Overlaying Time Scales in Financial Volatility Data

Eric Hillebrand

Louisiana State University*

September 18, 2004

Abstract

Apart from the well-known, high persistence of daily financial volatility data, there is also a short correlation structure that reverts to the mean in less than a month. We find this short correlation time scale in six different daily financial time series and use it to improve the short-term forecasts from GARCH models. We study different generalizations of GARCH that allow for several time scales. On our holding sample, none of the considered models can fully exploit the information contained in the short scale. Wavelet analysis shows a correlation between fluctuations on long and on short scales. Models accounting for this correlation as well as long memory models for absolute returns appear to be promising.

JEL-codes: C22, C51.

Key words: GARCH, volatility persistence, spurious high persistence, long memory, fractional integration, change-points, wavelets, time scales

*Department of Economics, 2126 CEBA Bldg., Baton Rouge, LA 70803, phone: (225) 578-3795, fax (225) 578-3807, e-mail erhil@lsu.edu

Contents

1 Introduction

GARCH usually indicates high persistence or slow mean reversion when applied to financial data (e.g., Engle and Bollerslev 1986, Ding, Granger, and Engle 1996, Engle and Patton 2001). This finding corresponds to the visually discernable volatility clusters that can be found in the time series of squared or absolute returns from financial objects. The presence of change-points in the data, however, prevents GARCH from reliably identifying the long time scale that is associated with the high persistence (Diebold 1986, Lamoureux and Lastrapes 1990, Mikosch and Starica 2004, Hillebrand 2004a). Recently, the possibility of the presence of several correlation time scales in financial volatility data has been discussed (Fouque et al. 2003, Chernov et al. 2003, Gallant and Tauchen 2001).

Using six different time series of daily financial data, we show that ARCH, GARCH, and Integrated GARCH (IGARCH) fail to capture a short correlation structure in the squared returns with a time scale of less than 20 days. This serial correlation with low persistence can be used to improve the short-term forecast error from GARCH models.

Several models that are being proposed to capture different time scales are discussed and estimated on our data set. Unfortunately, none is able to fully exploit the information contained in the short scale in our data set. In correspondence to the results of Ding and Granger (1996), ARFIMA estimation of absolute, as opposed to squared, returns yields a consistent improvement of the 1-day forecast on our holding sample. Wavelet transformation of the volatility time series from our data set reveals that there is correspondence between long and short correlation scales. This supports the findings of Müller et al. (1997) and Ghashgaie et al. (1996). The HARCH model of Müller et al. (1997), which allows for correlation between long and short scales, indeed uses the short scale information for the stock market part of our data set best.

In Section 2, we introduce our data set and discuss the ARCH, GARCH, and IGARCH results, which turn out to be very typical for financial data. Under number 2.3, we then identify the short correlation scale that is being missed by GARCH, using spectral methods as well as ARMA estimation, which provides a benchmark for forecast improvement over GARCH. In Section 3, we segment our time series using recent results for change-point detection in volatility data. GARCH is then estimated locally on the segments. As change-points cause GARCH to show spurious high persistence, the idea is that accounting for change-points may reveal the short scale within the segments. This is indeed the case, but does not help to improve forecasts for our holding sample. Section 4 discusses and estimates an array of models designed to capture several time scales in volatility. We begin with fractionally integrated models and estimate ARFIMA and FIGARCH on our data set. ARFIMA estimation of absolute returns yields a consistent improvement of the 1-day ahead forecast on our holding sample. Next, a two-scale version of GARCH by Engle and Lee (1999) is estimated. Under number 4.3, we demonstrate the correspondence between long and short scales with wavelet analysis. The HARCH model of Müller et al. (1997) is then estimated. We conclude that, on our data set, none of the considered models captures the short scale as well as the simple ARMA fit to the residual in Section 2.3. For the stock market section of our data, HARCH yields the relatively largest improvements in the short-term forecast errors.

2 Integrated Volatility

2.1 Data

We use an array of daily financial data. Four series measure the U.S. stock market: the S&P 500 index (**sp**) and the Dow Jones Industrial Average (**dj**), obtained from Datastream, as well as the CRSP equally weighted index (**cre**) and the CRSP value-weighted index (**crv**). Next, we consider the exchange rate of the Japanese Yen against the U.S. Dollar (**yd**, Datastream series BBJPYSP) and the U.S. federal funds rate (**ffr**). These series cover the period 4 Jan 1988 through 31 Dec 2003 and contain 4,037 observations, except for the Yen/Dollar exchange rate, which is recorded in London and therefore contains only 3,991 observations.

2.2 ARCH and GARCH

In the ARCH(p) model (Engle 1982) the log-returns of a financial asset are given by

$$r_t = m_t + \varepsilon_t \ , t = 1, \ldots, T, \tag{1}$$

$$\varepsilon_t|_{\mathcal{F}_{t-1}} \sim \mathcal{N}(0, h_t), \tag{2}$$

$$h_t = \sigma^2 + \sum_{i=1}^{p} \alpha_i(\varepsilon_{t-i}^2 - \sigma^2), \tag{3}$$

where $m_t = \mathbb{E}_{t-1} r_t$ is the conditional mean function, $\mathcal{F}_t$ denotes the filtration that models the information set, and $\sigma^2 = \omega/(1 - \sum_{i=1}^{p} \alpha_i)$ is the unconditional mean with ω being a constant. For the conditional mean m_t different models are used, for example ARMA processes, regressions on exogenous variables, or simply a constant. Since we focus on the persistence of the conditional volatility process, we will use a constant in this study.

The GARCH(p,q) model (Bollerslev 1986) adds the term $\sum_{i=1}^{q} \beta_i(h_{t-i} - \sigma^2)$ to the conditional variance equation (3). The unconditional mean becomes $\sigma^2 = \omega/(1 - \sum_{i=1}^{p} \alpha_i - \sum_{i=1}^{q} \beta_i)$. Empirically, GARCH models often turn out to be more parsimonous than ARCH models since these often need high lag orders to capture the dynamics of economic time series.

There are different ways to measure persistence in the context of GARCH models, as discussed, for example, in Engle and Patton (2001) and Nelson (1990). The most commonly used measure is the sum of the autoregressive parameters,

$$\lambda := \sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i. \tag{4}$$

The parameter λ can be interpreted as the fraction of a shock to the volatility process that is carried forward per unit of time. Therefore, $(1 - \lambda)$ is the fraction of the shock that is washed out per unit of time. Hence, $1/(1 - \lambda)$ is the average time needed to eliminate the influence of a shock. The closer λ is to unity, the more persistent is the effect of a change in h_t. The stationarity condition is $\lambda < 1$. A stationary but highly persistent process returns slowly to its mean, a process with low persistence reverts quickly to its mean. Different persistences therefore imply different times of mean reversion. In this study, the term "time scales" is understood in this sense.

Estimations of the conditional volatility process with GARCH models usually indicate high persistence (Engle and Bollerslev 1986, Bollerslev et al. 1992, Bollerslev and Engle 1993, Bollerslev et al. 1994, Baillie et al. 1996, Ding et al. 1993, Ding and Granger 1996, Andersen and Bollerslev 1997). The estimated sum λ of the autoregressive parameters is often close to one. In terms of time scales, $\lambda = 0.99$ corresponds to 100 days or 5 months mean reversion time for daily data, whereas $\lambda = 0.999$ corresponds to 1000 days or 4 years. Therefore, it is difficult to identify the long scale numerically and interpret it economically.

Engle and Bollerslev (1986) suggested the integrated GARCH, or IGARCH model to reflect the empirical fact of high persistence. In the IGARCH model, the likelihood function of the GARCH model is maximised subject to the constraint that $\lambda = 1$. In the IGARCH(1,1) specification this amounts to the parameter restriction $\beta = 1 - \alpha$. The estimated conditional volatility process is not stationary, a shock has indefinite influence on the level of volatility.

