**Forecasting the volatility of the Australian Dollar using high frequency data: Does estimator accuracy improve forecast evaluation?**

Abstract

We compare forecasts of the volatility of the Australian Dollar exchange rate to alternative measures of ex-post volatility. We develop and apply a simple test for the improvement in the ability of loss functions to distinguish between forecasts when the quality of a volatility estimator is increased. We find that both realized variance and the daily high-low range provide a significant improvement in loss function convergence relative to squared returns. We find that a model of stochastic volatility provides the best forecasts for models that use daily data, and the GARCH(1,1) model provides the best forecast using high-frequency data.

Keywords: Volatility forecasting, exchange rate, Australian Dollar, stochastic volatility, realized variance, high-low range.

December 2018

**1. Introduction**

A key purpose of volatility models is to produce forecasts of volatility, which are useful for the pricing of financial assets that depend heavily on the evolution of volatility, such as derivatives, and also for input into economic policy analysis and forecasting.[[1]](#footnote-1) Because 'true' volatility is unobservable, volatility estimators are required when one wishes to undertake a comparison of forecasted values. However, the early studies of volatility forecast evaluation,[[2]](#footnote-2) found that the forecasts performed poorly when compared to the standard volatility estimator, squared daily returns. Andersen and Bollerslev (1998) demonstrated that more sophisticated estimators would increase the explanatory power offered by the volatility models, and recommended the use of volatility measured from high-frequency (intra-day) returns, which has become known as realized variance.

A key consideration in the evaluation of forecasts, more generally, is the selection of the loss function. Hansen and Lunde (2005) examined 330 ARCH related models to an exchange rate and a stock return series and found that different loss functions selected different models as providing the best volatility forecasts.[[3]](#footnote-3) Combining the two issues, Hansen and Lunde (2006) and Patton (2011) explore the theoretical and practical interactions between the selected loss function and the quality of the volatility estimator. Both papers identify those loss functions that are less sensitive to the noise in volatility proxies and show that more sophisticated measures of volatility are better able to distinguish between forecasts.

In this paper we also focus on the interaction between the choice of volatility estimator and the selection of loss functions in the evaluation of volatility forecasts, but seek to answer the following specific question. Does the use of more sophisticated volatility proxies generate greater convergence amongst the forecast rankings across different loss functions? This is useful knowledge in forecast evaluation because, if the answer is yes, it means that the choice of loss function matters much less than would otherwise be the case.

We contribute and innovate in number of important ways. Both Hansen and Lunde (2006) and Patton (2011) examine the volatility of IBM stock returns, over similar time frames, ending in 2003. We examine an exchange rate, specifically the Australian Dollar / US Dollar (AUD/USD) rate, and over recent years. Although, Hansen and Lunde (2005) did examine an exchange rate in their comparison of forecasts, they use only high-frequency measures of volatility and, like many volatility forecasting studies, used the German Mark (or Euro) / US Dollar rate. The AUD/USD is the fourth highest traded currency pair (BIS, 2013) and has approximately double the trading activity of the next two highest currency pairs.[[4]](#footnote-4) The AUD itself is the fifth highest traded currency, after the US Dollar, Euro, British Pound and Japanese Yen. Using this currency pair, which is still highly traded, avoids the pitfalls from over-working a dataset and has greater potential to provide new empirical findings. However, we recognise that such a narrow empirical application runs the risk that our findings may be sample specific. So, as a control, we additionally examine the AUD/USD at a different point in time (2 years apart), the EUR/USD exchange rate (as a different asset in the same asset class) and the S&P 500 index (as an asset in a different asset class).

While Patton examined forecasts from past squared returns, using either a rolling window average or a partial adjustment mechanism, and Hansen and Lunde (2006) examine forecasts from a small number of GARCH models, we consider a model of stochastic volatility (Harvey et al, 1994) as well as a selection of GARCH models, also relaxing the distributional assumptions in Hansen and Lunde (2006). Harvey et al (1994) showed that the stochastic volatility model fitted well to four US Dollar exchange rates, and comparisons to GARCH models for exchange rate data have been conducted by Heynen and Kat (1994), Dunis et al (2003) and Lopez (2001). However, all of these studies focus on the cross rates of the US Dollar with the German Mark (or Euro), the Japanese Yen, the British Pound, the Canadian Dollar and the Swiss Franc, and none use high frequency data. By contrast, Chortareas et al (2011) do use high frequency data, but only high frequency data, in their comparison of stochastic volatility to (symmetric) GARCH models for exchange rate data, and also for a subset of the aforementioned currencies. Moreover, the balance of evidence from these studies is unable to clearly distinguish between the forecasts from these two classes of model. So, our study will compare the stochastic volatility model to a selection of GARCH models using a fresh data set, of comparable trading activity levels, that includes both low and high frequency data. To our knowledge, our study is the first to examine forecasts of the volatility of the AUD/USD using the GARCH model applied directly to high frequency data and the first to compare this to forecasts derived from low frequency data.

We use the set of loss functions examined in Hansen and Lunde (2005), which is broader than, but encompasses, those used by Patton (2011) and Hansen and Lunde (2006). Our selection of volatility estimators is similar to Patton (2011). As well as examining squared returns and realized variance from high frequency returns, he also considers the range based variance estimator of Parkinson (1980). While he calibrates the efficiency losses of this volatility measure relative to those of squared returns and realized variance measures, like Hansen and Lunde (2006), he only uses realized variance and squared returns in his empirical application. We believe our study to be the first to include all three measures in the empirical application. Moreover, we focus on a different asset class and a much more recent period of time.

The question of whether there is increased convergence across loss functions of relative forecasting performance when higher quality volatility estimators are used is examined only indirectly by Hansen and Lunde (2006) and Patton (2011). Hansen and Lunde (2006) provide comparative performance tables for the forecasts from different models across different loss functions using either squared returns or realized variance, while Patton (2011) formally tests differences between such forecasts using the tests proposed by Diebold and Mariano (1995) and West (1996). In this study, we address the question directly, by conducting paired t-tests of the mean (across forecasts/models) of the standard deviation of the loss function rankings (across loss functions). A smaller value of this mean indicates a greater convergence of the rankings across loss functions, since converged rankings would generate a zero standard deviation across the loss functions for each model. We supplement this, with a related procedure, similarly constructed, that tests whether the ability of loss functions to distinguish between the first and second best models is enhanced by using more sophisticated variance estimators.

Our key findings are as follows. Using our test for ranking convergence, we find that using either a range based estimator (Parkinson, 1980) or a realized variance estimator rather than squared returns as the measure of ex-post volatility results in a significant increase in the convergence of loss function rankings. With better quality measures of volatility, loss functions converge more strongly on the preferred forecast. This is the case both for loss functions classified as robust by Patton (2011) and those not regarded as robust. Moreover, we also find a significant increase in the ability of loss functions to distinguish between the best and second best models when either the range measure or the realized variance are used. Thus, the margin by which the best model is chosen is significantly increased by using the higher quality volatility measures.

We find that there are significant differential gains in terms of loss function ranking convergence by using the realized variance measure, such that forecast comparison tests are more able to facilitate the rejection of one or more competing forecasts against a benchmark forecast. If squared returns are used as the measure of ex-post volatility, the forecast comparison tests are less able to distinguish between any of the competing forecasts. Taken together, these results also suggest that while the range measure may provide an improvement upon the use of squared returns, and so be a valuable substitute for realized variance when high frequency data is unavailable or too costly to collect, it is still not going to be as reliable as using realized variance.

We find that the regression based tests are not enhanced by the use of realized variance for forecasts derived from daily data, which is in contrast to early studies, such as Andersen and Bollerslev (1998) but is in line with more recent studies, such as McMillan and Speight (2012) and Chortareas et al (2011). This is the case both in regards to the regression *R2* and tests of forecast unbiasedness.

Regarding the specific characteristics of the Australian Dollar / US Dollar exchange rate, we find among the models applied to daily data that the stochastic volatility model is ranked highest by the loss functions, and is significantly better than the next best model when realized variance is used as the ex-post volatility measure. The best performing GARCH model out of sample is the GJR-GARCH model, which reflects significant asymmetries in the volatility in sample, but that this result is not stable, as in the earlier sample of the AUD/USD the base-line GARCH(1,1) model is preferred, despite also there being asymmetries in the volatility in sample. These asymmetries imply that an unanticipated depreciation of the Australian Dollar will have a greater impact of the volatility of the exchange rate than an unanticipated appreciation, relative to the US Dollar. In contrast to earlier studies, we find little evidence that power ARCH models add value to the modelling or forecasting of the exchange rate, although again this result appears to have some sample specificity. We also find no differences between forecasts generated assuming normally distributed errors as opposed to those from *t*-distributed errors, although the latter provides a significantly better fit in sample.

We find that a GARCH model of daily volatility that is generated from intra-day data provides forecasts that are superior to the stochastic volatility model from daily data for most, but not all, loss functions and variance estimators. This is not only the case for the AUD/USD, in both sample periods, but also for our two control samples. This result indicates that, as has been found in studies of other exchange rates and for other asset classes that intra-day data is the most valuable for both evaluating forecasts and also generating forecasts of daily volatility.

The remainder of the paper proceeds as follows. Section 2 reviews related literature to provide some context for our work and enable comparisons to our findings to be drawn. Section 3 describes the stochastic volatility model and the GARCH models that will be used to forecast the exchange rate, describes the computation of the ex-post volatility measures, and explains the forecast evaluation testing procedures. In Section 4, we describe the data and report the results of the model estimation in sample and the out of sample forecast evaluations. Section 5 summarizes and concludes.

**2. Related literature**

The literature on volatility forecasting and the evaluation of these forecasts is extensive, and has grown alongside the development of models of volatility.[[5]](#footnote-5) However, the literature on evaluating forecasts of exchange rate volatility is relatively small, and so presents further opportunities for discovery.[[6]](#footnote-6) Our work contributes to a number of literatures including, but not limited to, the literature on forecasting exchange rates, both with low and high frequency data, forecast comparisons using high frequency data, and to studies of specific financial markets, in this case the AUD/USD exchange rate.

Evaluating forecasts of the volatility of exchange rates was first undertaken by Taylor (1986) who compared forecasts from his AMARCH model with the exponentially-weighted moving average (EWMA) model in Taylor and Kingsman (1979), using daily data for the £/$ between 1974 and 1982, and marginally favouring the EWMA model. Forecast comparison has thereafter broadened out to evaluate forecasts from a variety of models in the GARCH class, stochastic volatility models, autoregressive models of squared and absolute returns and models including implied volatility from options. Bera and Higgins (1997), who examined the weekly $/£ rate between 1985 and 1991, favoured a GARCH model over a bilinear model, while Cumby et al (1993), who examine the weekly Yen/$ between 1977 and 1990 favoured an EGARCH model over historical measures of volatility. By contrast, Figlewski (1997), who examine the daily DM/$ rate over the period 1971 to 1995, favoured an historical measure over a GARCH model. Jorion (1995) who examined both the DM/$ and the Yen/$ using daily data between 1985 and 1992, also found that GARCH models added little incremental explanation of volatility compared to historical and implied measures. Comparing models within the GARCH class and a nonparametric estimator, and for 5 weekly observed cross rates with the US $ over the period 1973 to 1989, Lee (1991) found little difference between the forecasting performance of the alternative GARCH specifications, but better performance for the nonparametric specification. By contrast, West and Cho (1995) using an almost identical data set, found that the GARCH models outperformed a nonparametric estimator, but was inferior to a constant variance assumption, but that the differences were not statistically significant. Studies that have introduced forecasts from stochastic volatility models into the comparisons include Heynan and Kat (1994) and Dunis et al (2003). They use daily data for the most active cross rates against the USD (1980-1992) and the USD and DM (1990-1998), respectively, and find that GARCH based models forecast volatility better than stochastic volatility models. By contrast, Lopez (2001) using daily data for four cross rates against the US Dollar between 1980 and 1995 could not distinguish between the forecasts of GARCH based models and a stochastic volatility model. Not only are these studies inconclusive in aggregate as to a preferred volatility forecasting model for exchange rates, the performance of any of the models in these studies is low. The typical value of the coefficient of variation (*R2*) in their regression tests of forecast unbiasedness was less than 5 percent. [[7]](#footnote-7)

Andersen and Bollerslev (1998) proved that regression methods would anyway give low values of *R2* because they are noisy estimates of volatility. This also explained why models that seemed to fit well in sample were being found to fail out of sample. Specifically, they show that the value of *R2* in a volatility unbiasedness regression is bounded above by the inverse of the kurtosis of the underlying returns series. They show that intra-day returns can be used to construct a realized variance (RV) measure (from the daily sum of squared intra-day returns) that eliminates the noise in measurements of daily volatility. Using 5 minute and 1 hour returns on the German Mark / US Dollar, they find that the *R2* for forecasts from a GARCH(1,1) model increased to 0.48 and 0.33, respectively. In a series of papers, Anderson et al (2001, 2003) and Barndorf-Nielsen and Shephard (2002a,b, 2004a,b) develop further the theoretical basis for using realized variance as a measure of the daily conditional variance and provide further empirical validation. Specifically, these papers show that the ex-post value of realized variance is an unbiased estimate of the conditional return variance of returns for log price processes that are special semi-martingales – in which class fit most financial models.

Forecasts of exchange rate volatility measured as intra-day realized variance have also been examined by, for example, Taylor and Xu (1997), Martens (2001), Li (2002), Pong et al (2004), Chortareas et al (2011) and McMillan and Speight (2012). Taylor and Xu (1997) find that realized variance from 5 minute returns on the DM/$ during 1992-93 contains incremental information for forecasting beyond that contained in implied volatility. Martens (2001) compares forecasts constructed using different intra-day intervals and finds that the higher the frequency the better the out of sample performance. Li (2002) finds that implied volatility has no incremental information relative to high frequency data, for forecasts of the volatility of the US Dollar against each of the British Pound, Japanese Yen and the German Mark, between 1994 and 1999. Pong et al (2004) find that high frequency data applied to relatively simple models can generate more accurate forecasts than those from more complex models that use lower frequency data. Chortareas et al (2011) examine four Euro exchange rates (US Dollar, Japanese Yen, British Pound and the Swiss Franc) between 2000 and 2004 and find better forecasting performance from those models that use intra-day (15 minute) returns than those that use daily returns. McMillan and Speight (2012) also examine three Euro exchange rates (against the US Dollar, British Pound and Japanese Yen) for the period 2002 to 2006. They find that forecasts from GARCH models that use intra-day data outperform those that use daily data. The contribution and focus of all these studies is, however, in showing that intra-day returns contain valuable information for exchange rate volatility forecasting, specifically that intra-day returns are a valuable input (to improving the volatility forecast) rather than (also) being a valuable output (as an improved measure of ex-post volatility).[[8]](#footnote-8) All of these studies take the use of realized variance as the ex-post measure of volatility as a given. By contrast, our focus is on showing that high frequency data can help exchange rate forecast evaluation by generating convergence among the rankings from different loss functions. Nonetheless, we also include forecasts generated from high frequency data to compare these to those from daily data.

By contrast to the US Dollar, the British Pound, the Euro (or the German Mark prior to 1999) and the Japanese Yen, the Australian Dollar has received much less attention. McKenzie (1997) documents the presence of ARCH effects in 21 bilateral Australian Dollar exchange rates, finding greater effects in daily data than lower frequency data. McKenzie and Mitchell (2002) feature the Australian Dollar / US Dollar rate among a study of 17 heavily traded exchange rates between 1986 and 1997, using GARCH models. They find an asymmetry in the impact of shocks to the volatility process such that an unanticipated depreciation of the Australian Dollar has a bigger impact on future volatility than an unanticipated depreciation of the US Dollar. A similar asymmetry is reported in Villar (2010) who examines this same bilateral exchange rate, including more recent data, for the period 1994 and 2007. Comparative volatility forecasting performance, but for the Australian Dollar versus the Malaysian Ringit for the period 2010 to 2011, is considered by Ramasamy and Munisgamy (2012). They find no asymmetry and report extremely large forecasting errors. McMillan and Speight (2004) include the Australian Dollar / US Dollar exchange rate among their data for 1990 to 1996 for forecast comparisons between three GARCH and two historical volatility measures and find that the preferred forecast is highly sensitive to the loss function, even with the use of a high frequency measure of ex-post volatility. Our paper adds to this evidence by considering alternative measures of ex-post volatility, including stochastic volatility as well as GARCH based models, generating forecasts from both daily and intra-day data, calculating a wide range of loss functions and focussing on the convergence in loss function rankings when using high-frequency ex-post measures of volatility.

**3. Methods**

**3.1 GARCH models**

Forecasts for the volatility of the USD/AUD exchange rate will be generated from a set of GARCH models and a model of stochastic volatility. The models employed are listed in Table 1. GARCH based models, introduced by Engle (1982) and Bollerslev (1986), are autoregressive models of the conditional variance of a time series that depend on the squared residuals from an underlying model of the conditional mean of the time series. It is typical to parameterize the conditional mean of a time series of log changes in an exchange rate by either a low order ARMA process or a constant mean white noise process, as these series display little if any autocorrelation, see, for example, Taylor (2005).

