High Frequency Data Do Improve Volatility and Risk Estimation

Fulvio Corsi, Michel Dacorogna,
Ulrich Müller, Gilles Zumbach

FCO.2000-03-18

January 6, 2009

An internal paper by the Olsen Ltd. Research Group

Contents

1 Introduction

Volatility is an essential ingredient for many applied issues in finance and financial engineering, ranging from asset pricing and allocation to risk management. In particular, due to the increasingly central role played by the Value-at-Risk (VaR) approach in risk assessment, it is becoming progressively more important to have a good definition, measure and forecast of short term volatility.

But despite its importance, volatility is still an ambiguous term for which there is no unique, universally accepted precise definition. The standard approaches currently compute volatilities either by fitting econometric models such as GARCH, or by studies volatilities implied from specific option pricing models such as Black-Scholes, or by studying historical indicators computed from daily squared (or absolute) returns. But, especially for short term volatility forecasts, all of those approaches suffer from noticeable drawbacks. Empirical evidence points towards misspecification of the GARCH and Black-Scholes models. GARCH models have aggregation, scaling and long memory properties which does not corresponds to that widely found in the real data. In Black-Scholes models implied volatility presents the so called smile and can yet have a too slowly changing dynamical behaviour. Using daily returns as a proxy for the volatility of the day in GARCH models and historical indicators, leads to very large measurement errors, as pointed out in (Andersen and Bollerslev, 1998)

In particular the large measurement errors of volatility deeply affects the results of VaR computation; it has been shown (Müller, 1999) that the daily computation of VaR suffers from an alarming sensitivity against the choice of the sampling daytime (i.e. the choice of the hour of the day at which the price is sampled and the volatility and VaR computed). Such excessive high sensitivity is primarily due to the stochastic noise introduced by the arbitrary choice of prices, which makes the results dependent on only one observation (which may not be representative of the full day dynamics). Using only one price per day may potentially lead to loose all the other information contained in the whole process of prices formation during the day. These serious shortcomings of conventional VaR computations have motivated (Zumbach and Müller, 2001) to develop alternative approaches based on new time series operators that efficiently compute statistical variables from inhomogeneous time series. This approach allows a tick-by-tick VaR computation using an (almost) continuously overlapping of daily returns (Müller, 1999) that significantly reduces the stochastic noise of the standard VaR computed only once a day. For instance, as theoretically shown by the same author (Müller, 1993), a continuously overlapping scheme applied to a Gaussian random walk can reduced the error variance of 1/3 of the value obtained without overlapping.

On the other hand (Taylor and Xu, 1997) and (Dacorogna *et al.*, 1998) motivated by the purpose of converting the volatility from a latent variable to an observable one, proposed a new measurement of the daily volatility based on a sum of many high frequency intraday squared returns. This approach, as claimed by (Andersen *et al.*, 2001) promise, in theory, to reach an error free estimation of the volatility. Treating the realized volatility as an observable object has the far reaching consequence of allowing to directly fit forecasting models rather than using much more complicated ARCH-type econometric models required when volatility is viewed as a latent variable. In practice however, in this study we found that for return intervals less than few hours, such definition is affected by a considerable systematic error. In other words, the expectation of the volatility computed with high frequency returns is not equal to the one obtained with daily returns. Since our target measure is the daily volatility we shall consider the expectation of the volatility from 1-day returns the suitable one; hence we term the difference between the two measures the bias of the high frequency volatility. To cope with this discrepancy we will therefore employ the bias correction methods proposed by (Müller *et al.*, 2000) with the aim of achieving a realized volatility measure of optimal precision.

We believe that, once the practical subtle pitfalls contained in the definition of high frequency realized volatility will be overcame, it will be extremely advantageous to integrate this better measurement of the target function (i.e. the realized volatility) in the quasi-continuously updated tick-by-tick VaR proposed by (Müller, 1999)

Viewing this work as a first step towards an alternative and superior VaR computation than that of

RiskMetrics, the purpose of this paper is twofold. First illustrate how using high frequency data in the measurement of the realized volatility significantly improve the standard forecasting models currently used in the conventional approaches to VaR. Second presents evidence that a wide spectrum of frequency ranging from few minutes to months allows to compute better forecasts of the realized volatility.

The paper is organized as follows. In section 2 a proper target function for volatility forecast, realized volatility and its bias are analyzed, together with a description of a bias correction procedure. In section 3, our volatility forecasting model based on a combination of time series operator with different horizons (EMA-HAR) is introduced, while the methodology and data employed described in section 4. In section 5 the empirical results of forecasting performances are summarized and section 6 concludes.

2 Realized volatility

2.1 High-frequency realized volatility

In the standard framework volatilities are usually computed by squaring returns obtained from an artificially regular time series of daily prices (indexed by the subscripts i). The daily return can thus be defined by the following point-wise price difference:

$$r(t_i) \equiv r(\Delta t; t_i) \equiv x(t_i) - x(t_i - \Delta t_{ref}) \ , \tag{1}$$

where $x(t)$ is the logarithmic price at time t and $\Delta t_{ref} = 1$ working day.

The idea of employing high frequency data in the computation of volatility trace back to the seminal intuition of Merton (1983) according to which higher frequency are not useful for the mean but essential for the variance. Yet only recently such idea has been exploited with intraday data: (Dacorogna *et al.*, 1998) use 1-hour returns to compute a 1-day volatility benchmark for other volatility forecasting models, Schwert (1998) utilizes 15-minute returns to estimate daily stock market volatilities, while Taylor and Xu (1997) and Andersen, Bollerslev, Deibold and Labys (1999) rely on 5-minute returns in the measurement of daily exchange rate volatilities. Furthermore (Andersen *et al.*, 2001) provide a formal justification for the use of realized volatility measure constructed by summing high frequency intraday returns [1].

Hence following (with only slightly differences) (Dacorogna *et al.*, 1998) and (Andersen *et al.*, 2001) we can define the 1-day realized volatility as the square root of the mean variance of high frequency returns

$$\sigma(t; \Delta t) = \sqrt{\frac{1}{m} \sum_{i=0}^{m-1} r^2(t - i\Delta t)} \quad \text{with} \quad m = \frac{\text{1-day}}{\Delta t} \ , \tag{2}$$

where m is here the aggregation factor of returns depending on the interval size of the high frequency returns Δt (for instance for a daily volatility computed with 5-minute returns we have $m = 288$) and r is assumed to be zero; and the normalized realized volatility

$$\sigma(t; \Delta t) = \sqrt{\frac{T_{ref}}{\Delta t} \frac{1}{m} \sum_{i=0}^{m-1} r^2(t - i\Delta t)} = \sqrt{\frac{T_{ref}}{\Delta t}} \cdot \sigma(t) \ , \tag{3}$$

where T_{ref} is a normalization period often chosen equal to one day or one year (in which case we term this operation, annualization). In the computation of this quantity we will use business time scale and previous tick interpolation.

[1] This definition is justified, as shown in (Andersen *et al.*, 2001), on the basis of a continuous time model for which the realized variance is a consistent estimator of the 1-day integrated one i.e. the one obtained integrating the instantaneous variance over one day.

