Forecasting Daily Return Densities from Intraday Data: a Multifractal Approach

Mark Hallam, *City University London & Koç University*
Jose Olmo, *University of Southampton*

October 2013

Abstract

This paper proposes a new approach for estimating and forecasting the moments and probability density function of daily financial returns from intraday data. This is achieved through a new application of the distributional scaling laws for the class of multifractal processes. Density forecasts from the new multifractal approach are typically found to provide substantial improvements in predictive ability over existing forecasting methods for the EUR/USD exchange rate and are also competitive with existing methods when forecasting the daily return density of the S&P500 and NASDAQ-100 equity index.

1 Introduction

Over the past decade there has been a dramatic increase in the availability of intraday financial data, resulting in an extensive literature on the use of high-frequency data in financial econometrics. These data obviously allow for the study of financial market behaviour at intraday timescales, but they also contain potentially valuable information for longer timescales, which are arguably of more interest for most market participants. As a result, there have been efforts to incorporate intraday data into the modelling and forecasting of financial variables at daily or even lower frequencies.

The most notable example is provided by the large literature on realised volatility, a concept that was introduced by Andersen & Bollerslev (1998) and then subsequently formalised by Andersen et al. (2001). The daily realised volatility is obtained by summing the squared intraday returns observed during that day and can then be used as an estimate of the unobservable daily volatility. It has been found (see for example Andersen et al., 1999, or Andersen et al., 2003) that the use of high-frequency data in the form of these realised volatility measures can provide significant improvements in the modelling and forecasting of daily return volatility compared to models using only daily data.

Whilst return volatility is undoubtedly a variable of substantial academic and practical interest, there are many situations encountered in finance in which information concerning just the first two moments of the distribution of returns is not sufficient. Perhaps most obviously, risk management problems, such as the calculation of value-at-risk and expected shortfall, require knowledge of particular quantiles of the return distribution. In addition, numerous studies including Harvey & Siddique (2000) and Brooks et al. (2005) have shown that higher moments such as skewness and kurtosis are time varying and there is empirical support for these higher moments being relevant in problems of portfolio allocation and asset pricing (see for example Harvey & Siddique, 2000, or Dittmar, 2002, amongst others).

However, as noted by Žikěs (2009), the use of intraday data to model and forecast characteristics of daily return distributions beyond the first two moments is not a subject that has yet received much attention. Notable exceptions include Andersen et al. (2003), Giot & Laurent (2004), Clements, Galvão, & Kim (2008) and Maheu & McCurdy (2010), all of which extend the use of realised volatility measures to either quantiles of daily returns or the entire daily return density. The methods used by all of these previous studies to link the realised volatility measures produced from intraday data to the density (or quantiles) of daily returns consist of two components. The first is a parametric time series model for volatility incorporating one or more realised volatility measures, which is used to model and produce point forecasts for daily volatility. The specific parametric volatility model varies, with Giot & Laurent (2004) using an ARCH-type model, Clements, Galvão, & Kim (2008) considering mixed data sampling (MIDAS) and heterogenous autoregressive (HAR) models and a bivariate VAR being used by Andersen et al. (2003). The second component is a parametric distributional assumption about daily returns, allowing density or quantile forecasts for daily returns to be produced from the point forecasts of daily realised volatility. This is typically based on the finding of Andersen et al. (2001), that daily returns are approximately normally distributed when standardised by their corresponding daily realised volatilities, although Clements, Galvão, & Kim (2008) also explore the use of an empirical distribution estimated from the data.

In response to these theoretical limitations, Hallam & Olmo (2013) propose a method for estimating and forecasting the probability density of daily returns from intraday data, based on a new application of distributional scaling laws for the class of unifractal processes. These processes possess a form of scale invariance, such that the distribution of the process at a given timescale is related to that at any other timescale through a distributional scaling law.

Under the assumption of unifractality, the form of this distributional scaling can be estimated for a given sample of data and it was demonstrated how these estimates can be used to appropriately rescale the intraday returns such that they are equal in distribution to daily returns; the density of daily returns can then be directly estimated from these rescaled intraday observations. It should however be noted that the applicability of the proposed approach only requires unifractal distributional scaling to be present locally over each estimation window and for the range of timescales that are of direct interest, rather than globally over all sampling intervals and sub-periods as for a true unifractal process in the traditional sense.

In contrast to existing methods, information concerning both the magnitude and sign of intraday returns can be incorporated into the estimates of the daily return density. Furthermore, this approach also allows the use of nonparametric density estimation methods, thus removing the need to impose a specific parametric form for the density of daily returns. The empirical application of Hallam & Olmo (2013) suggests that the proposed unifractal density forecasting method produces density forecasts that perform well when the true scaling behaviour of the return processes is sufficiently close to that of a unifractal process, even if it is not exactly unifractal. However, it also appears that the predictive ability of the unifractal approach can be adversely affected by larger deviations from the unifractal distributional scaling behaviour that is required for the method to be theoretically valid.

The current paper therefore proposes an alternative approach for producing density forecasts for daily returns from intraday data, based on distributional scaling laws for the more general class of multifractal processes. Compared to unifractal processes, multifractal processes allow for a more flexible scaling relationship between return distributions at different sampling frequencies, overcoming a key theoretical limitation of the previous method. However, whilst the multifractal approach of the current paper permits more flexible distributional scaling behaviour than the earlier unifractal approach, the implementation of the method is more restrictive in some respects, most notably requiring a parametric form to be selected for the daily return distribution. Nonetheless, the proposed method still allows the intraday data to directly influence properties of the daily return density beyond the second moment. In particular, the approach allows the kurtosis of daily returns to be estimated directly from the intraday data and incorporated into the forecasts of the density of daily returns.

The aim of the current work is to formalise this alternative multifractal approach and explore whether the additional flexibility it permits in terms of distributional scaling behaviour allows it to produce accurate density forecasts, despite the more restrictive implementation it requires compared to the competing unifractal approach. The density forecasting performance of the proposed multifractal approach is compared to that of benchmark models from the GARCH and realised volatility literature, in addition to the unifractal approach of Hallam & Olmo (2013) in an empirical application using a dataset of 5-minute intraday equity and exchange rate data.

The structure of the paper is as follows: Section 2 presents the relevant theory on unifractal and multifractal processes and describes how these results can be applied to link the properties of the return process at different sampling frequencies. Section 3 then discusses how these concepts can be applied in practice for the multifractal case to estimate and forecast the moments of daily returns and ultimately forecast the daily return density. Section 4 presents the empirical application of the new multifractal approach and finally, Section 5 concludes.

2 Unifractal & Multifractal Processes

In order to estimate the density of daily returns from intraday data, a method for formally linking the characteristics of return distributions across different sampling frequencies is required. In contrast to the existing work in the literature based on realised volatility measures, the proposed method relies instead on results from the theory of unifractal and multifractal processes.

On an intuitive level, such stochastic processes exhibit some form of *scale invariance*, such that the behaviour of the process observed at one timescale is, after an appropriate transformation, identical in a statistical sense to that observed at another time scale. Whilst much of the work with these processes originated in physics, a large number of empirical studies have subsequently confirmed the existence of this type of distributional scaling behaviour in a wide range of financial time series[1] and this has led to the development of several asset pricing models that explicitly reproduce this distributional scaling behaviour[2].

The current section begins with a brief summary of the theoretical properties of these processes (with more detailed treatments found in Mandelbrot, Fisher, & Calvet, 1997, Calvet & Fisher, 2002 or Kantelhardt, 2009), before exploring how these properties could be applied to relate the distributional properties of the return process at the intraday and daily sampling intervals.

[1]Empirical studies include Calvet & Fisher (2002), Fillol (2003), Matia, Ashkenazy, & Stanley (2003), Xu & Gençay (2003), Di Matteo, Aste, & Dacorogna (2005), Selçuk & Gençay (2006), Di Matteo (2007) and Onali & Goddard (2009)

[2]Key examples include the multifractal model of asset returns of Mandelbrot, Fisher, & Calvet (1997) and Calvet & Fisher (2002), the Markov switching multifractal model of Calvet & Fisher (2004) and Calvet, Fisher, & Thompson (2006) and the multifractal random walk of Muzy, Delour, & Bacry (2000)

2.1 A Review of Unifractal and Multifractal Processes

The distributional scaling behaviour of a unifractal or self-affine process can be defined by a simple expression that links the distribution of the process at different sampling intervals. Formally, unifractal or self-affine processes can be defined in the following way:

Definition 2.1. A process is said to be *self-affine* or *unifractal* if for some $H > 0$, all $c \geq 0$ and all $t_1, t_2, \ldots, t_k \geq 0$ it obeys the distributional scaling relationship

$$\{X(ct_1), X(ct_2), \ldots, X(ct_k)\} \stackrel{d}{=} \{c^H X(t_1), c^H X(t_2), \ldots, c^H X(t_k)\} \tag{2.1}$$

which can be expressed more compactly as:

$$X(ct) \stackrel{d}{=} c^H [X(t)] \tag{2.2}$$

If the increments of the process are stationary, then the distributional scaling law of (2.2) also holds at the local level:

$$X(t + c\Delta t) - X(t) \stackrel{d}{=} c^H [X(t + \Delta t) - X(t)] \tag{2.3}$$

The parameter H is known as the self-affinity index and can be estimated for a specific time series of data using a variety of methods. Common examples of unifractal processes in finance include the standard Brownian motion, for which $H = 1/2$, and also the more general fractional Brownian motion (and the corresponding increment series, the fractional Gaussian noise), for which H is constant but not constrained to be equal to $1/2$. Equations (2.2) and (2.3) state that the distribution of the process $X(t)$ and the corresponding increment series are, after an appropriate rescaling, identical when the time scale of the process is changed. In the current context this implies that the distribution of returns over different horizons or sampling intervals, for example 1 hour and 1 day returns, are identical after rescaling by a factor that depends on the characteristics of the particular return process (via H) and the difference between the two sampling intervals (via c).

One can also consider the more general class of multifractal processes, which allow for a more flexible relationship between distributions across different sampling frequencies. In the case of a multifractal process, equations (2.2) and (2.3) can be generalised to:

$$X(ct) \stackrel{d}{=} c^{H(c)} [X(t)] \tag{2.4}$$

and

$$X(t + c\Delta t) - X(t) \stackrel{d}{=} c^{H(c)} [X(t + \Delta t) - X(t)] \tag{2.5}$$

where the scaling factor c^H has been replaced by the more general function of $c, c^{H(c)}$, allowing for a more flexible scaling relationship between distributions over different sampling frequencies than in the unifractal case. An alternative characterisation of scaling behaviour is often used, particularly in the case of multifractal processes, for which equation (2.4) is perhaps somewhat less intuitive than the unifractal analogue of (2.2). It can be shown (see for example Mandelbrot et al., 1997) that a stochastic process $X(t)$ with increments $X(t + \Delta t) - X(t)$ is multifractal if these increments are stationary and satisfy:

$$E[\mid X(t + \Delta t) - X(t) \mid^q] = c(q)(\Delta t)^{\tau(q)+1} \tag{2.6}$$

where $c(q)$ and $\tau(q)$ are deterministic functions of q. The function $\tau(q)$ in (2.6) is referred to as the *scaling function* and describes the scaling behaviour for different moments (i.e. values of q) of the absolute increments of the process $X(t)$ for a given range of sampling intervals, Δt. It can be demonstrated (see Calvet & Fisher, 2002) that for a multifractal process the scaling function is non-linear (though always concave with intercept equal to -1), implying that different moments of the absolute increments scale differently with the sampling interval, Δt, than others. For a unifractal process (2.6) also holds, but the scaling function is linear and of the form $\tau(q) = Hq - 1$, where H is the same self-affinity index from equations (2.1) and (2.2). As with the self-affinity index, H, for a unifractal process, the scaling function can be estimated for a particular time series using various methods (see Kantelhardt, 2009, for a survey of several common estimators).