So, the general structure of the models that we consider is

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where is the log change in the daily closing value of the exchange rate, is the constant mean, and the residual has mean zero and conditional variance , given by one of the GARCH or stochastic volatility models in the Table 1. We assume that the residuals are independently and identically normally distributed for all models, and additionally consider the possibility that the residuals may be *t*-distributed for three of the models.[[9]](#footnote-9)

Our base line volatility model, is the GARCH(1,1) specification, denoted “GARCH”, in Table 1. In this model the conditional variance depends on the size of its past lagged squared residuals, through the parameter , and on its own past values, through the parameter . While most applications of the GARCH model consider the first lagged squared residuals and lagged own values, Hansen and Lunde (2005) did consider second lags in their forecast comparisons. However, they found that the higher order ARCH-type models rarely performed better out-of-sample than lower order lag alternatives, so we confine our study to the GARCH(1,1) specification.

In an early application of GARCH models to exchange rates, Bera and Higgins (1992), explored a non-linear version of the GARCH model, now known as the Power GARCH (or PGARCH) model. This permits the lagged residuals and the conditional standard deviation to take any positive power, including the special case of 2, which is the GARCH model, and where the power is a free parameter to be estimated. In their study of 6 US Dollar exchange rates (against the Canadian Dollar, French Franc, Swiss Franc, German Mark, British Pound and Japanese Yen) over the period 1973 to 1985, they found that the non-linear model fitted the data better than the linear (GARCH) model for some of the exchange rates.

A key consideration in the estimation of GARCH models is ensuring that the parameter estimates are constrained to prevent negative estimates of the conditional variance. Nelson (1991) proposed an exponential form of the GARCH model, now known as the EGARCH model, which models the natural log of the conditional variance rather than the conditional variance directly. The model also introduced the possibility of asymmetry in the volatility process, whereby negative shocks could have a different effect on future volatility than positive shocks. The EGARCH model captures effect of the sign of the past residual through the parameter and the effect of the size of the shock through the parameter . If , then negative shocks will cause the conditional variance to rise more than it does in response to positive shocks.

The so-called “leverage effect”, whereby an increase in corporate leverage can increase both the required return (which induces a decrease in the stock price) and the risk of equity, provides a theoretical basis for a negative relationship between stock returns and subsequent stock return variances, see Black (1976) and Christie (1982).[[10]](#footnote-10) By contrast, for exchange rates, the possibility of an asymmetric relationship between the sign of unanticipated returns and variance is largely an empirical matter.[[11]](#footnote-11) However, evidence of such an asymmetry can be interpreted as a differential volatility impact of domestic versus foreign unanticipated currency depreciations. What is striking among those studies that identify an asymmetry is that they all find that the impact of the local currency depreciating against a more major currency has a bigger impact on the exchange rate volatility than a depreciation of the major currency.Hu et al (1997) attribute this to the greater likelihood of there being a policy intervention in this situation.

We test for the presence of asymmetric effects of shocks to the conditional variances using the sign and size bias tests of Engle and Ng (1993), together with a preliminary examination of the correlation between returns and squared returns. The sign bias test indicates whether asymmetric effects may be present in the variance, while the size bias tests indicate whether the size of the shock influences the asymmetric effect of the shock. The coefficients in the following equations are estimated

|  |  |  |
| --- | --- | --- |
| Sign Bias Test:  |  | (2a) |
| Negative Size Bias Test: |  | (2b) |
| Positive Size Bias Test: |  | (2c) |

where is the residual from the Gaussian symmetric GARCH(1,1), is an indicator variable that takes the value 1 if and is zero otherwise, , and is an i.i.d error term.[[12]](#footnote-12) The presence of significant , coefficients would indicate that there are asymmetries present that are not accommodated by the symmetric GARCH model.

The asymmetric GJR-GARCH model, Glosten et al (1993), proposes adding an interaction variable to the standard GARCH(1,1) model. This variable comprises the lagged squared residual multiplied by an indicator variable that takes the value unity when the lagged residual is positive, and takes the value zero when the lagged residual is not positive. The asymmetry works through the parameter Similarly to the EGARCH model, if , then negative shocks will cause the conditional variance to rise more than it does in response to positive shocks.[[13]](#footnote-13)

The threshold GARCH (TGARCH) model of Zakoïan (1994) also captures the sign of shocks with an interaction term, but adds this to the AMARCH model of Taylor (1986), which models the conditional standard deviation (as a function of the absolute value of the residuals) rather than the conditional variance. Again, if the coefficient on this interaction term is negative, then negative shocks will cause the conditional variance to respond more than it does in response to positive shocks.

Building upon the work of Taylor (1986), Schwert (1989) and the PARCH model of Bera and Higgins (1992), Ding, Engle and Granger (1993) proposed an asymmetric power ARCH (APARCH) model. In their study of the S&P 500 daily closing prices from 1928 to 1991, they estimated the value for the power function to be 1.43 which means that the process lies between two models, namely the GARCH(1,1) d=2 and the TGARCH model, d=1. Validation of this model for exchange rate volatility is provided in studies by Tse and Tsui (1997) and McKenzie and Mitchell (2002). Both this model and the PARCH model have been shown to be able to capture well the long lag autocorrelation of some volatility series.[[14]](#footnote-14) This is the final GARCH based specification that we examine and, once again, a negative sign on the term capturing asymmetry indicates a greater response to negative shocks.

**3.2 Stochastic Volatility**

This paper will consider the simple SV model from Ruiz (1994) and Harvey, Ruiz and Sheppard (1994), in which the log of the conditional variance, , is assumed to follow a first order autoregressive process,

|  |  |
| --- | --- |
|  | (3) |

where the residual in equation (1) is now , where and where . Taking the log of the squared residuals, a state space form of the model can be generated, where the measurement equation is

|  |  |
| --- | --- |
|  | (4) |

and the transition equation is given by equation (3). In the measurement equation, and is a non-Gaussian, zero-mean white noise. It is assumed that and are uncorrelated (see Harvey, Ruiz and Shephard 1994, p.261), and the autoregressive parameter in the volatility equation satisfies, .[[15]](#footnote-15) Since we assume that , we know that the mean and variance of are and , respectively (Abramowitz and Stegun 1970), and so also has variance of . This model is therefore time invariant for the variances of and . Estimation of the state space form of the SV model is undertaken using the Kalman Filter, which is quasi-maximum likelihood.

**3.3 Ex-post Volatility**

We will use and compare three measures of ex-post volatility, squared daily returns, realized volatility calculated from intra-day returns and the range-based estimator of Parkinson (1980). For a trading day with daily returns, , that are the sum of intraday returns, the realized volatility for day , , is defined as,

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

where are the squared intraday returns, . In selecting the frequency for the intra-day observations, there is a trade-off between the increase in accuracy from increased frequency and the increase in microstructure frictions as the frequency increases. While Andersen et al (2003) in their study of two major exchange rates recommend 30 minute observations, more recent studies, such as Liu et al (2015) have suggested that 5 minute intervals are the best choice. Bandi and Russell (2006) suggest that the optimal sampling frequency for the UK is between 0.4 and 13.8 minutes. We select 15 minutes, as being close to the upper bound, and note that this interval has precedent for exchange rate forecast comparisons in Chortareas et al (2011), who use it in their forecasting models, and should be mostly free of microstructure noise.[[16]](#footnote-16)

The range-based estimator proposed by Parkinson (1980), which we denote by , uses the daily high and low prices and is defined as

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

where and are the daily high and low prices respectively.[[17]](#footnote-17)

**3.4 Volatility forecasts from Intra-day returns**

Although the focus of this paper is on determining whether volatility calculated from intra-day returns is the best measure of ex-post volatility against which to evaluate forecasts, in the sense that, as we will show, different loss functions are more likely to give consistent rankings of alternative forecasts, prior research also indicates that volatility forecasts generated from intra-day data are superior to those that are generated just from daily data. Therefore, we augment our analysis of forecasts based on daily data with a GARCH model that generate daily volatility forecasts from intra-day data.

 The GARCH model applied to intra-daily returns requires an accommodation of the fluctuations of volatility within the 24-hour interval. We use an adjustment proposed by Taylor and Xu (1997) that involves dividing each 15-minute return by the square root of the average squared return for that same interval, where the average is taken across the entire estimation sample period for the given 15-minute interval. These rescaled returns are used in the GARCH model to produce the forecasts of 15-minute variance. These forecasts are then transformed back into those from the original returns by reversing the rescaling procedure before being summed within each 24-hour period to generate daily variance forecasts. The volatility specification for the GARCH model is in Table 1, and the model additionally incorporates an MA(1) term, see equation (1), to control for any remaining microstructure noise elements.

**3.5** **Forecasting volatility and Forecast Evaluation**

The dynamics of the volatility processes in Table 1 provide for the estimation of one-step ahead forecasts of volatility in the (out-of-sample) evaluation period, using the estimated parameter values for the estimation period. We use four procedures for forecast evaluation: Mincer-Zarnowitz (MZ) regressions, loss function scores, forecast comparison tests and new tests of loss function ranking convergence.

In the regression test method of Mincer and Zarnowitz (1969), a measure of (ex-post) variance,, is regressed on a constant and the predicted variance from a postulated model,, equation (7), where the residual is assumed to be mean zero and serially uncorrelated.

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

An unbiased forecast will have and , while the value from the regression provides a simple measure of thelevel of predictability in the volatility process. We estimate this model for each of the GARCH and SV models in turn with each of the three different ex-post estimators, squared returns, the range measure and realized variance. Since returns are heteroscedastic, squared returns will be even more heteroscedastic and so robust standard errors are calculated, although this may not be adequate since both sides of the regression are affected.

Since estimation error in can cause downward bias in the estimates of , and the values do not penalize biased forecasts, these regressions are usually accompanied by loss function calculations. Loss functions measure the distance between the observation and forecast, and have been further adapted to penalize more extreme distances more heavily, by using quadratic functions, and to include asymmetry whereby downside and upside losses have different weightings. Patton (2011) has shown that the loss function rankings of forecasts produced by competing models are dependent on the quality of the estimator of ex-post volatility. Poor proxies for ex-post volatility can lead to the selection of not the best forecasting model. Our use of three different measures of ex-post volatility is designed to see whether the results (the rankings of models) from loss function evaluations show more convergence with more sophisticated volatility measures. We use the following set of loss functions, from Hansen and Lunde (2005):

|  |  |
| --- | --- |
|  | (8) |
|  | (9) |
|  | (10) |
|  | (11) |
|  | (12) |
|  | (13) |

where is the number of forecasts in the evaluation period. The MAE criteria (8) and (9) are likely to be more robust to outliers than the MSE criteria (10) and (11). The MSE criterion (11) is equivalent (only if the constant term, ) to using the from the MZ regression, equation (7). An MZ regression using logs (rather than levels) of variances is similarly equivalent to using . , criterion (12), corresponds to the loss function implied by a Gaussian likelihood. We are motivated to use a wide range of loss functions from the observation of Poon and Granger (2003) that when loss-functions use variances, it becomes increasingly difficult to find a significant difference between competing forecasts, and that this is compounded when they are squared (as for ).

While loss functions provide a ranking, they do not determine whether, say, the first best and second best models produce forecasts that are statistically significantly different. To determine whether this is the case, we apply the modified Diebold-Mariano (DM) test of Harvey, Leybourne and Newbold (1998) (HLN-DM) to the three highest ranked forecasts from the MZ regressions and the loss function tests. The test can itself use a variety of functional forms to measure the distance between two sets of forecasts and we report results of both a quadratic and an absolute distance function. The objective is to see whether this test is more often able to distinguish between competing forecasts when the accuracy of the volatility estimator increases. More recently, multiple comparison tests have been developed, such as the Model Confidence Set (MCS) test of Hansen et al (2003), and its forerunners, the test of Superior Predictive Ability (Hansen, 2003) and Reality check test (White, 2000).[[18]](#footnote-18) The MCS test determines whether, with a certain level of confidence, that a set of forecasts contains the forecasts from the best of the competing models. These tests though also require the loss function to be chosen and, as has been shown by Hansen et al (2003), the results of forecast comparison tests are not independent of the choice of the loss function.

**3.6 Testing the Convergence of Loss Function Rankings**

The dependence of the results of forecast comparison tests on the choice of loss function motivates our focus on quantifying the convergence of the rankings among different loss functions, and determining whether the degree of convergence is influenced by the quality of the ex-post volatility measure. For if there is greater convergence among loss function rankings, it will matter less which loss function is used in forecast comparison tests. To address this issue, we examine the mean (across forecasts/models) of the standard deviation of the loss function rankings (across loss functions). A smaller value of this mean indicates a greater convergence of the rankings across loss functions, since converged rankings would generate a zero standard deviation across the loss functions for each model. Our test is constructed as follows. Let the rank of loss function for model using volatility estimator be denoted by , , and and similarly for volatility estimator be denoted by . The null hypothesis is that volatility estimator provides the same amount of convergence among the loss functions as volatility estimator , and the alternative hypothesis is the volatility estimator provides for greater convergence than volatility estimator . Our null hypothesis is formally that , where

|  |  |
| --- | --- |
|  | (14) |

where , and is similarly defined. The null hypothesis can be tested using a one-side paired sample t-test. By applying this test across the different volatility proxies, we provide direct evidence on the improvements in convergence from using higher quality data.

We supplement this evidence, with further tests that determine whether the ability of loss functions to distinguish between the first and second best models is also enhanced by using more sophisticated variance estimators. We regard a volatility estimator as more able to facilitate loss functions to distinguish between two forecasts if the percentage difference between the loss function values is greater. Specifically, let be the percentage difference between the loss function rank for the first and second best models for loss function using volatility estimator , and be similarly defined for volatility estimator . Then, a test of the null hypothesis that two volatility estimators provide equal facility for loss functions to distinguish between two forecasts, is a test of , against the alternative that , where , and . The mean for volatility estimator is defined similarly. The null hypothesis is tested using a paired sample t-test.

**4. Data and Results**

**4.1 Data**

Daily data for the AUD/USD exchange rate (Australian Dollars per one US Dollar), the EUR/USD (Euros per one US Dollar) and the SPX (the S&P500 index) was collected from the Bloomberg Professional service. From these we created four samples of 1639 observations, divided into a 1540 observation estimation period and a 99 observation forecast evaluation period at the end of each sample. For the AUD/USD and EUR/USD the full samples run from 09/04/12 to 18/07/2018, with the evaluation period starting on 02/03/2018. For the SPX, the sample runs from 27/02/12 to 30/07/2018, with the evaluation period starting on 09/03/2018.[[19]](#footnote-19) For the AUD/USD, we also consider a sample period of the same length and calendar location but that is displaced in time by two years, that is, from 06/04/2010 to 18/07/2016.[[20]](#footnote-20) This provides a control for sample specificity in using only the 2018 evaluation period, and provides some indication of the stability of the results for the AUD/USD. These data are used to estimate the models (in the evaluation period) and generate their forecasts for the evaluation period, as well as construct two of the variance estimators for the evaluation period; the squared returns and the high-low range measure.

For the ex-post realized variance estimator of volatility, we collected 15-minute intra-day observations for the 99 days of the evaluation period, again from Bloomberg Professional, for each of the three assets considered. The intra-day period for the exchange rate series runs from 22:00 hours the prior day to 21.45 hours on the day, UK time (GMT+1 during this period), the latter time point coinciding with the daily closing observation, and comprises up to 9493 observations. We do not include observations between 22:00 hours on Friday to 21.45 hours on Sunday, to reduce any noise from weekend trading, and adjustments for the start and end of daylight savings times.[[21]](#footnote-21) The intra-day period for the SPX runs from 14:00 hours on the day until 21:00 hours on the day, UK time (GMT +1 during this period), and comprises 2672 observations over the 99 day evaluation period.

For the estimation samples for the GARCH model applied to the intra-day data, we were constrained by the availability of data from our data platform and so for the AUD/USD and EUR/USD exchange rate, we were able to collect just 3936 15-minute observations, while for the SPX we could only collect 999 15-minute observations.[[22]](#footnote-22)

**4.2 Summary Statistics**

Summary statistics for daily returns, log differences in the exchange rate or index value, are shown in Table 2. For the exchange rate series, positive returns reflect a strengthening of the US Dollar (weakening of the Australian Dollar or Euro), while negative returns reflect a strengthening of the Australian Dollar or Euro (weakening of the US Dollar). The daily mean returns for the exchange rates are not significantly different from zero (p>0.220), while the daily variances correspond to an annualized standard deviations ranging from of 8.20 percent for the EUR/USD to 12.10 percent for the SPX. The Jarque-Bera test (JB Test) indicates that the returns series are not normally distributed (p<0.001), displaying excess kurtosis and a small positive skew for the AUD/USD and a negative skew for the EUR/USD and the SPX. An augmented Dickey-Fuller test (ADF Test) strongly rejects the null hypothesis of non-stationarity in the returns series (p<0.001). A test for ARCH effects in the returns series also rejects the null hypothesis of no ARCH effects (p<0.044).

There is no significant autocorrelation in the daily returns at lags 1 to 40, in all the series, except for the SPX at some high lags, which can be seen in Table 3. By contrast, the autocorrelations of squared returns, also reported in Table 3, are statistically significant. The non-normality in returns, the significant autocorrelation in squared returns and the results of the ARCH test all point to the returns series being characterised by changing volatility. The correlation between returns and squared returns for the AUD/USD is 0.084 (0.064 for the 2016 ending sample period) and is significantly different from zero (p<0.011). For the EUR/USD and the SPX, the correlations are negative and significant, at -0.077 (p<0.001) and -0.162 (p<0.006).

