Whaley (1995), Christensen and Prabhala (1998), Fleming (1998) and Blair, Poon and Taylor (2001) all find that equity index *IV* dominate *MBF*.

While there is a degree of inconsistency in previous results, the general result that *IV* estimates often provide more accurate volatility forecasts than competing *MBF* is rationalised on the basis that *IV* should be based on a larger and timelier information set. That is, as *IV* is derived from the equilibrium market expectation of a future-dated payoff, rather than on the purely historical data *MBF* are estimated on, it is argued that *IV* contains all prior information garnered from historical data and also incorporates the additional information of the beliefs of market participants regarding future volatility; in an efficient options market, this additional information should yield superior forecasts.

In a related yet different context, Becker, Clements and White (2006) examine whether a particular *IV* index derived from S&P 500 option prices, the *VIX*, contains any information relevant to future volatility beyond that reflected in *MBF*. As they conclude that the *VIX* does not contain any such information, this result, *prima facie*, appears to contradict the previous findings summarised in Poon and Granger (2003). However, no forecast comparison is undertaken and they merely conjecture that the *VIX* may be viewed as a combination of *MBF*. Subsequently, Becker and Clements (2008) show that the *VIX* index produces forecasts which are statistically inferior to a number of competing *MBF*. Further, a combination of the best *MBF* is found to be superior to both the individual model based and *VIX* forecasts. They conclude that while it is plausible that *IV* combines information that is used in a range of different *MBF*, it is not the best possible combination of such information. This research provided an important contribution to the literature by allowing

for more robust conclusions to be drawn regarding comparative forecast performance, relative to the contradictory results of prior research. This was achieved by simultaneously examining a wide class of MBF and an IV forecast, rather than the typical pair-wise comparisons of prior work, using up-to-date forecast evaluation technology.

However, prior results cannot be viewed as definitive as, aside from the inconsistency of results, it may be argued that IV forecasts are inherently disadvantaged in the context of forecast evaluation. The pricing of financial derivatives is assumed to occur in a risk-neutral environment whereas MBF are estimated under the physical measure. Hence, prior tests of predictive accuracy have essentially been comparisons of risk-neutral versus real-world forecasts; as the target is also real-world, models estimated from like data may have an advantage. Without adjusting for the volatility risk premium, generally defined as the difference in expectations of volatility under the risk-neutral and physical measures, risk-neutral forecasts may be biased.

By matching the moments of model-free RV and IV, Bollerslev, Gibson, and Zhou (2008), hereafter BGZ, obtain an estimate of the volatility risk premium; it is then a straight-forward process to convert the risk-neutral IV into a forecast under the physical measure. This risk-adjusted IV can then be compared to MBF. Following Becker and Clements (2008), we conduct such a comparison using the model confidence set methodology of Hansen, Lunde and Nason (2003a) and find that the transformed *VIX* is statistically superior to the risk-neutral *VIX* in the majority of samples considered and is of equal predictive accuracy to MBF, particularly at the 22-day forecast horizon the *VIX* is constructed for.

The paper will proceed as follows. Section 2 will outline the data relevant to this study. Section 3 discusses the econometric models used to generate the various forecasts, along with the methods used to discriminate between forecast performance. Sections 4 and 5 present the empirical results and concluding remarks respectively.

## 2 Data

This study is based upon data relating to the S&P 500 Composite Index, from 2 January 1990 to 31 December 2008 (4791 observations) and extends the data set studied by Becker, Clements and White (2006) for the purposes of comparison and consistency. To address the research question at hand, estimates of both IV and future actual volatility are required. The *VIX* index constructed by the Chicago Board of Options Exchange from S&P 500 index options constitutes the estimate of IV utilised in this paper. It is derived from out-of-the-money put and call options that have maturities close to the target of 22 trading days. For technical details relating to the construction of the *VIX* index, see Chicago Board Options Exchange (CBOE, 2003). While the true process underlying option pricing is unknown, the *VIX* is constructed to be a model-free measure of the market's estimate of average S&P 500 volatility over the subsequent 22 trading days (BPT, 2001, Christensen and Prabhala, 1998 and CBOE, 2003).

Having a fixed forecast horizon is advantageous and avoids various econometric issues; hence, the 22-day length of the  $VIX$  forecast shall be denoted by  $\Delta$  hereafter. While this index has only been available since September 2003, when the CBOE replaced a previous IV index based on S&P 100 options, it can be calculated retrospectively back to January 1990. It's advantages in comparison to the previous IV index are that it no longer relies on IV derived from Black-Scholes option pricing models, it is based on more liquid options written on the S&P500 and is easier to hedge against (CBOE, 2003).

For the purposes of this study, estimates of actual volatility were obtained using the RV methodology outlined in ABDL (2001, 2003). RV estimates volatility by means of aggregating intra-day squared returns; it is important to note, as shall be discussed in more detail in the next section, that this measure is a model-free estimate of latent actual volatility. It should also be noted that the daily trading period of the S&P500 is 6.5 hours and that overnight returns were used as the first intra-day return in order to capture the variation over the full calendar day. ABDL (1999) suggest how to deal with practical issues relating to intra-day seasonality and sampling frequency when dealing with intra-day data. Based on the volatility signature plot methodology, daily RV estimates were constructed using 30 minute S&P500 index returns<sup>1</sup>. It is widely acknowledged (see e.g. Poon and Granger, 2003) that RV is a more accurate and less noisy estimate of the unobservable volatility process than squared daily returns. Patton (2006) suggests that this property of RV is beneficial when RV is used a proxy for observed volatility when evaluating forecasts.

Figure 1 shows the daily  $VIX$  and the 22-day mean daily S&P500 RV,  $\widehat{RV} = \frac{1}{\Delta} \sum_{i=1}^{\Delta} RV_{t+i}$ , for the sample period considered. While the RV estimates exhibit a similar overall pattern when compared to the  $VIX$ , RV reaches lower peaks than the  $VIX$ . This difference is mainly due to the fact that the  $VIX$  represents a risk-neutral forecast as opposed to RV that is a physical measure of volatility. In a perfectly efficient options market, the difference between the  $VIX$  and actual volatility should be a function of the volatility risk-premium alone.

### 3 Methodology

In this section the econometric models upon which forecasts are based will be outlined, followed by how the risk-neutral forecast provided by the  $VIX$  can be transformed into a forecast under the physical measure. This section concludes with a discussion of the technique utilised to discriminate between the volatility forecasts.

#### 3.1 Model based forecasts

While the true process underlying the evolution of volatility is not known, a range of candidate models exist and are chosen so that they span the space

---

<sup>1</sup>Intraday S&P 500 index data were purchased from Tick Data, Inc.