TABLE 1 ABOUT HERE.

Table 1 compares GARCH, ARCH, and IGARCH estimations on our data set. Within the class of model orders up to GARCH(5,5), the Bayes Information Criterion favors GARCH(1,1) for all series. Contrary to that, it indicates lag orders between 8 and 20 for ARCH estimations. For GARCH estimations, the sum of the autoregressive parameters is almost one for all series, with the CRSP equally weighted index showing the lowest persistence with a mean reversion time of 27 days and the volatility of the federal funds rate indicating non-stationarity.

The ARCH estimations show a very different picture of persistence, λ is of the order of 0.80 to 0.85 for the stock market, 0.5 for the yen-dollar exchange rate, and still close to one for the federal funds rate. This indicates that another, shorter scale of the order of a few days may be present in the data.

The last three rows of every panel show the mean absolute errors for a one-day, 20-day, and 60-day forecast of the conditional volatility process when compared to the squared returns of the holding sample Oct 7., 2003 through Dec 31., 2003. The mean absolute errors of the GARCH(1,1) forecast are taken as a benchmark. The forecast errors of other models are expressed as percentages of the GARCH(1,1) error. In terms of forecast performance on the holding sample considered here, GARCH dominates ARCH for the stock market data. For the federal funds rate and the exchange rate, ARCH seems to be a better forecast model.

One might suspect that the stock market volatility is in fact integrated and that GARCH(1,1) with λ close to one reflects this fact better, thereby providing better forecasts than ARCH. This interpretation, however, cannot hold as IGARCH(1,1) forecasts are worse than GARCH across all series and forecast horizons, with the single exception of the 1-day forecast for the exchange rate. Further, IGARCH performs worse than ARCH for the 20-day and 60-day forecasts of stock market volatility but better on the single day horizon. This is a counterintuitive result as the non-stationarity of IGARCH should reflect the long-term behavior better than the short term.

There are only two instances in Table 1 where the term structure of the mean absolute forecast error has the intuitively expected pattern of a monotonous increase with the forecast horizon: the IGARCH estimation of the equally weighted CRSP index and of the federal funds rate. In most cases, the forecast performance becomes better with increasing horizon.

2.3 The Missed Short Scale

The GARCH model aims to capture the entire correlation structure of the volatility process ε_t^2 in the conditional volatility process h_t. Therefore, we can define the "residual" $\nu_t := \varepsilon_t^2 - h_t$. This residual is white noise. From the distribution assumption (2) of the GARCH model, which can be written as $\varepsilon_t^2 = \eta_t^2 h_t$, where η_t is a standard normal random variable, we have that

$$\mathbb{E}\nu_t = \mathbb{E}(\mathbb{E}_{t-1}(\varepsilon_t^2 - h_t)) = \mathbb{E}(\mathbb{E}_{t-1}(\eta_t^2 h_t - h_t)) = 0, \tag{5}$$

$$\mathbb{E}\nu_t\nu_s = \mathbb{E}((\eta_t^2 - 1)(\eta_s^2 - 1)h_t h_s) = \begin{cases} 0 & \text{for } t \neq s, \\ 2\mathbb{E}h_t^2 & \text{for } t = s. \end{cases} \tag{6}$$

The latter is shown to exist in Theorem 2 of Bollerslev (1986). Therefore, if GARCH is an accurate model for financial volatility, the estimated residual process $\hat{\nu}_t = \hat{\varepsilon}_t^2 - \hat{h}_t$ should not exhibit any serial correlation.

FIGURES 1, 2, AND 3 ABOUT HERE.

Figures 1 to 3 show the estimated averaged periodograms of the residuals of the GARCH estimations in Table 1. The averaged periodogram is estimated by subsampling with a Tukey-Hanning window of 256 points length allowing for 64 points overlap. A Lorentzian spectrum model

$$h(w) = a + b/(c^2 + w^2), \tag{7}$$

is fitted to the periodogram, where w denotes the frequencies and (a, b, c) are parameters. The average mean reversion time is obtained as a function of the parameter c.[1] Except for the CRSP equally weighted index and the federal funds rate, the

[1] The relation between the mean reversion time $1/c$ estimated from the Lorentzian and the mean reversion time $1/(1-\lambda)$ from the GARCH model was found in simulations as $1/c = -86.74 +$

series show strong serial correlation in the residual. The persistence time scale is of the order of five days.

TABLE 2 ABOUT HERE.

This finding is confirmed when ARMA models are fitted to the residuals. Table 2 reports the parameter estimates and the mean absolute errors of the forecasts for the volatility process ε_t^2 when the serial correlation in the residual ν_t is taken into account. The errors are stated as percentages of the errors from the simple GARCH estimations in Table 1. The first order autoregressive parameters are estimated around 0.8, corresponding to the 5-day mean reversion time scale found from the Lorentzian model and corresponding to the estimated persistence from the ARCH models in Table 1. Note that the forecast errors can be substantially reduced by correcting for this correlation, which is exogenous to the GARCH model.

We are now able to explain the counterintuitive pattern of forecast errors from the GARCH model. The fact that the short correlation scale is not captured leads to inflated errors on short forecast horizons. The correction for the short scale therefore reduces the errors at the 1-day and at the 20-day horizon much more than at the 60-day horizon.

In the following sections, we will explore different ways to account for the short scale in GARCH-type models. We will evaluate the forecast performance of these models against the benchmark of the exogenous correction for the short scale in Table 2.

$61.20 \cdot 1/(1-\lambda), \quad R^2 = 0.93$. A motivation of the Lorentzian for GARCH models and the details of the simulations are available upon request.

3 Local GARCH Estimation

The apparent integrated or near-integrated behavior of financial volatility may be an artefact of changing parameter regimes in the data. Several authors have studied this phenomenon in a GARCH framework with market data and in simulations (Diebold 1986, Lamoureux and Lastrapes 1990, Hamilton and Susmel 1994, Mikosch and Starica 2004). Hillebrand (2004a) shows that if a GARCH model is estimated globally on data that were generated by changing local GARCH models, then the global λ will be estimated close to one, regardless of the value of the data-generating λ's within the segments. In other words, a short mean reversion time scale in GARCH plus parameter change-points exhibits statistical properties similar to a long mean reversion time scale.

TABLES 3, 4, AND 5 ABOUT HERE.

Tables 3 and 4 show the estimation of GARCH(1,1) models on segmentations that were obtained by using a change-point detector proposed by Kokoszka and Leipus (1999). The detector statistic is given by

$$U(t) = \sqrt{T}\frac{t(T-t)}{T^2}\left(\frac{1}{t}\sum_{j=1}^{t} r_j^2 - \frac{1}{T-t}\sum_{j=t+1}^{T} r_j^2\right). \tag{8}$$

The estimator of the single change-point in the sample of T observations of the squared returns r_t^2 is obtained as

$$\hat{\tau} = \min\left\{\tau \;:\; |U(\tau)| = \max_{t\in\{1,\dots T\}} |U(t)|\right\}. \tag{9}$$

Kokoszka and Leipus (1999) show that the statistic $U(t)/\sigma$ converges in distribution to a standard Brownian bridge, where $\sigma^2 = \sum_{j=-\infty}^{\infty} \mathrm{cov}(r_t^2, r_{t+j}^2)$. Therefore, confidence intervals can be calculated from the distribution of the maximum of a Brownian bridge. We follow Andreou and Ghysels (2002) and use the VARHAC estimator of den Haan and Levin (1997) for σ. Change-points at a significance level of 0.05 or less are considered in Tables 3 and 4.