The results of the sign bias and size bias tests are reported in Table 4 and confirm the characteristic that for the AUD/USD positive shocks are associated with a bigger impact on volatility, than negative shocks. This means that an unanticipated depreciation in the Australian Dollar has a bigger effect on volatility than an unanticipated depreciation in the US Dollar. Specifically, we find that the coefficient, , in the sign bias test is negative in both samples, although only significant in the 2016 ending sample, indicating a negative relationship between the sign of lagged shocks and current squared residuals. This result is supported by the positive size bias test, coefficient , that indicates that the size of the positive shocks (unanticipated depreciation of the Australian Dollar) also influences current exchange rate volatility. These results are consistent with earlier findings for the AUD/USD rate by McKenzie and Mitchell (2002) for the period1986 to 1997 and Villar (2010) for the period 1994 to 2007.[[23]](#footnote-23) By contrast, there is less evidence of asymmetry for the EUR/USD, where only the size of the negative shock seems to influence volatility, while for the SPX there is evidence of the “leverage” effect in equity returns.

**4.3 Volatility Model Estimation**

The coefficients of the competing forecasting models are given in Table 5, panels (a) to (d), estimated using data for the “estimation” period of the sample for each asset’s return series. We first describe the results of the models applied to daily data. Although the simplest model reported is the GARCH(1,1), we also examined ARCH models with up to 15 lags. Both the AIC and BIC selected the GARCH(1,1) as the preferred model against any of the ARCH alternatives. The GARCH likelihoods are maximized using the BHHH and BFGS algorithms in turn. For the Stochastic Volatility model, the estimation is done using a Kalman filter, with the quasi-likelihood maximized using the BHHH and Newton Raphson algorithms.

The estimated coefficients from all the models and series meet their respective conditions for stationarity of the variance processes, and are indicative of a high level of persistence in volatility across all specifications. The coefficients are stable across the alternative distributional assumptions, showing very little numerical differences and similar statistical significance, except in the case of the EUR/USD where also the t-distributed EGARCH model defied all attempts to achieve convergence of parameter values. However, likelihood ratio tests indicate that the t-distributed versions in general provide a significantly better fit to the data (p<0.01), in sample.

Likelihood ratio tests comparing the power models to their closest GARCH counterpart are unable to reject the null hypothesis that the models provide an equal fit to the data (p>0.65). This finding is similar to that of Tse and Tsui (1997), who were examining the Malasian Ringit and Singapore Dollar against the US Dollar with data from 1978 to 1994. However, this finding contrasts with those of Bera and Higgins (1992) who found that, with the exception of the GBP/USD rate, the values of the power function were all less than 1, significantly so in three out of the five US Dollar exchange rates considered. The GBP/USD rate was not significantly different from either 1 or 2, while the Swiss France / US Dollar rate was not significantly different from either 1 or 0. For data between 1986 and 1997, McKenzie and Mitchell (2002) report a power function value of 1.295 for the AUD/USD, while Villar (2010) for data between 1994 and 2007 estimates the power to be 1.279. Our results for the sample period ending in 2018, of a power function value of 1.207 (for the symmetric model) and of 1.688 (for the asymmetric model), are similar to these earlier studies. However, for the sample ending 2016, we find values much closer to 2, which is equivalent to the baseline GARCH model, and suggests that the applicability of power ARCH models for the AUD/USD exchange rate may have some sample dependency.

The asymmetric models all have positive coefficient values on the term that captures asymmetry, which is consistent with the findings from the sign and size bias tests and suggests that an unanticipated depreciation of the Australian Dollar has a bigger impact on exchange rate volatility than an unanticipated appreciation, against the US Dollar. We find this effect also in the earlier sample period for the Australian Dollar and in the EUR/USD. By contrast, the negative coefficients for the SPX are consistent with a “leverage” effect. The coefficients are all statistically significant, except for the APGARCH model, which again suggests that this non-linear specification is currently less well suited to the AUD/USD exchange rate. Indeed, for the SPX series neither the APGARCH nor the TGARCH model were able to achieve parameter convergence and so are not included in Table 5(d).

We explore the differences between the asymmetric specifications by calculating the news impact curves for each of the specifications. Pagan and Schwert (1990) first suggested news impact curves which graphically demonstrate how different models react to positive and negative shocks. The process involves substituting positive and negative values of the standardized residuals into an estimated model and then plotting the predicted variance against the lagged standardized residuals, where the lagged conditional variance is set equal to the unconditional variance. The news impact curves for AUD/USD are shown in Figure 1, where it can be seen that the APGARCH and GJR-GARCH models generate more asymmetry than the EGARCH and TGARCH models.[[24]](#footnote-24) Moreover, log likelihood tests cannot distinguish between the in sample fit of the APGARCH and GJR-GARCH models (p=0.75, and for the 2016 sample, p=0.53). However, the EGARCH model provide a significantly better in sample fits against any of the other three alternative asymmetric models. For the 2016 ending sample, the TGARCH model is not significantly different from the EGARCH model (p=0.32) and is significantly better than either of the APGARCH or GJR-GARCH models (p<0.01). These results indicate that while there is some asymmetry to be modelled, the models with more modest adjustment (relatively less asymmetry) are preferred, in sample. By contrast, for the EUR/USD, the GJR-GARCH model is significantly worse than the three other asymmetric models, that cannot themselves be distinguished at usual significance levels. For the SPX, the EGARCH model is preferred to the GJR-GARCH model.

The estimated coefficients of the stochastic volatility model applied to the daily data series and those for the GARCH model applied to the intra-daily data series are also given in Tables 5(a) to 5(d). Here, we can see that the GARCH models applied to intra-daily data display a lower degree of persistence in volatility than when the model is applied to the respective daily data.

**4.4 Volatility Forecast Evaluation**

Summary statistics of the measures of ex-post volatility are shown in Table 6, for the evaluation samples for each of the four series, and show that the three measures of ex-post volatility display the skewness and excess kurtosis characteristics that are well- established features of high frequency exchange rate data.

*4.4.1 Mincer-Zarnowitz Regression Tests*

The results of the MZ regressions are shown in Table 7, panels (a) to (d). Considering first the AUD/USD over the period ending 2018, then among the models applied to daily data there is no consistent picture as to the ranking of the models (by value, in parentheses below the value) emerging from the different measures of ex-post volatility. However, the finding that the GARCH model is preferred when using realized volatility as the variance estimator, while power models are preferred when using squared returns, is consistent with differences between prior studies, for example between McKenzie and Mitchell (2002) and McMillan and Speight (2004). However, all of the coefficients on the predictors take the wrong sign, with many significantly negative, and the goodness of fit is poor across all models and estimators. While the values for the GARCH models do improve when the ex-post volatility measure is changed from squared returns to the range measure, they do not improve with the use of realized volatility. This is in sharp contrast to earlier studies, such as Andersen and Bollerslev (1998), who use the observed increase in for their data to, in part, empirically validate the use of realized volatility in place of squared returns. However, studies using more recent exchange rate data, such as McMillan and Speight (2012) – who specifically note this result – and Chortareas et al (2011), document similarly small values of for GARCH models using realized volatility as the ex-post measure.[[25]](#footnote-25) By contrast, the for the GARCH model applied to intra-day data is remarkably better than any of the models applied to daily data, and shows a six fold increase when realized volatility is also the ex-post variance estimator. However, for the range based variance estimator and for realized volatility, the Wald test for forecast unbiasedness is rejected (p<0.002) for all the models, including the GARCH model applied to intra-day data. Only if squared returns are used as the variance estimator does the stochastic volatility model and the intra-day GARCH model show signs of unbiasedness, but this result is largely an artefact of relatively large standard errors on the coefficients of the MZ regression equation.

 The results for the AUD/USD sample ending 2018 are different to those for the 2016 ending sample in that there is more agreement in the earlier sample across the alternative measures of ex-post variance that an asymmetric model may be preferred. Again the increases in are modest when switching to use realized variance as the variance estimator, and a similar pattern is observed regarding forecast unbiasedness. Unbiasedness tends to be driven by large standard errors on the estimated coefficients and is found in the models that offer the least good fit according to the MZ regressions.

 The limitations of the MZ regression as a forecast evaluation tool is further evidenced by the results for the EUR/USD and the SPX. Again, among the models applied to daily data, the rankings of the models and the unbiasedness of the forecasts are heavily dependent on the choice of variance estimator. It is only for the GARCH model applied to intra-day data that a substantial increase in is observed when switching to use realized variance as the ex-post variance estimator. For the case of using realized variance, for the SPX, the forecasts from the intra-day GARCH model also appear to be unbiased, although this is not the case for the EUR/USD.

*4.4.2 Loss Function Scores*

The values of the loss functions for each of the estimators and for each of the predictor models are shown in Tables 8 (a) to (d) and should provide more robust results than the MZ regressions.[[26]](#footnote-26) Among the models using daily data for AUD/USD (2018 sample), the loss functions rank the stochastic volatility model as the best model, for all loss functions for realized variance and the high-low estimator, and for four of the six loss functions for the squared returns estimator. For the 2016 sample, the stochastic volatility model is again the best performing model, indicating stability in this finding for the AUD/USD. While this result is in contrast to the rankings implied by the R-squared values obtained from MZ regressions, it was only the stochastic volatility model that provided evidence of unbiased forecasts. In the 2018 evaluation sample, the next best performing models are the GARCH model and the GJR GARCH model, with the latter only preferred over the former for the more recent sample period and when using the realized variance estimator. Moreover, the symmetric GARCH model is preferred over the alternative asymmetric models. For the 2016 sample, there appears to be some interaction between the choice of variance estimator and the loss function rankings. For the realized variance estimator, the symmetric GARCH and PGARCH specifications produce smaller losses than the asymmetric EGARCH and TGARCH specifications across all loss functions. However, for the squared returns and the high-low estimators, the MAE loss functions produce smaller losses for the asymmetric models. However, since the realized variance estimator has the highest kurtosis (Table 6), consistent with a greater presence of outliers, the MAE loss function, which is more robust to outliers than the other loss functions, should be the most reliable in that instance. This suggests that the symmetric models are preferred for out-of-sample forecasting, even though the asymmetric models seemed to provide the better fit in-sample. This finding is echoed even within the asymmetric models themselves, when those models that permit a more pronounced asymmetry (see Figure 1) perform relatively less well out-of-sample. While this contrast between in-sample and out-of-sample properties is difficult to reconcile, we note that the differences in forecasting performance and in-sample fit between the symmetric and asymmetric GARCH specifications are small, that the superiority of the stochastic volatility model over the GARCH models indicates that more parsimonious models may be more effective for out-of-sample forecasting, and that the differences between the 2016 and 2018 findings indicate that the data properties out-of-sample can vary from those in-sample.

 By contrast, for both the EUR/USD and SPX data, there is more consistency with the in-sample findings. We find much more evidence of better forecasting performance among the asymmetric models, with the TGARCH (for the EUR/USD) and the EGARCH (for the SPX) being preferred over alternative asymmetric and symmetric specifications. Also, the stochastic volatility model performs poorly for both series, uniformly so when judged against the realized variance estimator. These contrasting results indicate the importance of examining a wide variety of models when forecasting data series that may not be regularly analysed, such as the AUD/USD, as their properties may be quite different and so be better modelled with alternative models. It appears that the stochastic volatility model is particularly well-suited to modelling the volatility of the AUD/USD using daily data.

 However, there is one result in Tables 8(a) to (d) that stands out from all others and that is that daily volatility forecasts generated by the GARCH model applied to the intra-day data dominate those produced by models using daily data, in most but not all cases. Given the relatively short estimation window used to generate the model parameters, this is an even more remarkable result.[[27]](#footnote-27) When realized variance is used as the ex-post variance estimator, all loss functions rank the intra-day GARCH model as the best model. For the AUD/USD, the two MAE loss function ranks the intra-day GARCH model second to the (daily) stochastic volatility model when the high-low variance estimator is used, and the R2LOG loss function has this same ranking when squared returns are the ex-post variance estimator. For EUR/USD, the same two MAE loss functions evaluating the forecasts against the high-low estimator also rank the intra-GARCH model second, this time to the daily GARCH model. For the SPX, the intra-day GARCH model is preferred for all loss functions and variance estimators.

*4.4.3 Modified Diebold-Mariano Tests*

We supplement the above loss function calculations with the forecast comparison test of Harvey, Leybourne and Newbold (1998), hereafter HLN, which is a small sample version of the test developed by Diebold and Mariano (1995). This test permits direct comparisons between pairs of forecasts, using a given loss function. The model rankings reported above suggest that for the AUD/USD the stochastic volatility is the most preferred model among the models using daily data and that the intra-day GARCH model is the most preferred model overall and so these are the models against which other models are tested. To provide for the most conservative test, and following the observation of Poon and Granger (2003) that forecasts comparisons are most difficult for metrics that are quadratic in the variance, we focus on the MSE loss function. But, to provide for a comparison and because the rankings were most different for this function, we also examine the MAE loss function.

By increasing the accuracy of the volatility estimator, the HLN tests, reported in Table 9, are able to reject more models in favour of the intra-day GARCH model. This result is echoed by the EUR/USD and SPX test results. The more conservative MSE loss function provides some evidence that for the more recent sample period, using the realized variance estimator, that the stochastic volatility model is only weakly significantly better than the competing models that use daily data. We can also see that for the earlier sample period for the AUD/USD that the realized variance estimator permits the HLN test to better distinguish between the stochastic volatility model and the competing daily GARCH models, although the evidence is much weaker during this sample period. This result is in contrast to Patton (2011) who did not find significant differences between these models, even using more accurate volatility estimators for these loss functions.

While the study by Hansen and Lunde (2005) found that nothing (within a wide class of GARCH models) could beat the base-line GARCH(1,1) model for exchange rate forecasting (of the DM/$), our study complements more recent evidence, such as Chortareas (2011) who find that the stochastic volatility model may outperform the GARCH(1,1) model for some exchange rates series when using daily data. Although for intra-day data, our results for the AUD/USD add further supporting evidence to their conclusions.

*4.4.4 Loss Function Convergence Tests*

The loss function rankings, also reported in Tables 8 (a) to (d), appear to show that the more sophisticated (ex-post) variance estimators generate more convergence among the loss functions rankings. We test this formally by conducting paired t-tests of the mean (across models) of the standard deviation of the loss function rankings (across loss functions), equation (14) above. A smaller value of this mean indicates a greater convergence of the rankings across loss functions, since (in the limit) converged rankings would generate a zero standard deviation across the loss functions for each model. The standard deviations of the ranks and the means of these standard deviations are reported alongside the loss function rankings in Tables 8 (a) to (d). By comparing the ranks in Table 8(a) and (b), we observe that the more recent sample period has generated loss functions rankings with a greater convergence (smaller mean values for each variance estimator) and so, conversely, the earlier sample indicates that loss function ranks can be in considerable disagreement. For the 2018 sample, we find that the squared returns and the high-low estimators display only weakly different average levels of ranking convergence (p=0.07), whereas the realized variance estimator generates loss functions that are signficicantly more converged than either the squared returns or high-low estimators, (p=0.03) and (p=0.02), respectively.[[28]](#footnote-28) For the 2016 sample, we find that the high-low estimator and realized variance estimators do not display different average levels of ranking convergence (p=0.31), but that the rankings using squared returns are significantly less converged than those for either the high-low estimator (p<0.01) or the realized variance estimator (p=0.02). These results, which are consistent with the study by Patton (2011), indicate that employing superior variance estimators can reduce the impact a loss function's characteristics have on forecast evaluation, thus reducing the importance of loss function selection.

We conduct a further test of the benefit from using a more sophisticated variance estimator, by examining the ability of the loss functions to distinguish between the best and second best models. Loss function values across models, for a given loss function, are often very similar in magnitude, a characteristic noted by, for example, Brailsford and Faff (1996) for the Australian stock market. We calculate the average (across loss functions) of the percentage difference between the loss function scores for the best and second best models, for each of the three variance estimators. Our results echo those relating to the convergence of the rankings. For the 2018 sample, we find no significant difference between the average differences for the range (high-low) and squared returns estimators (p=0.23), but significant differences between the smaller average differences for the realized variance estimator, and the larger average differences for both the high-low estimator (p=0.01) and the squared returns estimator (p=0.01). For the 2016 sample, we find no significant difference between the average differences for the range (high-low) and realized variance estimators (p=0.28), but significant differences between the smaller average differences for the squared returns estimator, and the larger average differences for both the high-low estimator (p=0.01) and the realized returns estimator (p=0.03). These results indicate that loss functions are much better able to distinguish between the first and second best models when either the high-low or realized variance estimators are used, than when the squared returns estimators is used.[[29]](#footnote-29)

**5. Summary and conclusions**

By forecasting the volatility of the Australian Dollar / US Dollar exchange rate using different proxies for ex-post volatility, we show how to measure and test the significance of the improvement in the convergence of loss function scores towards selecting the best forecast. Using our test for loss function ranking convergence, we find that using either a range based estimator or a realized variance estimator rather than squared returns as the measure of ex-post volatility results in a significant increase in the convergence of loss function rankings. With better quality measures of volatility, loss functions converge more strongly on the preferred forecast. Moreover, we also find a significant increase in the ability of loss functions to distinguish between the best and second best models when either the range measure or the realized variance are used. Thus, the margin by which the best model is chosen is significantly increased by using the higher quality volatility measures.