2.2 Application of the Distributional Scaling Laws

Before explaining how the distributional scaling laws for multifractal processes can be applied to estimate the density of daily return from intraday data, it is beneficial to begin with a brief summary of the method proposed by Hallam & Olmo (2013) under the assumption of a unifractal return process, in order to emphasise the differences between the methods and explain why an identical approach cannot be used in the multifractal context.

Assume that a series of intraday returns are observed over a given time period and we wish to use these intraday returns to estimate the density of daily returns over the same time period. Denote the density functions of the intraday and daily returns over the time period by $f(y_I)$ and $f(y_D)$ respectively. Under the assumption that the return process is unifractal, it must satisfy the distributional scaling laws of equation (2.2) and (2.3), which in the current context imply that:

$$f(y_D) = f(c^H y_I) \tag{2.7}$$

From (2.7), the density of daily returns is equal to the density of the intraday returns, when these intraday returns have been appropriately rescaled by a factor consisting of two components: c and the self-affinity index, H. From (2.1), the value of c is determined solely by the relative lengths of the two sampling intervals and the self-affinity index, H, can be estimated from the intraday data using various estimators[3].

The resulting estimate of the self-affinity index, denoted by $\hat{H}$, can then be combined with the appropriate value of c to rescale the intraday returns by the factor $c^{\hat{H}}$; the density of these rescaled intraday returns can then be estimated and from (2.7) the resulting estimate can then be viewed as an estimate of the density of daily returns over the same period. The large number of rescaled intraday returns obtained over even short time periods allows a wide range of methods to be used to estimate the daily return density from these observations, including nonparametric methods, such as kernel estimation, as used by Hallam & Olmo (2013).

It is therefore relatively straightforward to estimate the density of daily returns from intraday data when the return process is assumed to be unifractal. In the multifractal case the direct analogue of the distributional scaling rule in (2.3) employed above for the unifractal case is given by (2.5), with the simple scalar H replaced with the function $H(c)$. If this function could be estimated (as H can be in the unifractal case), then it would be possible to proceed in the same way as before. Unfortunately there is no existing method for estimating this function and so any application in the multifractal case must be based on an alternative representation of scaling behaviour.

Instead, we will employ the moment scaling property of multifractal processes given in equation (2.6), which has been widely used in empirical studies of multifractal processes in finance as the basis for estimating the scaling function, $\tau(q)$, for a given sample of time series data. However, at least within the finance literature, the resulting estimates of the scaling function have only been used to assess whether the distributional scaling properties of the return process are consistent with that of a multifractal process and have not been employed to estimate the moments of the time series process at one sampling interval from data observed at a different timescale.

One possible reason for this is that the moment scaling condition is not immediately applicable in the this context without making some additional assumptions; in particular, equation (2.6) describes how the non-central moments of the absolute increments of the price process scale with the sampling interval, but what is of more interest for financial returns is the scaling behaviour of the central moments (such as variance, skewness and kurtosis) of the untransformed returns (i.e. the increments of the log price process).

If it can be assumed that the expected value of returns is zero, the central and non-central moments are equal. Furthermore, for all even values of q in (2.6), the moments of the increments and absolute increments are equal. Therefore, under the assumption of a multifractal return process with mean of zero, from equation (2.6) the q-th central moment of the return process at sampling interval Δt, denoted by $m(q, \Delta t)$, is given by:

$$m(q, \Delta t) = c(q)\Delta t^{\tau(q)+1} \tag{2.8}$$

for all even values of q. The scaling function, $\tau(q)$, and the prefactor, $c(q)$, can both be

[3] A detailed survey of common estimators can be found in Kantelhardt (2009)

estimated from a given sample of intraday data using the method presented in Section 3.1 below. These values can then be used to produce an estimate of the q-th central moment of returns at any sampling interval for all even numbered values of q. In particular, this allows both the variance and kurtosis of daily returns for a given time period to be estimated from intraday data observed over the same period.

Unlike the unifractal approach of Hallam & Olmo (2013), this method does not produce a sample of rescaled intraday data from which the daily return density can be estimated. Instead, a parametric distributional form can be assumed that is uniquely determined by the moments estimated from the intraday data, with some possible candidates discussed in Section 3.3.

3 Estimating and Forecasting the Moments of Daily Returns from Intraday Data

The current section demonstrates how the theoretical results from the previous section can be used in practice to estimate the moments of daily returns from intraday data under the assumption that the return process is multifractal, before proceeding to the problem of forecasting the moments and density of daily returns. Section 3.1 begins by describing the chosen method for estimating the scaling function, $\tau(q)$, and the prefactor, $c(q)$, from a given sample of intraday data, which can then provide estimates of the moments of the return process at the daily sampling interval. Section 3.2 then moves to a dynamic context and considers how the moment estimates produced in this way can be used to produce out-of-sample forecasts for the daily return moments and finally Section 3.3 discusses a possible method for constructing density forecasts for daily returns from these point forecasts for the daily return moments.

3.1 Estimation of the Multifractal Scaling Function

Estimating the moments of daily returns from intraday data requires estimates of the scaling function, $\tau(q)$, and the prefactor, $c(q)$, for the relevant values of q. Whilst many methods have been proposed for estimating the scaling function, the majority do not provide a direct estimate of the prefactor, since this is typically not of direct interest in most studies of multifractal processes in finance or elsewhere, which focus almost exclusively on the scaling function. However, from the discussion in Section 2.2 it is clear that an estimate of $c(q)$ is indeed required for the current application, in order to estimate the moments via equation (2.8)

Initially the partition function estimator was employed for estimation; this is one of the simpler estimators for $\tau(q)$, but was selected primarily because it also provides a direct estimate of the prefactor, in addition to being one of the most commonly employed estimators in the multifractal finance literature (see for example Calvet & Fisher, 2002). Subsequently, more complex estimators for the scaling function were also tested, including the multifractal detrended fluctuation analysis method of Kantelhardt et al. (2002) and the multifractal detrended/centred moving average method of Schumann & Kantelhardt (2011). These more complex methods do not however estimate the prefactor and so the estimates of $\tau(q)$ obtained from these alternative estimators were combined with the corresponding partition function method estimate of $c(q)$. Interestingly however, despite being a simpler estimator for the scaling function, the density forecasting performance of the method when using the partition function approach was actually found to be superior than that obtained when using these alternative estimators.

The partition function method is based directly on the multifractal moment scaling condition of (2.6), which must be satisfied by any multifractal process. If the process $X(t)$ is observed over the interval $[0, T]$ and this interval is divided into N subintervals of length Δt then the q-th order *partition function* of $X(t)$ is defined as:

$$S_q(T, \Delta t) \equiv \sum_{i=0}^{N-1} |X(i\Delta t + \Delta t) - X(i\Delta t)|^q$$

From the stationarity of the increments of $X(t)$ it follows that:

$$E\left[S_q(T, \Delta t)\right] = N \cdot E\left[|X(i\Delta t + \Delta t) - X(i\Delta t)|^q\right]$$

Then from the multifractal moment scaling condition of (2.6) and the fact that $N\Delta t = T$:

$$\begin{aligned} E\left[S_q(T, \Delta t)\right] &= Nc(q)(\Delta t)^{\tau(q)+1} \\ \log E\left[S_q(T, \Delta t)\right] &= \log c(q) + \log T + \tau(q)\log(\Delta t) \\ \log E\left[S_q(T, \Delta t)\right] &= c^*(q) + \tau(q)\log(\Delta t) \end{aligned} \tag{3.1}$$

where $c^*(q) = \log c(q) + \log T$. Therefore, by calculating the value of $S_q(T, \Delta t)$ for a range of sampling intervals, Δt, it is possible to estimate the value of the scaling function, $\tau(q)$, for a given value of q via (3.1) from the slope of $\log S_q(T, \Delta t)$ plotted against $\log(\Delta t)$. The corresponding value of the prefactor, $c(q)$, can be estimated via the intercept. In practice this process requires a minimum and maximum sampling interval (i.e. value of Δt) to be selected. This is largely an empirical issue, with the optimal choices being dependent on the intended application of the estimated scaling function and also to some extent on the characteristics of the time series in question; as such, this issue will be discussed further during the empirical exercise of Section 4.

Fisher, Calvert & Mandelbrot (1997) suggest the use of OLS to obtain estimates of the slope and intercept for each partition function and this is generally the method that has been employed since. Typically this process is repeated for a range of values of q, producing estimates of a set of points on the scaling function; an estimate of the complete function $\tau(q)$ can then be obtained by fitting a curve to this set of points. For the current application this only needs to be performed for the values of q corresponding to the moments of interest. Assuming that we wish to estimate both the variance and kurtosis of daily returns, then the second and fourth central moments are required, which can be obtained from the estimated values of $c(q)$ and $\tau(q)$ for $q = 2$ and $q = 4$. In principle higher order moments can also be estimated in the same way, but it has been noted (see Schmitt et al., 1999) that estimates of the scaling function from a finite time series will become less reliable as the value of q increases. Given that these higher moments have less direct interpretation in finance, attention will be restricted to the second and fourth moments for now.

3.2 Forecasting the Moments of Daily Returns

The discussion so far has only considered the estimation of the moments of daily returns from intraday data in a static context, but given the final objective of forecasting, extending this to a dynamic environment is required. This can be achieved by applying the above estimation method to a rolling window of intraday data; by rolling this estimation window forward one day at a time, a time series of estimates for the daily return variance and kurtosis is obtained, with an estimate of each moment for every trading day.

More formally, it is assumed that a series of intraday returns are observed over a period of T days, together with a corresponding series of daily returns. At day m, estimates of the scaling function and prefactor are produced using the first m days of intraday data (from day 1, up to day m) and are then used to estimate the variance and kurtosis of daily returns for day m via equation (2.8). The m day window is then rolled forward by one day and the above procedure is repeated using the intraday data from day 2 up to day $m+1$ to produce estimates of daily return variance and kurtosis for day $m+2$. By repeating this process over the complete sample, a time series of $M = T - m + 1$ estimates for both the daily return variance and kurtosis are obtained.

Producing moment forecasts from these time series of estimated moments requires some form of dynamic structure to be imposed that describes the evolution of the daily return process over time. The simplest way of achieving this is to impose the dynamic structure directly onto the time series of estimated moments themselves; this is an approach previously employed in the realised volatility literature, where various time series models have been fitted to daily realised volatility measures obtained from intraday data in order to produce forecasts for daily volatility.