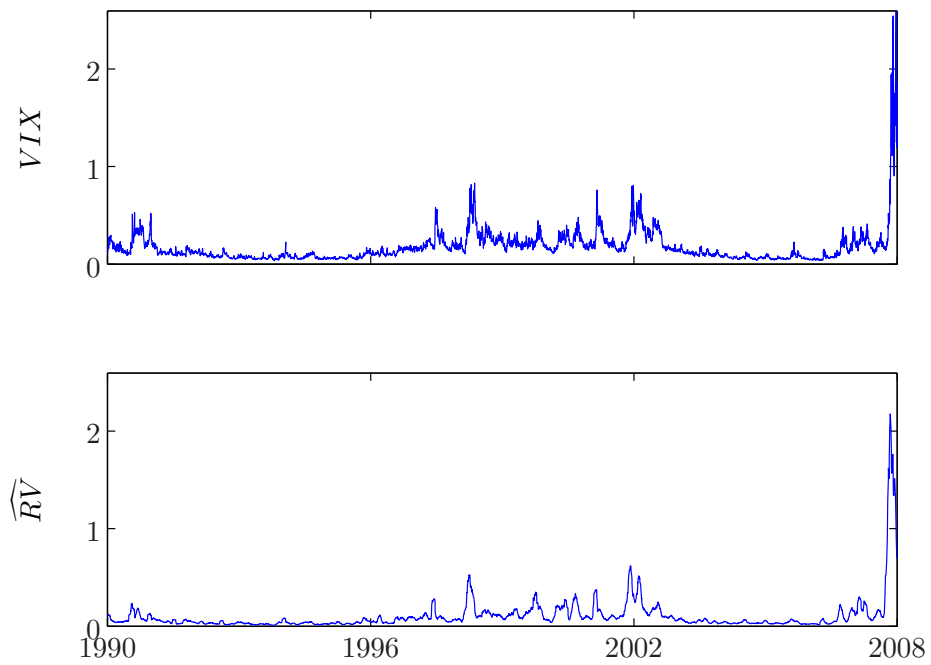


Figure 1: Daily VIX (top panel) from 2/01/1990 to 1/12/2008 and 22-day mean daily S&P500 Index RV estimate (bottom panel) from 2/01/1990 to 31/12/2008.

of available model classes. The set of models chosen are based on the models considered when the informational content of IV has been considered in Koopman, Jungbacker and Hol (2005), BPT (2001), and Becker, Clements and White (2006). The models chosen include those from the GARCH, stochastic volatility (SV), and RV classes.

GARCH style models employed in this study are similar to those proposed by BPT (2001). We begin with the original specification of Bollerslev (1986), where volatility is assumed to possess the following dynamics,

$$\begin{aligned} r_t &= \mu + \varepsilon_t, & \varepsilon_t &= \sqrt{h_t}z_t, & z_t &\sim N(0,1) \\ h_t &= \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned} \quad (1)$$

This simplest model specification is then extended to accomodate one of volatility's stylised facts, the asymmetric relationship between volatility and returns; this extension is provided by the GJR (see Glosten, Jagannathan and Runkle, 1993, Engle and Ng, 1991) process,

$$\begin{aligned} r_t &= \mu + \varepsilon_t, & \varepsilon_t &= \sqrt{h_t}z_t, & z_t &\sim N(0,1) \\ h_t &= \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2s_{t-1}\varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned} \quad (2)$$

with  $s_{t-1}$  at unity when  $\varepsilon_{t-1} < 0$  and 0 otherwise. This process nests the standard GARCH model when  $\alpha_2 = 0$ <sup>2</sup>.

Parameter estimates for the GARCH and GJR models are similar to those commonly observed for GARCH models based on various financial time series, reflecting strong volatility persistence, and are qualitatively similar to those reported in BPT (2001)<sup>3</sup>. Furthermore, allowing for asymmetries in conditional volatility is important, irrespective of the volatility process considered.

This study also proposes that an SV model may be used to generate forecasts. SV models differ from GARCH models in that conditional volatility is treated as an unobserved variable, and not as a deterministic function of lagged returns. The simplest SV model describes returns as

$$r_t = \mu + \sigma_t u_t, \quad u_t \sim N(0,1) \quad (3)$$

where  $\sigma_t$  is the time  $t$  conditional standard deviation of  $r_t$ . SV models treat  $\sigma_t$  as an unobserved (latent) variable, following its own stochastic path, the simplest being an AR(1) process,

$$\log(\sigma_t^2) = \alpha + \beta \log(\sigma_{t-1}^2) + w_t, \quad w_t \sim N(0, \sigma_w^2). \quad (4)$$

---

<sup>2</sup>The conditional mean of returns is modeled as a constant. As Cochrane (2001, p 388) points out, returns are predictable over very long horizons but at a daily frequency are virtually unpredictable.

<sup>3</sup>As the models discussed are re-estimated 3770 times in order to generate volatility forecasts for the 22 days following the respective estimation period, reporting parameter estimates is of little value. Here we will merely discuss the estimated model properties qualitatively. Parameter estimates for the rolling windows and the full sample are available on request.

In addition to GARCH and SV approaches, it is possible to utilise estimates of RV to generate forecasts of future volatility. These forecasts can be generated by directly applying time series models to daily measures of RV,  $RV_t$ . In following ADBL (2003) and Koopman *et al.* (2005), an ARMA(2,1) process is utilised where parameter estimates reflect the common feature of volatility persistence.

In order to generate MBF which capture the information available at time  $t$  efficiently, the volatility models were re-estimated for time-step  $t$  using data from  $t - 999$  to  $t$ . The resulting parameter values were then used to generate volatility forecasts for the subsequent  $\Delta$  business days ( $t + 1 \rightarrow t + \Delta$ ), corresponding to the period covered by the *VIX* forecast generated on day  $t$ . The first forecast covers the trading period from 13 December 1993 to 12 January 1994. For subsequent forecasts the model parameters were re-estimated using a rolling estimation window of 1,000 observations. The last forecast period covers 12 December 2008 to 31 December 2008, leaving 3,770 forecasts. For the shorter forecast horizons of 5- and 1-day ahead forecasts, the sample is shortened to also contain 3,770 forecasts.

### 3.2 A risk-adjusted VIX forecast

As discussed in Section 1, prior tests of the relative predictive accuracy of MBF and IV forecasts may be inherently biased as the IV forecasts are generated in a risk-neutral environment even though the target is under the physical measure. However, the recent work of BGZ has put forth an approach to estimate the volatility risk premium and then incorporate this risk premium to generate a risk-adjusted forecast of volatility from IV; the broad details of their approach are now given. BGZ make use of the fact that there exists a model-free estimate of volatility, RV, and a model-free forecast from IV, in this case the *VIX*. We begin with a brief revision of some of the properties of RV and the *VIX*, and then describe how the volatility risk premium is estimated within a Generalised Method of Moments (GMM) framework.

We let  $\mathcal{V}_{t,t+\Delta}^n$  be the RV computed by aggregating intra-day returns over the interval  $[t, t + \Delta]$ :

$$\mathcal{V}_{t,t+\Delta}^n \equiv \sum_{i=1}^n \left[ p_{t+\frac{i}{n}(\Delta)} - p_{t+\frac{i-1}{n}(\Delta)} \right]^2 \quad (5)$$

where  $p_t$  is the price at time  $t$ .