While we find that there are no significant differential gains in terms of loss function ranking convergence between the range measure and the realized variance measure in the 2016 sample, we do for the more recent 2018 sample, and we find that the realized variance measure is more often able to facilitate the rejection of one or more competing forecasts against a benchmark forecast. However, the forecast comparison tests are much less able to distinguish between any of the competing forecasts when squared returns are used as the measure of ex-post volatility. While the range measure does provide a significant improvement upon the use of squared returns, and so be a valuable substitute for realized variance when high frequency data is unavailable or too costly to collect, high frequency data should still be used to create a realized variance estimator when it is available.

By contrast to early studies that advocated the use of realized variance, we do not find that the regression based tests are enhanced by the use of realized variance for models that generate forecasts from daily data. However, other recent studies, such as McMillan and Speight (2012) and Chortareas et al (2011), have also noted that these regression based tests still perform poorly even with high frequency data, both in regards to the regression *R2* and tests of forecast unbiasedness. This suggests that for evaluating volatility forecasts, comparisons based upon loss functions are more likely to give reliable results. Our work supports studies that show that the use of loss functions is enhanced by the use of high frequency data for measuring ex-post volatility, and provides a simple test (using the standard deviation of the loss function rankings) of how much better the loss functions are at distinguishing between forecasts as the quality of the volatility estimator improves.

We choose the Australian Dollar / US Dollar exchange rate as this has had little empirical study yet is the fifth highest traded exchange rate; the four higher all having had their volatility extensively modelled and forecast. We find that, for daily data, a model of stochastic volatility provides superior forecasts of daily volatility compared to a range of GARCH models, and that, in general, the more simple GARCH formulations generate superior forecasts to those generated by more complex models featuring asymmetries and power functions. As prior studies of this exchange rate have found that asymmetric components and power functions can add to the modelling and forecasting of volatility, our results indicate a shift in the characteristics of this exchange rate in recent years. In particular, our study is the only study to consider the period post the 2008/09 financial crisis period.

We conjecture that the superiority of the stochastic volatility model may come from its natural ability to model information flow into a financial market.[[30]](#footnote-30) Unlike the GARCH models, the stochastic volatility model identifies a latent stochastic volatility process, which may better reflect the arrival flow of information than the backward looking GARCH process. This may be more important in some markets than others.

Finally, we also find that if intra-daily data are available for forecasting the daily volatility of the AUD/USD (as well as for forecast evaluation) that, in most cases, a GARCH(1,1) model applied to that intra-day data generates far superior forecasts of daily volatility than any of the models applied to daily data, including the stochastic volatility model. We believe that our study is the first to evaluate the forecasts of the AUD/USD volatility using intra-day data, and this suggests that further analysis of such data in the context of the AUD/USD is likely to be a fruitful future avenue for research.

**References**

Abdalla, Suliman Z. S. 2012. "Modelling Exchange Rate Volatility using GARCH Models: Empirical Evidence from Arab Countries." *International Journal of Economics and Finance* 4: 216-229.

Abramowitz, Milton, and Irene A. Stegun. 1970. *Handbook of Mathematical Functions.* New York: Dover Publications Inc.

Akgiray, Vedat. 1989. "Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts." *The Journal of Business* 62: 55-80.

Andersen, Torben G., and Tim Bollerslev. 1998. "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts."  *International Economic Review* 39: 885-905.

Andersen, T., T. Bollerslev, F.X. Diebold, P. Labys (2001). The Distribution of Realized Exchange Rate Volatility, Journal of the American Statistical Association 96, 42-55.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2003. "Modeling and forecasting realized volatility." *Econometrica* 71: 579-625.

Anderton, Robert, and Frauke Skudelny. 2001. *Exchange Rate Volatility and Euro Area Imports*. ECB Working Paper No. 64.

Baillie, Richard T., Tim Bollerslev. 1989. “The message in daily exchange rates: A conditional variance tale”, *Journal of Business and Economic Statistics* 7, 297-305.

Baillie, Richard T., Tim Bollerslev, and Hans O. Mikkelsen. 1996. " Fractionally integrated generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 74: 3-30.

Bandi, F.M., and J.R. Russell. 2006. “Separating microstructure noise from volatility”. *Journal of Financial Economics* 79, 655-692.

Bank for International Studies. 2013. *Triennial Central Bank Survey - Foreign exchange turnover in April 2013: preliminary global results*. Monetary and Economic Department.

Barndorff-Nielsen, O.E., and N. Shephard (2002a). “Estimating Quadratic Variation Using Realized Variance,” *Journal of Applied Econometrics*, 17, 457-477.

Barndorff-Nielsen, O.E., and N. Shephard (2002b). “Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society, Series B,* 64, 253-280.

Barndorff-Nielsen, O.E., and N. Shephard (2004a). “Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics,” *Econometrica*, 73(3), 885-926.

Barndorff-Nielsen, O.E., and N. Shephard (2004b). “How Accurate Is the Asymptotic Approximation to the distribution of Realized Volatility?,” in *Identification and Inference for Econometric Models. a Festschrift in Honour of T.J. Rothenberg*, ed. by D.W.K. Andrews, J. Powell, P.A. Ruud, and J.H. Stock. Cambridge: Cambridge University Press.

Bera, Anil. K., and Matthew L. Higgins. 1992. "A Class of Nonlinear ARCH Models." *International Economic Review* 33: 137-58.

Bera, Anil. K., and Matthew L. Higgins. 1997. “ARCH and bilinearity as competing models for nonlinear dependence”. *Journal of Business and Economic Statistics*, 15, 43-50.

Black, Fischer. 1976. "Studies of Stock Price Volatility Changes." Paper presented at the proceedings of the 1976 meetings of the American Statistical Association, Business and Economic Statistics Section, 177-181, Washington, D.C.

Blair, Bevan. J., Ser-Huang Poon, and Stephen J. Taylor*.* 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, Tim, Julia Litvinova, and George Tauchen. 2006. "Leverage and Volatility Feedback Effects in High-Frequency Data." *Journal of Financial Econometrics* 4: 353-384

Bollerslev, Tim. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31: 307-327.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner. 1992. "ARCH modeling in finance." *Journal of Econometrics* 52: 5-59

Bollerslev, Tim and Ian Domowitz. (1993). “Trading patterns and prices in the interbank foreign exchange market”, *Journal of Finance* 48, 1421-1443.

Bollerslev, Tim., Robert F. Engle., and Daniel B. Nelson. 1994. "ARCH models." In *Handbook of Econometrics, Volume 4*, edited by Robert F. Engle, and Daniel L. McFadden, 2961-3038. Amsterdam: Elsevier Science B.V.

Brailsford, Timothy, and Robert W. Faff. 1996. "An evaluation of volatility forecasting techniques." *Journal of Banking & Finance* 20: 419-438.

Campbell, J., and L. Hentschel, 1992. “No news is good news: An asymmetric model of changing volatility in stock returns.” *Journal of Financial Economics*, 31, 281-331.

Chortareas, Georgios, Ying Jiang, and John C. Nankervis. 2011. "Forecasting exchange rate volatility using high-frequency data: Is the euro different?" *International Journal of Forecasting* 27: 1089:1107.

Choudry, Taufiq, and Syed S. Hassan. 2015. "Exchange rate volatility and UK imports from developping countries: The effect of the global financial crisis." *Journal of International Financial Markets, Institutions and Money* 39: 89-101.

Christie, Andrew A. 1982. " The stochastic behavior of common stock variances: Value, leverage and interest rate effects." *Journal of Financial Economics* 10: 407-432

Clark, Peter K. 1973. "A subordinated stochastic process model with finite variance for speculative prices." *Econometrica* 41: 135–155.

Cumby, Robert, Stephen Figlewski, and Joel Hasbrouck. 1993. "Forecasting Volatilities and Correlations with EGARCH Models." *Journal of Derivatives* 1: 51-63

Day, Theodore E.,and Craig M. Lewis. 1992. "Stock market volatility and the information content of stock index options." *Journal of Econometrics* 52: 267–287

Dickey, D.A., and W.A. Fuller, “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” *Journal of the American Statistical Association* 74, 427-431.

Diebold, Francis X., 1988. *Empirical modelling of exchange rate dynamics.* (Springer Verlag: New York).

Diebold, Francis X., & Mariano, Robert. S. (1995). “Comparing predictive accuracy”. *Journal of Business and Economic Statistics* 13, 253–265.

Ding, Zhaunxin, Clive W. J. Granger, Robert F. Engle. 1993. "A long memory property of stock market returns and a new model." *Journal of Empirical Finance* 1: 83-106.

Duffee, Gregory R. 1995. "Stock returns and volatility A firm-level analysis." *Journal of Financial Economics* 37: 399-420.

Dunis, Christian., Jason Laws and Stephane Chauvin. 2003. “FX volatility forecasts and the informational content of market data for volatility.” *The European Journal of Finance,* 9, 242-272.

Engle, Robert F. 1982. "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation." *Econometrica* 50: 987-1007.

Engle R.F., and Lee, G.G.J. 1999. A permanent and transitory component model of stock return volatility. In *Cointegration,Causality and Forecasting: A Festschrift in Honour of Clive W J Granger*, Engle RF, White H (eds). Oxford University Press: Oxford; 475–497.

Engle, Robert F., and Victor K. Ng. 1993. "Measuring and Testing the Impact of News on Volatility." *The Journal of Finance* 48: 1749-1778.

Epps, T.W. and M. L. Epps. 1976. The stochastic dependence of security price changes and transaction volumes: implications for the mixture-of-distributions hypothesis. *Econometrica* 44:305–321.

Fama, Eugene F. 1965. "The behavior of stock-market prices." *Journal of Business* 38: 34–105.

Figlewski, Stephen, and Xiaozu Wang. 2001. "Is the ‘Leverage Effect’ a Leverage Effect?" Working Paper, Department of Finance, New York University.

Figlewski, Stephen., 1997. "Forecasting volatility." *Financial Markets, Institutions and Instruments* 6: 1-88.

French, K., G. Schwert and R. Stambaugh, 1987. “Expected stock returns and volatility”. *Journal of Financial Economics,* 19, 3-30.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. 1993. "On the Relation between the Expected Value and Volatility of the Nominal Excess Return on Stocks." *The Journal of Finance* 48: 1779-1801.

Gulay E., and H. Emec, 2015. “Comparison of forecasting performances: Does normalization and variance stabilization method beat GARCH(1,1)-type models? Empirical evidence from the stock markets”. *Journal of Forecasting* 37, 133-150.

Hansen, P., 2005. “A test for superior predictive ability.” *Journal of Business and Economic Statistics* 23: 365-380.

Hansen , P. A. Lunde and J. Nason, 2011. “The model confidence set.” *Econometrica* 79, 453-497.

Hansen, Peter R., and Asger Lunde. 2005. "A Forecast Comparison of Volatility Models: Does Anything Beat A GARCH(1,1)?" *Journal of Applied Econometrics* 20: 873-889.

Hansen, Peter R., and Asger Lunde. 2006. “Consistent ranking of volatility models”. *Journal of Econometrics* 131, 97-121.

Harvey, Andrew C., Esther Ruiz, and Neil Shephard. 1994. "Multivariate stochastic variance models." *Review of Economic Studies* 61: 247–264.

Harvey, David I., Stephen J. Leybourne, and Paul Newbold. 1998. "Test for Forecast Encompassing." *Journal of Business & Economic Statistics* 16: 254-259.

Heynen, Ronald C., and Harry M. Kat. 1994. "Volatility prediction: a comparison of the stochastic volatility, GARCH(1, 1) and EGARCH(1, 1) models." *Journal of Derivatives* 2: 50–65.

Hsieh, D.A. 1989. “Modeling Heteroscedasticity in daily Foreign-Exchange Rates”, *Journal of Business and Economic Statistics,* 7, 307 - 317.

Hu, Michael Y., Christine X. Jiang, and Christos Tsoukalas. 1997. "The European exchange rates before and after the establishment of the european monetary system." *Journal of International Financial Markets, Institutions and Money* 7: 235-253.

Huang, Fei-xue, Yan Zhao, and Tie-shan Hou. 2009. " Long Memory and Leverage Effect of Euro Exchange Rate Based on ARFIMA-FIEGARCH." Paper presented at the International Conference on Management Science & Engineering (16th), Moscow, Russia, September 14-16.

Jarque, Carlos M., and Bera, Anil K., 1987. "A Test for Normality of Observations and Regression Residuals," *International Statistical Review* 55, 163-172.

Jorion, Philippe. 1995. "Predicting Volatility in the Foreign Exchange Market." *Journal of Finance* 50: 507-528.

Lamoureux, Christopher G., and Wiliam D. Lastrapes. 1990. "Persistence in variance, structural change and the GARCH model." *Journal of Business and Economic Statistics* 8: 225-234.

Laurent, S., and F. Violante, 2017. “Volatility forecasts evaluation and comparison”. *WIREs Computational Statistics* 4, 1-12.

Lee, Keun Yeong. 1991. “Are the GARCH models best in out-of-sample performance?” *Economics Letters* 37, 305-308.

Li, Kai. 2002. “Long-memory versus option-implied volatility prediction”, *Journal of Derivatives* 9, 9-25

Lopez, Jose A. 2001. “Evaluating the predictive accuracy of volatility models”, *Journal of Forecasting* 20, 87-109.

Liu, Lily, Andrew J. Patton, and Kevin Sheppard. 2015. " Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes." *Journal of Econometrics* 187: 293-311.

G. M. Ljung; G. E. P. Box, 1978. "On a Measure of a Lack of Fit in Time Series Models". *Biometrika*. 65, 297–303.

Ma, F., Y. Li, L. Liu and Y. Zhang, 2018. “Are low-frequency data really uninformative? A forecasting combination perspective”. *North American Journal of Economics and Finance* 44, 92-108.

Mandelbrot, Benoit. 1963. "The Variation of certain speculative prices." *The Journal of Business* 36: 394-419.

Martens, M., 2001. “Forecasting daily exchange rate volatility using intraday returns.” *Journal of International Money and Finance* 20, 1–23.

McKenzie, Michael D., (1997). “ARCH modelling of Australian bilateral exchange rate data”, *Applied Financial Economics* 7, 147-164.

McKenzie, Michael D., (1998), “The Impact of Exchange Rate Volatility on Australian Trade Flows,” *Journal of International Financial Markets, Institutions and Money,* 8, pp.21-38.

McKenzie, Michael, and Heather Mitchell. 2002. "Generalised Asymmetric Power ARCH Modeling of Exchange Rate Volatility." *Applied Financial Economics* 12: 555-564.

McMillan, David, G., and Alan E.H. Speight (2004) “Daily Volatility Forecasts: Reassessing the Performance of GARCH Models”, *Journal of Forecasting* 23, 449–460.

McMillan, David, G., and Alan E.H. Speight (2012). “Daily FX Volatility Forecasts: Can the GARCH(1,1) Model be Beaten using High-Frequency Data?”, *Journal of Forecasting* 31, 330–343.

Mincer, Jacob A., and Victor Zarnowitz. 1969. "The Evaluation of Economic Forecasts." *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research

Nelson, Daniel B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59: 347-370.

Pagan, Adrian R., and G William Schwert. 1990. "Alternative Models for Conditional Stock Volatility." *Journal of Econometrics* 45: 267-290.

Parkinson, Michael. 1980. "The extreme value method for estimating the variance of the rate of return." *Journal of Business* 53: 61–65.

Patton, Andrew J. 2011. "Volatility Forecast Comparison Using Imperfect Volatility Proxies." *Journal of Econometrics* 160: 246-256.

Pong, Shiuyan Pong, Mark B Shackleton, Stephen J Taylor, Xinzhong Xu (2004) Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models, *Journal of Banking and Finance*, 28, 2541–2563.

Poon, Ser-Huang, and Clive W. J. Granger. 2003. "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature* 41: 478-539

Praetz, P. D. 1969. "Australian Share Prices and the Random Walk Hypothesis." *Australian & New Zealand Journal of Statistics* 11: 123-139.

Ramasamy, Ravindram, and Shanmugam Munisamy. 2012. "Predictive Accuracy of GARCH, GJR and EGARCH Models Select Exchange Rates Application." *Global Journal of Management and Business Research* 12: 88-100

Rodríguez Villar, Maria José. 2010. "Volatility Models with Leverage Effect." PhD thesis, Universidad Carlos III de Madrid.

Ruiz, Esther. 1994. "Quasi-maximum likelihood estimation of stochastic volatility models*." Journal of Econometrics* 63: 289-306.

Schwert, G.William. 1989. "Why does stock market volatility change over time?" *Journal of Finance* 44: 1115–1153.

Tauchen, G. E. and M. Pitts. 1983. The price variability-volume relationship on speculative markets. *Econometrica* 51:485–505.

Taylor, Stephen J. 1980. "Conjectured Models for Trends in Financial Prices, Tests and Forecasts." *Journal of the Royal Statistical Society* 143: 338-362.

Taylor, Stephen J. 1986. *Modelling financial time series*. Chichester:Wiley.

Taylor, Stephen J. 2005. *Asset Price Dynamics, Volatility, and Prediction*. New Jersey: Princeton University Press.