Numerous time series models have been employed for this purpose in the realised volatility literature, with some allowing for relatively complex dynamics, such as the Mixed Data Sampling (MIDAS) and Heterogeneous Autoregressive models used by Clements et al. (2008). Whilst these could also be employed here, two simpler autoregressive specifications will be considered initially that were previously used by Andersen et al. (2003) to model and forecast the realised volatility of three exchange rates.

The first possibility is to assume that the dynamics of the daily return variance and kurtosis can each be described separately by a standard univariate autoregressive (AR) model. Whilst the true values of the daily return variance and kurtosis for day t are not observable, they can be replaced by their corresponding multifractal estimates obtained from the intraday data using the method of the previous subsection. These multifractal moment estimates of the daily return variance and kurtosis are denoted by $\hat{\sigma}^2$ and $\hat{k}$ respectively, with the hats used to emphasise the fact that we are modelling the observable estimated daily return moments and not the true latent moments of the daily return process. The general form of the first model is then given by:

$$
\begin{aligned}
\log\hat{\sigma}^2_{t+1} &= \alpha + \sum_{i=0}^{p-1} \phi_i \log\hat{\sigma}^2_{t-i} + \epsilon_t \\
\log\hat{k}_{t+1} &= \beta + \sum_{j=0}^{q-1} \psi_j \log\hat{k}_{t-j} + \nu_t
\end{aligned}
\tag{3.2}
$$

where ϵ_t and ν_t are iid error terms. The specification in (3.2) will be referred to as the autoregressive multifractal variance and kurtosis, or AR-MFVK(p,q) model. Note that following Andersen et al. (2003), the logarithmic multifractal moment estimates are modelled in practice, rather than their levels, for two reasons. Firstly, the logarithmic multifractal moment estimates are much closer to being normally distributed than their levels and so should be easier to model using standard Gaussian time series methods; this is supported in practice when testing for dynamic misspecification using the standard Ljung-Box test for residual serial correlation, where models based on the levels of the estimated moments display more evidence of dynamic misspecification than equivalent models based on the logarithmic moments. Secondly, this guarantees that the resulting moment forecasts obtained are non-negative, as is required for the second and fourth standardised moments.

The parameters in (3.2) can then be estimated from the time series of estimated moments and used to produce one-step-ahead forecasts for the daily return variance and kurtosis. Denoting these parameter estimates by $\hat{\alpha}$, $\hat{\beta}$, $\{\hat{\phi}_i : 1 \leq i \leq p\}$ and $\{\hat{\psi}_j : 1 \leq j \leq q\}$, one-step-ahead out-of-sample forecasts for the moments at time $t+1$ are then given by:

$$
\begin{aligned}
\log\tilde{\sigma}^2_{t+1} &= \hat{\alpha} + \sum_{i=0}^{p-1} \hat{\phi}_i \log\hat{\sigma}^2_{t-i} \\
\log\tilde{k}_{t+1} &= \hat{\beta} + \sum_{j=0}^{q-1} \hat{\psi}_j \log\hat{k}_{t-j}
\end{aligned}
\tag{3.3}
$$

where tilde is used to distinguish the one-step-ahead out-of-sample forecasts for the moments, from the in-sample multifractal moment estimates. A slightly more general dynamic structure can be considered that allows for interdependence between the two moments by jointly modelling the daily return variance and kurtosis using a vector autoregression (VAR); this is again similar in spirit to the trivariate VAR specification previously used by Andersen et al. (2003) to jointly model and forecast the realised volatilities of three exchange rates. The general p-th order form of the model, expressed in terms of the estimated moments, is given by:

$$\begin{pmatrix} \log\hat{\sigma}^2_{t+1} \\ \log\hat{k}_{t+1} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \sum_{i=0}^{p-1} \begin{pmatrix} \phi_{11,i}\phi_{12,i} \\ \phi_{21,i}\phi_{22,i} \end{pmatrix} \begin{pmatrix} \log\hat{\sigma}^2_{t-i} \\ \log\hat{k}_{t-i} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \tag{3.4}$$

The dynamic specification of (3.4) was also tested in the empirical exercise of Section 4, but was found to produce nearly identical density forecasting performance to the simpler pair univariate AR models in equation (3.2) and as a result, this alternative VAR specification has been omitted when reporting the empirical results.

Consistent with the previous notation, it is assumed that M multifractal estimates can be produced for the variance and kurtosis from the complete sample of T days of data; for notational simplicity it will be assumed from this point onwards that the first of these moment estimates are produced for period 1[4].

A standard rolling estimation scheme is used for producing out-of-sample forecasts for the daily return variance and kurtosis, with an n-day in-sample window used to estimate the values of the parameters in (3.2) or (3.4). In period n the values of the parameters in (3.2) or (3.4) are estimated using the first n estimated moments of order 2 and 4 (from period 1 to period n) and these parameter estimates are then substituted into (3.2) or (3.4) to produce one-step-ahead forecasts of variance and kurtosis for use in period $n + 1$. The n-period in-sample window is then rolled forward by one day and the n moment estimates from period 2 up to period $n + 1$ are used to produce moment forecasts for use in period $n+2$. This process can be repeated to produce one-step-ahead moment forecasts for each day in the chosen out-of-sample period.

Finally, it should be noted that the use of the multifractal moment estimates to approximate the unobserved true daily return moments in the predictive regressions of (3.2) and (3.4) can potentially present difficulties for inference due to the 'generated regressor problem' of Pagan (1984). However, the use of a rolling estimation window means that this should not be a problem in the current context, since the length of the estimation sample does not grow to infinity.

3.3 Producing Density Forecasts from Point Forecasts of Moments

Whilst the forecasts of the variance and kurtosis of daily returns could be used directly in many financial applications, the aim of the current paper is to produce forecasts of the complete probability density. In the multifractal case, the obvious way to achieve this is to impose a specific parametric distribution for daily returns that is uniquely characterised by the forecasted moments.

As discussed previously in Section 2, one of the limitations of the multifractal approach is the inability to estimate the odd-numbered moments of daily returns, such as skewness, from intraday data. Initially attention will simply be restricted to symmetric distributions, as is common in financial econometrics, but non-zero values could be imposed for any odd-numbered moments based on estimates from daily data or other information. A key advantage the multifractal method possesses over existing methods

[4]Following the previous discussion, this will not automatically be the case unless $m = 1$

based on realised volatility measures is that it allows the daily return kurtosis to be estimated directly from intraday data, in addition to the variance. Therefore a symmetric parametric distribution is required for daily returns, which will also allow kurtosis to vary and be determined independently of the variance; this eliminates the normal distribution and the simple 1-parameter version of the t-distribution as suitable options. The generalised error distribution is also unsuitable, since recovering the distributional parameters from the moments requires inversion of the Gamma function[5].

An obvious choice commonly employed for modelling financial returns is the more general location-scale (or three-parameter) t-distribution, which has the density function:

$$f(x:\mu,\lambda,\nu)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\lambda\sqrt{\nu\pi}\Gamma(\nu/2)}\left[\frac{\nu+\left(\frac{x-\mu}{\lambda}\right)^2}{\nu}\right]^{-(\nu+1)/2}$$

where μ, λ and ν are the location, scale and degrees of freedom parameters respectively. The distribution has mean equal to μ and skewness equal to zero; the variance and the fourth central moment, denoted by σ^2 and m_4, are given by:

$$\sigma^2=\frac{\lambda^2\nu}{\nu-2}\quad\text{for}\quad\nu>2\quad\text{and}\quad m_4=\frac{3\lambda^4\nu^2}{(\nu-4)(\nu-2)}\quad\text{for}\quad\nu>4$$

Kurtosis, denoted by k, is then equal to:

$$k=\frac{m_4}{\sigma^4}=\frac{3(\nu-2)}{(\nu-4)}\quad\text{for}\quad\nu>4$$

The distribution automatically satisfies the assumption of symmetry, but the location parameter μ must also be set equal to zero to satisfy the assumption that the mean of returns is zero required by the multifractal approach. The degrees of freedom and scale parameters, ν and λ, can then be obtained from the estimates of daily return variance and kurtosis produced using the intraday data via:

$$\nu=\frac{4k-6}{k-3}\quad\text{for}\quad 3\leq k\leq 9 \tag{3.5}$$

and

$$\lambda=\sqrt{\frac{\sigma^2(\nu-2)}{\nu}} \tag{3.6}$$

From equations (3.5) and (3.6) the location scale t-distribution allows the values of the distributional parameters to be recovered easily from the multifractal estimates of the variance and kurtosis, although kurtosis is required to satisfy $3\leq k\leq 9$ in order for the distributional parameters to be well-defined. Whilst this restriction is generally satisfied

[5] This is only possible as an approximation and only then for values of the distributional parameters that produce an unsuitable density function for modelling asset returns.

for daily return kurtosis estimates from long samples of intraday data, the lower bound of $k = 3$ can be violated by the multifractal estimates of kurtosis produced from short windows of intraday data. The simplest way to overcome this problem is to impose a lower bound on kurtosis of $k = 3$, so that if the estimated value of kurtosis is strictly less than 3, then it is truncated and set equal to 3, resulting in a normal distribution (or equivalently a location-scale t-distribution with infinite degrees of freedom).

Given that there is no theoretical justification for this restriction and that imposing it discards information whenever the estimated value of kurtosis is less than 3, possible alternative distributional forms were also explored. The first notable alternative considered was the Pearson distribution family, which also results in a location-scale t-distribution when kurtosis is greater than 3, but a symmetric 4-parameter beta distribution otherwise; such a distribution has finite support and the density forecasting performance when using he Pearson family of distributions was found to be poorer than that obtained from the location-scale t-distribution.

The second alternative investigated was the use of Gram-Charlier expansions, which can be employed to obtain a semi-parametric approximation to the density function, in which the moments, such as skewness and kurtosis, appear as the parameters in a polynomial expansion. However, the polynomial expansion is not guaranteed to produce a valid density function unless restrictions are imposed on the moments; for a symmetric distribution this requires kurtosis to satisfy $3 \leq k \leq 7$ (see for example Jondeau & Rockinger, 2001), which is even more restrictive than for the truncated location-scale t-distribution approach proposed above. Again, the resulting density forecasts from the alternative Gram-Charlier specification were found to perform worse for the current dataset than those from the location-scale t-distribution, despite the slightly arbitrary restriction on kurtosis that is required for the parameters to be well defined. The original approach based on the t-distribution is therefore the distributional specification used for the remainder of the paper in order to construct daily return density forecasts from the values of variance and kurtosis.

3.4 Sources of Multifractality and an Extension to the Method

A fundamental property of multifractal processes is that the scaling behaviour of fluctuations of different sizes is characterised by a range of scaling exponents and cannot be described by a single scaling exponent as in the unifractal case. As discussed in the literature (see for example Kantelhardt et al., 2002 or Kantelhardt, 2009), these differences in scaling behaviour for different sized fluctuations can arise from two possible sources; the first is multifractality due to a broad probability density function for the process (such as a power-law probability density function), whilst the second is caused by small and large fluctuations of the process having different long-range correlations and is therefore a consequence of the temporal structure of the data. Multifractality of the second type will be eliminated if the time series is shuffled randomly, since any temporal dependence present in the original ordered time series will be destroyed. If the multifractality displayed by the original series is purely of the second type then the resulting shuffled series will display non-multifractal distributional scaling behaviour, if it is entirely of the first type then the scaling behaviour of the shuffled series will be unchanged and finally if both types are present in the original data then the shuffled series will still display multifractal scaling, but weaker than that of the original series.