In fact, the average estimated λ's within the segments indicate lower persistence than the global estimates. The implied mean reversion time scales range from about 3 days for the yen per dollar exchange rate to about 18 days for the S&P 500 index.

However, while the approach uncovers a short scale, its forecast success is disappointing. Table 5 shows the mean absolute error from forecasts of the GARCH(1,1) model of the last segment. Two offsetting effects are at work. While we can expect the local estimation approach to capture the short-run dynamics of the series better, the shorter sample sizes increase the estimation error. The MAEs of Table 5 are again stated as percentages of the global forecast errors and show that in terms of forecast success, the global approach yields better results on our holding sample.

4 Models of Multiple Time Scales

In the local estimations in Section 3 (Tables 3 and 4), we have seen that accounting for change-points and estimating GARCH on segments reduces the estimated persistence. The mean reversion time estimated from the average persistence across segments stays below 20 days as opposed to the global estimation in Table 1, where four out of six volatility time series display a mean reversion time of more than 100 days. These findings suggest the interpretation that financial volatility displays low persistence with occasional structural breaks, which lead to apparent high persistence in estimations that ignore the change-points. Another possible interpretation is that two or more persistence time scales influence the volatility processes simultaneously and continuously in daily data. We study this interpretation in Sections 4.1 and 4.2. A third possible interpretation is that the scales remain only for some limited time and then die out and are replaced by different scales. They may also influence each other. Section 4.3 deals with this case.

A short mean reversion time scale in financial volatility has been found in many different recent studies, mostly in the context of stochastic volatility models. Gallant and Tauchen (2001) estimate a stochastic volatility model with two different volatility drivers. Estimating their model for daily returns on the Microsoft stock, one driver assumes high persistence, the other assumes low persistence. Fouque et al. (2003) and Chernov et al. (2003) propose and discuss multi-scale stochastic volatility models. LeBaron (2001) simulates a stochastic volatility model where the volatility process is an aggregate of three different persistence time scales and shows that the simulated time series display long memory properties. This corresponds to Granger's (1980) finding that aggregation over several scales implies long memory behavior. These findings suggest that both, long and short scale act continuously and simultaneously in the data.

4.1 Fractional Integration

Assume that the time series x_t, when differenced d times, gives a time series y_t that has an ARMA representation. Then, the time series x_t is called integrated of order d, denoted $x_t \sim I(d)$. Fractionally integrated time series, where d is not an integer, are a discretized version of fractional Brownian motion.[2] Granger (1980) and Granger and Joyeux (1980) show that the autocorrelation function of a fractionally integrated x_t is given by

$$\rho_k = \frac{\Gamma(1-d)}{\Gamma(d)} \frac{\Gamma(k+d)}{\Gamma(k+1-d)} \equiv A_d k^{2d-1},$$

for $d \in (0, 1/2)$, where A_d is a constant. The sum over the autocorrelations does not converge, so that it is a suitable model for long memory. If $d \in (1/2, 1)$, the process has infinite variance and is therefore non-stationary but still mean reverting. If $d > 1$, the process is non-stationary and not mean reverting. For $d \in (-1/2, 0)$, the

[2] Let $Y_t = \int_{-\infty}^{t} (t-s)^{H-1/2} dW(s)$ be fractional Brownian motion, where $W(t)$ is standard Brownian motion. Then, the relation between the Hurst coefficient H and the fractional integration parameter d is given by $H = d + 1/2$, see Geweke and Porter-Hudak (1983).

process is anti-persistent.

Granger (1980) shows that aggregation of a large number of processes with different short mean reversion times yields a fractionally integrated process. In particular, he shows that the process

$$x_t = \sum_{j=1}^{N} y_{j,t},$$

where the $y_{j,t}$ are AR(1) models

$$y_{j,t} = \alpha_j y_{j,t-1} + \varepsilon_{j,t},\ j = 1, \ldots, N,$$

with $\varepsilon_{j,t}$ independent white noise processes, is a fractionally integrated process with order d depending on the statistical properties of the α_j. This is particularly appealing for economic time series which are often aggregates. The surprising result is that an aggregate of many short scales can exhibit long memory properties.

Geweke and Porter-Hudak (1983) provide the standard estimation method for d. The estimator is obtained as the negative of the estimate $\hat{b}$ of the slope in the regression

$$\log I(w_j) = a + b \log\left(4 \sin^2 (w_j/2)\right) + \varepsilon_j,$$

where $I(w_j)$ is the periodogram at the harmonic frequencies $w_j = j\pi/T$ and T is the sample size.

TABLE 6 ABOUT HERE.

Most financial time series exhibit significant long memory in the sense of the Geweke and Porter-Hudak (1983) estimator. Table 6 shows the estimated d for the squared and the absolute returns of the time series considered here. The estimates are significantly larger than zero and below 1/2, indicating stationary long memory. The only exception are the absolute returns of the CRSP equally weighted index, which may be non-stationary but still mean reverting. These findings may indicate the presence of multiple overlaying time scales in the data in the sense of Granger

(1980). The converse conclusion that a fractionally integrated process decomposes into several processes with finite persistence structures does not hold, however.

TABLE 7 ABOUT HERE.

For the absolute returns calculated from the time series considered here, we estimate an ARFIMA model using Ox (Doornik 2002) and the ARFIMA package for Ox (Doornik and Ooms 1999). We choose to model absolute returns since they exhibit an even richer correlation structure than squared returns (Ding and Granger 1996) and a model for absolute returns allows to forecast squared returns as well. The ARFIMA(p,d,q) model for the time series y_t is defined as

$$\Phi(L)(1-L)^d y_t = \Psi(L)\eta_t, \tag{10}$$

where L is the lag operator, η_t is white noise with mean zero and variance σ^2,

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p,$$

is the autoregressive lag polynomial, and

$$\Psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \ldots + \psi_p L^p,$$

is the moving-average lag polynomial. We assume that all roots of the lag polynomials are outside the unit circle. The fractional differencing operator $(1-L)^d$, where $d \in (-0.5, 0.5)$, is defined by the expansion

$$(1-L)^d = 1 - d\,L + \frac{d(d-1)}{2!}\,L^2 - \frac{d(d-1)(d-2)}{3!}\,L^3 + \ldots \tag{11}$$

In estimations, a truncation lag has to be defined; we use 100 in this study.

We determine the ARFIMA lag orders p and q by significance according to t-values. Table 7 reports the estimation results and the forecast errors in percent of the GARCH(1,1) forecast error. The estimates of the fractional integration parameter d are highly significant across all time series. Also, on the 1-day horizon the forecast error is smaller than the forecast error of the GARCH(1,1) model across all time series. Except for the yen per dollar exchange rate, the forecast performance on the holding sample is worse than GARCH(1,1) for the 20-day and 60-day forecast horizons.

Baillie et al. (1996) proposed a GARCH model that allows for a fractionally integrated volatility process. The conditional variance equation (3) can be rewritten as an ARMA process for ε_t^2

$$[(1-\alpha(L)-\beta(L)]\varepsilon_t^2 = \omega + [1-\beta(L)]\nu_t, \tag{12}$$

where $\nu_t = \varepsilon_t^2 - h_t$. The coefficients $\alpha(L) = \alpha_1 L + \ldots + \alpha_p L^p$ and $\beta(L) = \beta_1 L + \ldots + \beta_q L^q$ are polynomials in the lag operator L. The order of the polynomial $[(1-\alpha(L)-\beta(L)]$ is $\max\{p,q\}$.