Taylor, Stephen J., and Brian G. Kingsman. 1979. "An analysis of the variance and distribution of commodity price changes." *Australian Journal of Management* 4: 135–149.

Taylor, Stephen J., and Xinzhong Xu. 1997. “The incremental volatility information in one million foreign exchange quotations”. *Journal of Empirical Finance* 4, 317-40.

Tse, Yiu K., and Albert K. C. Tsui. 1997. "Conditional volatility in foreign exchange rates: evidence from the Malaysian ringgit and Singapore dollar." *Pacific-Basin Finance Journal* 5: 345-356.

Vilasuso, Jon. 2001. "Forecasting exchange rate volatility." *Economics Letters* 76: 59-64.

West, Kenneth. D., and Dongchul Cho. 1995. "The predictive ability of several models of exchange rate volatility." *Journal of Econometrics* 69: 367–391.

West, Kenneth D., 1996. “Asymptotic inference about predictive ability.” *Econometrica* 64, 1067–1084.

White, H., 2000. “Reality check for data-snooping”. *Econometrica* 68, 1097-1126.

Zakoïan, Jean-Michel. 1994. "Threshold heteroskedastic models." *Journal of Economic Dynamics and Control* 18: 931–955.

Zhou, Bin. 1996. "High-frequency data and volatility in foreign-exchange rates." *Journal of Business and Economic Statistics* 14: 45–52.

Table 1: Volatility Model Specifications

|  |  |
| --- | --- |
| Model | Variance Specification |
| GARCH |  |
| *t*-GARCH |
| GJR |  |
| *t*-GJR |
| EGARCH | where  |  |
| *t*-EGARCH |  |
| PGARCH |  |
| APGARCH |  |
| TGARCH |  |
| SV |  |

Notes: The variable is the degrees of freedom, which is the parameter governing the density function of the *t* distribution. The function is the gamma function, which is defined by an integral, , . All other variables and parameters are as defined in the text. The Intra-day Garch (ID GARCH) model uses the GARCH volatility specification except that it is applied to rescaled (by cross sectional standard deviation) returns and includes an MA(1) term in the conditional mean.

Table 2: Summary Statistics of the Daily Returns

|  |
| --- |
| Summary statistics for daily returns between 27/04/12 and 01/03/2018 (except 06/04/2010 to 29/02/2016 for AUD/USD 2016 and 27/02/2012 to 08/03/2018 for the SPX). Mean is the mean daily return. Std. Dev. Is the standard deviation of the returns. Skew is the Skewness and Kurt is the Kurtosis of the daily returns. JB is the Jarque-Bera (1987) test for normality. ADF is the Augmented Dickey Fuller (1979) test (no trend) using the SIC criterion to select lag length. ARCH is the Engle (1982) test for ARCH effects in the variance of the returns. P-values are in brackets. |
|  | Mean  | Std. Dev. | Skew. | Kurt. | JB | ADF | ARCH | Min | Q1 | Median | Q3 | Max |
| AUD/USD | 0.1927 | 0.0061 | 0.0790 | 3.873 | 49.79 | -40.74 | 4.06 | -0.0216 | -0.0036 | 0.0001 | 0.0039 | 0.0241 |
| 2018 | [0.220] |  |  |  | [0.000] | [0.000] | [0.044] |  |  |  |  |  |
| AUD/USD | 0.1655 | 0.0072 | 0.0813 | 4.916 | 237.3 | -40.00 | 8.86 | -0.0356 | -0.0040 | 0.0001 | 0.0042 | 0.0360 |
| 2016 | [0.369] |  |  |  | [0.000] | [0.000] | [0.003] |  |  |  |  |  |
| EUR/USD | 0.0430 | 0.0053 | -0.161 | 5.174 | 305.2 | -40.13 | 31.35 | -0.0301 | -0.0029 | -0.0001 | 0.0032 | 0.0247 |
|  | [0.752] |  |  |  | [0.000] | [0.000] | [0.000] |  |  |  |  |  |
| SPX | 0.4587 | 0.0078 | -0.473 | 5.937 | 601.9 | -39.04 | 96.78 | -0.0418 | -0.0030 | 0.0005 | 0.0047 | 0.0383 |
|  | [0.022] |  |  |  | [0.000] | [0.000] | [0.000] |  |  |  |  |  |

Table 3: Autocorrelations of Returns and Squared Returns

|  |
| --- |
| Autocorrelations for daily returns, and squared daily returns . Q stats are the Ljung Box (1978) statistics. P-values for the autocorrelations and Q statistics are given underneath in brackets. |
|  | Autocorrelations at lag | Q stats at lag |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 20 | 40 |
| AUD/USD 2018 |
|  | -0.045 | 0.016 | 0.007 | 0.004 | -0.012 | 0.020 | -0.002 | 0.036 | -0.035 | 0.009 | 16.08 | 55.91 |
|  | [0.076] | [0.170] | [0.304] | [0.455] | [0.565] | [0.613] | [0.723] | [0.595] | [0.502] | [0.584] | [0.712] | [0.050] |
|  | 0.052 | 0.102 | 0.028 | 0.078 | 0.065 | 0.046 | 0.047 | 0.088 | 0.060 | 0.047 | 135.6 | 273.2 |
|  | [0.041] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| AUD/USD 2016 |
|  | -0.0204 | 0.0321 | -0.0429 | 0.0147 | -0.0073 | 0.0069 | -0.0089 | 0.0163 | -0.0349 | -0.0049 | 17.53 | 48.08 |
|  | [0.422] | [0.327] | [0.166] | [0.247] | [0.358] | [0.473] | [0.575] | [0.635] | [0.534] | [0.625] | [0.619] | [0.178] |
|  | 0.0778 | 0.1797 | 0.0794 | 0.1005 | 0.1388 | 0.1304 | 0.0898 | 0.1304 | 0.0841 | 0.1453 | 380.10 | 574.31 |
|  | [0.002] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| EUR/USD |
|  | -0.031 | 0.016 | -0.019 | 0.005 | 0.030 | -0.031 | -0.009 | 0.016 | 0.010 | -0.038 | 23.90 | 50.52 |
|  | [0.222] | [0.390] | [0.483] | [0.646] | [0.566] | [0.498] | [0.601] | [0.662] | [0.736] | [0.602] | [0.247] | [0.123] |
|  | 0.144 | 0.081 | 0.082 | 0.089 | 0.097 | 0.057 | 0.023 | 0.104 | 0.057 | 0.063 | 153.9 | 330.5 |
|  | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| SPX |
|  | -0.003 | -0.038 | -0.005 | -0.062 | -0.047 | -0.004 | 0.007 | 0.012 | -0.048 | 0.003 | 30.79 | 62.49 |
|  | [0.894] | [0.338] | [0.531] | [0.087] | [0.042] | [0.074] | [0.115] | [0.160] | [0.081] | [0.119] | [0.058] | [0.013] |
|  | 0.251 | 0.215 | 0.260 | 0.206 | 0.100 | 0.115 | 0.071 | 0.094 | 0.092 | 0.097 | 447.3 | 480.2 |
|  | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |

**Table 4: Coefficient Estimates from the Sign and Size Bias Tests**

|  |  |  |  |
| --- | --- | --- | --- |
| Coefficient estimates from the following regressions

|  |
| --- |
|  |
|  |
|  |

where where is the residual from a GARCH(1,1) model applied to the daily returns. is an indicator variable that takes the value 1 if and is zero otherwise, , and is an i.i.d error term. |
| Sign | Negative Size | Positive Size |
|  |  |  |  |  |  |
| AUD/USD 2018 |
| 3.96 | -0.44 | 1.80 | -8.35 | 1.58 | 9.10 |
| [0.000] | [0.177] | [0.000] | [0.000] | [0.000] | [0.000] |
| AUD/USD 2016 |
| 5.95 | -1.45 | 5.50 | 1.07 | 4.56 | 2.44 |
| [0.000] | [0.006] | [0.000] | [0.085] | [0.000] | [0.000] |
| EUR/USD |
| 2.64 | 0.03 | 2.43 | -1.89 | 2.71 | 0.472 |
| [0.000] | [0.289] | [0.000] | [0.000] | [0.000] | [0.313] |
| SPX |
| 4.92 | 0.03 | 4.37 | -0.006 | 5.82 | 1.07 |
| [0.000] | [0.001] | [0.000] | [0.000] | [0.000] | [0.179] |

**Table 5a - Coefficient Estimates from the GARCH and Stochastic Volatility Models for the AUD/USD 2018**

|  |
| --- |
| Coefficient estimates for the models given in Table 1 above applied to the daily returns on the AUD/USD exchange rates between 09/04/2012 to 01/03/2018. P-values are given in parentheses. is for GARCH and GJR-GARCH, and is for PGARCH, APGARCH and TGARCH. Log L is the maximized value of the likelihood function. Kurt is the kurtosis of the standardized residusals. Q(40) is p-value of the Ljung-Box statistic (with 40 lags). |
|  |  |  |  |  |  |  |  | Log L | Kurt | Q(40) |
| GARCH | 0.138 | 0.339 | 0.036 | 0.955 |  |  |  | 5630.85 |  |  |
|  | (0.334) | (0.019) | (0.000) | (0.000) |  |  |  |  |  |  |
| GARCH(t) | 0.089 | 0.303 | 0.037 | 0.955 | 10.854 |  |  | 5640.60 | 3.66 | 0.175 |
|  | (0.533) | (0.091) | (0.000) | (0.000) |  |  |  |  |  |  |
| GJR | 0.187 | 0.266 | 0.012 | 0.966 |  | 0.029 |  | 5633.12 | 3.66 | 0.154 |
|  | (0.201) | (0.017) | (0.139) | (0.000) |  | (0.005) |  |  |  |  |
| GJR(t) | 0.138 | 0.225 | 0.011 | 0.967 | 10.997 | 0.032 |  | 5642.91 | 3.68 | 0.155 |
|  | (0.335) | (0.087) | (0.274) | (0.000) |  | (0.013) |  |  |  |  |
| EGARCH | 0.201 | -18.453 | -0.008 | -0.810 |  | 0.018 |  | 5581.64 | 3.88 | 0.047 |
|  | (0.204) | (0.000) | (0.783) | (0.011) |  | (0.346) |  |  |  |  |
| EGARCH(t) | 0.140 | -0.094 | 0.074 | 0.991 | 11.347 | 0.021 |  | 5641.2 | 3.62 | 0.176 |
|  | (0.323) | (0.061) | (0.000) | (0.000) |  | (0.064) |  |  |  |  |
| PGARCH | 0.147 | 0.301 | 0.045 | 0.953 |  |  | 1.207 | 5631.60 | 3.61 | 0.181 |
|  | (0.301) | (0.719) | (0.000) | (0.000) |  |  | (0.024) |  |  |  |
| APGARCH | 0.188 | 0.001 | 0.029 | 0.965 |  | 0.292 | 1.688 | 5633.17 | 3.64 | 0.158 |
|  | (0.201) | (0.765) | (0.002) | (0.000) |  | (0.071) | (0.009) |  |  |  |
| TGARCH | 0.188 | 0.059 | 0.027 | 0.961 |  | 0.020 |  | 5632.76 |  | 0.172 |
|  | (0.193) | (0.010) | (0.006) | (0.000) |  | (0.035) |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| ID-GARCH | -0.008 | 0.085 | 0.085 | 0.847 | -0.011 |  |  | -5857.71 | 5.88 | 0.055 |
|  | (0.618) | (0.000) | (0.000) | (0.000) | (0.537) |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Stochastic Volatility | -0.152 | 0.985 | 0.007 |  |  |  | -3332.13 | 3.16 | 0.000 |
|  |  | (0.058) | (<0.001) | (0.040) |  |  |  |  |  |  |

**Table 5b - Coefficient Estimates from the GARCH and Stochastic Volatility Models for the AUD/USD 2016**

|  |
| --- |
| Coefficient estimates for the models given in Table 1 above applied to the daily returns on the AUD/USD exchange rates between 06/04/2010 to 29/02/2016. P-values are given in parentheses. is for GARCH and GJR-GARCH, and is for PGARCH, APGARCH and TGARCH. Log L is the maximized value of the likelihood function. Kurt is the kurtosis of the standardized residusals. Q(40) is p-value of the Ljung-Box statistic (with 40 lags). |
|  |  |  |  |  |  |  |  | Log L | Kurt | Q(40) |
| GARCH | 0.143 | 0.483 | 0.056 | 0.936 |  |  |  | 5507.8 | 3.658 | 0.875 |
|  | (0.382) | (0.051) | (<0.001) | (<0.001) |  |  |  |  |  |  |
| GARCH(t) | 0.093 | 0.461 | 0.054 | 0.938 | 10.095 |  |  | 5518.6 | 3.660 | 0.873 |
|  | (0.554) | (0.055) | (<0.001) | (<0.001) |  |  |  |  |  |  |
| GJR | 0.262 | 0.344 | 0.004 | 0.958 |  | 0.064 |  | 5516.9 | 3.677 | 0.863 |
|  | (0.105) | (0.074) | (0.737) | (<0.001) |  | (<0.001) |  |  |  |  |
| GJR(t) | 0.182 | 0.297 | <0.001 | 0.960 | 10.79 | 0.067 |  | 5527.0 | 3.695 | 0.858 |
|  | (0.581) | (0.075) | (0.981) | (<0.001) |  | (<0.001) |  |  |  |  |
| EGARCH | 0.263 | -0.093 | 0.091 | 0.991 |  | 0.050 |  | 5513.4 | 3.686 | 0.861 |
|  | (0.002) | (0.033) | (<0.001) | (<0.001) |  | (<0.001) |  |  |  |  |
| EGARCH(t) | 0.185 | -0.089 | 0.090 | 0.991 | 10.56 | 0.052 |  | 5523.8 | 3.693 | 0.860 |
|  | (0.178) | (0.036) | (<0.001) | (<0.001) |  | (<0.001) |  |  |  |  |
| PGARCH | 0.144 | 0.175 | 0.052 | 0.935 |  |  | 2.200 | 5507.9 | 3.656 | 0.876 |
|  | (0.379) | (0.750) | (0.001) | (<0.001) |  |  | (<0.001) |  |  |  |
| APGARCH | 0.257 | 0.865 | 0.021 | 0.957 |  | 0.626 | 2.270 | 5517.1 | 3.672 | 0.866 |
|  | (0.110) | (0.791) | (0.365) | (<0.001) |  | (0.326) | (0.002) |  |  |  |
| TGARCH | 0.261 | 0.600 | 0.021 | 0.956 |  | 0.051 |  | 5512.9 | 3.709 | 0.849 |
|  | (0.098) | (0.040) | (0.039) | (0.001) |  | (<0.001) |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Stochastic Volatility | -0.116 | 0.989 | 0.007 |  |  |  | -3378.60 | 6.002 | 0.002 |
|  |  | (0.030) | (<0.001) | (0.015) |  |  |  |  |  |  |

**Table 5c - Coefficient Estimates from the GARCH and Stochastic Volatility Models for the EUR/USD**

|  |
| --- |
| Coefficient estimates for the models given in Table 1 above applied to the daily returns on the EUR/USD exchange rates between 09/04/2012 to 01/03/2018. P-values are given in parentheses. is for GARCH and GJR-GARCH, and is for PGARCH, APGARCH and TGARCH. Log L is the maximized value of the likelihood function. Kurt is the kurtosis of the standardized residusals. Q(40) is p-value of the Ljung-Box statistic (with 40 lags). |
|  |  |  |  |  |  |  |  | Log L | Kurt | Q(40) |
| GARCH | 0.000 | 0.128 | 0.027 | 0.968 |  |  |  | 5871.07 | 4.78 | 0.331 |
|  | (0.997) | (0.000) | (0.000) | (0.000) |  |  |  |  |  |  |
| GARCH(t) | 0.061 | 0.084 | 0.034 | 0.964 | 7.198 |  |  | 5902.93 | 4.99 | 0.350 |
|  | (0.599) | (0.222) | (0.000) | (0.000) |  |  |  |  |  |  |
| GJR | 0.045 | 0.063 | 0.006 | 0.978 |  | 0.028 |  | 5879.32 | 4.49 | 0.508 |
|  | (0.720) | (0.032) | (0.173) | (0.000) |  | (0.000) |  |  |  |  |
| GJR(t) | 0.085 | 0.042 | 0.016 | 0.971 | 7.609 | 0.024 |  | 5905.93 | 4.75 | 0.487 |
|  | (0.465) | (0.455) | (0.079) | (0.000) |  | (0.012) |  |  |  |  |
| EGARCH | 0.070 | -0.029 | 0.037 | 0.997 |  | 0.033 |  | 5882.26 | 4.31 | 0.571 |
|  | (0.573) | (0.020) | (0.000) | (0.000) |  | (0.000) |  |  |  |  |
| PGARCH | 0.004 | 0.002 | 0.033 | 0.967 |  |  | 1.533 | 5871.65 | 4.79 | 0.349 |
|  | (0.975) | (0.657) | (0.000) | (0.000) |  |  | (0.000) |  |  |  |
| APGARCH | 0.094 | 0.089 | 0.018 | 0.985 |  | 0.983 | 0.672 | 5882.63 | 4.20 | 0.580 |
|  | (0.448) | (0.460) | (0.000) | (0.000) |  | (0.000) | (0.003) |  |  |  |
| TGARCH | 0.068 | 0.014 | 0.002 | 0.983 |  | 0.033 |  | 5882.11 | 4.27 | 0.598 |
|  | (0.583) | (0.016) | (0.524) | (0.000) |  | (0.000) |  |  |  |  |
| ID-GARCH |  |  |  |  |  |  |  |  |  |  |
|  | -0.008 | 0.054 | 0.055 | 0.904 | -0.051 |  |  | -5993.67 | 8.96 | 0.029 |
|  | (0.616) | (0.000) | (0.000) | (0.000) | (0.003) |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Stochastic Volatility | -0.088 | 0.992 | 0.006 |  |  |  | -3270.73 | 5.65 | 0.000 |
|  |  | (0.066) | (0.000) | (0.020) |  |  |  |  |  |  |