This issue has previously been studied both for simulated multifractal processes (see again Kantelhardt et al., 2002) and also return series for various financial assets (see for example Onali & Goddard, 2009). In the case of financial data, it is generally found that the randomly shuffled returns display different multifractal scaling behaviour than the original ordered return series (as indicated by differences in the shape of the estimated scaling function) implying that at least some of the scaling present in financial data is due to the second source of multifractality. This in turn implies that the multifractal estimates of the daily return variance and kurtosis obtained from ordered and reshuffled financial data will generally differ; a potentially interesting extension is therefore to explore which of these moment estimates results in forecasts with the greatest predictive ability, by applying the proposed multifractal method to both ordered and randomly shuffled data. If the part of multifractal scaling due to the temporal structure of the data is not relevant for the current application, then randomly shuffling the intraday data before estimating the daily return moments may result in more accurate density forecasts. Conversely, if this component of scaling behaviour is informative for the current application, then eliminating it through reshuffling the data should reduce the forecasting performance of the multifractal method.

One potential problem with this extension to the method is that the partition function estimator of Section 3.1 is affected by the ordering of the observations and so each random shuffling of the data will produce a different estimate of the scaling function and therefore different estimates of the daily return moments. The predictive ability of density forecasts produced from these multifractal moment estimates will therefore become stochastic. This has not posed a problem in previous studies, since scaling functions were only estimated in a static context from very long time series of data, making the resulting estimates of the scaling function relatively insensitive to the specific ordering of the observations obtained from shuffling the data. Unfortunately, in the current context where the scaling function is estimated from a short rolling window of data this is no longer the case and estimated scaling functions obtained from successive reshufflings of each window of intraday data exhibit substantial variability.

The solution proposed here is to shuffle the sample of intraday data multiple times, each time producing a new estimate of the daily return variance and kurtosis, before taking an average of these moment estimates that could then be used as before to produce density forecasts for daily returns. However, for this approach to work in practice the average moment estimates obtained over the repetitions must have a tendency to converge to a particular value as the number of repetitions increases. Whether this is the case in practice will be investigated for the current dataset during the empirical exercise of Section 4.

4 Empirical Application

The current section compares the density forecasting performance of the new multifractal approach with that of existing methods when applied to both foreign exchange and equity data. Section 4.1 describes the dataset employed for the empirical analysis and Section 4.2 discusses the alternative density forecasting methods used as benchmarks to compare the multifractal method against. Section 4.3 outlines the methods used to formally compare the relative performance of these competing density forecasting models and finally Section 4.4 presents the empirical results.

4.1 Data

The data used throughout were obtained from Olsen Associates and consist of intraday 5-minute observations from 3rd January 2007 until 31st December 2010 on the Euro (EUR) and Japanese Yen (JPY) exchange rates against the US Dollar (USD) and the levels of the S&P500 and NASDAQ-100 equity indexes. The choice of 5-minute data was guided by the desire to exploit as much of the potentially valuable intraday information as possible, whilst at the same time avoiding the distortions caused by market microstructure effects typically encountered at very short sampling intervals[6].

The raw price or level data contains all 5-minute intervals in the sample period and so weekends and other non-trading days need to be removed. For the S&P500 and NASDAQ-100 data this is a relatively straightforward task, since the equity markets have well-defined trading hours with no trading taking place over weekends or on holidays (such as Christmas day and Thanksgiving). The list of non-weekend closures for the S&P500 and NASDAQ-100 was constructed from the historical list of holidays available on the NYSE website. Throughout the sample there were also 9 days for which the market was open, but trading took place for reduced hours (such as the day after Thanksgiving); the analysis was performed with these partial trading days both removed and included, but the choice did not have any significant effect on the results.

For the EUR/USD and JPY/USD exchange rate series trading takes place for 24 hours a day and 7 days a week, however over weekends and certain holidays trading slows substantially. Following Andersen et al. (2001), the end of each 24-hour trading day was taken to be 21:00 GMT and the 48-hour weekend periods between 21:05 GMT on each Friday and 21:00 on each Sunday were removed from the raw 5-minute price series. For most of the NYSE holidays during the sample period both the EUR/USD and JPY/USD markets were open for normal trading hours; only for Christmas Day and New Year's Day was trading noticeably slower than normal and so only these holidays were omitted from the exchange rate series. The analysis was also performed with a larger and more comprehensive list of holidays removed from the EUR/USD and JPY/USD series, but as with the partial trading days for the equity index data, this did not significantly influence the results.

[6] This problem is also encountered in the literature on realised volatility, where the 5-minute sampling interval has generally been found to be a good compromise between these two factors (see for example Andersen et al., 2001).

Removing the weekends and holidays from the original data series leaves a sample size of 1008 trading days for the equity index series and 1037 for the exchange rate series. Continuous 5-minute returns were then constructed from the first difference of the log-price series for each asset, with the first 5-minute return for each day calculated between the closing price in the previous trading day and the opening price in the current day (thus including any overnight or weekend effects). A daily return series was also constructed for both assets from the last 5-minute price observed in each trading day. This daily return series is required for estimating the GARCH models used for density forecast comparison and the statistical method used for forecast comparison.

Before proceeding it is worth checking that the assumptions of Section 2.2 required for the proposed multifractal method to be applicable are satisfied for the current dataset. The first requirement is that the distributional scaling properties of the data are consistent with that of a multifractal process[7], in order for the moment scaling condition of equation (2.6) to hold. It is important to note however that the rolling estimation scheme employed allows the values of the scaling function, $\tau(q)$, and the prefactor, $c(q)$, in equation (2.6) to vary over time; this implies that it is sufficient for the process to be locally multifractal within each of the rolling estimation windows, even if it may have more complex non-multifractal scaling properties when viewed globally. The same distinction between local and global scaling properties is also relevant for the earlier unifractal approach of Hallam & Olmo (2013), where local unifractality is sufficient within each estimation window.

Previous work on unifractal and multifractal processes in finance has relied on an informal graphical method based on the scaling function, $\tau(q)$, from equation (2.6) in order to distinguish between unifractal and multifractal scaling; as previously discussed in Section 2.1, $\tau(q)$ is strictly concave for a multifractal process and linear for a unifractal process with functional form $\tau(q) = Hq - 1$, where H is the self-affinity index of Section 2.1. A visual inspection of the sample estimate of $\tau(q)$ for a given sample of data can then be used to assess the distributional scaling properties of the series.

In financial applications, this informal graphical testing method is almost exclusively employed to study the global distributional scaling properties of return series, with a single sample estimate of the scaling function produced for the complete sample of data, thus requiring the visual inspection of a single estimated function. However, as stated above it is the local, rather than global, scaling properties that are relevant for the current application; in this context the graphical testing method becomes problematic, since an assessment of the local scaling properties of the series over the complete sample period may require the visual inspection of a very large number of estimates for the scaling function. In the absence of an appropriate method for examining the local scaling properties of a series, the estimated global scaling properties of the return series are briefly examined instead, however the distinction between local and global scaling properties should be kept in mind, with the global scaling properties not necessarily

[7] Again, multifractal scaling nests unifractal scaling as a special limiting case and so the method will also be applicable in this case. In this situation however, it may be more logical to employ the unifractal approach of Hallam & Olmo (2013), given the additional flexibility provided by that methodology.

being informative about the local scaling properties.

Figure 1 presents estimated scaling functions obtained from the complete sample of 5-minute returns for each asset; the solid lines are unrestricted estimates of $\tau(q)$ obtained using the partition function estimator of Section 3.1. The dashed lines in each sub-plot are the linear scaling functions obtained under the assumption of unifractality, with the self-affinity index (and therefore the slope of $\tau(q)$) estimated using the standard detrended moving average estimator of Alessio et al. (2002), previously employed by Hallam & Olmo (2013). These functions are plotted over the domain $0 \leq q \leq 5$, which is a common choice in empirical studies of scaling behaviour in asset returns. For Figure 1, the partition function and DMA estimators were implemented exactly as in Hallam & Olmo (2013), in order to maintain consistency between Figure 1 and the equivalent figure in the earlier work. This does however mean that the maximum sampling intervals used for the partition function estimates of $\tau(q)$ in Figure 1 are not identical to those used in the actual forecasting exercise of Section 4.4.

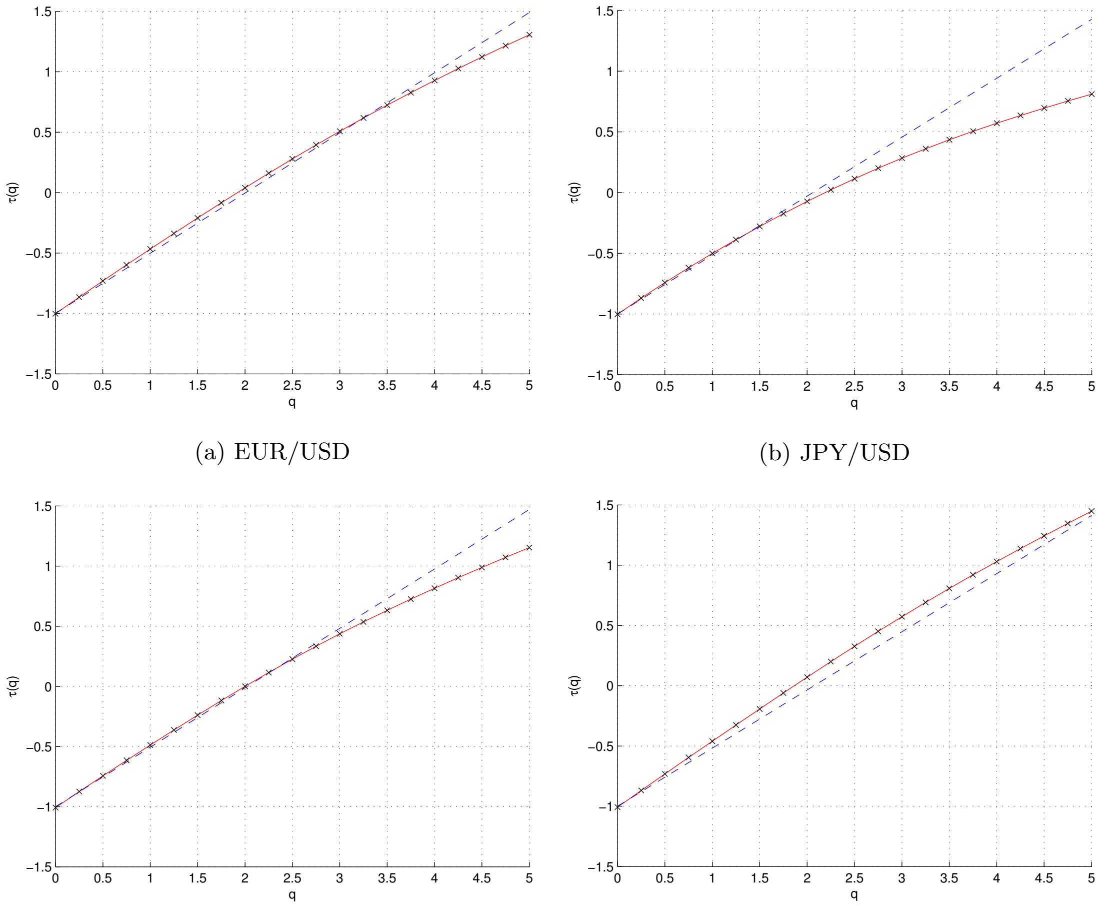
It can be seen from Figure 1 that all estimated scaling functions are strictly concave, indicating distributional scaling consistent with multifractal rather than unifractal processes, at least in a global sense. Whilst the degree of nonlinearity, and thus the strength of multifractality, varies from one return series to another, all return series appear to be consistent with the first requirement that the processes possesses either multifractal or unifractal distributional scaling properties.