As  $n$  increases asymptotically,  $\mathcal{V}_{t,t+\Delta}^n$  becomes an increasingly accurate measure of the latent, underlying, volatility by the theory of quadratic variation (see Barndorff-Nielson and Shephard (2004) for asymptotic distributional results when allowing for leverage effects). The first conditional moment of RV under the physical measure is given by (see Bollerslev and Zhou (2002), Meddahi(2002), and Anderson *et al.* (2004) )

$$\mathbb{E}(\mathcal{V}_{t+\Delta,t+2\Delta} | \mathcal{F}_t) = \alpha_{\Delta} \mathbb{E}(\mathcal{V}_{t,t+\Delta} | \mathcal{F}_t) + \beta_{\Delta} \quad (6)$$

where  $\mathcal{F}_t$  is the information set up to time  $t$ . The co-efficients  $\alpha_{\Delta} = e^{-\kappa\Delta}$  and  $\beta_{\Delta} = \theta(1 - e^{-\kappa\Delta})$  are functions of the underlying parameters of the general



continuous-time stochastic volatility model for the logarithm of the stock price of Heston (1993); specifically,  $\kappa$ , is the speed of mean reversion to the long-term mean of volatility,  $\theta$ . Having described the first moment of RV under the physical measure, we now detail the calculation of a model-free, risk-neutral, forecast of volatility derived from the options market and how it may be transformed into a risk-adjusted forecast.

A model-free estimate of IV equating to an option expiring at time  $t + \Delta$ ,  $IV_{t,t+\Delta}^*$ , has been described in a continuous framework by Britten-Jones and Neuberger (2000), and extended to the case of jump-diffusion processes by Jiang and Tian (2005), as the integral of a basket of options with time-to-maturity  $\Delta$ ,

$$IV_{t,t+\Delta}^* = 2 \int_0^\infty \frac{C(t + \Delta, K) - C(t, K)}{K^2} dK \quad (7)$$

where  $C(t, K)$  is the price of a European call option maturing at time  $t$  with an exercise price  $K$ . In this study we use the  $VIX$  as a proxy for  $IV_{t,t+\Delta}^*$ , the calculation of the  $VIX$  differs from the measure just given, see the CBOE (2003) for details. This measure is a model-free, risk-neutral forecast of volatility,

$$\mathbb{E}^*(\mathcal{V}_{t,t+\Delta} | \mathcal{F}_t) = IV_{t,t+\Delta}^* \quad (8)$$

with  $\mathbb{E}^*(\cdot)$  the expectation under the risk-neutral measure. To transform this risk-neutral expectation into its equivalent under the physical measure, we invoke the result of Bollerslev and Zhou (2002),

$$\begin{aligned} \mathbb{E}(\mathcal{V}_{t,t+\Delta} | \mathcal{F}_t) &= \mathcal{A}_\Delta IV_{t,t+\Delta}^* + \mathcal{B}_\Delta \\ \mathcal{A}_\Delta &= \frac{(1 - e^{-\kappa\Delta})/\kappa}{(1 - e^{-\kappa^*\Delta})/\kappa^*} \\ \mathcal{B}_\Delta &= \theta \left[ \Delta - (1 - e^{-\kappa\Delta})/\kappa \right] - \mathcal{A}_\Delta \theta^* \left[ \Delta - (1 - e^{-\kappa^*\Delta})/\kappa^* \right] \end{aligned} \quad (9)$$

where  $\mathcal{A}_\Delta$  and  $\mathcal{B}_\Delta$  depend on the underlying parameters,  $\kappa$ ,  $\theta$ , and  $\lambda$ , of the afore mentioned stochastic volatility model; specifically,  $\kappa^* = \kappa + \theta$  and  $\theta^* = \kappa\theta/(\kappa + \theta)$ . It is now possible to recover the volatility risk premium,  $\lambda$ .

### 3.3 Estimation of Volatility Risk-Premium

The unconditional volatility risk-premium,  $\lambda$ , is estimated in a GMM framework utilising the moment conditions (6) and (9), as well as a lagged instrument of  $IV^4$  to accommodate over-identifying restrictions, leading to the system of equations:

---

<sup>4</sup>While BGZ use lagged RV as their instrument, we find the use of IV dramatically improves forecast performance, with details available upon request.

$$f_t(\xi) = \begin{bmatrix} \mathcal{V}_{t+\Delta, t+2\Delta} - \alpha_\Delta \mathcal{V}_{t, t+\Delta} + \beta_\Delta \\ (\mathcal{V}_{t+\Delta, t+2\Delta} - \alpha_\Delta \mathcal{V}_{t, t+\Delta} + \beta_\Delta) IV_{t-\Delta, t}^* \\ \mathcal{V}_{t, t+\Delta} - \mathcal{A}_\Delta IV_{t, t+\Delta}^* - \mathcal{B}_\Delta \\ (\mathcal{V}_{t, t+\Delta} - \mathcal{A}_\Delta IV_{t, t+\Delta}^* - \mathcal{B}_\Delta) IV_{t-\Delta, t}^* \end{bmatrix} \quad (10)$$

where  $\xi$  is the parameter vector  $(\kappa, \theta, \lambda)'$ . We estimate  $\xi$  via standard GMM arguments such that  $\hat{\xi}_t = \arg \min g_t(\xi)' W g_t(\xi)$ , where  $g_t(\xi)$  are the sample means of the moment conditions, and  $W$  is the asymptotic co-variance matrix of  $g_t(\xi)$ . We follow BGZ and employ an autocorrelation and heteroscedasticity robust  $W$ , as per Newey and West (1987). A Monte Carlo experiment conducted by BGZ confirms that the above approach leads to an estimate of the volatility risk premium comparable to using actual (unobserved and infeasible) risk-neutral implied volatility and continuous-time integrated volatility. The optimisation process is constrained such that  $\kappa$  and  $\theta$  are positive, to ensure stationarity, and positive variance respectively. Once the elements of  $\hat{\xi}_t$  have been determined, substitution into (9) yields a risk-adjusted forecast of volatility, derived from a risk-neutral IV.

We recursively estimate  $\hat{\xi}_t$  with an expanding window beginning with an initial estimation period of 1000 observations to align with the estimation process of the MBF; this recursive estimation is updated daily. However, we cannot use all data points in the estimation period due to the 22-day window for calculating  $\mathcal{V}_{t, t+\Delta}$ . That is, with 1000 days of data there are 45 periods of 22 days, so we start at day 10; daily updating results in differing start dates, i.e. with 1001 there are still 45 periods of 22 days, so we begin on day 11.

### 3.4 Evaluating forecasts

Following Becker and Clements (2008), we employ the model confidence set (MCS) approach to examine the relative forecast performance of the transformed  $VIX$ . At the heart of the methodology (Hansen, Lunde and Nason, 2003a) as it is applied here, is a forecast loss measure. Such measures have frequently been used to rank different forecasts and the two loss functions utilised here are the MSE and QLIKE,

$$MSE^i = (\overline{RV}_{t+\Delta} - f_t^i)^2, \quad (11)$$

$$QLIKE^i = \log(f_t^i) + \frac{\overline{RV}_{t+\Delta}}{f_t^i}, \quad (12)$$

where  $f_t^i$  are individual forecasts (formed at time  $t$ ) obtained from the individual models,  $i$ , and both the risk-neutral and risk-adjusted  $VIX$  forecasts. The total number of candidate forecasts will be denoted as  $m_0$ ; therefore, the competing individual forecasts are given by  $f_t^i$ ,  $i = 1, 2, \dots, m_0$ . While there are many alternative loss functions, Patton (2006) shows that MSE and QLIKE belong to a family of loss functions that are robust to noise in the volatility proxy,  $\overline{RV}_{t+\Delta}$  in this case, and would give consistent rankings of models irrespective of the volatility proxy used. Each loss function has somewhat different

properties, with MSE weighting errors symmetrically whereas QLIKE penalises under-prediction more heavily than over-prediction. Regression based techniques proposed by Mincer and Zarnowitz (1969) are not used here as Patton (2006) shows that they are sensitive to the assumed distribution of the volatility proxy and can lead to changes in forecast ranking as the proxy changes<sup>5</sup>.