**Table 5d - Coefficient Estimates from the GARCH and Stochastic Volatility Models for the SPX**

|  |
| --- |
| Coefficient estimates for the models given in Table 1 above applied to the daily returns on the S&P 500 index (SPX) between 27/02/2012 to 08/03/2018. P-values are given in parentheses. is for GARCH and GJR-GARCH, and is for PGARCH, APGARCH and TGARCH. Log L is the maximized value of the likelihood function. Kurt is the kurtosis of the standardized residusals. Q(40) is p-value of the Ljung-Box statistic (with 40 lags). |
|  |  |  |  |  |  |  |  | Log L | Kurt | Q(40) |
| GARCH | 0.722 | 5.440 | 0.196 | 0.721 |  |  |  | 5350.45 | 4.72 | 0.124 |
|  | (0.000) | (0.000) | (0.000) | (0.000) |  |  |  |  |  |  |
| GARCH(t) | 0.769 | 3.600 | 0.204 | 0.761 | 4.955 |  |  | 5397.31 | 5.00 | 0.147 |
|  | (0.000) | (0.000) | (0.000) | (0.000) |  |  |  |  |  |  |
| GJR | 0.382 | 4.600 | 0.309 | 0.787 |  | -0.338 |  | 5387.84 | 5.71 | 0.133 |
|  | (0.026) | (0.000) | (0.000) | (0.000) |  | (0.000) |  |  |  |  |
| GJR(t) | 0.510 | 3.890 | 0.374 | 0.784 | 5.673 | -0.412 |  | 5433.89 | 6.25 | 0.152 |
|  | (0.001) | (0.000) | (0.000) | (0.000) |  | (0.000) |  |  |  |  |
| EGARCH | 0.311 | -0.858 | 0.164 | 0.913 |  | -0.258 |  | 5407.58 | 5.09 | 0.138 |
|  | (0.061) | (0.000) | (0.000) | (0.000) |  | (0.000) |  |  |  |  |
| EGARCH(t) | 0.442 | -0.722 | 0.180 | 0.927 | 6.203 | -0.264 |  | 5443.05 | 5.21 | 0.143 |
|  | (0.003) | (0.000) | (0.000) | (0.000) |  | (0.000) |  |  |  |  |
| PGARCH | 0.755 | 0.032 | 0.202 | 0.734 |  |  | 1.638 | 5350.88 | 4.75 | 0.136 |
|  | (0.000) | (0.474) | (0.000) | (0.000) |  |  | (0.000) |  |  |  |
| ID-GARCH |  |  |  |  |  |  |  |  |  |  |
|  | 0.036 | 0.001 | 0.106 | 0.892 | -0.006 |  |  | -1446.85 | 7.40 | 0.000 |
|  | (0.180) | (0.011) | (0.000) | (0.000) | (0.858) |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| Stochastic Volatility | -1.108 | 0.892 | 0.190 |  |  |  | -3585.35 | 3.77 | 0.000 |
|  |  | (0.036) | (0.000) | (0.048) |  |  |  |  |  |  |

**Table 6: Summary Statistics for Ex-post Volatility**

|  |
| --- |
| Summary statistics for three ex-post measures of volatility, all across 99 observations. Mean is the mean value. Std. Dev. is the standard deviation. of the returns. Skew is the Skewness and Kurt is the Kurtosis. ADF is the Augmented Dickey Fuller (1979) test (no trend) using the SIC criterion to select lag length. Q(40) is the Ljung-Box statistic (with 40 lags). P-values are in brackets. |
|  | Mean  | Std. Dev. | Skew. | Kurt. | ADF Test | Q(40) |
| AUD/USD 2018 (02/03/18-18/07/18) |
| Squared returns | 0.0282 | 0.0370 | 1.668 | 5.553 | -10.343 | 43.18 |
|  | [0.000] |  |  |  | [0.000] | [0.337] |
| Range (High-Low) | 0.0249 | 0.0171 | 1.232 | 3.782 | -8.443 | 39.05 |
|  | [0.000] |  |  |  | [0.000] | [0.513] |
| Realized variance | 0.0266 | 0.0105 | 1.337 | 5.574 | -8.036 | 78.56 |
|  | [0.000] |  |  |  | [0.000] | [0.000] |
| AUD/USD 2016 (01/03/16-15/07/16) |
| Squared returns | 0.0688 | 0.0996 | 2.695 | 11.26 | -9.474 | 35.16 |
|  | [0.000] |  |  |  | [0.000] | [0.688] |
| Range (High-Low) | 0.0644 | 0.0836 | 5.729 | 43.67 | -9.427 | 20.49 |
|  | [0.000] |  |  |  | [0.000] | [0.996] |
| Realized variance | 0.0700 | 0.0777 | 6.544 | 52.87 | -9.119 | 23.09 |
|  | [0.000] |  |  |  | [0.000] | [0.985] |
| EUR/USD (02/03/18-18/07/18) |
| Squared returns | 0.0203 | 0.0413 | 6.293 | 51.52 | -10.120 | 330.47 |
|  | [0.000] |  |  |  | [0.000] | [0.000] |
| Range (High-Low) | 0.0198 | 0.0211 | 6.266 | 52.11 | -9.969 | 21.28 |
|  | [0.000] |  |  |  | [0.000] | [0.993] |
| Realized variance | 0.0218 | 0.0131 | 2.746 | 14.27 | -7.170 | 51.25 |
|  | [0.000] |  |  |  | [0.000] | [0.109] |
| SPX (09/03/18-25/07/18) |
| Squared returns | 0.0753 | 0.1330 | 2.968 | 12.26 | -4.352 | 68.93 |
|  | [0.000] |  |  |  | [0.001] | [0.003] |
| Range (High-Low) | 0.0511 | 0.0726 | 2.701 | 10.26 | -3.152 | 130.4 |
|  | [0.000] |  |  |  | [0.026] | [0.000] |
| Realized variance | 0.0638 | 0.0638 | 2.199 | 9.292 | -2.251 | 167.9 |
|  | [0.000] |  |  |  | [0.190] | [0.000] |

**Table 7a: Coefficient Estimates, *R2* and unbiasedness test statistics from Mincer-Zarnowitz Regressions (AUD/USD 2018)**

|  |
| --- |
| Coefficient estimates (p-values in parentheses below) from the regression equations , where is one of the three measures of ex-post volatility and is the forecast of this volatility provided by one of the models in Table 1 (and using coefficients estimates from Table 5a). is the coefficient of determination of the regression, and “rank” is the rank of the model’s forecast ordered by . Wald p-value is the p-value of the joint hypothesis that , that is, that the forecast is unbiased. SV is the stochastic volatility model. |
|  | Squared Returns | High-low | Realized Variance |
| Model | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value |
| GARCH | 1.026 | -2.506 | 0.0306 | 0.045 | 0.686 | -1.472 | 0.0493 | 0.000 | 0.422 | -0.526 | 0.0168 | 0.000 |
|  | (0.018) | (0.083) | (8) |  | (0.001) | (0.027) | (2) |  | (0.001) | (0.201) | (2) |  |
| t-GARCH | 0.997 | -2.410 | 0.0306 | 0.044 | 0.668 | -1.411 | 0.0491 | 0.000 | 0.411 | -0.491 | 0.0158 | 0.000 |
|  | (0.017) | (0.083) | (8) |  | (0.001) | (0.028) | (3) |  | (0.001) | (0.215) | (3) |  |
| GJR | 1.07 | -2.622 | 0.0329 | 0.037 | 0.687 | -1.457 | 0.0474 | 0.000 | 0.348 | -0.273 | 0.0044 | 0.000 |
|  | (0.016) | (0.073) | (6) |  | (0.001) | (0.030) | (4) |  | (0.007) | (0.512) | (6) |  |
| t-GJR | 1.019 | -2.448 | 0.0323 | 0.036 | 0.654 | -1.348 | 0.0458 | 0.000 | 0.333 | -0.225 | 0.0034 | 0.000 |
|  | (0.015) | (0.075) | (7) |  | (0.001) | (0.034) | (6) |  | (0.006) | (0.567) | (7) |  |
| EGARCH | 1.247 | -2.580 | 0.0038 | 0.034 | 0.769 | -1.392 | 0.0051 | 0.000 | 0.469 | -0.545 | 0.0021 | 0.000 |
|  | (0.436) | (0.546) | (11) |  | (0.299) | (0.482) | (11) |  | (0.302) | (0.653) | (9) |  |
| t-EGARCH | 0.954 | -2.179 | 0.0336 | 0.022 | 0.561 | -1.012 | 0.0339 | 0.000 | 0.272 | -0.021 | 0.0000 | 0.000 |
|  | (0.011) | (0.069) | (5) |  | (0.001) | (0.068) | (9) |  | (0.012) | (0.951) | (11) |  |
| PGARCH | 0.97 | -2.287 | 0.0342 | 0.025 | 0.619 | -1.230 | 0.0461 | 0.000 | 0.386 | -0.400 | 0.013 | 0.000 |
|  | (0.011) | (0.067) | (3) |  | (0.000) | (0.033) | (5) |  | (0.000) | (0.262) | (4) |  |
| APGARCH | 1.069 | -2.590 | 0.0343 | 0.029 | 0.663 | -1.365 | 0.0444 | 0.000 | 0.328 | -0.204 | 0.0026 | 0.000 |
|  | (0.014) | (0.067) | (2) |  | (0.001) | (0.036) | (7) |  | (0.009) | (0.614) | (8) |  |
| TGARCH | 0.979 | -2.285 | 0.034 | 0.023 | 0.577 | -1.076 | 0.0353 | 0.000 | 0.282 | -0.054 | 0.0002 | 0.000 |
|  | (0.011) | (0.068) | (4) |  | (0.001) | (0.063) | (8) |  | (0.011) | (0.881) | (10) |  |
| SV | 0.781 | -1.859 | 0.0231 | 0.064 | 0.527 | -1.038 | 0.0335 | 0.001 | 0.344 | -0.292 | 0.0071 | 0.001 |
|  | (0.021) | (0.133) | (10) |  | (0.001) | (0.070) | (10) |  | (0.000) | (0.407) | (5) |  |
| ID GARCH | -4.325 | 2.276 | 0.1120 | 0.096 | -0.176 | 1.353 | 0.1121 | 0.000 | -0.252 | 1.649 | 0.7340 | 0.000 |
|  | (0.039) | (0.001) | (1) |  | (0.058) | (0.000) | (1) |  | (0.000) | (0.000) | (1) |  |

**Table 7b: Coefficient Estimates, *R2* and unbiasedness test statistics from Mincer-Zarnowitz Regressions (AUD/USD 2016)**

|  |
| --- |
| Coefficient estimates (p-values in parentheses below) from the regression equations , where is one of the three measures of ex-post volatility and is the forecast of this volatility provided by one of the models in Table 1 (and using coefficients estimates from Table 5d). is the coefficient of determination of the regression, and “rank” is the rank of the model’s forecast ordered by . Wald p-value is the p-value of the joint hypothesis that , that is, that the forecast is unbiased. SV is the stochastic volatility model. |
|  | Squared Returns | High-low | Realized Variance |
| Model | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value |
| GARCH | 1.592(0.013) | -1.376(0.147) | 0.0215(5) | 0.0438 | 1.407(0.009) | -1.162(0.144) | 0.0218(8) | 0.0268 | 1.027(0.040) | -0.499(0.503) | 0.0046(9) | 0.1165 |
| t-GARCH | 1.666(0.011) | -1.485(0.128) | 0.0237(4) | 0.0393 | 1.455(0.008) | -1.232(0.132) | 0.0232(7) | 0.0258 | 1.058(0.040) | -0.544(0.47i97) | 0.0052(7) | 0.1165 |
| GJR | 1.222(0.003) | -0.853(0.174) | 0.0190(7) | 0.0121 | 1.309(0.000) | -1.062(0.042) | 0.0419(4) | 0.0006 | 1.117(0.001) | -0.666(0.173) | 0.0190(4) | 0.0026 |
| t-GJR | 1.183(0.003) | -0.787(0.189) | 0.0177(8) | 0.0114 | 1.267(0.000) | -0.992(0.047) | 0.0400(5) | 0.0005 | 1.097(0.000) | -0.633(0.176) | 0.0188(5) | 0.0020 |
| EGARCH | 1.648(0.002) | -1.564(0.062) | 0.0354(3) | 0.0081 | 1.615(0.000) | -1.581(0.024) | 0.0515(3) | 0.0014 | 1.282(0.002) | -0.948(0.149) | 0.0214(3) | 0.0079 |
| t-EGARCH | 1.640(0.002) | -1.552(0.056) | 0.0371(2) | 0.0063 | 1.614(0.000) | -1.582(0.020) | 0.0548(1) | 0.0009 | 1.289(0.001) | -0.960(0.131) | 0.0234(1) | 0.0057 |
| PGARCH | 1.546(0.014) | -1.308(0.161) | 0.0201(6) | 0.0469 | 1.388(0.008) | -1.136(0.147) | 0.0216(9) | 0.0260 | 1.030(0.036) | -0.504(0.491) | 0.0049(8) | 0.1063 |
| APGARCH | 1.146(0.004) | -0.7188(0.222) | 0.0153(9) | 0.0142 | 1.241(0.000) | -0.938(0.056) | 0.0370(6) | 0.0006 | 1.070(0.001) | -0.581(0.206) | 0.0165(6) | 0.0026 |
| TGARCH | 1.731(0.002) | -1.716(0.050) | 0.0390(1) | 0.0067 | 1.661(0.000) | -1.674(0.022) | 0.0527(2) | 0.0015 | 1.307(0.002) | -0.999(0.145) | 0.0217(2) | 0.0084 |
| SV | 1.182(0.037) | -0.820(0.372) | 0.0082(10) | 0.0997 | 0.971(0.041) | -0.544(0.480) | 0.0051(10) | 0.1221 | 0.747(0.091) | -0.078(0.913) | 0.0001(10) | 0.1514 |

**Table 7c: Coefficient Estimates, *R2* and unbiasedness test statistics from Mincer-Zarnowitz Regressions (EUR/USD)**

|  |
| --- |
| Coefficient estimates (p-values in parentheses below) from the regression equations , where is one of the three measures of ex-post volatility and is the forecast of this volatility provided by one of the models in Table 1 (and using coefficients estimates from Table 5b). is the coefficient of determination of the regression, and “rank” is the rank of the model’s forecast ordered by . Wald p-value is the p-value of the joint hypothesis that , that is, that the forecast is unbiased. SV is the stochastic volatility model. |
|  | Squared Returns | High-low | Realized Variance |
| Model | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value |
| GARCH | 0.600 | -1.801 | 0.016 | 0.133 | 0.372 | -0.789 | 0.0117 | 0.029 | 0.265 | -0.215 | 0.0022 | 0.030 |
|  | (0.063) | (0.213) | (3) |  | (0.025) | (0.286) | (4) |  | (0.011) | (0.641) | (7) |  |
| t-GARCH | 0.513 | -1.407 | 0.0147 | 0.109 | 0.330 | -0.596 | 0.0101 | 0.017 | 0.249 | -0.142 | 0.0015 | 0.009 |
|  | (0.052) | (0.232) | (4) |  | (0.015) | (0.322) | (5) |  | (0.004) | (0.704) | (9) |  |
| GJR | 0.395 | -0.842 | 0.0055 | 0.231 | 0.233 | -0.152 | 0.0007 | 0.054 | 0.175 | 0.187 | 0.0027 | 0.059 |
|  | (0.141) | (0.465) | (6) |  | (0.090) | (0.797) | (9) |  | (0.042) | (0.610) | (6) |  |
| t-GJR | 0.436 | -1.017 | 0.0095 | 0.129 | 0.267 | -0.299 | 0.0031 | 0.019 | 0.196 | 0.096 | 0.0008 | 0.018 |
|  | (0.079) | (0.338) | (5) |  | (0.037) | (0.582) | (8) |  | (0.014) | (0.777) | (10) |  |
| EGARCH | 0.168 | 0.143 | 0.0001 | 0.503 | 0.053 | 0.596 | 0.0087 | 0.090 | 0.009 | 0.859 | 0.0467 | 0.139 |
|  | (0.591) | (0.911) | (10) |  | (0.737) | (0.359) | (7) |  | (0.927) | (0.032) | (4) |  |
| PGARCH | 0.698 | -2.163 | 0.0175 | 0.131 | 0.432 | -1.019 | 0.0149 | 0.020 | 0.269 | -0.222 | 0.0018 | 0.048 |
|  | (0.069) | (0.192) | (2) |  | (0.028) | (0.229) | (3) |  | (0.029) | (0.674) | (8) |  |
| APGARCH | 0.071 | 0.533 | 0.0028 | 0.513 | 0.023 | 0.712 | 0.0193 | 0.062 | 0.024 | 0.786 | 0.0607 | 0.062 |
|  | (0.780) | (0.602) | (7) |  | (0.861) | (0.171) | (2) |  | (0.764) | (0.014) | (2) |  |
| TGARCH | 0.161 | 0.174 | 0.0002 | 0.528 | 0.045 | 0.633 | 0.0093 | 0.104 | -0.003 | 0.911 | 0.0496 | 0.170 |
|  | (0.615) | (0.894) | (8) |  | (0.781) | (0.343) | (6) |  | (0.979) | (0.027) | (3) |  |
| SV | 0.232 | -0.109 | 0.0001 | 0.221 | 0.171 | 0.101 | 0.0002 | 0.002 | 0.080 | 0.513 | 0.0115 | 0.000 |
|  | (0.575) | (0.943) | (9) |  | (0.417) | (0.898) | (10) |  | (0.540) | (0.290) | (5) |  |
| ID GARCH | -0.670 | 3.415 | 0.2220 | 0.000 | -0.320 | 2.030 | 0.2999 | 0.000 | -0.295 | 2.009 | 0.7953 | 0.000 |
|  | (0.000) | (0.000) | (1) |  | (0.000) | (0.000) | (1) |  | (0.000) | (0.000) | (1) |  |

