SFB 649 Discussion Paper 2007-027

Long Memory Persistence in the Factor of Implied Volatility Dynamics

Wolfgang Härdle*
Julius Mungo*

* Humboldt-Universität zu Berlin, Germany

SFB 649 ECONOMIC RISK BERLIN

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Long Memory Persistence in the Factor of Implied Volatility Dynamics

Wolfgang Karl Härdle[1], Julius Mungo[2]

[1]CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany

[2]CASE – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany; e-mail: mungo@wiwi.hu-berlin.de; phone: +49(0)30 2093-5654

Abstract

The volatility implied by observed market prices as a function of the strike and time to maturity form an Implied Volatility Surface (IVS). Practical applications require reducing the dimension and characterize its dynamics through a small number of factors. Such dimension reduction is summarized by a Dynamic Semiparametric Factor Model ($DSFM$) that characterizes the IVS itself and their movements across time by a multivariate time series of factor loadings. This paper focuses on investigating long range dependence in the factor loadings series. Our result reveals that shocks to volatility persist for a very long time, affecting significantly stock prices. For appropriate representation of the series dynamics and the possibility of improved forecasting, we model the long memory in levels and absolute returns using the class of fractional integrated volatility models that provide flexible structure to capture the slow decaying autocorrelation function reasonably well.

JEL classification: C14, C32, C52, C53, G12

Keywords: Implied Volatility, Dynamic Semiparametric Factor Modeling, Long Memory, Fractional Integrated Volatility Models

Acknowledgement: This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 'Economic Risk'

1 Introduction

As a measure of the standard deviation of the daily range of price movements, volatility is an important determinant of the riskiness of an asset, a crucial parameter in derivative pricing such as options. Findings across several asset markets have reported high persistence of volatility shocks and that over sufficiently long periods of time, volatility is typically stationary with "mean reverting" behavior, Bollerslev and Jubinski (1999). Such series are characterized by distinct but nonperiodic cyclical patterns and their behavior is such that current values are not only influenced by immediate past values but values from previous time periods, allowing for persistence or long memory. Long memory describes the correlation structure of a series at long lags.

It is well known that the volatility implied by observed option prices as a function of the strike and time to maturity form an Implied Volatility Surface (IVS). For each day the IVS forms a high dimensional object that has unknown stochastic behavior that needs to be analyzed. For practical applications such as in risk management, it is desirable to reduce the dimension of this object and characterize its dynamics through a small number of factors. Such dimension reduction may be summarized by a Dynamic Semiparametric Factor Models ($DSFM$) that characterize the IVS and their movements across time by a multivariate time series of factor loadings, Borak et al. (2005), Fengler et al. (2007) and Borak et al. (2007).

The $DSFM$s approximate the implied volatility surface by regressing log-implied volatility on a two-dimensional covariate containing moneyness and time-to-maturity. To introduce this model, denote by $Y_{t,j} = \log\{\hat{\sigma}_{t,j}(\kappa,\tau)\}$, the log-implied volatility where $t = 1,\ldots,I$ is an index of time, in this case the number of the day, and $j = 1,\ldots,J_t$ is the number of IV observations on day t. Let $X_{t,j} = (\kappa_{t,j},\tau_{t,j})$ be a two-dimensional covariate where $\kappa_{t,j}$ is a moneyness matrix and $\tau_{t,j}$ denotes time-to-maturity. Moneyness is defined as $\kappa_{t,j} = \frac{K_{t,j}}{F_{t,j}}$ where $K_{t,j}$ is a strike and $F_{t,j} = S_t e^{(r_{t,j},\tau_{t,j})}$ the underlying futures price belonging to the option trade (t,j). The model is expressed as:

$$Y_{t,j} = \sum_{k=0}^{K} z_{t,k} m_k(X_{t,j}) + \varepsilon_{t,j} \tag{1}$$

where $z_{t,0} = 1$, m_k are smooth basis functions ($k = 0,\ldots,K$) and $z_{t,k}$ are time dependent weights or factor loadings. The IVS is assumed to be a weighted sum of the smooth functional factors, m_k and its dynamics is explained by the stochastic behavior of the factor loadings, $z_{t,k}$. Approximations of the factor loadings are obtained by fitting model (1) to the implied volatility observations and the functions m_k are estimated by orthogonal series estimators so that they have zero correlations among each other, Borak et al. (2005). The factor loadings $z_t = (z_{t,1}, \dots, z_{t,K})^\top$ forms an unobserved multivariate time series.