The residual ν_t is shown to be white noise in (5) and thus, (12) is indeed an ARMA process. If the process ε_t^2 is integrated, the polynomial $[(1-\alpha(L)-\beta(L)]$ has a unit root and can be written as $\Phi(L)(1-L)$ where $\Phi(L) = [(1-\alpha(L)-\beta(L)](1-L)^{-1}$ is of order $\max\{p,q\}-1$. Motivated by these considerations the fractionally integrated GARCH, or FIGARCH model defines the conditional variance as

$$\Phi(L)(1-L)^{\delta}\varepsilon_t^2 = \omega + [1-\beta(L)]\nu_t, \tag{13}$$

where $\delta \in (0,1)$.

TABLE 8 ABOUT HERE.

Table 8 reports the parameter estimates and mean absolute errors of the forecast from a FIGARCH(1,δ,1) model. The estimation was carried out using the GARCH 2.2 package of Laurent and Peters (2002) for Ox.

The interpretation of persistence in a FIGARCH model is difficult. The estimates of the parameter of fractional integration are highly significant across all series. Laurent and Peters (2002) note, however, that in the FIGARCH model the parameter δ

does not have the same interpretation as d in the ARFIMA model. The FIGARCH process is non-stationary if $\delta > 0$, thus our estimates indicate non-stationary rather than stationary long memory. Also, the autoregressive parameters α and β do not have a straightforward interpretation as persistence parameters. Their sums are substantially below one for all considered series but we cannot extract a persistence time scale from this as in the case of GARCH.

In summary, long memory models describe financial volatility data well and the parameter of fractional integration is usually highly significant. This may indicate that there are several persistence time scales present in the data. Fractionally integrated models do not allow, however, for an identification of the different scales.

4.2 Two-Scale GARCH

In the GARCH framework, Engle and Lee (1999) proposed a model that allows for two different overlaying persistence structures in volatility. According to (3), the conditional variance in the GARCH(1,1) model is given by

$$h_t = \sigma^2 + \alpha(\varepsilon_{t-1}^2 - \sigma^2) + \beta(h_{t-1} - \sigma^2).$$

Engle and Lee (1999) generalize the unconditional variance σ^2 from a constant to a time varying process q_t, which is supposed to model the highly persistent component of volatility:

$$h_t - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \beta(h_{t-1} - q_{t-1}), \tag{14}$$

$$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - h_{t-1}). \tag{15}$$

Then, $\lambda = \alpha + \beta$ can capture the short time scale. The authors show that the model can be written as a GARCH(2,2) process.

TABLE 9 ABOUT HERE.

Table 9 reports the estimation of the Engle and Lee (1999) model on the time series considered in this paper. The results show that there is a problem with assigning the two different scales to the correct driver: The sum $\lambda = \alpha + \beta$ is close to one for all series while the autoregressive parameter ρ, which is supposed to capture the slowly varying component, reflects the short scale. Apart from this identification problem, however, the results clearly show two different scales. The results for ρ indicate a short scale between 2 and 18 days. As before, the last three rows report the mean absolute forecast errors for the holding period as a fraction of the errors of the GARCH(1,1) forecast. Except for three instances at the 1-day horizon, the forecast errors are higher than those of the GARCH(1,1) model.

4.3 Wavelet Analysis and Heterogeneous ARCH (HARCH)

The presence of different time scales in a time series (or "signal") is a well-studied subject in the physical sciences. A standard approach to the problem is spectral analysis, that is, Fourier decomposition of the time series. The periodogram estimation carried out in Section 2.3 rests on such a Fourier decomposition. The time series $x(t)$ is assumed to be a stationary process and for ease of presentation, assumed to be defined on continuous time. Then, the process is decomposed into a sum of sines and cosines with different frequencies ω and amplitudes $\hat{x}(\omega)$:

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{x}(\omega) e^{i\omega t} d\omega. \tag{16}$$

The amplitudes $\hat{x}(\omega)$ are given by the Fourier transform of $x(t)$:

$$\hat{x}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt. \tag{17}$$

The periodograms in Figures 1 through 3 are essentially plots of the $\hat{x}(\omega)$ against the frequencies ω.

We saw that spectral theory can be used to identify the short scale. However, this result pertains to the entire sample period only. The integration over the time dimension in (17) compounds all local information into a global number, the amplitude $\hat{x}(\omega)$. The sinusoids $e^{i\omega t} = \cos \omega t + i \sin \omega t$ have support on all $t \in \mathbb{R}$. The coefficient in the Fourier decomposition (16), the amplitude $\hat{x}(\omega)$, has therefore only global meaning.

It is an interesting question, however, if there are time scales in the data that are of local importance, that is, influence the process for some time and then die out or are replaced by other scales. This problem can be approached using the wavelet transform instead of the Fourier transform (Mallat 1999). The wavelet transform decomposes a process into wavelet functions which have a frequency parameter, now called scale, and also a position parameter. Thereby, the wavelet transform results in coefficients which indicate a time scale *plus* a position in time where this time scale was relevant. Contrary to that, the Fourier coefficients indicate a globally relevant time scale (frequency) only.

A wavelet $\psi(t)$ is a function with mean zero:

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0,$$

which is dilated with a scale parameter s (corresponding to the frequency in Fourier analysis), and translated by u (this is a new parameter that captures the point in time where the scale is relevant):

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right).$$

In analogy to the Fourier decomposition (16), the wavelet decomposition of the process $x(t)$ is given by

$$x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Wx(u,s)\, \psi_{u,s}(t)\, du\, ds. \tag{18}$$

The coefficients $Wx(u,s)$ are given by the wavelet transform of $x(t)$ at the scale s and position u, which is calculated by convoluting $x(t)$ with the complex conjugate of the wavelet function:

$$Wx(u,s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt. \tag{19}$$

This is the analog to the Fourier transform (17).

Figures 4 and 5 plot the wavelet coefficients $Wx(u,s)$ against position $u \in \{1, \ldots, 4037\}$ in the time series and against time scale $s \in \{2, 4, \ldots, 512\}$ days. The wavelet function used here is of the Daubechies class with 4 vanishing moments (Mallat 1999, pp. 249 ff.) but other wavelet functions yield very similar results. The plots for the CRSP data are left out for brevity, they are very similar to Figure 4. These are three dimensional graphs; the figures present a bird's eye view of the coefficient surface. The color encodes the size of the coefficient $Wx(u,s)$: the higher the coefficient, the brighter its color.[3]

If there were a long scale and a short scale in the data, which influenced the process continuously and simultaneously, we would expect to see two horizontal ridges, one at the long and one at the short scale. According to the GARCH estimation in Table 1, the ridge of the long scale of the S&P500 should be located at $s = 1/(1-0.995) = 200$ days, in the case of the Dow Jones at $s = 1/(1-0.99) = 100$ days, and so on. From the results in Section 2.3, we would expect the ridge of the short scale to be located at around 5 days.

Figures 4 and 5 do not exhibit such horizontal ridges. Instead, there are vertical ridges and braid-like bulges that run from longer scales to shorter scales. A pronounced example is in the upper panel of Figure 5, where a diagonal bulge runs from observation 500 and the long end of the scale axis to observation 1250 and the short end of the scale axis. A high vertical ridge is located around observation 750. The corresponding event in the federal funds rate series is the period of high fluc-

[3]The color plots can be downloaded at

`http://www.bus.lsu.edu/economics/faculty/ehillebrand/personal/`.

tuations during the end of the year 1990 and early 1991, when the Federal Reserve phased out of minimum reserve requirements on non-transaction funds and the first gulf war was impending.