While these loss functions allow forecasts to be ranked, they give no indication of whether the top performing model is statistically superior to any of the lower-ranked models; the MCS approach does allow for such conclusions to be drawn. The construction of a MCS is an iterative procedure that requires sequential testing of equal predictive accuracy (EPA); the set of candidate models is trimmed by deleting models that are found to be statistically inferior. The interpretation attached to a MCS is that it contains the best forecast with a given level of confidence; although the MCS may contain a number of models which indicates they are of EPA. The final surviving models in the MCS are optimal with a given level of confidence and are statistically indistinguishable in terms of their forecast performance.<sup>6</sup>

The procedure starts with a full set of candidate models  $\mathcal{M}_0 = \{1, \dots, m_0\}$ . The MCS is determined by sequentially trimming models from  $\mathcal{M}_0$ , reducing the number of models to  $m < m_0$ . Prior to starting the sequential elimination procedure, all loss differentials between models  $i$  and  $j$  are computed,

$$d_{ij,t} = L(\overline{RV}_{t+\Delta}, f_t^i) - L(\overline{RV}_{t+\Delta}, f_t^j), \quad i, j = 1, \dots, m_0, \quad i \neq j, \quad t = 1, \dots, T - \Delta \quad (13)$$

where  $L(\cdot)$  is one of the loss functions described above. At each step, the EPA hypothesis

$$H_0 : \mathbb{E}(d_{ij,t}) = 0, \quad \forall i > j \in \mathcal{M} \quad (14)$$

is tested for a set of models  $\mathcal{M} \subset \mathcal{M}_0$ , with  $\mathcal{M} = \mathcal{M}_0$  at the initial step. If  $H_0$  is rejected at the significance level  $\alpha$ , the worst performing model is removed and the process continued until non-rejection occurs with the set of surviving models being the MCS,  $\widehat{\mathcal{M}}_\alpha^*$ . If a fixed significance level  $\alpha$  is used at each step,  $\widehat{\mathcal{M}}_\alpha^*$  contains the best model from  $\mathcal{M}_0$  with  $(1 - \alpha)$  confidence<sup>7</sup>.

At the core of the EPA statistic is the  $t$ -statistic

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}}$$

---

<sup>5</sup>Andersen, Bollerslev and Meddahi (2005) have pointed out that, in order to reveal the full extent of predictability,  $R^2$  from Mincer-Zarnowitz regressions ought to be pre-multiplied with a correction factor allowing for the approximation error embodied in  $RV$ . As this work seeks to evaluate relative forecast performance, and all models are compared against the same volatility proxy, no such correction is sought for the  $MSE$  and  $QLIKE$ .

<sup>6</sup>As the MCS methodology involves sequential tests for EPA, Hansen, Lunde and Nason (2003a, 2003b) utilised the testing principle of Pantula (1989) to avoid size distortions.

<sup>7</sup>Despite the testing procedure involving multiple hypothesis tests this interpretation is a statistically correct one. See Hansen *et al.* (2003b) for a detailed discussion of these aspects.

where  $\bar{d}_{ij} = \frac{1}{T} \sum_{t=1}^T d_{ij,t}$ .  $t_{ij}$  provides scaled information on the average difference in the forecast quality of models  $i$  and  $j$ , the scaling parameter,  $\widehat{var}(\bar{d}_{ij})$ , is an estimate of  $var(\bar{d}_{ij})$  and is obtained from a bootstrap procedure described below. In order to decide whether the MCS must at any stage be further reduced, the null hypothesis in (14) is to be evaluated. The difficulty being that for each set  $\mathcal{M}$ , the information from  $(m-1)m/2$  unique  $t$ -statistics needs to be distilled into one test statistic. Hansen, et al. (2003a, 2003b) propose a range statistic,

$$T_R = \max_{i,j \in \mathcal{M}} |t_{ij}| = \max_{i,j \in \mathcal{M}} \frac{|\bar{d}_{ij}|}{\sqrt{\widehat{var}(\bar{d}_{ij})}} \quad (15)$$

and a semi-quadratic statistic,

$$T_{SQ} = \sum_{\substack{i,j \in \mathcal{M} \\ i < j}} t_{ij}^2 = \sum_{\substack{i,j \in \mathcal{M} \\ i < j}} \frac{(\bar{d}_{ij})^2}{\widehat{var}(\bar{d}_{ij})} \quad (16)$$

as test statistics to establish EPA, both of which indicate a rejection of the EPA hypothesis for large values. The actual distribution of the test statistic is complicated and depends on the covariance structure between the forecasts included in  $\mathcal{M}$ . Therefore,  $p$ -values for each of these test statistics have to be obtained from a bootstrap distribution (see below). When the null hypothesis of EPA is rejected, the worst performing model is removed from  $\mathcal{M}$ . The latter is identified as  $\mathcal{M}_i$  where

$$i = \arg \max_{i \in \mathcal{M}} \frac{\bar{d}_i}{\sqrt{\widehat{var}(\bar{d}_i)}} \quad (17)$$

and  $\bar{d}_i = \frac{1}{m-1} \sum_{j \in \mathcal{M}} \bar{d}_{ij}$ . The tests for EPA are then conducted on the reduced set of models and continues to iterate until the null hypothesis of EPA is not rejected.

Bootstrap distributions are required for the test statistics  $T_R$  and  $T_{SQ}$ . These distributions will be used to estimate  $p$ -values for the  $T_R$  and  $T_{SQ}$  tests and, hence, calculate model specific  $p$ -values. At the core of the bootstrap procedure is the generation of bootstrap replications of  $d_{ij,t}$ . In doing so, the temporal dependence in  $d_{ij,t}$  must be accounted for. This is achieved by the block bootstrap, which is conditioned on the assumption that the  $\{d_{ij,t}\}$  sequence is stationary and follows a strong geometric mixing assumption<sup>8</sup>. The basic steps of the bootstrap procedure are now described.

Let  $\{d_{ij,t}\}$  be the sequence of  $T - \Delta$  observed differences in loss functions for models  $i$  and  $j$ .  $B$  block bootstrap counterparts are generated for

---

<sup>8</sup>As discussed by White (2000) in a related context, a number of different block bootstrapping procedures are available. They differ chiefly in whether they use a constant or random block length. The latter methodology has the advantage of guaranteeing stationarity of the resulting bootstrap realisations (Politis and Romano, 1994) but will also produce a larger variance for the bootstrap statistics (Lahiri, 1999). Therefore, we follow the lead of Hansen *et al.* (2003, 2005) and use, in the terminology of Lahiri (1999), the circular block bootstrap with constant block size.

all combinations of  $i$  and  $j$ ,  $\{d_{ij,t}^{(b)}\}$  for  $b = 1, \dots, B$ . Values with a bar, e.g.  $\bar{d}_{ij} = (T - \Delta)^{-1} \sum d_{ij,t}$ , represent averages over all  $T - \Delta$  observations. First we will establish how to estimate the variance estimates  $\widehat{var}(\bar{d}_{ij})$  and  $\widehat{var}(\bar{d}_i)$ , which are required for the calculation of the EPA test statistics in (15), (16) and (17):

$$\begin{aligned}\widehat{var}(\bar{d}_{ij}) &= B^{-1} \sum_{b=1}^B (\bar{d}_{ij}^{(b)} - \bar{d}_{ij})^2 \\ \widehat{var}(\bar{d}_i) &= B^{-1} \sum_{b=1}^B (\bar{d}_i^{(b)} - \bar{d}_i)^2\end{aligned}$$

for all  $i, j \in \mathcal{M}$ . In order to evaluate the significance of the EPA test, a  $p$ -value is required. That is obtained by comparing either  $T_R$  or  $T_{SQ}$  with bootstrap realisations  $T_R^{(b)}$  or  $T_{SQ}^{(b)}$ .

$$\hat{p}_\tau = B^{-1} \sum_{b=1}^B \mathbb{I}(T_\tau^{(b)} > T_\tau) \quad \text{for } \tau = R, SQ$$

where  $\mathbb{I}(\cdot)$  is the indicator function.