The estimates $\widehat{z}_{t,k}$ and $\widehat{m}_k$ are obtained in (1) as minimizers of the following least squares criterion:

$$\sum_{t=1}^{I} \sum_{j=1}^{J_t} \int \left\{ Y_{t,j} - \sum_{k=0}^{K} \widehat{z}_{t,k} \widehat{m}_k \right\}^2 K_h(u - X_{t,j})\, du, \tag{2}$$

where K_h denotes a two-dimension kernel function, chosen as a product of one-dimensional kernels $K_h(u) = k_{h_1}(u_1) \times k_{h_2}(u_2)$, where $h = (h_1, h_2)^\top$ are bandwidths and $k_h(v) = k(h^{-1}v)/h$ is a one-dimensional kernel function. The minimization procedure is iterative, searching through all functions $\widehat{m}_k : \mathbb{R}^2 \longrightarrow \mathbb{R}$ $(k = 0, ..., K)$ and time series $\widehat{z}_{tk} \in \mathbb{R}$ $(t = 1, ..., I; k = 1, ..., K)$. The estimation procedure can be seen as a combination of functional principal component analysis, nonparametric curve estimation and backfitting for additive models. The $DSFM$ therefore simultaneously estimate the factor functions and fits the surface.

This paper applies the $DSFM$ on the German DAX index market from 04.01.1999 to 25.02.2003. Figure 1 displays the implied volatility surface from the $DSFM$ fit for the DAX-Option on 2 May 2000, with moneyness between 0.8 and 1.12 and time to maturity between 0 and 0.5 years. Figure 2 shows three volatility-driving factors that could be interpreted in terms of level, slope and curvature factor. $z1$ governs movements in the general level, $z2$ is largely associated with changes in the slope and $z3$ is closely related to dynamic changes in the curvature of the IVS.

The aim of this paper is to investigate dependence in the factor loadings of implied volatility strings because information on persistence can guide the search for economic explanation of the movements in asset returns as well as in risk management applications. Several research involving the autocorrelation functions of various volatility measures (squared, log-squared and absolute returns) have reported decay at a very slow mean-reverting hyperbolic rate, Ding et al. (1993), Bollerslev and Wright (2000) and Sibbertsen (2004). Our analysis follow this line of research on long range dependence investigation and modeling.

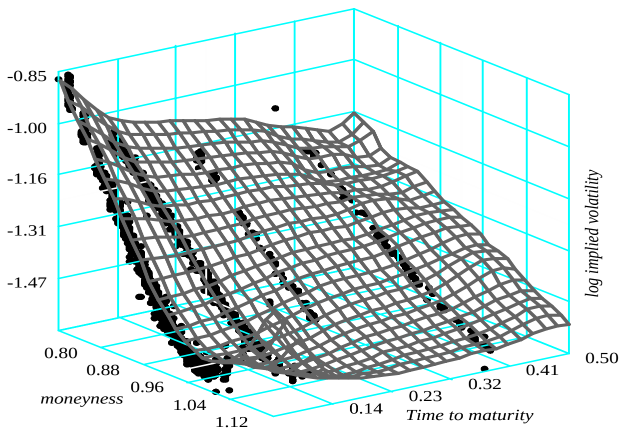

Figure 1: Implied volatility surface from DSFM fit for the DAX-Option on 2 May 2000, with moneyness between 0.8 and 1.12 and time to maturity between 0 and 0.5 years.