**Table 7d: Coefficient Estimates, *R2* and unbiasedness test statistics from Mincer-Zarnowitz Regressions (SPX)**

|  |
| --- |
| Coefficient estimates (p-values in parentheses below) from the regression equations , where is one of the three measures of ex-post volatility and is the forecast of this volatility provided by one of the models in Table 1 (and using coefficients estimates from Table 5c). is the coefficient of determination of the regression, and “rank” is the rank of the model’s forecast ordered by . Wald p-value is the p-value of the joint hypothesis that , that is, that the forecast is unbiased. SV is the stochastic volatility model. |
|  | Squared Returns | High-low | Realized Variance |
| Model | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value | (p-value) | (p-value) | (rank) | Wald p-value |
| GARCH | 0.020 | 1.008 | 0.2031 | 0.978 | -0.038 | 0.755 | 0.384 | 0.000 | 0.113 | 0.721 | 0.453 | 0.000 |
|  | (0.919) | (0.000) | (6) |  | (0.678) | (0.000) | (4) |  | (0.136) | (0.000) | (6) |  |
| t-GARCH | 0.065 | 0.855 | 0.1991 | 0.641 | -0.007 | 0.643 | 0.3798 | 0.000 | 0.144 | 0.614 | 0.4477 | 0.000 |
|  | (0.727) | (0.000) | (7) |  | (0.941) | (0.000) | (6) |  | (0.053) | (0.000) | (7) |  |
| GJR | 0.027 | 0.920 | 0.2763 | 0.827 | 0.044 | 0.592 | 0.3849 | 0.000 | 0.136 | 0.636 | 0.5751 | 0.000 |
|  | (0.872) | (0.000) | (5) |  | (0.594) | (0.000) | (2) |  | (0.027) | (0.000) | (5) |  |
| t-GJR | 0.086 | 0.762 | 0.2773 | 0.093 | 0.084 | 0.489 | 0.3845 | 0.000 | 0.177 | 0.527 | 0.5768 | 0.000 |
|  | (0.588) | (0.000) | (4) |  | (0.296) | (0.000) | (3) |  | (0.003) | (0.000) | (4) |  |
| EGARCH | -0.097 | 1.178 | 0.284 | 0.623 | -0.022 | 0.739 | 0.3761 | 0.000 | 0.048 | 0.817 | 0.5959 | 0.000 |
|  | (0.586) | (0.000) | (2) |  | (0.812) | (0.000) | (7) |  | (0.454) | (0.000) | (2) |  |
| t-EGARCH | -0.046 | 1.040 | 0.2810 | 0.964 | 0.005 | 0.660 | 0.3806 | 0.000 | 0.082 | 0.724 | 0.5935 | 0.003 |
|  | (0.792) | (0.000) | (3) |  | (0.953) | (0.000) | (5) |  | (0.189) | (0.000) | (3) |  |
| PGARCH | 0.032 | 0.987 | 0.1865 | 0.982 | -0.042 | 0.757 | 0.3698 | 0.000 | 0.109 | 0.723 | 0.437 | 0.001 |
|  | (0.872) | (0.000) | (8) |  | (0.655) | (0.000) | (8) |  | (0.163) | (0.000) | (8) |  |
| SV | 0.097 | 1.333 | 0.0635 | 0.110 | -0.049 | 1.139 | 0.1561 | 0.841 | 0.072 | 1.149 | 0.2059 | 0.032 |
|  | (0.737) | (0.012) | (9) |  | (0.742) | (0.000) | (9) |  | (0.570) | (0.000) | (9) |  |
| ID GARCH | -0.176 | 1.447 | 0.3528 | 0.047 | -0.148 | 1.027 | 0.5987 | 0.018 | 0.507 | 0.986 | 0.7136 | 0.970 |
|  | (0.295) | (0.000) | (1) |  | (0.042) | (0.000) | (1) |  | (0.925) | (0.000) | (1) |  |

**Table 8a: Loss Function Values and Ranks (AUD/USD 2018)**

|  |
| --- |
| The estimated values of the loss functions given in equations (8) to (13) for forecasts of daily volatility of the AUD/USD exchange rate between 02/03/2018 to 18/07/2018, using each of the models in Table 1 (normally distributed errors) against each of the three measures of ex-post volatility given in Table 6. SV is the stochastic volatility model. The rank of the loss function value, ordered across the models, are given in parentheses below the loss function value. S.Dev is the standard deviation of the ranks across loss functions for that model. Average is the average of the column of standard deviations. |
|  | Squared Returns | High-low | Realized variance |
|  | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev |
| GARCH | 1.317 | 1.397 | -9.452 | 7.807 | 3.175 | 2.930 |  | 0.327 | 0.340 | -9.573 | 0.624 | 1.582 | 1.584 |  | 0.118 | 0.132 | -9.523 | 0.186 | 0.899 | 0.939 |  |
|  | 3 | 2 | 2 | 3 | 3 | 3 | 0.516 | 3 | 3 | 3 | 3 | 4 | 3 | 0.408 | 4 | 4 | 5 | 4 | 4 | 4 | 0.408 |
| GJR | 1.325 | 1.399 | -9.451 | 7.849 | 3.182 | 2.939 |  | 0.331 | 0.343 | -9.572 | 0.632 | 1.581 | 1.584 |  | 0.117 | 0.131 | -9.524 | 0.186 | 0.895 | 0.936 |  |
|  | 4 | 3 | 3 | 4 | 4 | 4 | 0.516 | 4 | 4 | 4 | 4 | 3 | 4 | 0.408 | 3 | 3 | 3 | 3 | 3 | 3 | 0.000 |
| EGARCH | 1.511 | 1.444 | -9.438 | 8.543 | 3.425 | 3.222 |  | 0.449 | 0.450 | -9.528 | 0.833 | 1.868 | 1.917 |  | 0.203 | 0.227 | -9.483 | 0.308 | 1.233 | 1.330 |  |
|  | 8 | 8 | 8 | 8 | 8 | 8 | 0.000 | 8 | 8 | 8 | 8 | 8 | 8 | 0.000 | 8 | 8 | 8 | 8 | 8 | 8 | 0.000 |
| PGARCH | 1.336 | 1.408 | -9.446 | 7.872 | 3.196 | 2.954 |  | 0.336 | 0.348 | -9.569 | 0.638 | 1.599 | 1.604 |  | 0.122 | 0.137 | -9.521 | 0.193 | 0.921 | 0.965 |  |
|  | 6 | 6 | 6 | 5 | 6 | 6 | 0.408 | 6 | 6 | 6 | 5 | 6 | 6 | 0.408 | 7 | 7 | 7 | 7 | 7 | 7 | 0.000 |
| APGARCH | 1.335 | 1.403 | -9.450 | 7.886 | 3.195 | 2.953 |  | 0.336 | 0.346 | -9.570 | 0.640 | 1.595 | 1.599 |  | 0.119 | 0.133 | -9.523 | 0.189 | 0.906 | 0.949 |  |
|  | 5 | 5 | 4 | 6 | 5 | 5 | 0.632 | 5 | 5 | 5 | 6 | 5 | 5 | 0.408 | 5 | 5 | 4 | 5 | 5 | 5 | 0.408 |
| TGARCH | 1.346 | 1.410 | -9.447 | 7.917 | 3.204 | 2.964 |  | 0.339 | 0.350 | -9.569 | 0.645 | 1.601 | 1.607 |  | 0.120 | 0.134 | -9.523 | 0.191 | 0.917 | 0.962 |  |
|  | 7 | 7 | 5 | 7 | 7 | 7 | 0.816 | 7 | 7 | 7 | 7 | 7 | 7 | 0.000 | 6 | 6 | 6 | 6 | 6 | 6 | 0.000 |
| SV | 1.265 | 1.399 | -9.442 | 7.564 | 3.095 | 2.843 |  | 0.300 | 0.322 | -9.579 | 0.564 | 1.461 | 1.454 |  | 0.107 | 0.123 | -9.526 | 0.165 | 0.827 | 0.859 |  |
|  | 2 | 4 | 7 | 1 | 2 | 2 | 2.191 | 2 | 2 | 2 | 2 | 1 | 1 | 0.516 | 2 | 2 | 2 | 2 | 2 | 2 | 0.000 |
| ID GARCH | 1.245 | 1.261 | -9.523 | 7.742 | 3.072 | 2.818 |  | 0.282 | 0.283 | -9.608 | 0.558 | 1.461 | 1.456 |  | 0.062 | 0.065 | -9.558 | 0.107 | 0.676 | 0.691 |  |
|  | 1 | 1 | 1 | 2 | 1 | 1 | 0.408 | 1 | 1 | 1 | 1 | 2 | 2 | 0.516 | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 |
| Average |  |  |  |  |  |  | 0.686 |  |  |  |  |  |  | 0.333 |  |  |  |  |  |  | 0.102 |

**Table 8b: Loss Function Values and Ranks (AUD/USD 2016)**

|  |
| --- |
| The estimated values of the loss functions given in equations (8) to (13) for forecasts of daily volatility of the AUD/USD exchange rate between 01/03/2016 and 15/07/2016,, using each of the models in Table 1 (normally distributed errors) against each of the three measures of ex-post volatility given in Table 5. SV is the stochastic volatility model. The rank of the loss function value, ordered across the models, are given in parentheses below the loss function value. S.Dev is the standard deviation of the ranks across loss functions for that model. Average is the average of the column of standard deviations. |
|  | Squared Returns | High-low | Realized variance |
|  | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev |
| GARCH | 2.678 | 10.28 | -8.525 | 5.779 | 4.197 | 6.543 |  | 1.261 | 7.281 | -8.608 | 0.748 | 2.576 | 4.369 |  | 0.853 | 6.129 | -8.538 | 0.538 | 1.815 | 3.326 |  |
|  | 3 | 3 | 1 | 4 | 5 | 5 | 1.517 | 2 | 2 | 2 | 2 | 4 | 4 | 1.033 | 2 | 1 | 1 | 2 | 2 | 2 | 0.516 |
| GJR | 2.758 | 10.57 | -8.491 | 5.839 | 4.212 | 6.590 |  | 1.375 | 7.703 | -8.548 | 0.797 | 2.657 | 4.517 |  | 0.969 | 6.630 | -8.475 | 0.414 | 1.955 | 3.550 |  |
|  | 6 | 6 | 7 | 6 | 6 | 6 | 0.408 | 6 | 6 | 6 | 6 | 6 | 6 | 0.000 | 6 | 6 | 7 | 6 | 6 | 6 | 0.408 |
| EGARCH | 2.706 | 10.49 | -8.496 | 5.773 | 4.159 | 6.482 |  | 1.307 | 7.513 | -8.569 | 0.760 | 2.570 | 4.356 |  | 0.907 | 6.463 | -8.497 | 0.377 | 1.844 | 3.364 |  |
|  | 5 | 5 | 4 | 3 | 3 | 3 | 0.983 | 5 | 5 | 5 | 5 | 3 | 3 | 1.033 | 5 | 5 | 5 | 5 | 5 | 5 | 0.000 |
| PGARCH | 2.678 | 10.28 | -8.525 | 5.782 | 4.196 | 6.541 |  | 1.263 | 7.290 | -8.606 | 0.748 | 2.577 | 4.371 |  | 0.866 | 6.229 | -8.537 | 0.360 | 1.823 | 3.339 |  |
|  | 2 | 2 | 2 | 5 | 4 | 4 | 1.549 | 3 | 3 | 3 | 3 | 5 | 5 | 1.033 | 3 | 3 | 2 | 3 | 3 | 4 | 0.632 |
| APGARCH | 2.781 | 10.59 | -8.492 | 5.875 | 4.242 | 6.645 |  | 1.396 | 7.749 | -8.547 | 0.812 | 2.693 | 4.583 |  | 0.978 | 6.647 | -8.477 | 0.420 | 1.977 | 3.593 |  |
|  | 7 | 7 | 6 | 7 | 7 | 7 | 0.408 | 7 | 7 | 7 | 7 | 7 | 7 | 0.000 | 7 | 7 | 6 | 7 | 7 | 7 | 0.408 |
| TGARCH | 2.694 | 10.48 | -8.494 | 8.753 | 4.141 | 6.451 |  | 1.297 | 7.492 | -8.570 | 0.753 | 2.553 | 4.325 |  | 0.902 | 6.453 | -8.498 | 0.373 | 1.820 | 3.337 |  |
|  | 4 | 4 | 5 | 2 | 2 | 2 | 1.329 | 4 | 4 | 4 | 4 | 2 | 2 | 1.033 | 4 | 4 | 4 | 4 | 4 | 3 | 0.408 |
| SV | 2.595 | 10.23 | -8.525 | 5.583 | 4.053 | 6.310 |  | 1.194 | 7.179 | -8.610 | 0.690 | 2.427 | 4.122 |  | 0.827 | 6.214 | -8.530 | 0.331 | 1.672 | 3.088 |  |
|  | 1 | 1 | 3 | 1 | 1 | 1 | 0.816 | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 | 1 | 2 | 3 | 1 | 1 | 1 | 0.816 |
| Average |  |  |  |  |  |  | 0.970 |  |  |  |  |  |  | 0.590 |  |  |  |  |  |  | 0.456 |

**Table 8c: Loss Function Values and Ranks (EUR/USD)**

|  |
| --- |
| The estimated values of the loss functions given in equations (8) to (13) for forecasts of daily volatility of the EUR/USD exchange rate between 02/03/2018 to 18/07/2018, using each of the models in Table 1 (normally distributed errors) against each of the three measures of ex-post volatility given in Table 6. SV is the stochastic volatility model. The rank of the loss function value, ordered across the models, are given in parentheses below the loss function value. S.Dev is the standard deviation of the ranks across loss functions for that model. Average is the average of the column of standard deviations. |
|  | Squared Returns | High-low | Realized variance |
|  | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev |
| GARCH | 0.944 | 1.732 | -9.764 | 4.327 | 2.485 | 2.170 |  | 0.287 | 0.468 | -9.801 | 0.552 | 1.235 | 1.158 |  | 0.158 | 0.183 | -9.723 | 0.281 | 0.989 | 0.945 |  |
|  | 2 | 6 | 8 | 2 | 2 | 2 | 2.658 | 2 | 6 | 6 | 2 | 1 | 1 | 2.366 | 6 | 7 | 7 | 6 | 6 | 6 | 0.516 |
| GJR | 0.961 | 1.732 | -9.772 | 4.395 | 2.527 | 2.215 |  | 0.290 | 0.467 | -9.807 | 0.559 | 1.265 | 1.189 |  | 0.155 | 0.180 | -9.729 | 0.274 | 0.979 | 0.942 |  |
|  | 4 | 7 | 5 | 3 | 4 | 4 | 1.378 | 3 | 5 | 5 | 3 | 3 | 3 | 1.033 | 5 | 5 | 5 | 5 | 5 | 5 | 0.000 |
| EGARCH | 0.982 | 1.714 | -9.789 | 4.498 | 2.587 | 2.271 |  | 0.293 | 0.459 | -9.814 | 0.582 | 1.329 | 1.250 |  | 0.147 | 0.169 | -9.738 | 0.268 | 0.962 | 0.923 |  |
|  | 6 | 4 | 4 | 6 | 6 | 6 | 1.033 | 5 | 4 | 4 | 6 | 6 | 6 | 0.983 | 3 | 3 | 4 | 4 | 3 | 3 | 0.516 |
| PGARCH | 0.960 | 1.731 | -9.768 | 4.407 | 2.522 | 2.205 |  | 0.294 | 0.470 | -9.801 | 0.575 | 1.274 | 1.195 |  | 0.159 | 0.181 | -9.726 | 0.287 | 0.996 | 0.953 |  |
|  | 3 | 5 | 7 | 4 | 3 | 3 | 1.602 | 6 | 7 | 7 | 4 | 4 | 4 | 1.506 | 7 | 6 | 6 | 7 | 7 | 7 | 0.516 |
| APGARCH | 0.990 | 1.709 | -9.793 | 4.514 | 2.604 | 2.291 |  | 0.296 | 0.458 | -9.816 | 0.587 | 1.351 | 1.275 |  | 0.148 | 0.169 | -9.738 | 0.268 | 0.969 | 0.931 |  |
|  | 7 | 2 | 2 | 7 | 7 | 7 | 2.582 | 7 | 3 | 2 | 7 | 7 | 7 | 2.345 | 4 | 4 | 3 | 3 | 4 | 4 | 0.516 |
| TGARCH | 0.979 | 1.713 | -9.790 | 4.484 | 2.579 | 2.264 |  | 0.291 | 0.457 | -9.815 | 0.577 | 1.322 | 1.244 |  | 0.146 | 0.168 | -9.739 | 0.266 | 0.955 | 0.916 |  |
|  | 5 | 3 | 3 | 5 | 5 | 5 | 1.033 | 4 | 2 | 3 | 5 | 5 | 5 | 1.265 | 2 | 2 | 2 | 2 | 2 | 2 | 0.000 |
| SV | 1.060 | 1.743 | -9.768 | 4.788 | 2.724 | 2.409 |  | 0.347 | 0.496 | -9.786 | 0.701 | 1.492 | 1.415 |  | 0.184 | 0.196 | -9.718 | 0.346 | 1.126 | 1.088 |  |
|  | 8 | 8 | 6 | 8 | 8 | 8 | 0.816 | 8 | 8 | 8 | 8 | 8 | 8 | 0.000 | 8 | 8 | 8 | 8 | 8 | 8 | 0.000 |
| ID GARCH | 0.885 | 1.531 | -9.892 | 4.320 | 2.483 | 2.166 |  | 0.238 | 0.376 | -9.867 | 0.496 | 1.246 | 1.169 |  | 0.081 | 0.088 | -9.790 | 0.168 | 0.770 | 0.723 |  |
|  | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 | 1 | 1 | 1 | 1 | 2 | 2 | 0.516 | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 |
| Average |  |  |  |  |  |  | 1.388 |  |  |  |  |  |  | 1.252 |  |  |  |  |  |  | 0.258 |