These patterns indicate that there is correlation between fluctuations with long mean reversion time and fluctuations with short mean reversion time. This correlation has been established by Müller et al. (1997). Ghashgaie et al. (1996) relate this finding to hydrodynamic turbulence, where energy flows from long scales to short scales. Müller et al. (1997) propose a variant of ARCH, called heterogeneous ARCH, or HARCH, to account for this correlation. The idea is to use long term fluctuations to forecast short term volatility. This is achieved by adding squared j-day returns to the conditional variance equation. That is, (2) is replaced by

$$h_t = \omega + \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{j} \varepsilon_{t-i} \right)^2 , \tag{20}$$

where all coefficients are positive. The number of possible combinations of n and j renders information criterion searches for the best specification infeasible, in particular since high j's are desirable to capture the long mean reversion scales. The stationarity condition is $\lambda = \sum_{j=1}^{n} j\alpha_j < 1$. Müller et al. (1997) also propose a GARCH-like generalization of HARCH, adding lagged values of h to (20). We refrain from using such a specification, since it incurs spurious persistence estimates due to change-points (Hillebrand 2004b).

We use $n = 5$, that is, we include five different returns at horizons 1, 20, 60, 100, and 250 days. Thereby, we nest the simple ARCH(1) specification but also let 1-month, 3-month, 5-month, and 1-year returns influence daily fluctuations. Table 10 reports the estimation results. Remarkable are the consistently high significance of 3-month returns and the low estimates of the persistence parameter λ. HARCH yields consistent improvements on the 1-day forecast horizon for the four stock market series, but performs badly on the federal funds rate and the exchange rate.

5 Conclusion

This study shows that apart from the well-known, high persistence in financial volatility, there is also a short correlation, or fast mean reverting time scale. We find it in six different daily financial time series: four capture the U.S. stock market, the federal funds rate and the yen-dollar exchange rate. This short scale has a mean reversion time of less than 20 days. An ARMA fit to the residual $\varepsilon_t^2 - h_t$ from a GARCH estimation improves the short-term forecast of the GARCH model and provides a benchmark.

We estimate several generalizations of GARCH that allow for multiple correlation time scales, including segmentation of the data and local GARCH estimation. Unfortunately, none of the considered models is able to fully exploit the information contained in the short scale and improve over the benchmark from the simple ARMA fit.

Wavelet analysis of the volatility time series reveals that there is correlation between fluctuations on long scales and fluctuations on short scales. The heterogeneous ARCH model of Müller et al. (1997), which allows for correlation of this kind, can exploit some of the short scale information from the stock market part of our data set.

Acknowledgments

This paper benefited from discussions with and comments from George Papanicolaou, Knut Solna, Mordecai Kurz, Caio Almeida, Doron Levy, and Carter Hill. Remaining errors of any kind are mine.

References

Andersen, T. G. and T. Bollerslev. 1997. Intraday Periodicity and Volatility Persistence in Financial Markets. Journal of Empirical Finance 4: 115–158.

Andreou, E. and E. Ghysels. 2002. Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics 17: 579–600.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen. 1996. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 74: 3–30.

Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31: 307–327.

Bollerslev, T., R. Y. Chou, and K. F. Kroner. 1992. ARCH Modeling in Finance: A Review of Theory and Empirical Evidence. Journal of Econometrics 52: 5–59.

Bollerslev, T. and R. F. Engle. 1993. Common Persistence in Conditional Variances. Econometrica 61(1): 167–186.

Bollerslev, T., R. F. Engle, and D. B. Nelson. 1994. GARCH Models. in: Engle, R. F. and D. L. McFadden (eds.) Handbook of Econometrics, Vol. 4, Elsevier: Amsterdam.

Chernov, M., Gallant, A.R., Ghysels, E. and G. Tauchen. 2003. Alternative models of stock-price dynamics. Journal of Econometrics 116: 225-257.

Den Haan, W. and A. Levin. 1997. A Practitioner's Guide to Robust Covariance Matrix Estimation. In: Maddala, G. and C. Rao (eds.) Handbook of Statistics Vol. 15: Robust Inference. North-Holland: Amsterdam.

Diebold, F. X. 1986. Modeling the persistence of conditional variances: A comment. Econometric Reviews 5: 51–56.

Ding, Z., C. W. J. Granger, and R. F. Engle. 1993. A Long Memory Property of Stock Market Returns and a New Model. Journal of Empirical Finance 1: 83–106.

Ding, Z. and C. W. J. Granger. 1996. Modeling Volatility Persistence of Speculative Returns: A New Approach. Journal of Econometrics 73: 185–215.

Doornik, J. A. 2002. Object-Oriented Matrix Programming Using Ox. 3rd ed. Timberlake Consultants Press: London. www.nuff.ox.ac.uk/Users/Doornik.

Doornik, J.A. and M. Ooms. 1999. A Package for Estimating, Forecasting and Simulating Arfima Models. www.nuff.ox.ac.uk/Users/Doornik.

Engle, R. F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50(4): 987–1007.

Engle, R. F. and T. Bollerslev. 1986. Modelling the Persistence of Conditional Variances. Econometric Reviews 5(1): 1–50.

Engle, R. F. and G. G. J. Lee. 1999. A long-run and short-run component model of stock return volatility. In: Engle, R. F. and H. White. (eds.) Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger. Oxford University Press.

Engle, R. F. and A. J. Patton 2001. What good is a volatility model? Quantitative Finance 1(2): 237–245.

Fouque, J. P., Papanicolaou, G., Sircar, K. R., and K. Sølna. 2003. Short timescale in S&P 500 volatility. Journal of Computational Finance 6: 1–23.

Gallant, A. R. and G. Tauchen. 2001. Efficient Method of Moments. mimeo. http://www.unc.edu/∼arg.

Geweke, J. and S. Porter-Hudak. 1983. The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis 4(4): 221–238.

Ghashgaie, S., Breymann, W., Peinke, J., Talkner, P. and Y. Dodge. 1996. Turbulent cascades in foreign exchange markets. Nature 381: 767–770.

Granger, C. W. J. 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227–238.

Granger, C. W. J. and R. Joyeux. 1980. An Introduction to Long-memory Time Series Models and Fractional Differencing. Journal of Time Series Analysis 1(1): 15-29.

Hamilton, J. D. and R. Susmel. 1994. Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64: 307–333.

Hillebrand, E. 2004a. Neglecting Parameter Changes in GARCH Models. Forthcoming in the Journal of Econometrics, Annals Issue on Modeling Structural Breaks, Long Memory, and Stock Market Volatility.

Hillebrand, E. 2004b. Neglecting Parameter Changes in Autoregressive Models. Louisiana State University Working Paper.

`www.bus.lsu.edu/economics/faculty/ehillebrand/personal/research.htm`

Kokoszka, P. and R. Leipus. 1999. Testing for parameter changes in ARCH models. Lithuanian Mathematical Journal 39: 182–195.

Lamoureux, C. G. and W. D. Lastrapes. 1990. Persistence in variance, structural change, and the GARCH model. Journal of Business and Economic Statistics 8: 225–234.

Laurent, S. and J.-P. Peters. 2002. Garch 2.2: An Ox Package for Estimating and Forecasting Various ARCH Models. Journal of Economic Surveys 16(3): 447–485.

LeBaron, B. 2001. Stochastic Volatility as a Simple Generator of Apparent Financial Power Laws and Long Memory. Quantitative Finance 1(6): 621–631.

Mallat, S. 1999. A Wavelet Tour of Signal Processing. 2nd ed. Academic Press: San Diego.

Mikosch, T. and C. Starica. 2004. Change of structure in financial time series, long range dependence and the GARCH model. Forthcoming: The Review of Economics and Statistics. `http://www.math.ku.dk/ mikosch/preprint.html`

Müller, U. A., Dacorogna, M. M., Dave, R. D., Olsen, R. B., Pictet, O. V., and J. E. von Weizsaecker. 1997. Volatilities of different time resolutions - Analyzing the dynamics of market components. Journal of Empirical Finance 4: 213–239.

Nelson, D. B. 1990. Stationarity and Persistence in the GARCH(1,1) Model. Econometric Theory 6: 318–334.