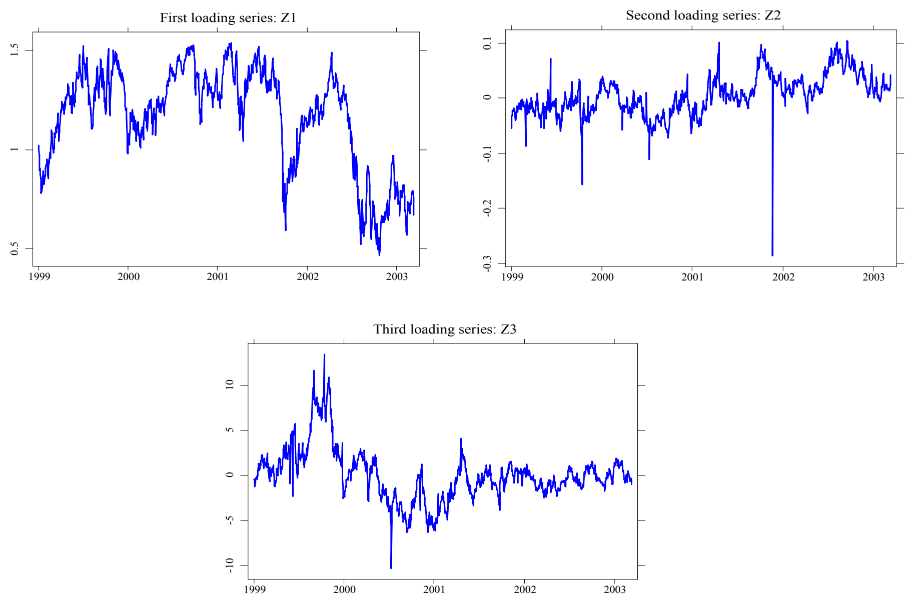

Figure 2: *Time series plots in levels of three loading series from a DSFM fit for the DAX-Option analyzed from* 04.01.1999 – 25.02.2003

First, we consider model independent tests for stationarity, $I(0)$ against fractional alternatives $I(d)$. We apply the rescaled variance test (V/S) of Giraitis et al. (1999) that uses the heteroscedastic and autocorrelation consistent (HAC) estimator of the variance, Newey and West (1987) for normalization and the semiparametric (*LobRob*) test of Lobato and Robinson (1998), that does not depend on a specific parametric form of the spectrum in the neighborhood of the zero frequency. We also apply the log-periodogram regression estimator (GPH) of Geweke and Porter-Hudak (1983) and the Gaussian Semiparametric estimator (GSP) of Robinson (1995a) in estimating the degree of long memory in the factor loadings series. Results are indicative of long-range dependence in the factor loadings series in levels and absolute returns. The first factor, $z1$ can be interpreted as highly persistent and influences all options similarly, irrespective of maturity. The second factor $z2$ gradually diminishes for longer maturities and the third factor $z3$ governs large volatility changes in relatively short maturities.

Second, for appropriate representation of the series dynamics and the possibility of improved forecasting, we model long memory in volatility using the $ARFIMA$, $FIGARCH$ and $HYGARCH$ models. These models provide flexible structure that captures slow decaying autocorrelation reasonably well. In comparison, models in absolute returns have better performance, confirming the findings of Ding et al. (1993), that absolute returns are most appropriate indicator to represent the long memory volatility processes. Our results imply that shocks to the volatility will persist for long time, affecting the DAX stock prices significantly.

Such dependence or persistence will have importance economic consequences for short-term trading and long range investment strategies. Better option pricing may results from models that price and hedge derivative securities when there is prior information on long-memory volatility in terms of expectation on the potential level of volatility and the rate at which volatility changes. In the presence of long memory, Granger and Joyeux (1980), Geweke and Porter-Hudak (1983) have shown the possibility for improved price forecasting performance within a linear time series framework than with traditional procedures. Option pricing have also been shown to be significantly different when standard models are applied as compared to models allowing for long memory.