**Table 8d: Loss Function Values and Ranks (SPX)**

|  |
| --- |
| The estimated values of the loss functions given in equations (8) to (13) for forecasts of daily volatility of the SPX between 02/03/2018 to 18/07/2018, using each of the models in Table 1 (normally distributed errors) against each of the three measures of ex-post volatility given in Table 6. SV is the stochastic volatility model. The rank of the loss function value, ordered across the models, are given in parentheses below the loss function value. S.Dev is the standard deviation of the ranks across loss functions for that model. Average is the average of the column of standard deviations. |
|  | Squared Returns | High-low | Realized variance |
|  | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev | MSE1 | MSE2 | QLIKE | R2LOG | MAE1 | MAE2 | S.Dev |
| GARCH | 2.999 | 13.978 | -8.697 | 5.856 | 4.451 | 7.101 |  | 1.340 | 3.882 | -9.042 | 1.500 | 3.086 | 4.570 |  | 0.763 | 2.554 | -8.775 | 0.671 | 2.241 | 3.575 |  |
|  | 4 | 4 | 4 | 5 | 5 | 5 | 0.548 | 4 | 2 | 4 | 5 | 5 | 4 | 1.095 | 4 | 3 | 4 | 4 | 4 | 4 | 0.408 |
| GJR | 2.848 | 12.740 | -8.796 | 5.658 | 4.309 | 6.946 |  | 1.364 | 4.924 | -9.093 | 1.346 | 3.042 | 4.774 |  | 0.700 | 2.697 | -8.827 | 0.558 | 2.103 | 3.519 |  |
|  | 3 | 3 | 3 | 4 | 3 | 3 | 0.408 | 5 | 6 | 3 | 3 | 4 | 6 | 1.378 | 3 | 5 | 3 | 3 | 3 | 3 | 0.816 |
| EGARCH | 2.753 | 12.678 | -8.810 | 5.621 | 4.238 | 6.782 |  | 1.185 | 3.935 | -9.111 | 1.283 | 2.946 | 4.516 |  | 0.567 | 1.817 | -8.853 | 0.513 | 1.944 | 3.099 |  |
|  | 2 | 2 | 2 | 3 | 2 | 2 | 0.408 | 2 | 3 | 2 | 2 | 3 | 3 | 0.548 | 2 | 2 | 2 | 2 | 2 | 2 | 0.000 |
| PGARCH | 3.056 | 14.270 | -8.687 | 5.918 | 4.486 | 7.177 |  | 1.376 | 3.960 | -9.034 | 1.536 | 3.136 | 4.650 |  | 0.786 | 2.610 | -8.769 | 0.692 | 2.293 | 3.660 |  |
|  | 5 | 5 | 5 | 6 | 6 | 6 | 0.548 | 6 | 4 | 5 | 6 | 6 | 5 | 0.816 | 5 | 4 | 5 | 5 | 5 | 5 | 0.408 |
| SV | 3.170 | 17.171 | -8.446 | 5.518 | 4.390 | 7.030 |  | 1.312 | 4.410 | -8.966 | 1.378 | 2.888 | 4.210 |  | 0.951 | 3.423 | -8.654 | 0.742 | 2.350 | 3.708 |  |
|  | 6 | 6 | 6 | 2 | 4 | 4 | 1.633 | 3 | 5 | 6 | 4 | 2 | 2 | 1.633 | 6 | 6 | 6 | 6 | 6 | 6 | 0.000 |
| ID GARCH | 2.362 | 12.062 | -8.920 | 5.015 | 3.906 | 6.334 |  | 0.671 | 2.261 | -9.247 | 0.821 | 2.219 | 3.245 |  | 0.309 | 1.154 | -8.943 | 0.295 | 1.395 | 2.181 |  |
|  | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 | 1 | 1 | 1 | 1 | 1 | 1 | 0.000 |
| Average |  |  |  |  |  |  | 0.591 |  |  |  |  |  |  | 0.912 |  |  |  |  |  |  | 0.272 |

**Table 9: HLN Forecast Comparison Tests**

|  |
| --- |
| P-values of the HLN (1998) forecast comparison test applied to forecasts from models drawn from Table 1 for each of the three ex-post volatility measures. The test is conducted using both an MSE and an MAE criterion. SV is the stochastic volatility model. IDG is the intra-day GARCH model. |
|  | Squared returns | High-low | Realized variance |
| Model comparison | MSE | MAE | MSE | MAE | MSE | MAE |
| AUD/USD 2018 |
| IDG Vs. GARCH | 0.070 | 0.044 | 0.023 | 0.017 | 0.003 | 0.000 |
| IDG Vs. EGARCH | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| IDG Vs. PGARCH | 0.025 | 0.018 | 0.010 | 0.008 | 0.000 | 0.000 |
| IDG Vs. TGARCH | 0.011 | 0.008 | 0.004 | 0.006 | 0.000 | 0.000 |
| IDG Vs. SV | 0.677 | 0.690 | 0.452 | 0.995 | 0.001 | 0.007 |
| AUD/USD 2018 |
| SV Vs. GARCH | 0.005 | 0.007 | 0.009 | 0.000 | 0.090 | 0.025 |
| SV Vs. EGARCH | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| SV Vs. PGARCH | 0.000 | 0.001 | 0.001 | 0.000 | 0.040 | 0.007 |
| SV Vs. TGARCH | 0.000 | 0.001 | 0.001 | 0.000 | 0.093 | 0.021 |
| AUD/USD 2016 |
| SV Vs. GARCH | 0.814 | 0.260 | 0.457 | 0.131 | 0.970 | 0.052 |
| SV Vs. EGARCH | 0.267 | 0.578 | 0.073 | 0.384 | 0.011 | 0.121 |
| SV Vs. PGARCH | 0.822 | 0.297 | 0.477 | 0.148 | 0.917 | 0.035 |
| SV Vs. TGARCH | 0.242 | 0.633 | 0.066 | 0.427 | 0.00 | 0.130 |
| EUR/USD |
| IDG Vs. GARCH | 0.409 | 0.976 | 0.243 | 0.877 | 0.001 | 0.000 |
| IDG Vs. EGARCH | 0.057 | 0.045 | 0.065 | 0.132 | 0.001 | 0.000 |
| IDG Vs. PGARCH | 0.255 | 0.533 | 0.151 | 0.673 | 0.000 | 0.000 |
| IDG Vs. TGARCH | 0.068 | 0.060 | 0.076 | 0.160 | 0.001 | 0.000 |
| IDG Vs. SV | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| SPX |
| IDG Vs. GARCH | 0.010 | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 |
| IDG Vs. EGARCH | 0.086 | 0.033 | 0.000 | 0.000 | 0.000 | 0.000 |
| IDG Vs. PGARCH | 0.006 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000` |
| IDG Vs. TGARCH | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| IDG Vs. SV | 0.079 | 0.033 | 0.001 | 0.003 | 0.000 | 0.000 |

1. . McKenzie (1998), Anderton and Skudelny (2001) and Choudhry and Hassan (2015), for example, demonstrate the significant impact of exchange rate volatility on international trade flows. [↑](#footnote-ref-1)
2. . Examples, using exchange rate data, include Cumby et al (1993), Jorion (1995) and Figlewski (1997). [↑](#footnote-ref-2)
3. . A further example of the sensitivity of forecast rankings to loss functions is Brailsford and Faff (1996) who show that asymmetric loss functions are more likely to favour forecasts from asymmetric volatility models. [↑](#footnote-ref-3)
4. . The US Dollar against the Canadian Dollar and the Swiss Franc. [↑](#footnote-ref-4)
5. . Early studies of the volatility of financial asset returns by, for example, Mandelbrot (1963), Fama (1965), Praetz (1969), Clark (1973), and Taylor and Kingsman (1979), found that the variances of financial asset prices vary across time. The development of the ARCH models of conditional variances of time series, Engle (1982), whose properties also matched closely the empirical distributions of many asset return classes, led to a huge growth in the development of variance modelling. Early examples the application of these models to exchange rate data include Taylor (1980), Diebold (1988), Hsieh (1989), Baillie and Bollerslev (1989) and Bera and Higgins (1992). Surveys include Bollerslev et al (1992) and Bollerslev et al (1994). [↑](#footnote-ref-5)
6. . In the survey paper by Poon and Granger (2003) of 93 papers available up to that time, fewer than 20 papers focussed on the volatility of spot exchange rates, with all focussing on some subset of the cross rates between the US Dollar, Canadian Dollar, Japanese Yen, German Mark, French Franc, Swiss Franc and Italian Lira. [↑](#footnote-ref-6)
7. . This result is not specific to exchange rates, see for example, Akgiray (1989), Pagan and Schwert (1990), and Day and Lewis (1992) for US stock indexes and Brailsford and Faff (1996) for individual Australian stocks. Somewhat higher values are reported by Blair et al (2002) for a US stock index but, for non-combined forecasts, are under 23 percent for GARCH and historical models. [↑](#footnote-ref-7)
8. . Although with an empirical application to stock volatility indices, Ma et al (2018) demonstrate that lower frequency data can add to the forecast performance of models that otherwise only use high frequency data, while Gulay and Emec (2017) explore a model-free (normalization and variance stabilization) method to generate forecasts. [↑](#footnote-ref-8)
9. 9. Baillie and Bollerslev (1989) first recommended the use of the t-distribution for modelling exchange rates with GARCH processes and this recommendation is repeated in the study by Hansen and Lunde (2005). [↑](#footnote-ref-9)
10. . There is an alternative explanation. If volatility is priced, it can feedback into returns requiring a decrease in price, see, for example French et al (1987) and Campbell and Hentschel (1992). Moreover, Duffee (1995) shows that the “leverage” effect is due to a contemporaneous positive relationship between returns and volatility that is strongest in firms with little actual leverage, and Figlewski and Wang (2001) suggests that it is a “down-market” effect that has little connection to leverage. Bollerslev et al (2006) show that high frequency data facilitates the distinction between alternative explanations of asymmetry. [↑](#footnote-ref-10)
11. 11. Asymmetric effects are modelled in exchange rate volatility by, for example, Tse and Tsui (1997), Hu, Jiang and Tsoukalas (1997), McKenzie and Mitchell (2002), McMillan and Speight (2004), Hansen and Lunde (2005), Huang et al (2009), Villar (2010), Abdalla (2012) and Ramasamy and Munisamy (2012) although the relative forecasting performance of models that accommodate asymmetry is mixed. [↑](#footnote-ref-11)
12. . The repeated notation for the coefficients and the error term across the four different equations, (2a) to (2d), is for notational simplicity and does not imply that the coefficients nor error term values would be expected to be the same across the four equations. [↑](#footnote-ref-12)
13. . In the original formulation of the GJR GARCH model, the indicator variable takes the value of one when the shock is negative, and is zero otherwise. The reversed formulation here (indicating positive shocks) enables the interpretation of the sign of $γ$ to remain the same across all the models incorporating asymmetry. [↑](#footnote-ref-13)
14. . Long memory in volatility, which gives rise to such autocorrelations, can also be modelled using the Component GARCH model of Engle and Lee (1999) and the Fractionally Integrated GARCH model of Baillie et al (1996). Applications to exchange rate volatility include Vilasuso (2002), Chortareas et al (2011) and McMillan and Speight (2012). Since Lamoureux and Lastrapes (1990) have shown that structural shifts can generate spurious persistence in volatility, and extensive simulations in Ding el al (1993) show that simple GARCH models are capable of capturing the autocorrelation structure in financial data, we examine only the PARCH and APARCH models. [↑](#footnote-ref-14)
15. . As we find that empirically we cannot reject the null hypothesis that all of the ARMA coefficients and the constant term in equation (1) are zero, the residual from equation (1) is $u\_{t}=y\_{t}$, and so the state space form uses log squared returns in the measurement equation (3). [↑](#footnote-ref-15)
16. . Although Liu, Patton and Sheppard (2015) do not examine spot exchange rate volatility, they also find that 5 minute and 15 minute return intervals produce the best forecasts among over 400 models applied to data between 2000 and 2010 for individual stocks, stock indexes and futures contracts on indexes, interest rates and currencies. [↑](#footnote-ref-16)
17. . The scaling of the squared log range depends on the underlying assumed data generating process. The value here, 4ln(2), is consistent with a diffusion having zero mean and constant conditional daily volatility, see Parkinson (1980) and Patton (2011). [↑](#footnote-ref-17)
18. . A recent application of these tests, to USD/EUR exchange rate volatility forecasts, is Laurent and Violante (2012). [↑](#footnote-ref-18)
19. . The overlapping sample dates for the SPX are caused in part by the higher number of market closings for the equity market than the currency market, and also that the data collection for the SPX was one week later due to a later decision to include this asset class in the revised paper. [↑](#footnote-ref-19)
20. . In the original draft of this paper, we examined only this earlier time period. In the revised draft, we lead with the more recent data sample, and use the original sample as a control. We find similar results in both samples. [↑](#footnote-ref-20)
21. . A similar approach to removing slower weekend trading periods was used by McMillan and Speight (2012) following the direction in Bollerslev and Domowitz (1993). For the 2016 ending sample for the AUD/USD, the 24-hour day ending was 16.45 UK time (GMT +1). [↑](#footnote-ref-21)
22. We were unable to obtain intra-day data for the estimation sample ahead of the 2016 AUD/USD evaluation sample from our data provider and so for this series we consider forecasts only derived from daily data. [↑](#footnote-ref-22)
23. . Evidence of a greater impact on volatility of local currency depreciations (against a major currency) can also be found in Hu et al (1997) and Tse and Tsui (1997). [↑](#footnote-ref-23)
24. . Although this Figure is for the sample period ending 2016, a similar picture emerges for the sample period ending 2018. [↑](#footnote-ref-24)
25. . McMillan and Speight (2012) question the robustness of earlier results, but we conjecture that it is more likely that this weak testing framework is highly susceptible to the specifics of a given sample. By contrast to McMillan and Speight (2012), Chortareas et al (2011) document this finding for some but not all of the exchange rate series analysed. [↑](#footnote-ref-25)
26. . As the loss function scores for the predictor models with *t*-distributed errors are mostly indistinguishable from the same model applied assuming normally distributed errors, we report in Table 8 only the latter, to make the comparison across different model specifications more apparent. The former results are available on request to the authors. [↑](#footnote-ref-26)
27. . Although the estimation sample for the AUD/USD is almost 4000 observations, this only equates to around 36 trading days prior to the forecast evaluation period that is 99 days in duration. [↑](#footnote-ref-27)
28. . A similar result emerges for the EUR/USD. By contrast, for the SPX, the high-low estimator results in loss functions must less converged than either the squared returns (p<0.05) or the realized variance (p<0.02). [↑](#footnote-ref-28)
29. . The results for the SPX echo those of the 2016 AUD/USD sample, while the EUR/USD has significant differences between each variance estimator pairings and also with the more accurate estimators better able to distinguish the first and second best models. [↑](#footnote-ref-29)
30. . The link between information arrivals and volatility of financial returns was originally proposed and demonstrated in papers by Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983) and a stochastic volatility model was extensively applied to various financial markets by Taylor (1986). [↑](#footnote-ref-30)