The second assumption that must be satisfied is that mean value of the return process is equal to zero, to guarantee that the central and non-central moments are equal and equation (2.8) holds. The validity of the assumption was checked for the current dataset using the standard test statistic for testing that the population mean is equal to some hypothesised value and for none of the 5-minute return series could the null that the population mean is equal to zero be rejected at any conventional significance level.

4.2 Benchmark Density Forecasting Models

For the empirical exercise it is necessary to have one or more existing density forecasting methods to compare the performance of the new multifractal method against. Whilst there are many possibilities, two simple but established examples from the literature have been used initially as benchmarks, in addition to the closely related unifractal density forecasting model of Hallam & Olmo (2013).

The first is a standard autoregressive conditional heteroskedasticity (ARCH) model based on daily data; the mathematical details of this approach will not be presented here, since they are discussed in detail elsewhere. Whilst standard ARCH models are somewhat simplistic in nature and utilise only daily data, they remain popular in empirical work due largely to this simplicity. Given this motivation, attention was restricted to standard commonly used specifications for the mean and volatility equations and dis-



(a) EUR/USD

(b) JPY/USD

(c) NASDAQ-100

(d) S&P500

Figure 1: Estimated unifractal and multifractal scaling functions. Solid lines correspond to estimated scaling functions for the multifractal case (obtained using the partition function estimator) and dashed lines correspond to the estimates under the assumption of unifractality (obtained using the DMA estimator).

tribution of the error terms[8]. Given that the current focus is on producing forecasts for the complete density of daily returns, rather than just the conditional variance, the specification of the ARCH models was chosen in order to maximise the performance of the resulting density forecasts for the current dataset in terms of the CRPS-based test outlined in Section 4.3. For both the S&P500 and EUR/USD data, this resulted in GARCH(1,1) volatility equations, AR(1) mean equations and t-distributed errors.

[8]These included ARMA(p,q) models for the mean equation and ARCH and GARCH specifications for the volatility equation. For the equity index data, EGARCH and GJR specifications were also included, to allow for the possibility of leverage effects. For the error distribution, the standard choices of normal, generalised error and Student's t distributions were all tested.

In order to include a comparison with established density forecasting methods employing intraday data, the second benchmark density forecasting model selected was the autoregressive realised volatility (AR-RV) model of Andersen et al. (2003), which fits a univariate autoregressive model to the time series of (logarithmic) daily realised volatility measures. Density forecasts for daily returns can then be produced by combining these point forecasts of volatility with the empirical observation that daily returns are approximately normally distributed if standardised by their corresponding (time-varying) realised volatilities for each day and their constant sample mean. Following Andersen et al. (2003), a 5th order AR-RV(5) model was used initially for the empirical exercise of Section 4.4 and this choice was also found to produce the best average density forecasting performance for the dataset employed here. Further details of the AR-RV method can be found in Andersen et al. (2003).

Finally, the unifractal density forecasting model of Hallam & Olmo (2013) is included as a third benchmark model. A brief summary of this method can be found in Section 2.2, with a more detailed treatment found in the original paper. As should be clear from the earlier discussion of Section 2.2, the unifractal method of Hallam & Olmo (2013) imposes a more restrictive distributional scaling structure than the current multifractal method, but allows more flexible specifications to be employed for the daily return density; as such, the inclusion of this unifractal benchmark model allows the relative importance of these factors for predictive ability to be investigated.

The unifractal benchmark method is implemented as described in Hallam & Olmo (2013), with the self-affinity index H estimated from the intraday data using the detrended moving average method of Alessio et al. (2002) and daily return densities estimated from the rescaled intraday returns using a standard non-parametric kernel density function estimator. The simpler variant of the unifractal method is employed for which the autoregressive parameter values used to produce density forecasts are fixed rather than time varying, since this was typically found to maximise density forecasting performance.

The set of benchmark forecasting methods, and indeed the proposed multifractal approach, could be extended through the use of more general dynamic structures or more flexible specifications for the daily return density[9]. Such extensions are not considered here, since the chosen benchmark methods, whilst simple, allow two key comparisons of forecasting performance to be made; the first is between the forecasting performance attainable when incorporating intraday data via the new multifractal approach and that from using existing approaches based on realised volatility. The second is an assessment of the benefits from allowing for more general multifractal distributional scaling rather than the more restrictive unifractal scaling imposed by the unifractal approach of Hallam & Olmo (2013). As stated above, the final GARCH benchmark has largely been included due to the popularity it maintains in empirical work; given that this popularity follows

[9] As noted by an anonymous referee, the latter could include specifications for the daily return density allowing for time varying skewness and kurtosis through the use of Gram-Charlier expansions, generalised t-distributions or other methods (see for example Jondeau & Rockinger, 2001, Guermat & Harris, 2002, Jondeau & Rockinger, 2003 or León et al., 2005).

largely from the simplicity of the approach, extensions beyond the basic variants were not considered.

For all benchmark density forecasting methods, the same rolling window estimation scheme as described in Section 3 was employed for producing density forecasts: the parameters of the models are estimated using an n-day rolling window of data (daily data in the case of the GARCH model and 5-minute intraday data for the case of the AR-RV model) and these parameter estimates are then used to produce one-step-ahead point forecasts for the relevant moments of daily returns. When combined with the relevant parametric form assumed for the return distribution, this allows one-step-ahead out-of-sample density forecasts to be produced for daily returns.

4.3 Methods for Density Forecast Comparison

The first method used for comparing the out-of-sample density forecasting performance of the methods is the statistical test for equal predictive ability proposed by Gneiting & Ranjan (2011) based on the continuous ranked probability score, which is similar in spirit to the earlier weighted likelihood ratio (WLR) test of Amisano & Giacomini (2007). Both of these testing methodologies provide tests of relative forecasting performance, sometimes referred to as forecast 'comparison', enabling a direct comparison of the predictive ability of forecasts from the proposed multifractal method and those from alternative approaches[10].

The CRPS-based test employed here assumes that two competing forecasting models are used to produce one-step-ahead out-of-sample density forecasts for the variable of interest, y. It is assumed that N density forecasts are produced by each forecasting method and the forecasts produced by the two models at time t (for use at time $t+1$) are denoted by $\tilde{f}_t(y)$ and $\tilde{g}_t(y)$, respectively.

The loss function employed by the test is the continuous ranked probability score (CRPS), generalised to allow more importance to be placed on forecast accuracy in particular regions of the density via the use of a weighting function. The value of the weighted CRPS for the forecast produced by the first model for use in period $t+1$, denoted by $\mathrm{S}(\tilde{f}_t, y_{t+1})$, is given by:

$$\mathrm{S}(\tilde{f}_t, y_{t+1}) = 2\int_0^1 \left(\mathrm{I}\{y_{t+1} \leq \tilde{F}_t^{-1}(\alpha)\} - \alpha\right)\left(\tilde{F}_t^{-1}(\alpha) - y_{t+1}\right) w(\alpha)\,\mathrm{d}\alpha \tag{4.1}$$

where $\tilde{F}_t(y)$ is the CDF forecast produced at time t obtained from the PDF forecast for the same period $\tilde{f}_t(y)$, $\mathrm{I}\{.\}$ is an indicator function and $w(\alpha)$ is a weighting function; the authors suggest several possible forms for $w(\alpha)$, which allow more weight to be placed on forecast accuracy in different regions of the density, such as the centre or tails.

[10] This differs from methods providing tests of absolute forecasting performance of a single method in isolation, sometimes referred to as forecast 'evaluation', such as that of Diebold et al. (1998) and related work. See Amisano & Giacomini (2007) for a short discussion of this issue.

Whenever a closed form expression for (4.1) is unavailable, it can be approximated easily to any degree of accuracy using the method outlined by Gneiting & Ranjan (2011). The average value of the weighted CRPS in (4.1) can be calculated for each of the two density forecasting models over the N out-of-sample periods (for period $m+1$ until period T) as:

$$\overline{\mathrm{S}}^f = \frac{1}{N}\sum_{t=m}^{T-1} \mathrm{S}(\tilde{f}_t, y_{t+1}) \quad \text{and} \quad \overline{\mathrm{S}}^g = \frac{1}{N}\sum_{t=m}^{T-1} \mathrm{S}(\tilde{g}_t, y_{t+1}) \tag{4.2}$$

A formal test can then be based on the following test statistic:

$$t = \frac{\overline{\mathrm{S}}^f - \overline{\mathrm{S}}^g}{\hat{\sigma}_n/\sqrt{N}} \tag{4.3}$$

where $\hat{\sigma}_n^2$ is a standard heteroskedasticity and autocorrelation consistent estimator for the asymptotic variance of $\sqrt{N}(\overline{\mathrm{S}}^f - \overline{\mathrm{S}}^g)$.

Under the null hypothesis that the two density forecasting models have equal predictive ability, the test statistic in (4.3) is asymptotically normally distributed, with the null rejected at the $\alpha\%$ significance level if $|t| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution. Given that lower values of the CRPS correspond to better forecasts, in the case of rejection, the forecasting model f should be chosen when the sample value of the test statistic is negative and model g when it is positive.

The CRPS-based test above is a purely statistical measure of predictive ability and as such can give no indication of the economic gains or losses that would be realised by applying the various density forecasting methods in practice. Therefore, the CRPS-based test is supplemented by a second density forecast comparison method based on the problem of optimal portfolio allocation between a risky and a risk-free asset.