The  $B$  bootstrap versions of the test statistics  $T_R$  or  $T_{SQ}$  are calculated by replacing  $|\bar{d}_{ij}|$  and  $(\bar{d}_{ij})^2$  in equations (15) and (16) with  $|\bar{d}_{ij}^{(b)} - \bar{d}_{ij}|$  and  $(\bar{d}_{ij}^{(b)} - \bar{d}_{ij})^2$  respectively. The denominator in the test statistics remains the bootstrap estimate discussed above.

This model elimination process can be used to produce model specific  $p$ -values. A model is only accepted into  $\widehat{\mathcal{M}}_\alpha^*$  if its  $p$ -value exceeds  $\alpha$ . Due to the definition of  $\widehat{\mathcal{M}}_\alpha^*$ , this implies that a model which is not accepted into  $\widehat{\mathcal{M}}_\alpha^*$  is unlikely to belong to the set of best forecast models. The model specific  $p$ -values are obtained from the  $p$ -values for the EPA tests described above. As the  $k$ th model is eliminated from  $\mathcal{M}$ , save the (bootstrapped)  $p$ -value of the EPA test in (15) or (16) as  $p(k)$ . For instance, if model  $\mathcal{M}_i$  was eliminated in the third iteration, i.e.  $k = 3$ . The  $p$ -value for this  $i$ th model is then  $\hat{p}_i = \max_{k \leq 3} p(k)$ . This ensures that the model eliminated first is associated with the smallest  $p$ -value indicating that it is the least likely to belong into the MCS<sup>9</sup>.

## 4 Empirical results

We break the forecasting results down into two periods: the entire sample period of 13 December 1993 to 31 December 2008 and a shortened sample of 13 December 1993 to 31 December 2007. The justification for this is simply to compare the forecasting ability in what one might term to “normal” conditions to the more extreme period which includes the fluctuations of 2008. Further, while the  $VIX$  is constructed to be a 22-day-ahead forecast, it may still contain

<sup>9</sup>See Hansen et al. (2003b) for a detailed interpretation for the MCS  $p$ -values.

information relevant to shorter forecast horizons. To examine this issue, we also consider 1- and 5-day ahead forecast performance in addition to the 22-day horizon.

#### 4.1 22 day-ahead forecasts

Generally, in forecasting the 22-day-mean of daily realised volatility for the entire sample period, we are left with a large group of models which are statistically indistinguishable. In fact, as Table 1 shows, only three models of the 11 are not included in the MCS under both loss functions<sup>10</sup>. Importantly in the context of this paper, the risk-neutral *VIX* has a  $p$ -value of 0.059 of being included in the MCS under the QLIKE loss function using the range statistic, and a  $p$ -value of 0.179 with the semi-quadratic statistic, while the transformed *VIX* is inseparable from the majority of competing MBF with  $p$ -values of 0.768 and 0.827 respectively.

When one examines Table 2, which relates to the shortened sample, it is clear that the transformed *VIX* outperforms the risk-neutral *VIX*, and there is a clearer distinction across models generally. In particular, the raw *VIX* has  $p$ -values of 0.032 and 0.037 under QLIKE for the range and semi-quadratic statistics respectively, implying rejection of inclusion in the MCS at most confidence levels, while the transformed *VIX* is included in the MCS almost surely. Under the MSE loss function, the risk-neutral *VIX* is least likely to be included in the MCS while the risk-adjusted *VIX* has a  $p$ -value of unity for both test statistics, implying rejecting the null is an error almost surely.

As the IV from an efficient options market should make use of a larger and timelier information set, such a forecast should be able to perform at least as well as models based purely on historical data. The results of this study would suggest, however, that MBF outperform risk-neutral IV based forecasts; although, adjusting for the risk-premium reverses that result.

#### 4.2 5 day-ahead forecasts

With regard to the shorter forecast horizon over the full sample, as Table 3 shows, the transformed *VIX* does less well relative to competing MBF but is still included in the MCS. The  $p$ -values for inclusion into the MCS for the full sample period are less than 0.4 under both loss measures and test statistics and is also statistically inseparable from the risk-neutral *VIX* under MSE. Under the QLIKE loss function, the risk-adjusted *VIX* forecast is statistically superior to a wide range of models, including the GARCH, EGARCH and SV class of models, as well as the risk-neutral *VIX*.

However, over the less volatile shorter sample period, with details provided in Table 4, the transformed *VIX* is still included in the MCS under both loss measures and test statistics, all  $p$ -values are above 0.89. The risk-adjusted *VIX* offers significant improvement over the raw *VIX* forecast, which is excluded from the MCS under both measures and test statistics, it's highest  $p$ -value is

---

<sup>10</sup>We generally defer to the results according to QLIKE as Patten and Shephard (2007) has shown it possesses superior power to MSE.

<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$	<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$
	$\hat{p}_i$	$\hat{p}_i$		$\hat{p}_i$	$\hat{p}_i$
<i>GVIX</i>	0.110	0.176	<i>GVIX</i>	0.001	0.003
<i>RVMID</i>	0.110	0.204	<i>ARMA</i>	0.001	0.007
<i>VIX</i>	0.116	0.242	<i>VIX</i>	0.059	0.179
<i>ARMA</i>	0.265	0.303	<i>RVMID</i>	0.768	0.827
<i>G</i>	0.314	0.343	<i>BGZ</i>	0.768	0.827
<i>BGZ</i>	0.377	0.455	<i>GRJRV</i>	0.768	0.827
<i>GJRVIX</i>	0.377	0.455	<i>SV</i>	0.768	0.827
<i>GJR</i>	0.377	0.455	<i>G</i>	0.776	0.827
<i>SV</i>	0.377	0.517	<i>GRV</i>	0.776	0.827
<i>GRV</i>	0.878	0.878	<i>GJR</i>	0.776	0.827
<i>GJRRV</i>	1.000	1.000	<i>GJRVIX</i>	1.000	1.000

Table 1: MCS of 22-day ahead forecasts for whole sample. The first row represents the first model removed, down to the best performing model in the last row.