Table 1: ARCH and GARCH Estimation of the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Oct. 6, 2003. The MAE's are calculated using the holding sample Oct. 7, 2003 through Dec 31, 2003. The MAE for the ARCH and IGARCH models are stated in percentages of the forecast error from the GARCH(1,1) model.

			Panel A: GARCH(1,1)			
			$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$			
	sp	dj	cre	crv	ffr	yd
$\alpha + \beta$	0.995	0.990	0.963	0.991	1.0	0.976
MAE(1)	7.24e-5	4.63e-5	3.36e-5	7.08e-5	0.0254	3.35e-5
MAE(20)	6.24e-5	5.33e-5	3.99e-5	5.76e-5	0.0182	3.18e-5
MAE(60)	4.59e-5	4.06e-5	3.95e-5	4.48e-5	0.0158	2.89e-5
			Panel B: ARCH			
			$h_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2$			
	sp	dj	cre	crv	ffr	yd
p	14	11	8	14	20	9
$\sum \alpha_i$	0.855	0.796	0.831	0.866	0.989	0.494
MAE(1)	1.48	1.58	1.44	1.34	0.64	0.85
MAE(20)	1.03	1.03	1.02	1.04	0.90	0.98
MAE(60)	1.03	1.02	1.00	1.02	0.91	1.03
			Panel C: IGARCH(1,1)			
			$h_t = \omega + \alpha\varepsilon_{t-1}^2 + (1-\alpha)h_{t-1}$			
	sp	dj	cre	crv	ffr	yd
α	0.047	0.058	0.013	0.072	0.169	0.050
MAE(1)	1.05	1.03	1.09	1.07	0.98	0.94
MAE(20)	1.31	1.39	1.31	1.41	2.86	1.06
MAE(60)	1.64	1.63	1.39	1.74	6.77	1.15

Table 2: ARMA parameter estimates and mean absolute errors for the 1-day, 20-days, and 60-days forecast of the global GARCH(1,1) model correcting for the short scale found in the residual $\nu_t = \varepsilon_t^2 - h_t$. The ARMA specification was found by including only parameters significant at the 0.05 level or lower. The errors are reported in percent of the global GARCH(1,1) forecast error in Table 1.

$\nu_t = \sum_{j=1}^{p} \phi_j \nu_{t-j} + \sum_{j=1}^{q} \psi_j \eta_{t-j} + \eta_t$, η white noise

	sp	dj	cre	crv	ffr	yd
spec	(1,1)	(1,1)	(1,2)	(1,1)	(2,2)	(1,1)
ϕ_1	0.82	0.80	0.79	0.78	0.73	0.53
ϕ_2					0.08	
ψ_1	-0.78	-0.78	-0.91	-0.75	-0.48	-0.44
ψ_2			0.26		-0.44	
MAE(1)	0.10	0.44	0.42	0.19	0.20	0.41
MAE(20)	0.76	0.81	0.95	0.81	0.79	0.97
MAE(60)	0.89	0.91	0.98	0.92	0.92	0.99

Table 3: Local GARCH(1,1) estimations on segments identified by the change-point detector of Kokoszka and Leipus (1999).

sp			**dj**			**cre**		
segment	prob	λ	segment	prob	λ	segment	prob	λ
01/04/88			01/04/88			01/04/88		
12/30/91	0.00	0.925	05/16/91	0.00	0.930	10/16/97	0.00	0.849
01/08/96	0.00	0.948	10/13/92	0.01	0.517	10/06/03		0.964
12/06/96	0.01	0.909	12/15/95	0.04	0.905			
08/14/97	0.00	0.989	03/26/97	0.00	0.942			
10/06/03		0.953	10/06/03		0.960			
$\bar{\lambda}$		0.944			0.901			0.892

Table 4: Local GARCH(1,1) estimations on segments identified by the change-point detector of Kokoszka and Leipus (1999).

crv			**ffr**			**yd**		
segment	prob	λ	segment	prob	λ	segment	prob	λ
01/04/88			01/04/88			01/04/88		
12/30/91	0.00	0.916	09/04/90	0.02	0.991	05/07/97	0.02	0.959
12/15/95	0.00	0.900	02/22/91	0.00	0.522	09/27/99	0.00	0.902
03/26/97	0.04	0.872	01/21/93	0.00	0.933	10/06/03		0.010
10/15/97	0.00	0.937	10/06/03		0.704			
10/06/03		0.973						
$\bar{\lambda}$		0.928			0.764			0.706

Table 5: Mean absolute errors for the 1-day, 20-days, and 60-days forecast of the GARCH(1,1) model for the last segment identified by the Kokoszka and Leipus (1999) change-point detector. The MAE's for the holding sample Oct 7, 2003 through Dec 31, 2003 are expressed in percent of the global GARCH(1,1) forecast error as reported in Table 1.

	sp	dj	cre	crv	ffr	yd
MAE(1)	1.56	1.64	1.36	1.41	1.06	1.06
MAE(20)	1.21	1.24	1.02	1.21	1.37	1.11
MAE(60)	1.25	1.24	0.97	1.18	1.52	1.21

Table 6: Results of the Geweke and Porter-Hudak (1983) test for the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Dec 30, 2003. The null hypothesis is $d = 0$, the alternative $d \neq 0$. All estimates are significant according to all common significance levels. The smoothed periodogram was estimated using a 1,024 points window with 256 points overlap. The $g(T)$ function in Theorem 2 of Geweke and Porter-Hudak (1983) was set to $[T^{0.55}] = 96$. The numbers in parentheses are asymptotic standard errors.

			Panel A: r_t^2			
	sp	dj	cre	crv	ffr	yd
d	0.354	0.328	0.390	0.387	0.238	0.319
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)
			Panel B: $\lvert r_t \rvert$			
	sp	dj	cre	crv	ffr	yd
d	0.418	0.428	0.507	0.436	0.343	0.422
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)

Table 7: ARFIMA Estimation of absolute returns of the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Oct. 6, 2003. All parameter estimates are significant according to all common confidence levels. The MAE's for the holding sample Oct. 7, 2003 through Dec. 31, 2003 are expressed in percent of the GARCH(1,1) error as reported in Table 1.

$(1-\phi L)(1-L)^d|r_t| = \sum_{j=0}^{q} \psi_j \eta_{t-j} + \eta_t$, η white noise

	sp	dj	cre	crv	ffr	yd
spec	(1,d,1)	(1,d,1)	(1,d,0)	(1,d,1)	(1,d,2)	(1,d,1)
d	0.407	0.400	0.360	0.404	0.415	0.300
ϕ	0.228	0.247	-0.244	0.153	0.558	0.343
ψ_1	-0.605	-0.616		-0.537	-0.515	-0.567
ψ_2					-0.244	
MAE(1)	0.81	0.64	0.62	0.79	0.86	0.43
MAE(20)	1.06	1.11	1.04	1.10	2.38	0.91
MAE(60)	1.31	1.33	1.05	1.32	3.48	0.99

Table 8: FIGARCH Estimation of the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Oct. 6, 2003. The MAE's for the holding sample Oct. 7, 2003 through Dec. 31, 2003 are expressed in percent of the GARCH(1,1) error as reported in Table 1. * indicates the only parameter estimate that is not significant according to all common confidence levels.

$$(1-(\alpha+\beta)L)(1-L)^{-1}(1-L)^{\delta}\varepsilon_t^2 = \omega + (1-\beta L)\nu_t,\ \nu_t = \varepsilon_t^2 - h_t.$$

	sp	dj	cre	crv	ffr	yd
δ	0.428	0.417	0.708	0.404	0.289	0.251
α	0.199	0.247	0.553	0.118	0.256	0.330
β	0.580	0.608	0.005*	0.478	0.106	0.498
MAE(1)	1.11	1.03	0.87	0.97	0.91	0.66
MAE(20)	1.06	1.08	0.91	1.06	2.80	0.92
MAE(60)	1.10	1.08	0.93	1.08	4.24	0.96