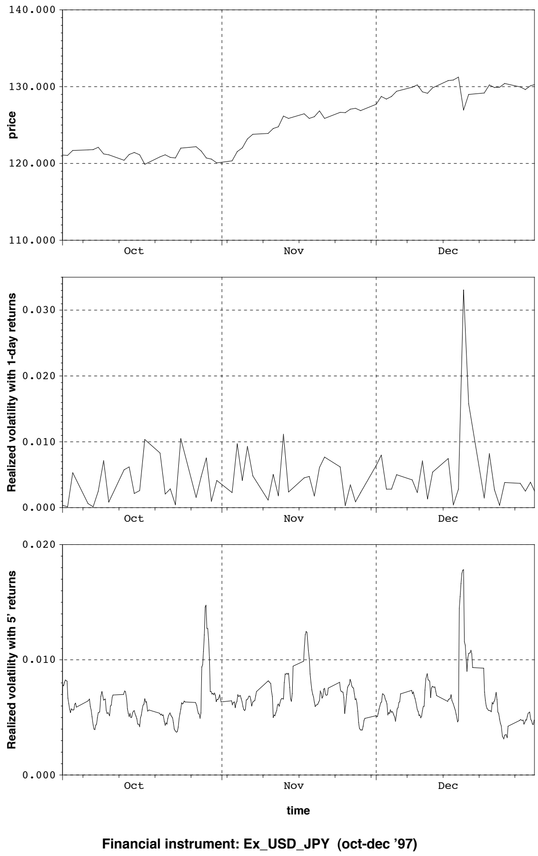

Figure 1: Sample comparison between volatilities evaluated from daily returns (middle panel) and corrected realized volatility computed with 5-minute ones (bottom); note the difference in scale.

2.2 Bias and anomalous scaling

Since the realized volatility is our target function for the forecast, it should be as much as possible free from measurement errors and noise. From this point of view, the volatility computed with daily returns is a very loose measure since it relies only on one observation per day, taken at a certain daytime; all the other information on the price process of the day is thrown away. For instance, as shown by (Andersen and Bollerslev, 1998) the variance of the measurement error of the volatility measured with daily squared returns, is often twenty time the unconditional variance of the squared volatility. This contrasts sharply with the theoretical result reached by (Andersen *et al.*, 2001) under the assumption of a standard continuous-time diffusion process [2] for the logarithmic price; in such a model the stochastic error of the measure can be arbitrary reduced by simply increasing the sampling frequency of returns (using the notation of equation (2) the measurement error goes to zero as $m \to \infty$).

Unfortunately, it turns out that in practice the assumption that log asset prices evolve as a diffusion

[2] Specifically, an Ito process with zero expected drift rate

process (i.e. with independent increments) becomes progressively less and less realistic as the time scale reduces. In fact, when the return interval shrinks market microstructures effects arise. Departures from an i.i.d. diffusion process makes the realized volatility computed with very short interval returns no longer an unbiased and consistent estimator of the daily volatility. To be more precise, indicating (with a slight abuse of notation [3]) the volatility computed with daily returns with $\sigma(\Delta t_{ref})$ and the one obtained with high frequency returns with $\sigma(\Delta t)$, the difference

$$B = \mathrm{E}[\sigma(\Delta t)] - \mathrm{E}[\sigma(\Delta t_{ref})] \tag{4}$$

is no longer equal to zero. We will term this non zero difference the *bias* of the realized volatility (and conventionally speaks of positive bias when $B > 0$ and negative bias for $B < 0$).

We justify the use of this term on the basis of the simple observation that for almost the totality of the agents currently operating on financial markets, the relevant variable of interest is the daily volatility (or a longer one) and not the volatility observed at a say, 5-minute level. In other words, operators are interested in measuring and forecasting the "one day risk" regardless of the detailed behaviour happening at very small time scale. The fact that we use high frequency return to compute the daily volatility is merely a measurement issue. We employ short term return because we want to reduce the estimation error of our measure, not because we are interested in evaluating the risk existing at such fine time frame. Not considering the bias would imply to contaminate the measure of daily risk with risk components that are present only at very short time scale, and that would never be perceived by an operator having a 1-day horizon.

Taking into account the existence of this bias, leads to a trade-off between two opposite types of limitations which precludes the possibility to have an infinitively precise, easy measure of the realized volatility. In fact, if on the one hand, statistical consideration would impose to use a very high number of return observations to reduce the stochastic error of the measurement, on the other hand, after a certain threshold market microstructure frictions come into play introducing a bias that usually increases as the sampling frequency increases.

A significant bias of the volatility can only be explained by a significant departure from the i.i.d. hypothesis of the price changes over a very short horizon. This bias represent nothing but the deviation of the scaling law of the realized volatility from that of an i.i.d process. Then the actual amount of bias for any given interval return can be analyzed by looking at the scaling behaviour of the average normalized volatility computed at different frequencies. Such scaling behaviour can be conveniently represented plotting the average normalized volatility $\mathrm{E}[\sigma(\Delta t)]$ against the logarithm of Δt. Using normalized volatility [4] the scaling behaviour of a standard i.i.d. diffusion process will appear as a flat line, meaning that the average rescaled volatility doesn't change with the change of return frequency. In other words, using our notation, if the price followed an i.i.d. random walk the realized volatility would be independent of Δt and therefore the function $\sigma(\Delta t)$ should have a 0 slope in the plan $\sigma - \log \Delta t$.

If this doesn't happen to be true, we are facing an anomalous scaling of the realized volatility. In this case the vertical distance between the constant line passing through the volatility computed with 1-day returns and the empirical value of the volatility obtained with any other return interval, directly gives us the amplitude of the bias at any given time scale. Empirical research performed at O&A on a quite large variety of currencies, stock indexes, and interest rates, shown a surprisingly rich set of different behaviours of the bias as a function of Δt. Plotting the normalized realized volatility against $\log \Delta t$ the general behaviour of the bias for each instrument can be very coarsely summarized as follow:

- FX : positive bias ranging between 30-80% of the realized volatility at $\Delta t = 1$ day; such rather strong positive bias seems mostly given by a short-term mean-reverting price changes that leads to a spurious addition to the volatility (Goodhart, 1989) (Goodhart and Figliuoli, 1991) (Guillaume *et al.*, 1997). The mean-reverting behaviour existing at very short lags, can be reasonably explained