By applying the $GARCH$, $EGARCH$, $FIEGARCH$ and $IEGARCH$ models, Bollerslev and Mikkelsen (1996) have shown that the price of an option increases with the degree of integration. This means that $GARCH$ models give the lowest price whereas the highest option price is obtained for the $IGARCH$ model. For long memory alternative, Herzberg and Sibbertsen (2004) have shown that prices for the $FIGARCH$, $HYGARCH$ are inbetween the $GARCH$ and $IGARCH$ prices. In addition to documented studies of the economic implications of long memory, Cheung and Lai (1995), Wilson and Okunev (1999), revealed that portfolio diversification decisions in the case of strategic asset allocation may become extremely sensitive to the investment horizon. There may be diversification benefits in the short and medium term, but not if the assets are held together over the long term if long memory is present. e.g., in a market that exhibits antipersistence, asset prices tend to reverse its trend in the short term thus creating short-term trading opportunities. In addition Mandelbrot (1971) has shown that in the presence of long memory the arrival of new market information cannot be fully arbitraged away. It is also known that the possibility of speculative profits as a result of superior long-range dependence model forecast would cast doubt on the basic tenets of market efficiency.

Motivated by evidence of long range dependence in the factor loadings levels and absolute returns, we perform estimation and prediction using the $ARFIMA$, $FIGARCH$ and $HYGARCH$ models that are known to provide flexible structure to capture slow decaying autocorrelation reasonably well than with traditional $ARMA$ procedures.

The rest of our work is structured as follows. Section 2, introduces fractional integration and Long-memory processes. Here we examine some methodology for testing and estimating long range dependence that we apply in our analysis. Section 3 introduces the structure of the class of models we apply to analyze the long memory in the factor loading series. In section 4 we report and discuss our results for the series in levels and absolute returns. A summary of our analysis results and conclusions is given in section 5.

2 Fractional integration and long-memory

The framework of fractional integration yields convenient modeling of long range dependence, Granger and Joyeux (1980), Baillie (1996). A time series process z_t is integrated of order d, $I(d)$ if

$$(1-L)^d z_t = \varepsilon_t \tag{3}$$

where $\varepsilon_t \in I(0)$ and L is the lag operator ($Lz_t = z_{t-1}$). The non-integer parameter d is the difference parameter and $(1-L)^d$ is the fractional filter defined by its binomial expansion $(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j$ where $\Gamma(z) =$

$\int_0^\infty t^{z-1}e^{-t}dt$ is the gamma function. The autocorrelation function of such a series is given by

$$\rho_k = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)} \sim Ck^{2d-1} \tag{4}$$

for d in the range of $(0, 0.5)$, where C is a strictly positive constant. In such case z_t is said to exhibit long memory. For $0 < d < 0.5$, the series is stationary. For $d = 0$, the series is an $I(0)$ process and said to have no long-memory. For $0.5 < d < 1$ the process is mean reverting as there is no long run impact of an innovation to future values of the process. In the case where $0 < d < 1$, not only the immediate past value of z_t influence the current value, but also values from previous time periods as well. The sum over the autocorrelation does not converge, so that it is a suitable model for long memory, Granger and Joyeux (1980).

2.1 Tests and estimators of long memory processes

We consider two model independent tests for stationarity, $I(0)$ against fractional alternatives $I(d)$. The tests include the rescaled variance test (V/S) of Giraitis et al. (1999) that uses the heteroscedastic and autocorrelation consistent (HAC) estimator of the variance, Newey and West (1987) for normalization and the semiparametric (*LobRob*) test of Lobato and Robinson (1998), that does not depend on a specific parametric form of the spectrum in the neighborhood of the zero frequency.

The Rescaled Variance test is applied by centering the $KPSS$ statistic based on the partial sum of the deviations from the mean:

$$V/S(q) = \frac{1}{T^2\hat{\sigma}_T^2(q)}\left[\sum_{k=1}^{T}\left\{\sum_{j=1}^{k}(z_j - \overline{z}_T)\right\}^2 - \frac{1}{T}\left\{\sum_{k=1}^{T}\sum_{j=1}^{k}(z_j - \overline{z}_T)\right\}^2\right] \tag{5}$$

where $S_k = \sum_{j=1}^{k}(z_j - \overline{z}_T)$ are the partial sums of the observations and $\hat{\sigma}_T^2(q) = \hat{\gamma}_0 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{1+q}\right)\hat{\gamma}_j$, is the heteroscedastic and autocorrelation consistent (HAC) estimator of the variance, $(q < T)$. $\hat{\gamma}_0$ is the variance of the process, and the sequence $\{\hat{\gamma}_j\}_{j=1}^{q}$ denotes the autocovariances of the process up to the order q. Giraitis, Kokoszka and Leipus (2000) have shown that this statistic can detect long range dependence in the volatility for the class of $ARCH(\infty)$ processes.