<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$	<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$
	$\hat{p}_i$	$\hat{p}_i$		$\hat{p}_i$	$\hat{p}_i$
<i>VIX</i>	0.084	0.027	<i>GVIX</i>	0.000	0.000
<i>G</i>	0.084	0.027	<i>ARMA</i>	0.000	0.000
<i>GVIX</i>	0.084	0.048	<i>VIX</i>	0.032	0.037
<i>GJR</i>	0.084	0.050	<i>GRJRV</i>	0.307	0.227
<i>GJRVIX</i>	0.084	0.065	<i>GJR</i>	0.394	0.227
<i>ARMA</i>	0.084	0.065	<i>G</i>	0.513	0.290
<i>GRJRV</i>	0.084	0.081	<i>SV</i>	0.513	0.336
<i>RVMID</i>	0.305	0.244	<i>RVMID</i>	0.513	0.451
<i>SV</i>	0.305	0.309	<i>GRV</i>	0.759	0.728
<i>GRV</i>	0.354	0.354	<i>GJRVIX</i>	0.759	0.746
<i>BGZ</i>	1.000	1.000	<i>BGZ</i>	1.000	1.000

Table 2: MCS of 22-day ahead forecasts for shortened sample.

<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$	<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$
<i>GVIX</i>	0.304	0.260	<i>ARMA</i>	0.000	0.000
<i>ARMA</i>	0.304	0.278	<i>VIX</i>	0.000	0.000
<i>BGZ</i>	0.304	0.339	<i>G</i>	0.006	0.016
<i>VIX</i>	0.304	0.339	<i>GJR</i>	0.014	0.081
<i>G</i>	0.304	0.445	<i>SV</i>	0.092	0.156
<i>RVMID</i>	0.704	0.670	<i>GVIX</i>	0.368	0.316
<i>GRV</i>	0.821	0.805	<i>BGZ</i>	0.368	0.397
<i>SV</i>	0.930	0.898	<i>GRJRV</i>	0.368	0.397
<i>GJRVIX</i>	0.930	0.898	<i>RVMID</i>	0.562	0.519
<i>GRJRV</i>	0.930	0.898	<i>GRV</i>	0.562	0.519
<i>GJR</i>	1.000	1.000	<i>GJRVIX</i>	1.000	1.000

Table 3: MCS of 5-day ahead forecasts for whole sample

0.073 under MSE and the range statistic while has  $p$ -values of zero for both test statistics under QLIKE.

### 4.3 1 day-ahead forecasts

When examining results for one day-ahead forecasts, as shown in Tables 5 and 6, the transformed *VIX* is excluded from the MCS under QLIKE and either of the test statistics for both the shortened and full sample; the risk-adjusted *VIX* is, however, included in the MCS under the semi-quadratic statistic and MSE loss function for both the shortened and full sample. Unexpectedly, the risk-neutral *VIX* is included in the MCS for the full sample under MSE and both test statistics.

For the shortened sample, the raw *VIX* again drops out of the MCS under both loss functions and test statistics while the transformed *VIX* is only included in the MCS under MSE using the semi-quadratic statistic. The decline in relative forecasting ability of the transformed *VIX* as the time-horizon shortens is perhaps unsurprising due to the *VIX* being constructed as a 22-day-ahead forecast which may incorporate information not relevant to shorter forecasting periods. Hence, it seems that while the risk-adjusted *VIX* yields promising results for the forecast horizon the *VIX* is constructed for, it is most likely less useful over shorter horizons.

## 5 Conclusion

Issues relating to forecasting volatility have attracted a great deal of attention in recent years, with such interest undoubtedly piquing given the extreme variations observed in late 2008. As a result, many studies into the relative merits of



<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$	<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$
<i>VIX</i>	0.073	0.067	<i>ARMA</i>	0.000	0.000
<i>G</i>	0.073	0.072	<i>VIX</i>	0.000	0.000
<i>ARMA</i>	0.073	0.171	<i>G</i>	0.002	0.002
<i>GJR</i>	0.324	0.406	<i>GJR</i>	0.006	0.015
<i>GJRVIX</i>	0.699	0.606	<i>SV</i>	0.131	0.154
<i>SV</i>	0.699	0.666	<i>GRJRV</i>	0.232	0.425
<i>GRJRV</i>	0.699	0.803	<i>GRV</i>	0.953	0.938
<i>BGZ</i>	0.892	0.926	<i>BGZ</i>	0.953	0.938
<i>GVIX</i>	0.892	0.926	<i>RVMID</i>	0.953	0.938
<i>RVMID</i>	0.892	0.926	<i>GVIX</i>	0.953	0.938
<i>GRV</i>	1.000	1.000	<i>GJRVIX</i>	1.000	1.000

Table 4: MCS of 5-day ahead forecasts for shortened sample

<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$	<i>Model</i>	$MCS(T_R)$ $\hat{p}_i$	$MCS(T_{SQ})$ $\hat{p}_i$
<i>BGZ</i>	0.121	0.378	<i>ARMA</i>	0.000	0.000
<i>G</i>	0.139	0.428	<i>VIX</i>	0.000	0.000
<i>GVIX</i>	0.415	0.733	<i>G</i>	0.000	0.000
<i>GJRVIX</i>	0.421	0.765	<i>SV</i>	0.000	0.001
<i>GJR</i>	0.421	0.765	<i>GJR</i>	0.000	0.007
<i>SV</i>	0.689	0.871	<i>GRV</i>	0.016	0.049
<i>RVMID</i>	0.689	0.871	<i>BGZ</i>	0.051	0.074
<i>ARMA</i>	0.689	0.871	<i>RVMID</i>	0.051	0.074
<i>GRJRV</i>	0.942	0.934	<i>GRJRV</i>	0.051	0.074
<i>GRV</i>	0.942	0.934	<i>GVIX</i>	0.224	0.224
<i>VIX</i>	1.000	1.000	<i>GJRVIX</i>	1.000	1.000

Table 5: MCS of 1-day ahead forecasts for whole sample

<i>MSE</i>			<i>QLIKE</i>		
<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$	<i>Model</i>	$MCS(T_R)$	$MCS(T_{SQ})$
	$\hat{p}_i$	$\hat{p}_i$		$\hat{p}_i$	$\hat{p}_i$
<i>G</i>	0.019	0.086	<i>ARMA</i>	0.000	0.000
<i>ARMA</i>	0.086	0.151	<i>VIX</i>	0.000	0.000
<i>VIX</i>	0.120	0.240	<i>G</i>	0.000	0.000
<i>GJR</i>	0.120	0.372	<i>GJR</i>	0.000	0.001
<i>SV</i>	0.120	0.377	<i>SV</i>	0.001	0.002
<i>GJRVIX</i>	0.120	0.499	<i>GRV</i>	0.010	0.010
<i>RVMID</i>	0.120	0.499	<i>GRJRV</i>	0.012	0.020
<i>BGZ</i>	0.120	0.499	<i>RVMID</i>	0.064	0.038
<i>GRV</i>	0.657	0.643	<i>BGZ</i>	0.192	0.235
<i>GRJRV</i>	0.657	0.643	<i>GJRVIX</i>	0.941	0.941
<i>GVIX</i>	1.000	1.000	<i>GVIX</i>	1.000	1.000

Table 6: MCS of 1-day ahead forecasts for shortened sample

implied and model based volatility forecasts have been conducted, and although it has often been found that implied volatility offers superior performance, many studies disagreed. Recently, Becker & Clements (2008) showed that the *VIX* was statistically inferior to a combination of model based forecasts, inferring that the *VIX* does not represent an optimal combination of volatility forecasts. However, it was argued in this paper that these comparisons may not have been “fair” given that they have involved comparisons of risk-neutral implied volatility forecasts while model based forecasts are generated under the physical measure. Using the methodology of Bollerslev, Gibson & Zhou (2008), a transformed *VIX* forecast that incorporated the volatility risk-premium was generated and its forecast performance compared with model based forecasts of S&P 500 Index volatility via the model confidence set technology of Hansen *et al.* (2003a, 2003b)

The transformed *VIX* offered significant improvement in forecasting ability over the risk-neutral *VIX* in four of the six samples considered. When compared under the QLIKE loss function with model based forecasts of 22-day-ahead volatility, the risk-adjusted *VIX* was included in the model confidence set over both the shortened and full-length sample periods, implying it performs at least as well as model based forecasts. We put forth that this is a significant addition to the volatility forecasting literature relating to implied volatility.