It is assumed that at time t an investor has total wealth of 1 to allocate between a single risky asset and a risk-free asset. The proportion invested in the risky asset is given by ω_t, with the remainder invested in the risk-free asset. Denoting the risky and risk-free returns from time t to $t+1$ by r_{t+1}^r and r_t^f respectively, the value of the portfolio at time $t+1$, denoted W_{t+1}, is then given by:

$$W_{t+1} = 1 + \omega_t r_{t+1}^r + (1-\omega_t) r_t^f$$

The utility of the investor at time $t+1$ is assumed to depend on final wealth W_{t+1} according to a power utility function, with a coefficient of relative risk aversion γ:

$$\begin{aligned} U(W_{t+1}) &= \frac{W_{t+1}^{1-\gamma}}{1-\gamma} \\ &= \frac{1}{1-\gamma}\left[1 + \omega_t r_{t+1}^r + (1-\omega_t) r_t^f\right]^{1-\gamma} \end{aligned}$$

When choosing w_t at time t, r_{t+1}^r the rate of return on the risky asset from time t to $t+1$, is unknown and so the investor chooses the portfolio weight in order to maximise the expected utility obtained at $t+1$. Formally, the optimal weight ω_t^* at time t is obtained as the solution to:

$$\begin{aligned}\omega_t^* &= \arg\max_{\omega_t} E_t\left[U(W_{t+1})\right] \\ &= \arg\max_{\omega_t} E_t\left[\frac{1}{1-\gamma}\left[1+\omega_t r_{t+1}^r + (1-\omega_t) r_t^f\right]^{1-\gamma}\right]\end{aligned} \tag{4.4}$$

Rewriting (4.4) using the standard expression for the expectation of a random variable gives:

$$\omega_t^* = \arg\max_{\omega_t} \int \frac{1}{1-\gamma}\left[1+\omega_t r_{t+1}^r + (1-\omega_t) r_t^f\right]^{1-\gamma} \tilde{f}_t(r_{t+1}^r)\, dr_{t+1}^r \tag{4.5}$$

where, consistent with previous notation, $\tilde{f}_t(r_{t+1}^r)$ is the density forecast produced at time t for the risky return r_{t+1}^r. From (4.5) it is clear that different density forecasts for r_{t+1}^r will lead to different portfolio allocations a time t and therefore different realised utilities at time $t+1$.

It is assumed that the investor holds the portfolio defined by the weight ω_t^* for a single period, before readjusting the portfolio weights based on new information in the following period. In the current empirical exercise this portfolio readjustment is performed daily and the portfolio allocation is made between one of the risky assets discussed in Section 4.1 and a risk-free asset, which is represented by the 3-month Treasury bill rate (with the rate converted to a daily return). Solving the portfolio allocation problem in equation (4.5) in each of the N days in the out-of-sample period results in a time series of portfolios, which in turn produces a time series of N realised utilities once the true risky return for the following period is observed.

The relative performance of the portfolios obtained from the density forecasting methods is then compared using the certainty equivalent return (CER) of the portfolio, which is defined as follows:

$$CER = \left[(1-\gamma)\frac{1}{N}\sum_{t=1}^{N} RU_t\right]^{\frac{1}{(1-\gamma)}} - 1$$

where RU_t is the realised utility obtained from the portfolio in period t. The CER gives the risk free rate of return that would provide the same average level of realised utility as the portfolio over the out-of-sample period, implying that higher CER values are preferable to lower values.

4.4 Empirical Results

For the empirical results presented in this section a 250-day rolling in-sample window (or value of n in the notation of Section 3) is used for parameter estimation and the density forecasts are compared over a 750 working day evaluation period, from the 250$^{\text{th}}$ until the 1000$^{\text{th}}$ working day in the sample for each series[11].

Following Andersen et al. (2003), the initial order for the autoregressive component of the AR-MFVK(p,q) specification was set equal to 5 (i.e. one working week). The adequacy of these initial dynamic specifications was then checked by applying the standard Ljung-Box test for residual autocorrelation to the residuals obtained from fitting each time series model over the complete sample period. In almost all cases, for these initial 5$^{\text{th}}$ order models the null of no residual autocorrelation could not be rejected at any conventional significance level, suggesting that the 5$^{\text{th}}$ order specifications are adequate for modelling the dynamic structure of the estimated moment series. The only exceptions were for the two equity index series, where the null of no residual autocorrelation was rejected for the variance component (but not the kurtosis component) at the 10% level for the NASDAQ-100 and the 5% level for the S&P500.

As previously noted in Section 3, implementation of the partition function estimator requires the selection of a maximum sampling interval (Δt in the notation of equation (3.1)). Given that the final aim of the current empirical exercise is to estimate and forecast the moments and probability density functions of daily returns from the intraday data, the maximum sampling interval used for the partition function estimator was initially set to just over 1 trading day for all series[12]. Alternative values were also explored, but these initial choices were generally found to produce the most accurate density forecasts (as measured by the CRPS-based criteria of Section 4.3) and were used throughout the current empirical exercise.

A final issue of model specification that must be investigated for the multifractal method is the optimal choice of size for the window of intraday data used to produce each rolling estimate of the scaling function and prefactor (m in the notation of Section 3.2). Table 1 contains sample values for the simple unweighted CRPS-based test statistic comparing the predictive ability of the multifractal method using various window sizes against the different benchmark methods.

It can be seen from Table 1 that on average the optimal window size is around 15 working days across the four assets; 10 working days appears to be approximately optimal for the two equity index series and 20 days for the exchange rate series, with these values used for the remainder of the empirical exercise. Longer windows increase the number of intraday observations available, but do not result in an improvement in forecasting performance, presumably because older intraday data are no longer informative about the current properties of the return process. Equally, shorter windows reduce density forecasting performance, either because some degree of smoothing produces superior

[11]Because of the difference in trading days, the start and end dates of this period differ slightly for the two series: for the EUR/USD it spans 19th Dec 2007 - 11th Nov 2010 and 31st Dec 2007 21st Dec 2010 for the S&P500 data

[12]For the two exchange rate series the number of 5-minute returns in each trading day is 288 and the maximum value of Δt used was 300. The equity index series average around 80 5-minute returns per trading day and the maximum value of Δt used was 100.

Table 1: Sensitivity of Predictive Ability to Changes in Estimation Window Length

Window Size (Working Days)	**2 days**	**5 days**	**10 days**	**15 days**	**20 days**	**25 days**	**50 days**
EUR/USD:							
GARCH(1,1) benchmark	4.28	0.75	-1.72	-2.23	-2.71	-2.75	-2.76
AR-RV(5) benchmark	5.45	3.13	0.22	-0.32	-0.93	-0.87	-0.56
AR-UF(5) benchmark	5.93	4.71	2.26	1.80	1.18	1.26	1.24
JPY/USD:							
GARCH(1,1) benchmark	3.03	1.58	0.39	-0.63	-0.66	-0.65	-0.59
AR-RV(5) benchmark	3.38	2.70	1.70	0.75	0.69	0.61	0.52
AR-UF(5) benchmark	3.39	2.22	0.83	-0.65	-0.70	-0.69	-0.57
NASDAQ-100:							
GARCH(1,1) benchmark	4.87	2.00	0.44	1.04	1.15	1.22	1.73
AR-RV(5) benchmark	5.37	2.62	0.32	1.07	1.14	1.16	1.60
AR-UF(5) benchmark	5.36	3.91	2.24	3.17	3.06	2.91	2.78
S&P500:							
GARCH(1,1) benchmark	4.59	1.79	0.53	0.74	0.82	1.11	1.97
AR-RV(5) benchmark	5.26	2.42	0.50	0.71	0.73	0.95	1.66
AR-UF(5) benchmark	5.51	3.98	2.41	2.68	2.74	2.83	3.11

Values correspond to the sample values of the simple unweighted CRPS-based test statistic of Section 4.3. The CRPS-based test statistic is asymptotically normally distributed under the null of equal predictive ability and the test statistic is constructed such that significant negative values imply the multifractal method is superior to the benchmark model. See Section 4.3 for further details.

estimates of daily return moments[13], or because of limited finite sample performance of the chosen partition function estimator.

Table 2 presents a comparison of density forecasting performance between the multifractal method and the benchmark models using the CRPS-based test outlined in Section 4.3. Considering first the more established GARCH and AR-RV benchmarks, it is clear that the density forecasts from the multifractal model perform well for the EUR/USD data, frequently providing highly statistically significant improvements in predictive ability over the GARCH benchmark method. Compared to the more competitive AR-RV benchmark utilising intraday data, the sample values for the EUR/USD data are generally negative, implying that the multifractal method provides superior predictive ability, but the gains are not large enough to be statistically significant. In addition, the gains in forecasting performance from the multifractal method appear to vary across the regions of the density function; the performance of the unifractal method is particularly strong in the centre and left tail of the EUR/USD return density, suggesting that it should perform well in risk management applications, such as the calculation of Value at Risk

[13]Noise in the observed intraday return process could make estimates calculated from short periods less informative about the true behaviour of the underlying process. This is an issue previously encountered in the realised volatility literature (see for example Andersen et al., 2003), where various methods have been proposed to mitigate the problem.

Table 2: Out-of-sample Density Forecast Comparison Using CRPS-based Test Statistic

Weighting Function	**None**	**Centre**	**Left Tail**	**Right Tail**
EUR/USD:				
GARCH(1,1) benchmark	-2.71***	-2.91***	-3.73***	-0.57
AR-RV(5) benchmark	-0.93	-1.57	-1.15	-0.35
AR-UF(5) benchmark	1.18	0.62	0.46	1.03
JPY/USD:				
GARCH(1,1) benchmark	-0.66	-0.12	-1.31	0.22
AR-RV(5) benchmark	0.69	0.65	-1.16	2.22**
AR-UF(5) benchmark	-0.70	-0.67	-1.75*	0.64
NASDAQ-100:				
GARCH(1,1) benchmark	0.44	0.34	-0.24	1.01
AR-RV(5) benchmark	0.32	0.98	0.15	0.41
AR-UF(5) benchmark	2.24**	2.12**	2.03**	1.81*
S&P500:				
GARCH(1,1) benchmark	0.53	0.29	-0.34	1.14
AR-RV(5) benchmark	0.50	1.14	-0.32	1.23
AR-UF(5) benchmark	2.41**	2.85***	1.69*	2.48**

The CRPS-based test statistic is asymptotically normally distributed under the null of equal predictive ability and the test statistic is constructed such that significant negative values imply the multifractal method is superior to the benchmark model. See Section 4.3 for further details.

or expected shortfall.

From the later sections of Table 2, the multifractal method is clearly less competitive with the GARCH and AR-RV benchmarks for the other return series than for the EUR/USD data, with the method unable to provide a statistically significant improvement in predictive ability over the benchmark methods. Nonetheless, in all but one case, the null of equal predictive ability cannot be rejected, implying that the multifractal method is again competitive with these benchmark established forecasting methods.

The final unifractal AR-UF benchmark model of Hallam & Olmo (2013) typically provides the strongest density forecasting performance of the three benchmark methods. For the equity index data in particular, the unifractal AR-UF method consistently provides gains in predictive ability over the new multifractal method that are highly significant across the whole domain of the density. This finding suggests that when modelling and forecasting the distribution of equity returns, the ability to employ nonparametric specifications for the daily return density is more beneficial than the additional flexibility in distributional scaling properties permitted by moving from a unifractal to a multifractal context. For the exchange rate data the relative performance of the unifractal and multifractal methods is closer, with the null of equal predictive ability not rejected in most cases.

The clearly observed differences in relative density forecasting performance for the new multifractal method that are observed across the different return series cannot simply be due to differences in the type of distributional scaling across the included assets (as appeared to be the case for the unifractal method of Hallam & Olmo, 2013) since the multifractal approach of the current paper should be valid for data exhibiting either unifractal or multifractal scaling behaviour.