[3] $\sigma(t; \Delta t) \equiv \sigma(\Delta t)$

[4] That is volatility rescaled to one day

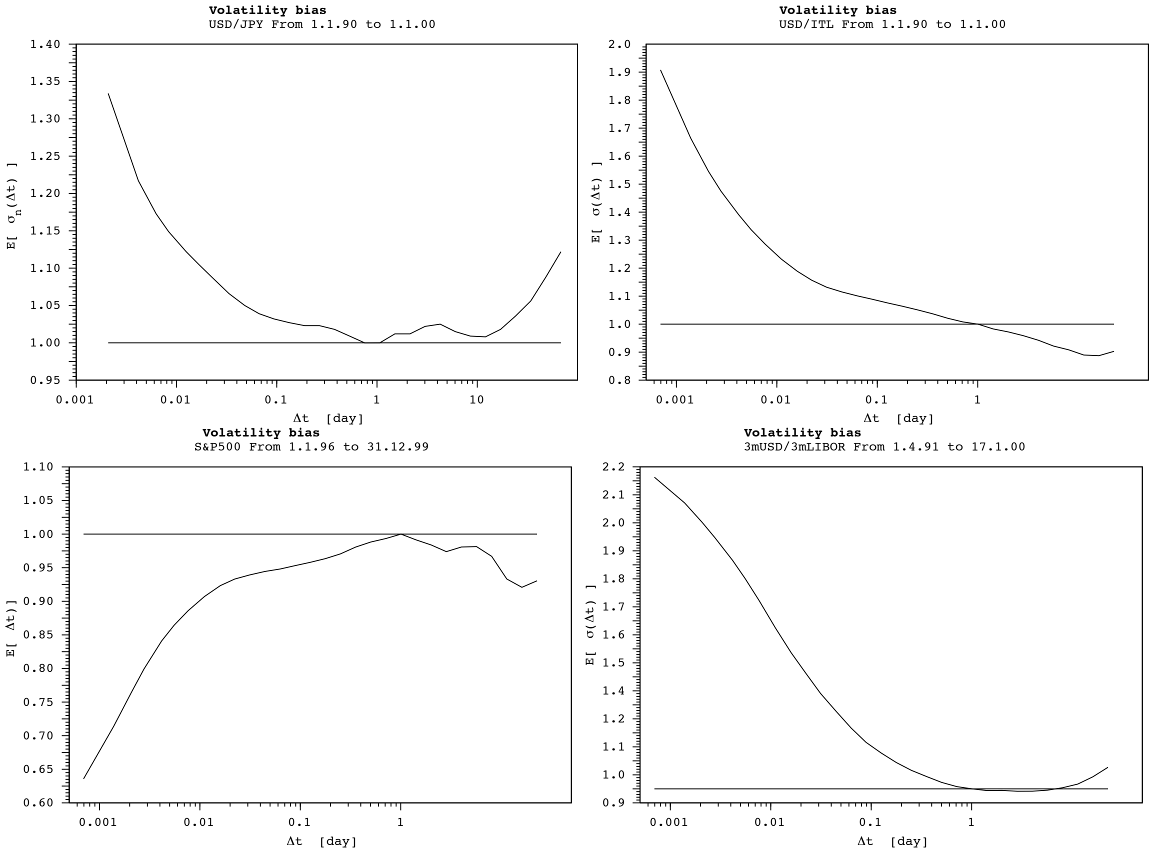

Figure 2: Volatility bias of different instruments for Δt ranging from 1 minute to 1 week. Volatilities are normalized so that $\sigma(\Delta t_{ref}) = 1$

(Roll, 1984) by the random hitting of the transaction value of the bid or ask price. This leads to a negative serial correlation in returns that is positively related to the size of the spread. The fact that, FX are OTC quoted prices, obviously enlarges such mean-reverting behaviour (and hence the size of the bias) by means of the so called "bid-ask inventory boucing" operated by the market-maker [5] (Bollerslev and Domowitz, 1993) (Flood, 1994). We also found that the bias tends to be higher for exchange rate with lower liquidity (like for instance the USD/ITL which has a bias of more than 80% at 3-minute level) than that between major currencies which present higher liquidity and hence lower bid-ask spreads. The dynamical behaviour of the bias (as for the other financial instruments analyzed) remain quite constant over time. This means that the booms and bursts of volatility result in quasi-parallel shifts of the curve depicted in figure 2.2 (cf. fig 3). This quite general empirical property of the bias will play a central role in the bias correction method we will propose in the next paragraph.

- Interest rates: highly variable results ranging from slightly negative bias to very strong positive ones; possible explanations of this instability and very high level of the bias can be related to the relative thiner thickness of those markets and to the much more severe rounding problem of the prices (since only two digits after the point are considered). Both these features point to the direction of increasing the size of the price movements observed at very high frequency and the presence and impact of the gaps.

[5] The tendency to skew the spread in a particular direction when they have order imbalance

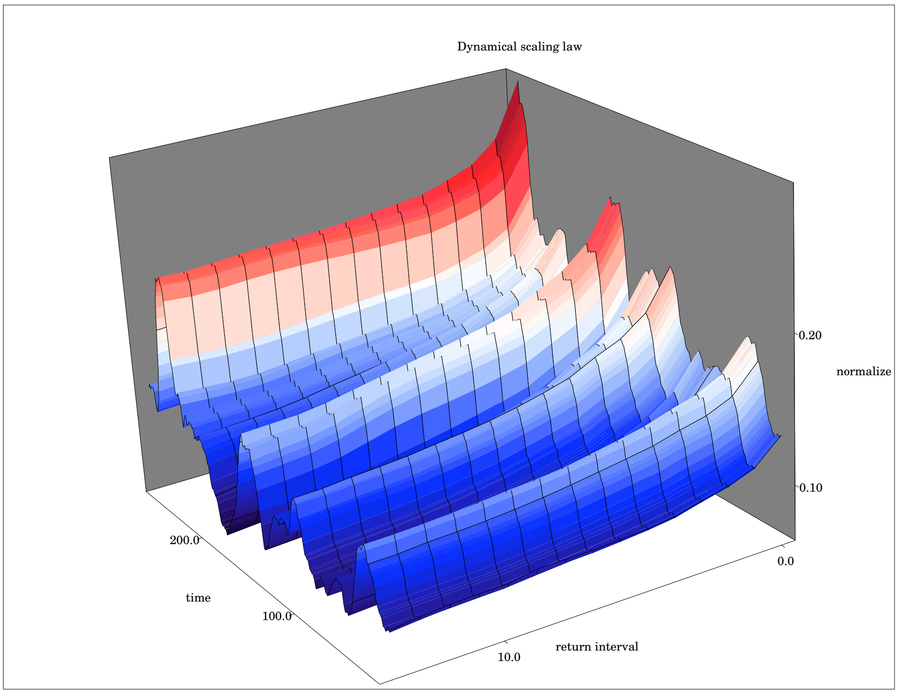

Figure 3: Dynamical behaviour of the volatility bias of USD/JPY for a 10 year period from 1990 to 2000 and for return interval ranging from 1 minute to 1 week.

- Stock indices: slightly or strongly negative biased. A negative bias implies a positive slope of the scaling function $\sigma(\Delta t)$ meaning that the price process is not i.i.d., but slightly persistent, at least for small time interval. The empirical evidence that the stock indices returns present a significant positive autocorrelation for intraday horizon (as also found by many other authors [6]) could be explained by the so-called lagged adjustment models. According to these models, inside indices, there are stocks (the "leading" one) that reacts more quickly to aggregated information and are followed by the others with a certain delay.

All these empirical results may be blended in a consistent picture of the dynamical properties of the price process, by noting that the scaling behaviour of the volatility is related to the correlation structures of returns. Such relation stems directly in a very simple way from the well-known expression of the variance of a sum of stationary random variables. This formula implicitly establishes an exact relation between the autocorrelation function of a process and its scaling behaviour. In fact the variance at a time scale of one day $\Delta t_{ref} = m\Delta t$ (where Δt could be for example 5 minutes) may be expressed as:

$$\sigma^2(\Delta t_{ref}) = \sigma^2(m\Delta t) = m\sigma^2(\Delta t) + 2\sum\sum_{k>l} cov[k-l] \tag{5}$$

$$= m\sigma^2(\Delta t) + 2\sigma^2(\Delta t)\sum_{k>l} \rho((k-l)\Delta t) \tag{6}$$

[6] See for example (Bouchaud, 1998)

$$= m\sigma^2(\Delta t) + 2\sigma^2(\Delta t)\sum_{k=1}^{m}(m-k)\rho(k\Delta t) \quad (7)$$

$$= m\sigma^2(\Delta t)\left[1 + 2\sum_{k=1}^{m}(1-\frac{k}{m})\rho(k\Delta t)\right] \quad (8)$$

where $\rho(k\Delta t)$ is the autocorrelation function at lag $k\Delta t$.

Hence knowing the autocorrelation structure from lags Δt to Δt_{ref} it would be possible, for any given stationary process, to exactly rescale the high frequency variance $\sigma^2(\Delta t)$ into the aggregated 1-day variance $\sigma^2(\Delta t_{ref})$. Note that the absence of autocorrelation implies $\sigma^2(m\Delta t) = m\sigma^2(\Delta t)$ i.e. linear scaling of the variance. Since at short time scale (usually less than an hour) the autocorrelation of returns becomes not negligible, the scaling behaviour of the variance strongly deviates from that of a standard i.i.d. process.

2.3 Bias correction

Given the trade-off between measurement error and bias previously described, the choice of the underlying return frequency becomes a critical issue if no specific treatment of the bias is employed. In a recent unpublished manuscript, Andersen et al. propose a direct graphical inspection of the scaling law behaviour of the realized volatility (which they call "volatility signature plot") as a guidance for the choice of the underlying sampling frequency of returns. The idea is simply to determine, for each financial instruments, the shortest return interval at which the resulting realized volatility is still not affected by the bias. They found a Δt of only 30-20 minutes for highly liquid exchange rates and even a positive bias for the less liquid ones.