Overall, this paper shows that if one correctly accounts for the volatility risk-premium, the market determined forecast of volatility over a 22-day horizon is in fact of equal predictive accuracy to a small number of model based forecasts. This result clarifies confusion from many prior studies which have, for the most part, typically conducted pair-wise comparisons and have neglected adjusting for the volatility risk-premium.

## References

- Andersen, T.G., Bollerslev, T., Diebold, F.X. & Labys, P. (1999). (Understanding, optimizing, using and forecasting) Realized Volatility and Correlation. Working Paper, University of Pennsylvania.
- Andersen T.G., Bollerslev T., Diebold F.X. & Labys P. (2001). The distribution of exchange rate volatility. *Journal of the American Statistical Association* 96, 42-55.
- Andersen T.G., Bollerslev T., Diebold F.X. & Labys P. (2003). Modeling and forecasting realized volatility. *Econometrica*. 71, 579-625.
- Andersen T.G., Bollerslev T. & N. Meddahi. (2004). Analytical Evaluation of Volatility Forecasts. *International Economic Review*, 45, 1079-1110.
- Andersen T.G., Bollerslev T. & N. Meddahi. (2005). Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities. *Econometrica* 73, 279-296.
- Barndorff-Nielson, O. & Shephard, N. (2004). A Feasible Central Limit Theory for Realised Volatility under Leverage. Manuscript, Nuffield College, Oxford University.
- Becker, R., & Clements, A. (2008). Are combination forecasts of S&P 500 volatility statistically superior? *International Journal of Forecasting.*, 24, 122-133.
- Becker, R., Clements, A. & White, S. (2006). Does implied volatility provide any information beyond that captured in model-based volatility forecasts?, forthcoming in *Journal of Banking and Finance*.
- Blair B.J., Poon S-H. & Taylor S.J. (2001). Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, 105, 5-26.
- Bollerslev, T., (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics.*, 31, 307-27
- Bollerslev, T., Gibson, M., & Zhou, H. (2008). Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. unpublished manuscript, Duke University.
- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion Using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, 109, 3365.
- Britten-Jones, M., & Neuberger, A. (2000) Option Prices, Implied Price Processes, and Stochastic Volatility. *Journal of Finance*, 55, 839-866.
- Canina, L. & Figlewski, S. (1993). The information content of implied volatility. *Review of Financial Studies*, 6, 659-681.

- Campbell, J.Y., Lo, A.W. & MacKinlay, A.G. (1997). *The Econometrics of Financial Markets*, Princeton University Press, Princeton NJ.
- Chicago Board of Options Exchange (2003) VIX, CBOE Volatility Index.
- Christensen B.J. & Prabhala N.R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, 50, 125-150.
- Clements, A.E., Hurn, A.S. & White, S.I. (2003). *Discretised Non-Linear Filtering of Dynamic Latent Variable Models with Application to Stochastic Volatility*, Discussion Paper No 179, School of Economics and Finance, Queensland University of Technology.
- Cochrane, J.H. (2001). *Asset Pricing*, Princeton University Press: Princeton, NJ.
- Day, T.E. & Lewis, C.M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics*, 52, 267-287.
- Ederington, L.H. & Guan, W. (2002). Is implied volatility an informationally efficient and effective predictor of future volatility?, *Journal of Risk*, 4, 3.
- Engle, R.F. & Ng, V.K. (1991). Measuring and testing the impact of news on volatility, *Journal of Finance*, 48, 1749-1778.
- Fleming J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices, *Journal of Empirical Finance*, 5, 317-345.
- Fleming, J., Ostdiek, B. & Whaely, R.E. (1995). Predicting stock market volatility: a new measure, *Journal of Futures Markets*, 15, 265-302.
- Glosten, L.R., Jagannathan, R. and Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.
- Gourieroux C. & Jasiak J. (2001). *Financial Econometrics*. Princeton University Press: Princeton NJ.
- Hansen, P.R., Lunde, A. & Nason, J.M. (2003). Choosing the best volatility models: the model confidence set approach, *Oxford Bulletin of Economics and Statistics* 65, 839-861.
- Hansen, P.R., Lunde, A. & Nason, J.M. (2003). *Model confidence sets for forecasting models approach*, Working Paper 2003-5, Brown University.
- Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6, 327-343.
- Jiang, G., & Yisong, T. (2005). Model-Free Implied Volatility and Its Information Content. *Review of Financial Studies*, 18, 1305-1342.

- Jorion P. (1995). Predicting volatility in the foreign exchange market. *Journal of Finance*, 50, 507-528.
- Koopman, S.J. and Jungbacker, B. & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12, 445-475.
- Lahiri, S.N. (1999) Theoretical comparison of block bootstrap methods, *Annals of Statistics*, 27, 386-404.
- Lameroueux, C.G. & Lastrapes, W. (1993). Forecasting stock-return variance: toward an understanding of stochastic implied volatilities, *Review of Financial Studies*, 6, 293-326.
- Meddahi, N. (2002). Theoretical Comparison Between Integrated and Realized Volatility. *Journal of Applied Econometrics*, 17, 479-508.
- Mincer, J. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts, In: Zarnowitz, J. (Ed.) *Economic Forecasts and Expectations*, National Bureau of Economic Research, New York.
- Newey, W. K., & West, K. D. (1987). A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703-708.
- Pantula, S.G. (1989). Testing for unit roots in time series data, *Econometric Theory*, 5, 256-271.
- Patton, A.J. (2005). Volatility forecast comparison using imperfect volatility proxies. Unpublished Manuscript.
- Patton, A. and Sheppard, K. (2007). Evaluating Volatility Forecasts, in *Handbook of Financial Time Series*, Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, T. eds., Springer-Verlag.
- Politis, D.N. & J.P. Romano (1994) The Stationary Bootstrap. *Journal of the American Statistical Association*, 89, 1303-1313.
- Poon S-H. & Granger C.W.J. (2003). Forecasting volatility in financial markets: a review. *Journal of Economic Literature*, 41, 478-539.
- Poon S-H. & Granger C.W.J. (2005). Practical Issues in forecasting volatility. *Financial Analysts Journal*, 61, 45-56.
- Taylor, S. J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press: Princeton NJ.
- White, H. (2000). A Reality Check For Data Snooping. *Econometrica*, 68, 1097-1127.