A possible alternative explanation is provided by differences in the strength, rather than the type, of distributional scaling across the different return series. This can be measured to some extent by considering the standard errors of the estimated values of $\tau(q)$ in the regressions of $\log S_q(T, \Delta t)$ on $\log(\Delta t)$ that follow from equation (3.1). Whilst these standard errors could be calculated for the scaling function estimates from the complete sample of data, given the rolling estimation method used to obtain the dynamic estimates of the scaling properties it is more relevant to calculate standard errors for estimates of $\tau(q)$ obtained from rolling windows.

Adjusting the lengths of the rolling windows employed to compensate for the difference in trading hours per day between the two types of asset[14], there is indeed evidence of differences in the strength of distributional scaling across the assets. The rolling estimates of $\tau(q)$ for the EUR/USD data have the lowest standard errors on average over the sample period, followed by the JPY/USD data. The standard errors for both equity index series are approximately double those for the EUR/USD data, suggesting that they do indeed exhibit weaker distributional scaling than the exchange rate series.

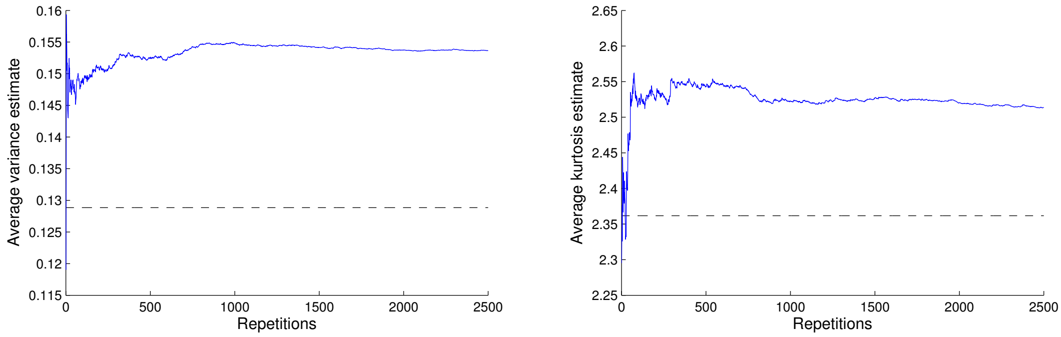
We next investigate the extension to the standard multifractal method proposed in Section 3.4, which modifies the basic method by randomly shuffling the intraday data in each rolling window before estimating the daily return variance and kurtosis. As previously discussed, one potential problem is that each time the data are randomly shuffled different estimates of daily return moments will be obtained, making the density forecasts produced by the model stochastic. The solution proposed for this problem in Section 3.4 is to repeat the shuffling process numerous times, producing multiple multifractal estimates of the daily return moments for each window of intraday data, before taking an average of these estimates over all of the repetitions. However, this solution will only be effective if the moment estimates averaged over the repetitions converge to a particular value as the number of repetitions is increased and so whether this holds in practice should be investigated.

Figure 2 contains plots of the average moment estimates obtained using 1 to 2500 repetitions of the shuffling process described above for the EUR/USD data[15]. To ensure that the exercise is as relevant for the current context as possible, the samples of intraday data used have the same length of 15 working days used previously to estimate the daily return moments for each trading day; 3 different 15-day windows were tested from arbitrary points in the sample, beginning on the 1$^{\text{st}}$, 500$^{\text{th}}$ and 1000$^{\text{th}}$ trading days respectively.

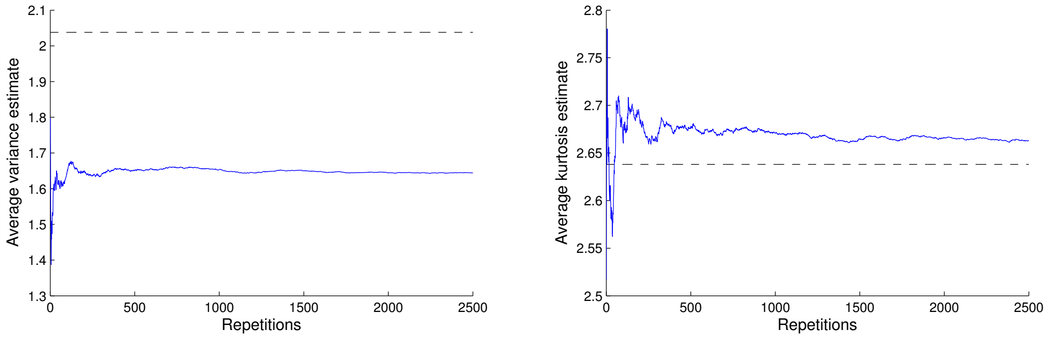
(FIGURE 2 TO BE PLACED HERE)

[14] A rolling window of 20 days was used for the exchange rate series and 65 days for the equity index series. This is necessary to ensure that each rolling estimate is calculated from approximately the same number of intraday observations, minimising the effects of sample size on the standard errors.

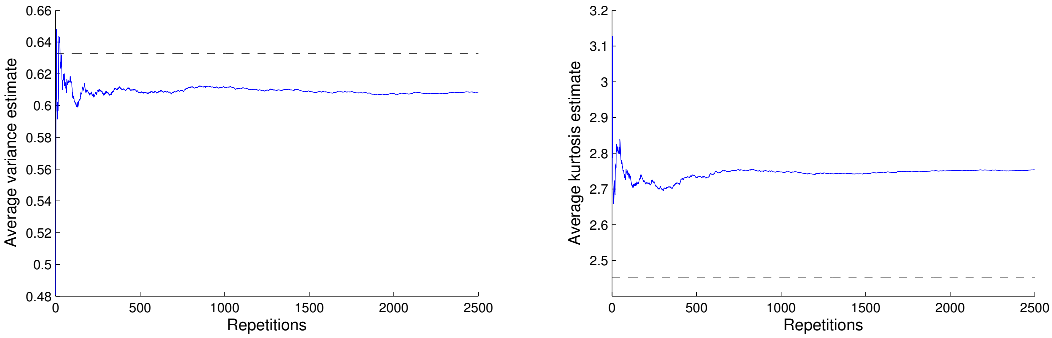
[15] Equivalent figures for the other series have been omitted to conserve space, but similar results are observed for each.



(a) 15-day window beginning on 1st trading day



(b) 15-day window beginning on 500th trading day



(c) 15-day window beginning on 1000th trading day

Figure 2: Multifractal moment estimates from shuffled and ordered intraday data
Dashed lines are multifractal moment estimates obtained from the original ordered 5-minute intraday data in the specified window. Solid lines are multifractal moment estimates obtained from the same intraday data shuffled randomly, averaged over the number of repetitions shown on the horizontal axis. Variance estimates are presented in the left column and kurtosis estimates in the right column.

From Figure 2 it can be seen that although the average moment estimates from the shuffled intraday data do not converge entirely to specific values as the number of repetitions increases (at least up to the maximum of 2500 repetitions considered here), they do typically converge to a narrow range of values. Whilst these figures represent just one possible realisation for each of the 3 arbitrarily chosen windows, the same pattern of convergence was observed for other windows of intraday data chosen from the complete sample of EUR/USD data and also more generally for data from the other 3 asset return series. Furthermore, in none of the cases from Figure 2 do the average moment estimates from the shuffled data converge to the estimated values obtained from the original ordered data, with the differences in many cases being substantial. This implies that at least some of the multifractal scaling present in the original ordered data is due to the second source of multifractality mentioned in Section 3.4 (small and large fluctuations of the process having different long-term correlations) and so this modification to the method should produce noticeable changes in the resulting density forecasts.

Having established that the modified version of the multifractal method should be valid for the current dataset, the density forecasting performance of the method can now be investigated. Whilst Figure 2 shows that some variation may remain in the average moment estimates beyond 2500 repetitions, in practice it was found that repetition numbers as low as 500 resulted in consistent density forecasting performance over the out-of-sample period[16] and so as a compromise the number of repetitions was set to 1000.

Table 3: Out-of-sample Density Forecast Comparison Using CRPS-based Test Statistic - Shuffled Data with 1000 Repetitions Used for Multifractal Approach

Weighting Function	**None**	**Centre**	**Left Tail**	**Right Tail**
EUR/USD:				
GARCH(1,1) benchmark	-3.75 ***	-3.42***	-4.00 ***	-1.83*
AR-RV(5) benchmark	-2.78 ***	-2.47 **	-2.30**	-1.98 *
AR-UF(5) benchmark	-1.49	-1.02	-1.07	-0.93
JPY/USD:				
GARCH(1,1) benchmark	-0.27	0.21	-0.90	0.49
AR-RV(5) benchmark	1.16	1.29	-1.02	2.75***
AR-UF(5) benchmark	-0.18	0.05	-1.15	0.98
NASDAQ-100:				
GARCH(1,1) benchmark	-0.36	-0.07	-0.84	0.26
AR-RV(5) benchmark	-0.85	0.39	-0.68	-0.76
AR-UF(5) benchmark	1.23	1.43	1.27	0.61
S&P500:				
GARCH(1,1) benchmark	-0.45	-0.29	-0.64	-0.16
AR-RV(5) benchmark	-0.82	0.27	-0.69	-0.76
AR-UF(5) benchmark	1.19	1.73*	1.43	0.03

[16]The remaining variation in the moment estimates for a given trading day will become less significant when comparing density forecasting performance in practice, since the CRPS-based test of equal predictive ability compares average forecasting accuracy over the complete length of the out-of-sample period (750-days in the current context).

Table 3 contains equivalent results for the modified multifractal method to those in Table 2; from the sample values it can be seen that this modification to the multifractal method consistently improves the predictive ability of the multifractal method for the EUR/USD, NASDAQ-100 and S&P500 data. The changes in the sample values of the test statistics are often substantial in size, with the outcome of the test often changing as a result. Most notably, the unifractal AR-UF benchmark previously provided statistically significant improvements in predictive ability over the standard multifractal approach, but when compared to the modified multifractal approach the null of equal predictive ability cannot typically be rejected at any conventional significance level. However, the same improvement in density forecasting performance is not found for the JPY/USD data for which the modification to the multifractal approach typically reduces performance, although only in the right tail of the return density are these changes large enough to alter the outcome of the test for equal predictive ability at any conventional significance level.

The improvements in predictive ability that are typically obtained from shuffling the data in this way suggest that of the two sources of multifractality highlighted in Section 3.4, multifractality due to a broad probability distribution is more relevant for the current application than that due to small and large fluctuations of the process having different long-range correlations. A possible explanation for this follows from the discussion and explanation of the multifractal method in Sections 3.1 to 3.3. It can be seen that at the point when the multifractal moment scaling law is applied to estimate the daily return moments, no dynamic structure has actually been imposed on the return process at any timescale; the dynamic structure required to produce forecasts for the moments (and ultimately the density function) of daily returns is imposed at a later stage onto the time series of estimated daily return moments. Intuitively therefore, it seems possible that the component of scaling that is due to the temporal structure of the ordered data is less relevant in the current application than the scaling of the unconditional distribution of returns at different timescales.