We argue that those results are distorted by the probable use of the linear interpolation scheme, employed by the authors to convert the original inhomogeneous time series of prices in to an homogeneous one. For high frequency volatility computation, linear interpolation should be avoided because it implicitly assumes minimum volatility between original ticks and induces an artificial positive correlation of returns in presence of gaps. This leads to a systematic underestimation of volatility that becomes more and more pronounced as the tick frequency decreases and the gaps increases. This could explain the positive bias encountered for the less liquid rates. In fact, in the case of exchange rate where the bias is mostly given by the bid-ask boucing, it can be theoretically shown (Müller *et al.*, 2000) that in absence of gaps the bias can not be positive: only an artificially positive autocorrelation induced by linear interpolation inside the gaps can produce such positive bias.

For these reasons it seems theoretically more correct to use the "previous-tick interpolation" scheme [7] that does not generate any underestimation of volatilities or distortion of autocorrelations. Employing previous-tick interpolation we found significant evidences of large bias at 30-20-minute level even for major exchange rates. In our analysis the return interval at which the bias is no longer significant, occurs only at the level of some hours. This means that even for the most highly liquid assets, the shortest return interval from which compute an unbiased volatility measure, is of the order of 2-3 hours, that means only 8-12 observations per day. Furthermore this "unbiased return interval" changes considerably from asset to asset.

So, without any specific treatment for the bias the stochastic error of the volatility measure cannot be significantly reduced and cannot be computed with the same return interval for all the instruments. Hence, if we want to have a general and homogeneous definition of a realized volatility measure of optimal precision, an explicit treatment of the bias is required.

As we wrote, in theory, knowing precisely the current autocorrelation structure of high frequency returns, we could directly employ equation (8) to exactly correct for the bias of the high frequency volatility.

[7] In which each tick remains valid until a new tick arrives.

Obviously in practice we don't have such information, and we would therefore need to empirically estimate the autocorrelation function from historical data. But then, since equation (8) prescribes to sum all the autocorrelation values from lag 1 to m, we will end up with a long sum of m estimated values each with its own error, which will lead to a relevant stochastic error of the overall expression.

Given those difficulties (which we hope to overcome in future works), we decide to adopt the simple bias correction method proposed in (Müller *et al.*, 2000). The bias correction advanced there is a purely multiplicative one, based on the use of quantities estimated from a rather long "calibration sample". Namely, we average the high frequency and daily volatilities over the calibration period T_{calib} (of, let say, one year). Labeling with $\sigma^2_{calib}(\Delta t)$ the average squared volatility computed with small return interval Δt over the calibration sample and $\sigma^2_{calib}(\Delta t_{ref})$ the one obtained from daily returns, we thus correct the current local observation of $\sigma^2(\Delta t)$ by means of the following multiplicative formula:

$$\sigma^2_{corr}(\Delta t) = \frac{\sigma^2_{calib}(\Delta t_{ref})}{\sigma^2_{calib}(\Delta t)} \sigma^2(\Delta t) \tag{9}$$

The effectiveness of such bias correction procedure relies on the following assumptions which also summarize the weak points of this approach:

1. Absence of structural breaks.
2. Relatively stable data frequency (that affects the size of the bias as explained in (Müller *et al.*, 2000)).
3. A reasonable stability of the correction factor with respect to the market state (excited or dull)

But it also presents remarkable advantages due to its simplicity, easy acceptability for the practitioner community and very limited model assumptions.

On the other hand, it is important to underline that the correction procedure suggested does not entirely solve all the problems associated with the definition of the realized volatility. In fact all the three variables appearing on the right-hand side of eq.(9) are stochastic and it turns out that the expectation of $\sigma^2_{corr}(\Delta t)$ is not exactly the same as that of $\sigma^2_{corr}(\Delta t_{ref})$; hence the bias is not completely eliminated though substantially reduced. Furthermore the stochastic nature of eq. (9) implies that $\sigma^2_{corr}(\Delta t)$ has also a stochastic error related to that of $\sigma^2_{calib}(\Delta t_{ref})$ and $\sigma^2(\Delta t)$ (while that of $\sigma^2_{calib}(\Delta t)$ can be safely neglected). Taking into account both the stochastic error and the bias, (Müller *et al.*, 2000) shows how a bias-corrected realized volatility computed with 5-minute returns presents a soundly reduced error compared to a non corrected one evaluated with hourly returns. In fact the relative error in the latter case of no bias correction is roughly 20%. A quite large error that causes serious problems when one tries to optimize a more precise volatility forecast using this definition of realized volatility as target function of her forecast. Indeed, such an erratic target function permits to optimize only predictors with a forecasting error larger than 20%, hence, confining ourselves to bad forecasting models. When, instead, the bias correction of eq (9) is employed, the expected error of the realized volatility shrinks to less than 9% opening the way for testing and optimizing better volatility forecasting models.

3 Forecasting model of volatility

3.1 Time series operators

The O&A approach is based on new time series operators, the theory of which is derived and described in (Zumbach and Müller, 2001); here only a minimum description is given. A time series operator Ω,

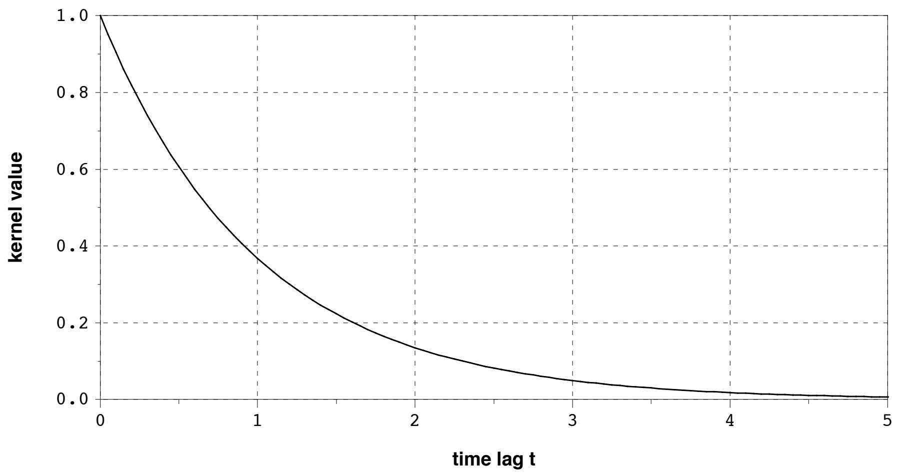

Figure 4: Kernel of the simple EMA operator $\mathrm{EMA}[\tau; z]$. Here: $\tau = 1$.

mapping the space of a time series z into a new time series denoted by $\Omega[z]$,can be represented [8] by a convolution with kernel $w(t)$

$$\Omega[z](t) = \int_{-\infty}^{t} w(t - t')z(t')\mathrm{d}t' \tag{10}$$

Being defined as an integral of the time series $z(t)$ to which the operator applies it should be defined in continuous time. Therefore, since actual data are known only at discrete sampling times, some forms of interpolation procedures need to be used. If the kernel $w(t)$ is non-negative and normalized, $\Omega[z]$ is a *weighted moving average* of z where the weight given to some past event coincides with the positive value of the kernel at its corresponding time. So in the particular case in which the kernel has a rectangular shape we obtain the "usual" moving average operator. The characteristic depth of the past covered by the kernel is called the *range* of an operator and is defined as the first moment of its kernel:

$$r = \int_{-\infty}^{\infty} w(t)t\mathrm{d}t \tag{11}$$

The basic building block of the time series operators employed in this analysis is the *Exponential Moving Average* (written $\mathrm{EMA}[\tau, z]$) characterized by an exponentially decaying kernel

$$w(t) = \mathrm{ema}(t) = \frac{\mathrm{e}^{-t/\tau}}{\tau} \tag{12}$$

with range $r = \tau$. The EMA operator is the corner stone of our time series operators because it can be computed very efficiently by the following simple iterative formula

$$\mathrm{EMA}[\tau; z](t_n) = \mu \mathrm{EMA}[\tau; z](t_{n-1}) + (1 - \mu)z(t_n) + (\mu - v)[z(t_n) - z(t_{n-1})] \tag{13}$$

where $\mu = \mathrm{e}^{-\alpha}$, $\alpha = (t_n - t_{n-1})/\tau$ and $v = \begin{cases} 1 & \text{previous point} \\ (1-\mu)/\alpha & \text{linear interpolation} \end{cases}$

and because other more complex operators with a wide variety of kernels can be built with it:

[8]Provided that the operator is linear and translation-invariant.