## List of NCER Working Papers

**No. 44** ([Download full text](#))

Adam Clements and Annastiina Silvennoinen

**On the economic benefit of utility based estimation of a volatility model**

**No. 43** ([Download full text](#))

Adam Clements and Ralf Becker

**A nonparametric approach to forecasting realized volatility**

**No. 42** ([Download full text](#))

Uwe Dulleck, Rudolf Kerschbamer and Matthias Sutter

**The Economics of Credence Goods: On the Role of Liability, Verifiability, Reputation and Competition**

**No. 41** ([Download full text](#))

Adam Clements, Mark Doolan, Stan Hurn and Ralf Becker

**On the efficacy of techniques for evaluating multivariate volatility forecasts**

**No. 40** ([Download full text](#))

Lawrence M. Kahn

**The Economics of Discrimination: Evidence from Basketball**

**No. 39** ([Download full text](#))

Don Harding and Adrian Pagan

**An Econometric Analysis of Some Models for Constructed Binary Time Series**

**No. 38** ([Download full text](#))

Richard Dennis

**Timeless Perspective Policymaking: When is Discretion Superior?**

**No. 37** ([Download full text](#))

Paul Frijters, Amy Y.C. Liu and Xin Meng

**Are optimistic expectations keeping the Chinese happy?**

**No. 36** ([Download full text](#))

Benno Torgler, Markus Schaffner, Bruno S. Frey, Sascha L. Schmidt and Uwe Dulleck

**Inequality Aversion and Performance in and on the Field**

**No. 35** ([Download full text](#))

T M Christensen, A. S. Hurn and K A Lindsay

**Discrete time-series models when counts are unobservable**

**No. 34** ([Download full text](#))

Adam Clements, A S Hurn and K A Lindsay

**Developing analytical distributions for temperature indices for the purposes of pricing temperature-based weather derivatives**

**No. 33** ([Download full text](#))

Adam Clements, A S Hurn and K A Lindsay

**Estimating the Payoffs of Temperature-based Weather Derivatives**

**No. 32** ([Download full text](#))

T M Christensen, A S Hurn and K A Lindsay

**The Devil is in the Detail: Hints for Practical Optimisation**

**No. 31** ([Download full text](#))

Uwe Dulleck, Franz Hackl, Bernhard Weiss and Rudolf Winter-Ebmer

**Buying Online: Sequential Decision Making by Shopbot Visitors**

**No. 30** ([Download full text](#))

Richard Dennis

**Model Uncertainty and Monetary Policy**

**No. 29** ([Download full text](#))

Richard Dennis

**The Frequency of Price Adjustment and New Keynesian Business Cycle Dynamics**

**No. 28** ([Download full text](#))

Paul Frijters and Aydogan Ulker

**Robustness in Health Research: Do differences in health measures, techniques, and time frame matter?**

**No. 27** ([Download full text](#))

Paul Frijters, David W. Johnston, Manisha Shah and Michael A. Shields

**Early Child Development and Maternal Labor Force Participation: Using Handedness as an Instrument**

**No. 26** ([Download full text](#))

Paul Frijters and Tony Beaton

**The mystery of the U-shaped relationship between happiness and age.**

**No. 25** ([Download full text](#))

T M Christensen, A S Hurn and K A Lindsay

**It never rains but it pours: Modelling the persistence of spikes in electricity prices**

**No. 24** ([Download full text](#))

Ralf Becker, Adam Clements and Andrew McClelland

**The Jump component of S&P 500 volatility and the VIX index**

**No. 23** ([Download full text](#))

A. S. Hurn and V.Pavlov

**Momentum in Australian Stock Returns: An Update**

**No. 22** ([Download full text](#))

Mardi Dungey, George Milunovich and Susan Thorp

**Unobservable Shocks as Carriers of Contagion: A Dynamic Analysis Using Identified Structural GARCH**

**No. 21** ([Download full text](#)) (forthcoming)

Mardi Dungey and Adrian Pagan

**Extending an SVAR Model of the Australian Economy**

**No. 20** ([Download full text](#))

Benno Torgler, Nemanja Antic and Uwe Dulleck

**Mirror, Mirror on the Wall, who is the Happiest of Them All?**

**No. 19** ([Download full text](#))

Justina AV Fischer and Benno Torgler

**Social Capital And Relative Income Concerns: Evidence From 26 Countries**

**No. 18** ([Download full text](#))

Ralf Becker and Adam Clements

**Forecasting stock market volatility conditional on macroeconomic conditions.**

**No. 17** ([Download full text](#))

Ralf Becker and Adam Clements

**Are combination forecasts of S&P 500 volatility statistically superior?**

**No. 16** ([Download full text](#))

Uwe Dulleck and Neil Foster

**Imported Equipment, Human Capital and Economic Growth in Developing Countries**

**No. 15** ([Download full text](#))

Ralf Becker, Adam Clements and James Curchin

**Does implied volatility reflect a wider information set than econometric forecasts?**

**No. 14** ([Download full text](#))

Renee Fry and Adrian Pagan

**Some Issues in Using Sign Restrictions for Identifying Structural VARs**

**No. 13** ([Download full text](#))

Adrian Pagan

**Weak Instruments: A Guide to the Literature**

**No. 12** ([Download full text](#))

Ronald G. Cummings, Jorge Martinez-Vazquez, Michael McKee and Benno Torgler

**Effects of Tax Morale on Tax Compliance: Experimental and Survey Evidence**

**No. 11** ([Download full text](#))

Benno Torgler, Sascha L. Schmidt and Bruno S. Frey

**The Power of Positional Concerns: A Panel Analysis**

**No. 10** ([Download full text](#))

Ralf Becker, Stan Hurn and Vlad Pavlov

**Modelling Spikes in Electricity Prices**

**No. 9** ([Download full text](#))

A. Hurn, J. Jeisman and K. Lindsay

**Teaching an Old Dog New Tricks: Improved Estimation of the Parameters of Stochastic Differential Equations by Numerical Solution of the Fokker-Planck Equation**



**No. 8** ([Download full text](#))

Stan Hurn and Ralf Becker

**Testing for nonlinearity in mean in the presence of heteroskedasticity.**

**No. 7** ([Download full text](#)) (published)

Adrian Pagan and Hashem Pesaran

**On Econometric Analysis of Structural Systems with Permanent and Transitory Shocks and Exogenous Variables.**

**No. 6** ([Download full text](#)) (published)

Martin Fukac and Adrian Pagan

**Limited Information Estimation and Evaluation of DSGE Models.**

**No. 5** ([Download full text](#))

Andrew E. Clark, Paul Frijters and Michael A. Shields

**Income and Happiness: Evidence, Explanations and Economic Implications.**

**No. 4** ([Download full text](#))

Louis J. Maccini and Adrian Pagan

**Inventories, Fluctuations and Business Cycles.**

**No. 3** ([Download full text](#))

Adam Clements, Stan Hurn and Scott White

**Estimating Stochastic Volatility Models Using a Discrete Non-linear Filter.**

**No. 2** ([Download full text](#))

Stan Hurn, J. Jeisman and K.A. Lindsay

**Seeing the Wood for the Trees: A Critical Evaluation of Methods to Estimate the Parameters of Stochastic Differential Equations.**

**No. 1** ([Download full text](#))

Adrian Pagan and Don Harding

**The Econometric Analysis of Constructed Binary Time Series.**