Table 9: Estimation of the Engle and Lee (1999) model for the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Oct. 6, 2003. The MAE's for the holding sample Oct. 7, 2003 through Dec. 31, 2003 are expressed in percent of the GARCH(1,1) error as reported in Table 1.

$$h_t - q_t = \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \beta(h_{t-1} - q_{t-1})$$
$$q_t = \omega + \rho q_{t-1} + \phi(\varepsilon_{t-1}^2 - h_{t-1}).$$

	sp	dj	cre	crv	ffr	yd
α	0.011	0.009	0.007*	0.011*	0.075*	0.028***
	(0.007)	(0.006)	(0.004)	(0.006)	(0.040)	(0.008)
β	0.989***	0.991***	0.992***	0.989***	0.923***	0.958***
	(0.008)	(0.007)	(0.005)	(0.007)	(0.031)	(0.012)
ρ	0.946***	0.947***	0.893***	0.933***	0.560***	0.780***
	(0.010)	(0.008)	(0.001)	(0.006)	(0.007)	(0.022)
ϕ	0.058	0.065	0.178*	0.073	0.430***	0.050*
	(0.218)	(0.248)	(0.104)	(0.179)	(0.008)	(0.285)
ω	0.002	0.003	0.001***	0.002**	0.185	0.003***
	(0.002)	(0.645)	(2e-4)	(7e-4)	(0.693)	(5e-4)
MAE(1)	1.53	1.82	0.92	1.36	0.67	0.87
MAE(20)	1.88	2.18	1.23	1.77	2.25	1.08
MAE(60)	2.59	2.98	1.30	2.30	3.80	1.24

Table 10: HARCH Estimation of the daily S&P 500 index (sp), Dow Jones Industrial Average (dj), CRSP equally-weighted index (cre), CRSP value-weighted index (crv), federal funds rate (ffr), and yen per dollar exchange rate (yd) from Jan 4, 1988 through Oct. 6, 2003. The MAE's for the holding sample Oct. 7, 2003 through Dec. 31, 2003 are expressed in percent of the GARCH(1,1) error as reported in Table 1.

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \left(\sum_{i=1}^{20} \varepsilon_{t-i}\right)^2 + \alpha_3 \left(\sum_{i=1}^{60} \varepsilon_{t-i}\right)^2 + \alpha_4 \left(\sum_{i=1}^{100} \varepsilon_{t-i}\right)^2 + \alpha_5 \left(\sum_{i=1}^{250} \varepsilon_{t-i}\right)^2$$

	sp	dj	cre	crv	ffr	yd
ω	6e-5***	6e-5***	2e-5***	5e-5***	0.047***	3e-5***
	(2e-7)	(2e-46)	(7e-7)	(2e-6)	(3e-4)	(8e-7)
α_1	0.132***	0.120***	0.327***	0.153***	0.556***	0.089***
	(0.013)	(0.012)	(0.023)	(0.017)	(0.011)	(0.009)
α_2	0.010***	0.012***	0.003***	0.011***	1e-10	0.007***
	(0.001)	(9e-4)	(3e-4)	(0.001)	(4e-4)	(8e-4)
α_3	0.002***	0.001***	4e-4***	0.002***	1e-10	0.001***
	(3e-4)	(3e-4)	(6e-5)	(3e-4)	(4e-4)	(2e-4)
α_4	1.5e-4	1e-10	1e-10	3e-4***	1e-10	2e-4***
	(9.5e-5)	(1e-4)	(2e-5)	(8e-5)	(1e-4)	(9e-5)
α_5	4.5e-6	1e-5***	1e-10	1e-10	1e-5***	1e-10
	(3.9e-6)	(4e-6)	(3e-6)	(4e-6)	(5e-6)	(5e-6)
λ	0.45	0.43	0.42	0.49	0.56	0.30
MAE(1)	0.62	0.62	0.25	0.50	1.84	1.59
MAE(20)	1.00	1.16	0.88	0.99	2.51	1.31
MAE(60)	1.17	1.31	0.92	1.09	2.83	1.23

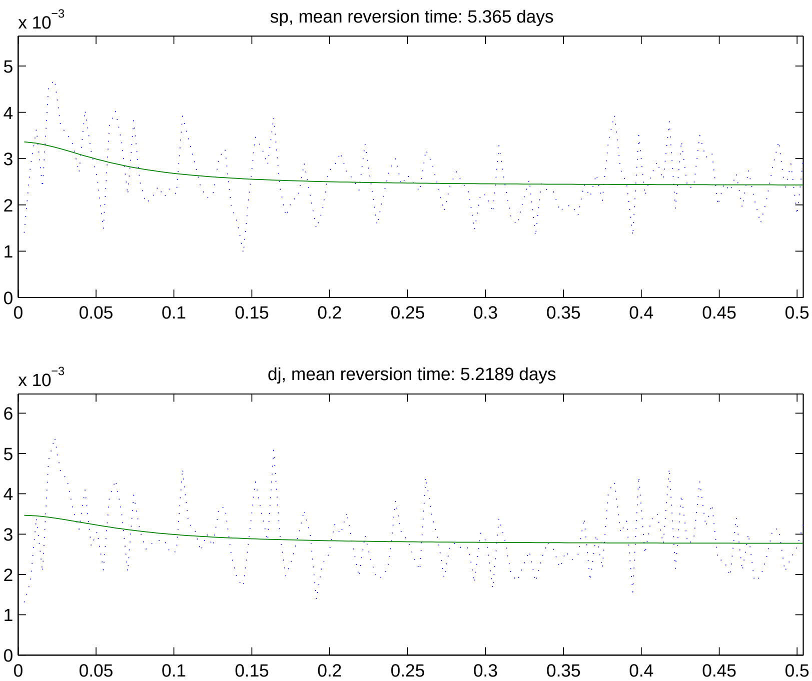

Figure 1: Estimation of the power spectra (dotted line) of the residual processes $\hat{\nu}_t = \hat{\varepsilon}_t^2 - \hat{h}_t$ of the S&P500 and Dow Jones series and nonlinear least squares fit of a Lorentzian spectrum (solid line). The estimate of the average mean reversion time is computed from the Lorentzian.

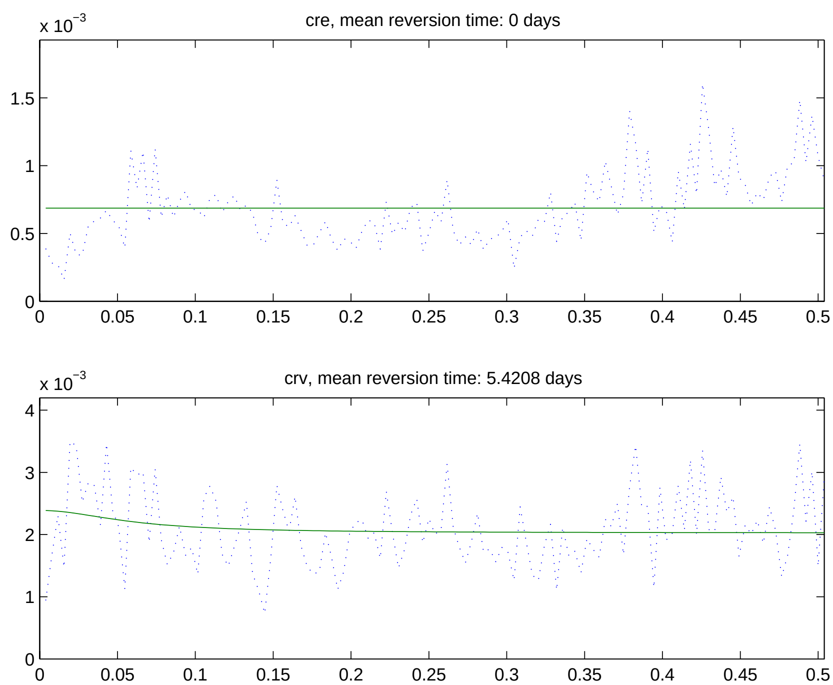

Figure 2: Estimation of the power spectra (dotted line) of the residual processes $\hat{\nu}_t = \hat{\varepsilon}_t^2 - \hat{h}_t$ of the CRSP equally weighted index and the CRSP value-weighted index and nonlinear least squares fit of a Lorentzian spectrum (solid line). The estimate of the average mean reversion time is computed from the Lorentzian.