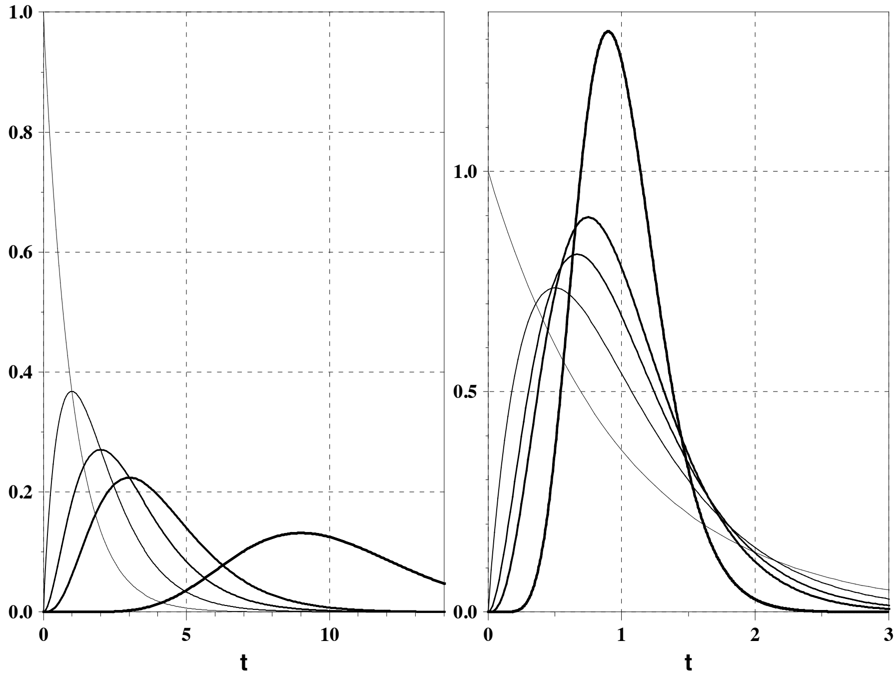

Figure 5: Kernels of $\mathrm{EMA}[\tau, n; z]$. Curves for $n = 1$ (thin), 2, 3, 4, 10 (bold). Left graph: for $\tau = 1$; right graph: $r = n\tau = 1$.

- The $\mathrm{EMA}[\tau, n; z]$ operator, obtained by repeatedly applying n time the same simple EMA operator, so that

$$\mathrm{EMA}[\tau, k; z] = \mathrm{EMA}[\tau; \mathrm{EMA}[\tau, k-1; z]] \tag{14}$$

  This operator has a kernel $\mathrm{ema}[\tau, n](t) = \frac{1}{(n-1)!} \left(\frac{t}{\tau}\right)^{n-1} \frac{\mathrm{e}^{-t/\tau}}{\tau}$ and range $r = n\tau$.

- The *Moving Average* $\mathrm{MA}[\tau, n]$ operator, constructed as a sum of $\mathrm{EMA}[\tau, n; z]$ operators:

$$\mathrm{MA}[\tau, n] = \frac{1}{n} \sum_{k=1}^{n} \mathrm{EMA}[\tau', k] \text{ with } \tau' = \frac{2\tau}{n+1} \tag{15}$$

  where the parameter τ' is chosen so that the range of $\mathrm{MA}[\tau, n]$ is $r = \tau$, independently of n. This combination of EMAs provides a family of more rectangular shaped kernels with expression:

$$\mathrm{ma}[\tau, n](t) = \frac{n+1}{n} \frac{\mathrm{e}^{-t/\tau'}}{2\tau} \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{t}{\tau'}\right)^{k} \tag{16}$$

  this permits to move, for increasing n, from a simple EMA operator with exponentially decaying kernel for $n = 1$, to an exactly rectangular shaped kernel for $n \to \infty$.

Using the MA operator we can now define the *nonlinear moving norm operator MNorm* as:

$$\mathrm{MNorm}[\tau, p; z] = \mathrm{MA}[\tau; |z|^{p}]^{1/p} \tag{17}$$

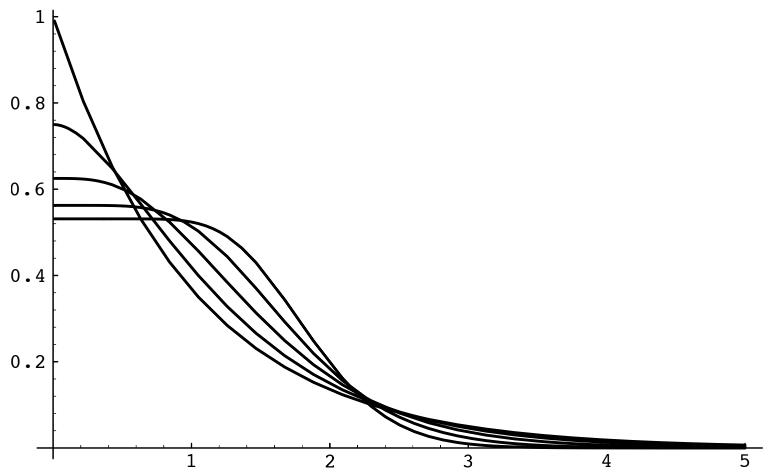

Figure 6: Kernels of $\mathrm{MA}[\tau, n; z]$; $n = 1, 2, 4, 8, 16$, for $\tau = 1$.

Having defined EMA, MA and MNorm operators, we can finally write a more general and widely applicable definition of volatility [9]:

$$\mathrm{Volatility}[\Delta t, T, p, x] = \mathrm{MNorm}\left[\frac{T}{2}, p; D[\Delta t; x]\right] = \left\{\mathrm{MA}\left[\frac{T}{2}, n; |D[\Delta t; x]|^p\right]\right\}^{1/p} \tag{18}$$

where the parameters appearing in the formula have the following meaning:

- Δt is the size of the return measurement intervals.
- T is the length of the total moving sample which is defined as the double of the range so that the MA operator has a range of $T/2$.
- p is the exponent of the moving norm operator. Usually a $p = 2$ is taken even though some researchers have shown that a lower value for p provide a more robust estimate.

The operator $D[\Delta t, x]$ is a *differential operator* which has the function to provide a more general definition of returns computed from a not necessarily homogeneous time series of logarithmic prices x with a given time interval Δt; i.e. $D[\Delta t, x]$ corresponds to $r(\Delta t)$. The differential operator also allows to compute a more smoothed return significantly diminishing the usual, rather noisy behaviour of the simple returns defined in equation (1). In theory, it would be possible to obtain smoothed returns applying a large variety of differential operators of different class. Nevertheless since we are particularly interested in having operators with high speed of adjustment to news, we will confine ourself to the differential operators [10] of the kind:

$$D[\Delta t, x] = x(t) - \mathrm{EMA}[T', n'; x](t) \tag{19}$$

with range $r = T' \cdot n' = \Delta t$.