Finally, Table 4 contains the results for the portfolio allocation exercise discussed in Section 4.3. The reported values are the certainty equivalent returns (CER) expressed as an annualised percentage return for the expected utility maximising portfolio using the density forecasts from the GARCH and AR-RV benchmarks, plus the new multifractal AR-MFVK method[17]. The portfolio allocation exercise has been performed with several different values of the coefficient of relative risk aversion γ, in order to assess whether the optimal forecasting method varies with the level of investor risk aversion.

From Table 4 it can be seen that the patterns observed when assessing density forecasting performance in the context of portfolio allocation are consistent with those previously observed in Table 2 in terms of the CRPS-based test statistic, thus reinforcing the previous empirical findings. For the exchange rate series the portfolios obtained from the multifractal approach provide the highest CER values, with those from the AR-RV

[17]The unifractal approach of Hallam & Olmo (2013) has been omitted from this comparison, since the use of a non-parametric specification for the daily return density makes evaluating the integral of the density forecast in equation (4.5) more complex.

Table 4: Certainty Equivalent Returns From Portfolio Allocation Exercise

Value of CRRA	**1**	**2.5**	**5**	**10**
EUR/USD:				
GARCH(1,1)	-9.29	-9.94	-10.31	-9.37
AR-RV(5)	-6.15	-6.29	-5.71	-4.84
AR-MFVK(5)	0.78	0.72	0.63	0.58
JPY/USD:				
GARCH(1,1)	-4.44	-3.65	-2.63	-3.02
AR-RV(5)	-1.38	-0.94	-0.15	0.23
AR-MFVK(5)	-0.27	-0.05	0.09	0.26
NASDAQ-100:				
GARCH(1,1)	27.07	23.95	17.34	8.99
AR-RV(5)	3.72	4.48	0.34	-1.15
AR-MFVK(5)	2.21	1.51	1.21	0.89
S&P500:				
GARCH(1,1)	23.04	18.24	10.72	6.01
AR-RV(5)	7.63	5.44	-0.92	-3.43
AR-MFVK(5)	1.57	1.20	0.99	0.78

Reported values are certainty equivalent returns (CERs) of the expected utility maximising portfolio for each density forecasting method, with various levels of investor risk aversion. All CER values are expressed as annualised % rates of return. The negative CER values observed in some cases may be due to the degree of risk aversion implicit in the investors' utility function, or the uncertainty around the density forecasts yielding portfolios with negative average rates of return. For the equity index series the first of these two factors is sufficient to explain all negative CER values, but for the exchange rate series both factors are relevant.

benchmark in second place. For the EUR/USD series the gains from the multifractal approach over the benchmark methods are substantial, again confirming the previous finding of strong performance for the EUR/USD data. For the JPY/USD series the performance of the multifractal and AR-RV methods are typically closer, with the largest differences observed at lower levels of risk aversion. For the equity index series the ranking of the forecasting methods is typically reversed, with the GARCH method producing portfolios with much higher CER values than either the AR-RV benchmark or the new multifractal method. It should however be noted the multifractal density forecasts produce portfolios with lower return volatility than the benchmark models and so for higher levels of risk aversion the multifractal method is actually able to provide higher a CER than the realised volatility approach.

5 Conclusion

The current paper has proposed a new method for estimating and forecasting the moments and probability density function of daily financial returns using intraday data.

The method is based on a new application of results from the theory of multifractal processes that provide a formal statistical link between the moments of the return process at different sampling intervals, allowing the variance and kurtosis of daily returns to be estimated directly from high-frequency intraday data. In the current application, these moment estimates are incorporated into density forecasts of daily returns, however in other financial applications the variance and kurtosis of returns are also variables of substantial interest in their own right.

In principle, the incorporation of relevant information contained in the intraday data can provide gains when estimating daily return moments, compared to methods based purely on daily data. At the same time, in comparison to existing methods utilising intraday data in the realised volatility literature, the multifractal approach preserves a greater proportion of the information contained in intraday returns by allowing the data to be used to directly estimate both the variance and kurtosis of daily returns.

The predictive ability of density forecasts produced by the new multifractal method was compared to existing methods in an empirical application using 5-minute intraday data on Euro (EUR) and Japanese Yen (JPY) exchange rates against the US Dollar (USD) and the S&P500 and NASDAQ-100 equity indexes. For the EUR/USD data the multifractal method provides large improvements in predictive ability over the GARCH benchmark model and is competitive with existing realised volatility based methods. This strong performance is improved further when considering the modified multifractal method proposed in Section 3.5 using randomly shuffled observations from each window of intraday data; this modification further increases the existing gains in predictive ability over the GARCH benchmark and also allows the method to provide statistically significant improvements over the realised volatility based benchmark.

For the remaining asset return series, density forecasting performance of the multifractal approach is competitive with existing methods, with the null of equal predictive ability unable to be rejected in the majority of cases. As with the EUR/USD data, the modified multifractal approach using shuffled intraday data provides consistent improvements in predictive ability for both the S&P500 and NASDAQ-100 data; in this case, in no situations can the null of equal predictive ability be rejected for the equity data, in contrast to the standard multifractal approach using ordered data for which the benchmark methods were found to be superior in some situations.

These empirical findings are reinforced by the results of a portfolio allocation exercise, in which the density forecasts from the competing methods are employed to optimally allocate funds between a risky and risk-free asset. In this context it was again found that the new multifractal approach can provide substantial gains over existing methods for the EUR/USD data, when measured in terms of the certainty equivalent return of the resulting portfolio. For the other assets the multifractal approach is found to outperform the realised volatility based method for higher levels of investor risk aversion, although for the equity index series both of the intraday methods are outperformed by the GARCH benchmark method.

A possible explanation identified for this variation in forecasting performance across the various return series is provided by differences in the strength of distributional scaling for the return series; the EUR/USD data seem to exhibit much stronger distributional scaling than the other series, with the JPY/USD and NASDAQ-100 data having the weakest scaling and the S&P500 data in between these extremes. Thus it seems that there is some positive relation between the strength of the distributional scaling exhibited by a given time series of data and the resulting density forecasting performance of the multifractal method.

There are several possible changes that could be made to the current implementation of the multifractal method that could potentially improve density forecasting performance further. The first is to identify an alternative parametric form for the daily return density that does not require restrictions to be placed on the daily return kurtosis, as is necessary with the current location-scale t-distribution. Secondly, alternative estimators for the scaling function could be investigated to replace the current partition function method; this would possibly allow the daily return moments for each day to be more accurately estimated for return series possessing weaker distributional scaling, such as the JPY/USD data. Finally, more flexible dynamic specifications could be tested for modelling and forecasting the daily return moments to replace the simple autoregressive models currently used; given the ability of the multifractal approach to estimate moments of returns at any chosen timescale from the same intraday data, specifications employing data at different sampling intervals could be employed, such as the mixed data sampling (MIDAS) and heterogenous autoregressive (HAR) models previously applied by Clements, Galvão, & Kim (2008) to the problem of producing quantile forecasts for daily returns from realised volatility measures.

References

Alessio, E., Carbone, A., Castelli, G., & Frappietro, V. (2002). Second-order moving average and scaling of stochastic time series. *The European Physical Journal B-Condensed Matter and Complex Systems*, *27*(2), 197–200.

Amisano, G. & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, *25*(2), 177–190.

Andersen, T. G. & Bollerslev, T. (1998). Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance*, *53*(1), 219–265.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-ā-vis forecast horizon. *Journal of Empirical Finance*, *6*(5), 457–477.

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

Calvet, L. & Fisher, A. (2002). Multifractality in asset returns: theory and evidence. *Review of Economics and Statistics*, *84*(3), 381–406.

Calvet, L. & Fisher, A. (2004). How to forecast long-run volatility: regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, *2*(1), 49–83.

Calvet, L., Fisher, A., & Thompson, S. (2006). Volatility comovement: a multifrequency approach. *Journal of Econometrics*, *131*(1-2), 179–215.

Clements, M. P., Galvão, A., & Kim, J. (2008). Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance*, *15*(4), 729–750.

Di Matteo, T. (2007). Multi-scaling in finance. *Quantitative Finance*, *7*(1), 21–36.

Di Matteo, T., Aste, T., & Dacorogna, M. (2005). Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance*, *29*(4), 827–851.

Diebold, F. X., Gunther, T., & Tay, A. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, *39*(4), 863–883.

Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance*, *57*(1), 369–403.

Fillol, J. (2003). Multifractality: theory and evidence, an application to the French stock market. *Economics Bulletin*, *3*(31), 1–12.

Giot, P. & Laurent, S. (2004). Modelling daily Value-at-Risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, *11*, 379–398.

Gneiting, T. & Ranjan, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *Journal of Business & Economic Statistics*, *29*(3), 411–244.

Guermat, C. & Harris, R. (2002). Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns. *International Journal of Forecasting*, *18*(3), 409–419.

Hallam, M. & Olmo, J. (2013). Semiparametric density forecasts of daily financial returns from intraday data. *Forthcoming in Journal of Financial Econometrics.*

Harvey, C. & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, *55*(3), 1263–1295.

Jondeau, E. & Rockinger, M. (2001). Gram-Charlier densities. *Journal of Economic Dynamics and Control*, *25*(10), 1457–1483.

Jondeau, E. & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control*, *27*(10), 1699–1738.

Kantelhardt, J. (2009). Fractal and multifractal time series. In: *Encyclopedia of Complexity and Systems Science*, Springer.

Kantelhardt, J., Zschiegner, S., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, *316*(1-4), 87–114.

León, Á., Rubio, G., & Serna, G. (2005). Autoregresive conditional volatility, skewness and kurtosis. *The Quarterly Review of Economics and Finance*, *45*(4-5), 599–618.

Maheu, J. & McCurdy, T. (2010). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, *160*(1), 69–76.

Mandelbrot, B., Fisher, A., & Calvet, L. (1997). A multifractal model of asset returns. *Cowles Foundation Discussion Papers, no. 1164, Available at SSRN: http://ssrn.com/abstract=78588.*

Matia, K., Ashkenazy, Y., & Stanley, H. (2003). Multifractal properties of price fluctuations of stocks and commodities. *Europhysics Letters*, *61*(3), 422–428.

Muzy, J., Delour, J., & Bacry, E. (2000). Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B-Condensed Matter*, *17*, 537–548.

Onali, E. & Goddard, J. (2009). Unifractality and multifractality in the Italian stock market. *International Review of Financial Analysis*, *18*(4), 154–163.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, *25*(1), 221–247.

Schmitt, F., Schertzer, D., & Lovejoy, S. (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models and Data Analysis*, *15*(1), 29–53.

Schumann, A. & Kantelhardt, J. (2011). Multifractal moving average analysis and test of multifractal model with tuned correlations. *Physica A*, *390*(14), 2637–2654.

Selçuk, F. & Gençay, R. (2006). Intraday dynamics of stock market returns and volatility. *Physica-Section A*, *367*, 375–387.

Žikěs, F. (2009). Semiparametric conditional quantile models for financial returns and realized volatility. *Mimeo, Imperial College London.*

Xu, Z. & Gençay, R. (2003). Scaling, self-similarity and multifractality in FX markets. *Physica-Section A*, *323*, 578–590.