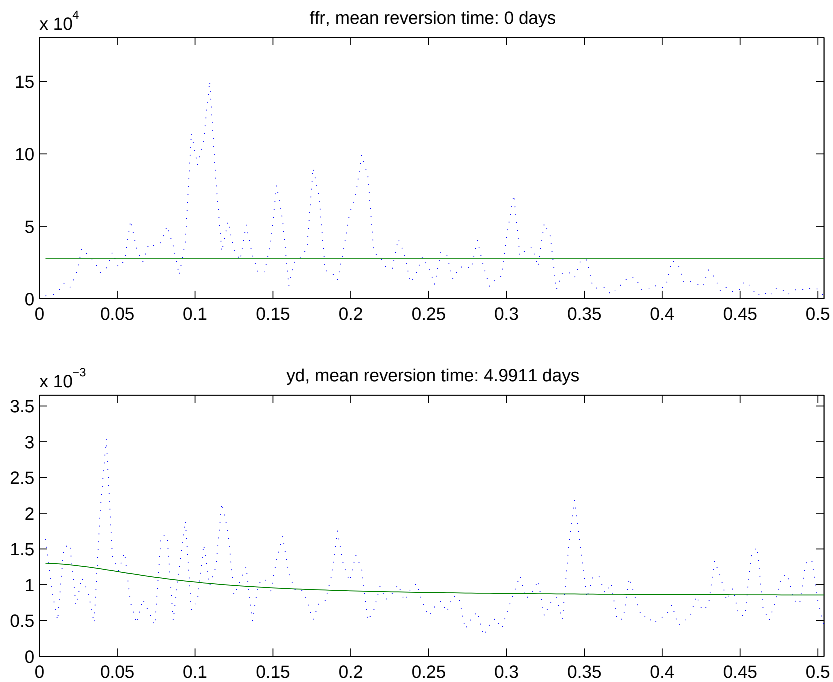

Figure 3: Estimation of the power spectra (dotted line) of the residual processes $\hat{\nu}_t = \hat{\varepsilon}_t^2 - \hat{h}_t$ of the federal funds rate and yen-dollar exchange rate series and nonlinear least squares fit of a Lorentzian spectrum (solid line). The estimate of the average mean reversion time is computed from the Lorentzian.

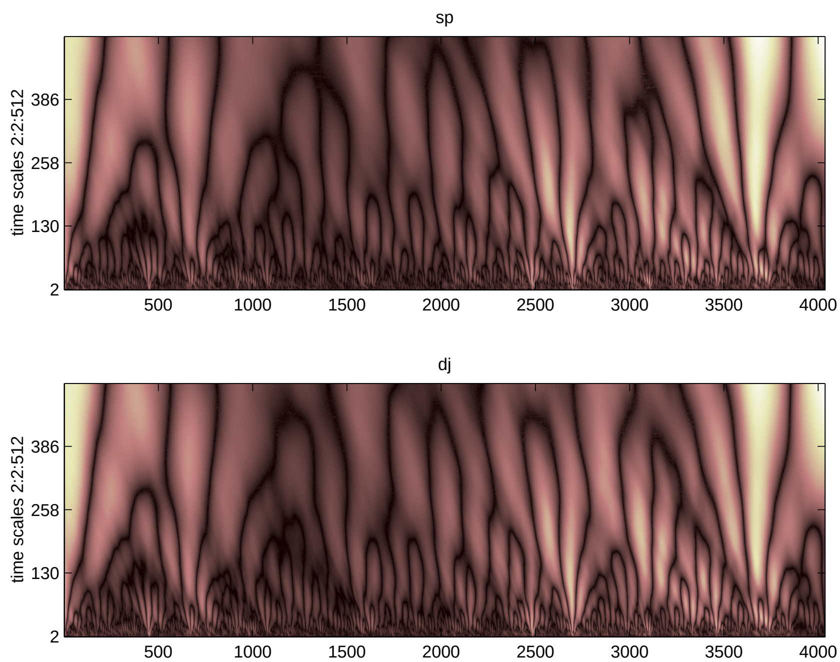

Figure 4: Plot of the wavelet coefficients of the absolute returns series of the S&P500 and the Dow Jones. The wavelet coefficients $Wx(u, s)$ are defined for a specific position u, here on the abscissa, and a specific time scale s, here on the ordinate. The panels give a bird's-eye view of a three dimensional graph that plots the coefficients against position and scale. The different values of the coefficients are indicated by the color: the brighter the color, the higher the coefficient at that position and scale.

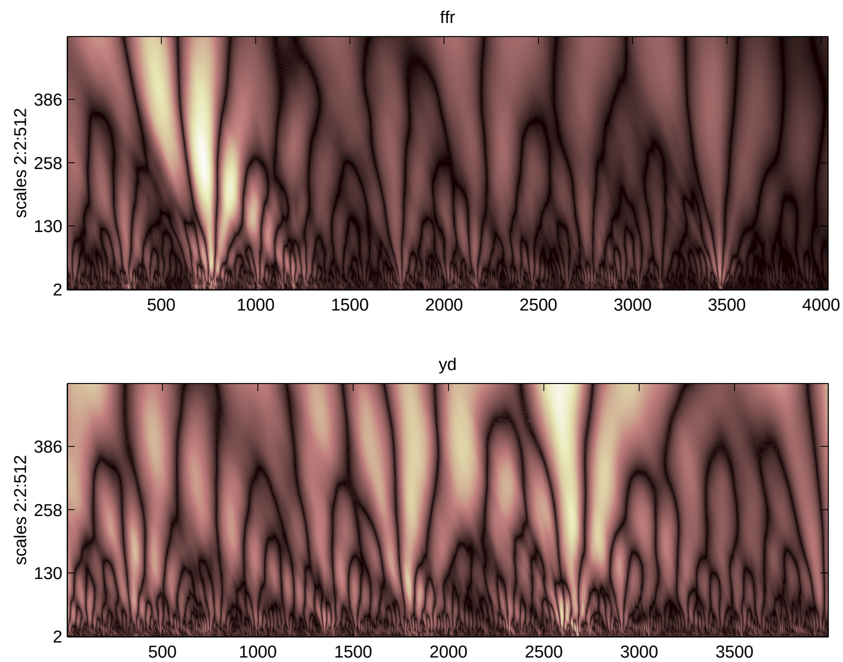

Figure 5: Plot of the wavelet coefficients of the absolute returns series of the federal funds rate and the yen-dollar exchange rate. The wavelet coefficients $Wx(u, s)$ are defined for a specific position u, here on the abscissa, and a specific time scale s, here on the ordinate. The panels give a bird's-eye view of a three dimensional graph that plots the coefficients against position and scale. The different values of the coefficients are indicated by the color: the brighter the color, the higher the coefficient at that position and scale.