Even with this simplification we still have a number of possibility of calculating a broad range of different measures of volatility changing the value of the remaining free parameters:

1. the exponent of the moving norm operator p

[9] Here we assume, as the empirical evidences shows, that $E[r] = 0$.

[10] Often called momentum or oscillator in finance.

2. the size of the return measurement intervals Δt
3. the length of the total moving sample T
4. the kernel, parametrized by n', that appear in the EMA$[T', n', x]$ operator used to compute smoothed returns
5. the kernel of the MA$[T, n]$ operator parametrized by n.

3.2 Tick-by-tick volatility computation

Using such new time series operators we will compute an estimation for the next day volatility applying the following formula at the arrival of each new tick:

$$\sigma_{ema}(t) = \left\{ \text{MA}\left[\frac{T}{2}, n, ; C \cdot |x(t) - \text{EMA}[\Delta t, n'; x(t)|^p\right]\right\}^{1/p} \tag{20}$$

where the constant C is a correction factor. Defining the tick frequency ratio as $\omega \equiv \frac{\Delta\text{tick}}{\Delta t}$, where Δtick is the time interval between two subsequent ticks and Δt the range of the EMA operator that computes the returns, C compensate for the two following kinds of biases:

1. when having a high density of ticks (that means a small value of the ratio ω) C compensates for the reduction of volatility caused by the use of the smoothed returns $x(t) - \text{EMA}[\Delta t, n'; x(t)]$ instead of the simple $x(t) - x(t - \Delta t)$. Under the Gaussian random walk assumption a theoretically correct value c for this quantity can be exactly computed for each EMA$[n']$ operators; for instance the correct value for an EMA[4] is $c = 128/93$.
2. while having a low density of ticks (i.e. frequent presence of holes in the data flow that result in large values of ω) C compensates for the underestimation of the unobserved volatility by rescaling the squared return Δx^2 with ω. In fact it can be shown that for a Gaussian random walk observed with gaps, the volatility inside a gap can be estimated by $\omega \cdot \Delta x^2$.

Hence, in order to compensate for both sources of biases the formula for the correction factor C takes the form:

$$C = c - k + \sqrt{k^2 + \omega^2} \tag{21}$$

where k=0.65 guarantees a smooth transition of the correction factor from high density data (small holes) to low density ones (large gaps).

3.3 Forecasting model

For the real time VaR, we will consider (as discussed in the previous section) the realized volatility as an observable quantity and hence our approach will be to directly forecast it [11] by means of a new kind of autoregressive model: the heterogeneous autoregressive (HAR) model with exponential moving averages.

This EMA-HAR model is the result of a combination of many exponential moving averages with different time horizon and thus with different memory. The memory decay of the overall model, being a sum of many exponential decay, become very close to the hyperbolic decay broadly found in real data. It therefore represents a valid (and probably simpler and more flexible) alternative approach to the fractionally integrated process commonly utilized to model long memory process. The EMA-HAR is also in

[11] as also proposed by (Andersen *et al.*, 2001) and empirically implemented by Beltratti and Morana 99

line with the findings that led to the formulation of the HARCH-type models (Dacorogna *et al.*, 1998): the heterogeneity of agents horizons that affects the volatility and the empirical evidence that the coarse volatility predicts the fine volatility better than the other way round. In addition it also allows to take into account the results of (Blair *et al.*, 1999), (Deparis, 1999) according to which the out-of-sample forecasting performance of historical volatility models are significantly improved by using information contained in the high frequency returns.

We start defining a partial volatility σ_j to be the contribution of the jth horizon (T_j) to the total market volatility σ; and we will estimate each of them by applying the (20) formula posing $T = T_j$ and $\Delta t = \Delta t_j$

$$\hat{\sigma}_j(t) = \left\{ \text{MA}\left[\frac{T_j}{2}, n, ; C \cdot |x(t) - \text{EMA}[\Delta t_j, n'; x(t)|^p\right]\right\}^{1/p} \tag{22}$$

Then in order to exploit the different information carried by volatilities computed at various time scales we linearly combine partial volatilities with different horizons so that the volatility forecast will take the form

$$\hat{\sigma}(t) = \sum_{j=1}^{J} h_j \hat{\sigma}_j(t) \tag{23}$$

where J is the number of different horizons employed in the model.

But unlike in the HARCH-type models, since we assume that the realized volatility is an observable quantity, we won't use the maximum likelihood estimation technics usually employed in ARCH process to compute the $h_1 \dots h_J$ parameters. Here a RiskMetrics like optimization approach will be followed; which means that we will search for just one set of parameters that will provide reasonably good results for a wide range of different assets, and then simply use them in all the volatility forecasts. To simplify even further the computation of the forecasts, we will express the h_j as function (a power-law one) of few parameters. This low-profile optimization procedure has been chosen in view of our future target of developing a new RiskMetrics like VaR approach based on high frequency data. Alternatively, if we wished to have a more flexible and dynamic optimization, we could estimate the vector of weights by means of the following simple multiple linear regression:

$$\hat{\sigma}(t) = \sum_{j=1}^{J} h_j \hat{\sigma}_j(t) + \epsilon \tag{24}$$

performed at regular time interval (for instance every day or every hour).

3.4 Bias Correction of high frequency partial volatilities forecasts

As previously described our forecasting model of volatility will be a combination of different partial volatilities computed at various time scale (cf. equation 23). Since for the smallest time scales the partial volatility forecast is computed using interval returns deeply intraday (typically up to few minutes), we will again encounter the problem of the high frequency bias, discussed in the previous section for the realized volatility. Analogously to what we did for the realized volatility, our forecasts of high frequency partial volatilities will be corrected adopting a simple multiplicative formula similar to that of equation 9. The only differences will be that :

- the forecast will be directly calibrated to the high frequency realized volatility $\sigma^2_{corr}(\Delta t)$ as define in equation 9. This permits to consistently reduce the stochastic error of the hole expression (or alternatively the length of the calibration period) and have all the intraday partial volatilities automatically calibrated to our relevant target function (the realized volatility).

- we will compute moving averages of volatilities instead of variances because, being the unconditional distribution of volatilities less skewed than that of variances, the mean value of the former will better capture the center of the distribution.

Hence the bias correction formula for the normalized forecast of volatility of the jth horizon will read:

$$\hat{\sigma}_{corr}(\Delta t_j) = \frac{\sigma_{corr}(\Delta t)}{\hat{\sigma}_{calib}(\Delta t_j)}\hat{\sigma}(\Delta t_j) \tag{25}$$

where $\sigma_{corr}(\Delta t)$ is the realized volatility, $\hat{\sigma}(\Delta t_j)$ is the biased forecast of the jth partial volatility, $\hat{\sigma}_{calib}(\Delta t_j)$ an average of the later over the calibration period and $\hat{\sigma}_{corr}(\Delta t_j)$ the corrected partial volatility forecast.

4 Methodology and Data

4.1 The benchmark model

Since in this paper volatility forecast is primarily considered as a first step towards a superior VaR approach with respect to the RiskMetrics one, our natural benchmark model will be the J.P.Morgan's RiskMetrics volatility forecast. This model is basically an integrated GARCH model (an IGARCH(1,1)), with coefficients predefined by the J.P.Morgan technology:

$$\sigma_t^2 = 0.94\sigma^2(t - \Delta t_{ref}) + 0.06r^2(t - \Delta t_{ref}) \tag{26}$$

where Δt_{ref} is one working day. But to be fair in the comparison, we will compute RiskMetrics volatility not only from daily returns but also from high frequency ones rewriting equation (26) in the equivalent operator notation (Müller, 1999):

$$\sigma^2(t) = \mathrm{EMA}[\tau = 15.67\mathrm{wday}; r^2(t - \Delta t)] \tag{27}$$

4.2 Choice of the time scale

In the computation of the aformentioned models we also have the possibility to choose among different types of time scale: physical time, business time, ϑ-time. For the business time scale we will adopt the definition used in (Müller, 1999) in which the 49 weekend hours [12] are not completely dropped but they are compressed in one working-day hour (that in turns is 1,4 hours in physical time). The ϑ-time (Dacorogna *et al.*, 1993) scale is instead a more sophisticated business time scale designed to remove intra-day and intra-week seasonalities by compressing period of inactivity while expanding periods of higher activity. The business time scale will be employed for the realized volatility computation while the ϑ-time will be used in our forecasts.

4.3 Performance measures

To describe the performance measures lets first define the volatility forecasting error simply as the difference between the forecasted volatility $\widehat{\sigma}(t)$ and the realized one $\sigma(t)$

$$e_1(t) = \widehat{\sigma}(t) - \sigma(t) \tag{28}$$

[12] From Friday 8 pm GMT to Sunday 9 pm GMT.

But in presence of holes, since we average convex transformation (the square root of the sum of returns) of a very wide range of values, the mean of the realized volatility tends to underestimate the true one (that is the one that we would obtain if we had information on the behaviour of the process during holes). To overcome this problem we will also construct measures of performance based on the variance error

$$e_2(t) = \widehat{\sigma^2}(t) - \sigma^2(t) \ . \tag{29}$$

Then using these two definitions in theory it would be possible to construct four types of performance measures depending on whether we take the sum of the square or the sum of the absolute value of each type of error. But due to the well documented non-existence of the 4^{th} moment (Müller *et al.*, 1998) of many financial time series we won't compute the root mean square forecasting error of variances (RMSFE($\hat{\sigma}^2$)); while the remain three performance measure will be analyzed, namely:

1. the root mean square forecast error of volatility RMSFE($\hat{\sigma}$)

$$\text{RMSFE}(\widehat{\sigma}) = \sqrt{\text{E}(\widehat{\sigma}(t) - \sigma(t))^2} = \sqrt{\frac{\sum_t (\widehat{\sigma}(t) - \sigma(t))^2}{\widehat{N}}} = \sqrt{\frac{\sum_t e_1(t)^2}{\widehat{N}}} \tag{30}$$

2. the mean absolute prediction error of volatility MAPE($\hat{\sigma}$)

$$\text{MAPE}(\widehat{\sigma}) = \text{E}|\widehat{\sigma}(t) - \sigma(t)| = \frac{\sum_t |\widehat{\sigma}(t) - \sigma(t)|}{\widehat{N}} = \frac{\sum_t |e_1(t)|}{\widehat{N}} \tag{31}$$

3. the mean absolute prediction error of squared volatility MAPE($\hat{\sigma}^2$)

$$\text{MAPE}(\widehat{\sigma}^2) = \text{E}|\widehat{\sigma}^2(t) - \sigma^2(t)| = \frac{\sum_t |\widehat{\sigma}^2(t) - \sigma^2(t)|}{\widehat{N}} = \frac{\sum_t |e_2(t)|}{\widehat{N}} \tag{32}$$

where $\widehat{N}$ is the number of one step ahead forecasts.

In addition to that forecasting measures, we will compute non parametric tests that evaluate the ability of the model to forecast the direction of the change of volatility (Dacorogna *et al.*, 1998). Before doing that we need to define quantities that can take positive and negative values contrary to the volatilities which are positive definite quantities; these quantities with sign are the forecasting signals

$$s_f(t) = \hat{\sigma}(t) - \sigma(t) \tag{33}$$

and the real signal

$$s_r(t) = \sigma(t + \Delta t) - \sigma(t) \tag{34}$$

Now we can define:

1. the direction quality index that expresses the fraction of correct sign prediction

$$Q_d = \frac{\mathcal{N}(\{\hat{\sigma}(t)|s_f \cdot s_r > 0\})}{\mathcal{N}(\{\hat{\sigma}(t)|s_f \cdot s_r \neq 0\})} \tag{35}$$

   where $\mathcal{N}$ is a counter function.

2. the realized potential (that combine the size of the movements with the direction quality)

$$Q_r = \frac{\sum \text{sign}(s_f \cdot s_r)|s_r|}{\sum |s_r|} \tag{36}$$

3. the comparison of the absolute error of a model to a benchmark model

$$Q_f = 1 - \frac{\sum |s_r - s_f|}{\sum |s_r - s_f^{\text{benchmark}}|} \tag{37}$$

here the benchmark is represented by the last available observation of the realized volatility.

4.4 Data

1. Stock index prices for DJ, NIKKEI, DAX and MIB30 defined as

$$x^s(t_j) \equiv \ln p^s(t_j) \tag{38}$$

where $p^s(t_j)$ is the price of the index s (with $s = 1 \ldots 4$) at time j and $x^s(t_j)$ the sequence of unequally spaced logarithmic prices.

2. Foreign exchange rates for USD/DEM, USD/JPY, JPY/DEM, USD/ITL defined as the arithmetic mean of the logarithmic bid and ask quotations (p_{bid} and p_{ask} respectively)

$$x^e(t_j) \equiv [\ln p^e_{ask}(t_j) + \ln p^e_{bid}(t_j)]/2 \tag{39}$$

with $e = 1 \ldots 4$ being the index for the four exchange rates examined

5 Results

In order to assess the effects of the different time resolutions of the realized volatility on the accuracy of the forecasts, we present the results of the forecasting performances for different measures of the realized volatility (i.e. realized volatility computed from different return intervals). In the analysis of the results two separately groups of performance measures are considered: those based on a loss function of the prediction error ((RMSFE($\hat{\sigma}$), MAPE($\hat{\sigma}$) and MAPE($\hat{\sigma}^2$)) that can be generically called global measure; and those based on local forecasting signal (Q_d, Q_r and Q_f) termed local measure.

The impact on the global measures of the choice of the grid-size with which the realized volatility is computed is quite significant (see Table 1). In general for all the exchange rate examined, the use of a finer grid for the realized volatility leads itself to a considerable improvement of the global measures of forecasting performances of both models. For example, passing from a grid-size of the realized volatility of 2-hour to a 5-minute one permits a reduction of the prediction error measures of the order of 20-30% for both models! For the most traded currencies for which it is rewarding to go below the 5-minute return intervals, the improvement could be even greater (up to 40%). These results clearly demonstrate the usefulness of employing high frequency data in the construction of a low-noise target function, that allows better forecasts. Needless to say that the improvement with respect to the volatility computed with 1-day returns is even much higher. Moreover the EMA-HAR models are, at any grid, always 5-10% better than the RiskMetrics computed with the same grid. This fact indicates that it is advisable to employ many different horizons (ranging from very short to very long time intervals) in forecasting the volatility. Therefore, it is evident that combining a less noisy definition of volatility with a many horizons forecasting models leads to a substantial improvement of the standard volatility forecasts currently used.

For what concerns the local measures the results seems to be some what reversed. Though the impact of the different grid-size of the realized volatility does not considerably affect the local measures, it seems that increasing the grid slightly worse those measures. We claim that such finding is simply an artificial result of the fact that, in average, a very noisy process appear to have a behaviour fairly similar to that of strong mean reverting ones (see figure 1). This means that a slow moving forecast that always predict a reversal to the long term mean is most of the time correct, even without having any real forecasting power of the target process. When, instead, using finer grids, the level of noise reduces, this mean reverting behaviour disappears and it becomes progressively more difficult to correctly forecast the sign of the next direction. Nevertheless, also for these kind of performance measures, the EMA-HAR model is most of the times superior (though sometime only slightly) to the RiskMetrics one.

6 Outlook and Conclusions

6.1 Outlook

In our opinion the direction for future research is twofold: (i) a better understanding and possibly modeling of the volatility scaling behaviour at high frequency, (ii) extension of this type of approach to the multivariate case. The former direction, will imply a deeper study of the presence and nature of the short term autocorrelation which in turns would allow to build a superior bias correction method of the realized volatility (unbiased and with better dynamical behaviour). The latter, which could have far reaching implications for risk management, extends the nature of observable variables also to covariances and correlations. This would permit, on the one hand, to considerably improve the standard RiskMetrics VaR and on the other hand to employ much more sophisticated model even for highly diversified portfolio, since it dramatically reduces the difficulties arising in multivariate GARCH-type models.

A possible way to define a high frequency correlation based on time series operators capable of solving the problem of asynchronicity of ticks, could be the following: adopting the O&A approach (Dacorogna and Ballocchi, 1996) we can define a real time correlation coefficient in term of repeated applications of EMA operators. Using the differential operator defined in (19), we can first construct the generic time series $\tilde{r}_h(t)$ as:

$$\tilde{r}_h(t) \equiv D[\tilde{\Delta t}, x] = x(t) - \text{EMA}[\tilde{T}', \tilde{n}'; x_h](t) \tag{40}$$

then we could define the real time covariance [13] $\tilde{\sigma}_{hk}(t)$ between two generic series x_h and x_k as:

$$\tilde{\sigma}_{hk}(t) = \text{MA}[\frac{\tilde{T}}{2}, \tilde{n}; \tilde{r}_h \cdot \tilde{r}_k] \tag{41}$$

and then the real time correlation $\tilde{\rho}_{hk}(t)$ as:

$$\tilde{\rho}_{hk}(t) \equiv \frac{\text{MA}[\frac{\tilde{T}}{2}, \tilde{n}; \tilde{r}_h \cdot \tilde{r}_k]}{\sqrt{\text{MA}[\frac{\tilde{T}}{2}, \tilde{n}; \tilde{r}_h^2] \cdot \text{MA}[\frac{\tilde{T}}{2}, \tilde{n}; \tilde{r}_k^2]}} = \frac{\tilde{\sigma}_{hk}(t)}{\sqrt{\tilde{\sigma}_k \cdot \tilde{\sigma}_h}} \ . \tag{42}$$

6.2 Conclusions

The recognition that, using high frequency data, the volatility could be usefully treated as an observable variable, open entirely new avenues of studies for VaR computation. First it allows a much less noisy measurements of the target function of volatility forecast. Having a target function not contaminated with high level of noise permits to better extract the real underlying signal and hence improves the forecasting performance of any given forecasting model. In fact, as we showed, the use of high frequency data in the computation of the realized volatility can dramatically reduce the prediction error of the forecast. Second it allows the use of more sophisticated dynamic models that can be directly fitted to the realized volatility without having to rely on the much more complicated estimation procedures needed

[13] Assuming $E[\tilde{x}_h] = E[\tilde{x}_k] = 0$.

when the volatility is assumed to be unobserved. In this paper a new kind of autoregressive model, the heterogeneous autoregressive model with exponential moving average (EMA-HAR), has been applied demonstrating the utility of employing many different horizons ranging from very long to very short time scale. Third they promise to be extremely useful in multivariate cases which have been proven to be difficult to analyze with multivariate versions of ARCH-family models.

Nevertheless those potentially enormous advantages come only at the cost of the necessity to use high frequency data (of the order of few minutes); this often means having to deal with microstructure frictions, data gaps and other pitfalls we typically encounter at very short time scale. Those distortions give rise to anomalous scaling behaviour of the volatility (via a non negligible autocorrelation of returns) which would bias a correct measurement of the realized volatility. In order to overcome such difficulties we propose a bias correction procedure that, though not perfect, allows the use of arbitrary short interval returns in the computation of the realized volatility. This bias correction has proven to be necessary to really reduce the stochastic noise of the realized volatility to a level at which the volatility can be effectively treated as an observable quantity.

References

Andersen T. G. and Bollerslev T., 1998, *Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts*, International Economic Review, **39**, 885–905.

Andersen T. G., Bollerslev T., Diebold F. X., and Labys P., 2001, *The Distribution of Realized Exchange Rate Volatility*, Journal of the American Statistical Association, **96**, 42–55.

Blair B., Poon S.-H., and Taylor S. J., 1999, *Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High Frequency Index Returns*, FEN-List Econometric Abstracts, Working Paper Series, **4**(23), 1–25.

Bollerslev T. and Domowitz I., 1993, *Trading patterns and prices in the interbank foreign exchange market*, The Journal of Finance, **48**, 1421–1443.

Bouchaud J.-P., 1998, *Elements for a Theory of Financial Risks*, To appear in 'Order, Chance and Risk', Les Houches (March 1998), to be published by Springer/EDP Sciences, 1–22.

Dacorogna M. M. and Ballocchi G., 1996, *Design of the Correlation Service for Interest Rate Markets*, Internal document MMD.1996-11-16, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Dacorogna M. M., Müller U. A., Nagler R. J., Olsen R. B., and Pictet O. V., 1993, *A geographical model for the daily and weekly seasonal volatility in the FX market*, Journal of International Money and Finance, **12**(4), 413–438.

Dacorogna M. M., Müller U. A., Olsen R. B., and Pictet O. V., 1998, *Modelling Short-Term Volatility with GARCH and HARCH Models*, published in "Nonlinear Modelling of High Frequency Financial Time Series" edited by Christian Dunis and Bin Zhou, John Wiley, Chichester, 161–176.

Deparis S., 1999, *Study of the Empirical Properties of Implied Volatility and its Relations with Historical Volatility*, Internal document SDE.1999-09-17, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Flood M. D., 1994, *Market structure and inefficiency in the foreign exchange market*, Journal of International Money and Finance, **13**(2), 131–158.

Goodhart C. A. E., 1989, *'News' and the foreign exchange market*, Proceedings of the Manchester Statistical Society, 1–79.

Goodhart C. A. E. and Figliuoli L., 1991, *Every minute counts in financial markets*, Journal of International Money and Finance, **10**, 23–52.

Guillaume D. M., Dacorogna M. M., Davé R. D., Müller U. A., Olsen R. B., and Pictet O. V., 1997, *From the Bird's Eye to the Microscope: A Survey of New Stylized Facts of the Intra-Daily Foreign Exchange Markets*, Finance and Stochastics, **1**, 95–129.

Müller U. A., 1993, *Statistics of variables observed over overlapping intervals*, Internal document UAM.1993-06-18, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Müller U. A., 1999, *Volatility Computed by Time Series Operators at High Frequency*, Internal document UAM.1999-09-01, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Müller U. A., Corsi F., Dacorogna M. M., and Zumbach G. O., 2000, *Modeling the Bias of Realized and Predicted Volatility*, Internal document UAM.2000-01-11, Olsen & Associates, Seefeldstrasse 233, 8008 Zürich, Switzerland.

Müller U. A., Dacorogna M. M., and Pictet O. V., 1998, *Heavy tails in high-frequency financial data*, published in the book "A practical guide to heavy tails: Statistical Techniques for Analysing Heavy Tailed Distributions", edited by Robert J. Adler, Raisa E. Feldman and Murad S. Taqqu and published by Birkhäuser, Boston 1998, 55–77.

Roll R., 1984, *A simple implicit measure of the effective bid-ask spread in an efficient market*, The Journal of Finance, **39**(4), 1127–1139.

Taylor S. J. and Xu X., 1997, *The incremental volatility information in one million foreign exchange quotations*, the Journal of Empirical Finance, **4**(4), 317–340.

Zumbach G. O. and Müller U. A., 2001, *Operators on Inhomogeneous Time Series*, International Journal of Theoretical and Applied Finance, **4**(1), 147–178.