The Semiparametric test is based on the approximation of the spectrum of a long memory process. This test allows to discriminate between $d > 0$ and $d < 0$. In the univariate case the test statistic for $I(0)$ against $I(d)$ is given by

$$t_{LobRob} = \sqrt{(m)}\frac{\widehat{C}_1}{\widehat{C}_0} \tag{6}$$

with $\widehat{C}_k = \frac{1}{m}\sum_{j=1}^{m}\zeta_j^k I(\lambda_j)$ and $\zeta_j = \log(j) - \frac{1}{m}\sum_{i=1}^{m}\log(i)$, where $I(\lambda) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} z_t e^{\mathbf{i}t\lambda}\right|^2$, $(\mathbf{i} = \sqrt{-1})$ is the estimated periodogram. $\lambda_j = \frac{2\pi j}{T}$, $j = 1, \ldots, m << [T/2]$ is a degenerate band of Fourier frequencies with bandwidth parameter m. Under the null hypothesis the test statistic is asymptotically normally distributed. If the statistic is in the lower fractile of the standardized normal distribution, the series exhibit long-memory whilst if in the upper fractile of that distribution, the series is antipersistent.

To estimate the memory parameter d, we apply two frequently used estimators, the log-periodogram regression estimator (GPH) of Geweke and Porter-Hudak (1983) and the Gaussian Semiparametric estimator (GSP) of Robinson (1995a).

The **log-periodogram regression estimator** is based on the periodogram of a time series z_t, $(t = 1, \ldots, T)$ defined by

$$I(\lambda_j) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} z_t e^{-\mathbf{i}\lambda t}\right|^2 \tag{7}$$

where $\lambda_j = \frac{2\pi j}{T}$, $j = 1, \ldots, m$ (m is a positive integer). The memory parameter d is estimated from a linear regression of the $\log I(\lambda_j)$ on a constant and the variable $X_j = \log\left\{4\sin^2(\lambda_j/2)\right\}$:

$$\widehat{d}_{GPH} = -\frac{\sum_{j=1}^{m}(X_j - \bar{X})\log\{I(\lambda_j)\}}{2\sum_{j=1}^{m}(X_j - \bar{X})} \tag{8}$$

We consider only harmonic frequencies $\lambda_j = \frac{2\pi j}{T}$, (the j^{th} Fourier frequency) with $j \in (l, m]$, where l is a trimming parameter discarding the lowest frequencies and m is a bandwidth parameter. The cut-off parameter ensures robustness of the estimator. For the Gaussian case with $d \in (-0.5, 0.5)$, the estimator is consistent and asymptotically normal with standard error of $\pi/\sqrt{24m}$, Robinson (1995b).

Validity of the GPH estimator for an enlarged interval has been demonstrated by Velasco (1999). More precisely, he shows that in the interval

$[0.5, 0.75)$, where the time series is nonstationary, asymptotic Normality and consistency is preserved as in the original interval $(-0.5, 0.5)$, while for values of d in the interval $[0.75, 1)$ the estimator is still consistent. Deo and Hurvich (2001) have shown that this estimator is also valid for some non-Gaussian time series.

The **Gaussian Semiparametric estimator** is based on the approximation, $\lim_{\lambda_i \to 0^+} f(\lambda_i) = C\lambda_i^{-2d}$ of a long memory process in the Whittle approximate maximum likelihood estimator, $L_W(\theta)$. For $m^* = [\frac{T}{2}]$, an approximation to the Gaussian likelihood, Beran (1994) is given by

$$L_W(\theta) = -\frac{1}{2\pi}\sum_{j=1}^{m^*} \log f_\theta(\lambda_j) + \frac{I_T(\lambda_j)}{f_\theta(\lambda_j)} \qquad (9)$$

for a given parametric spectral density $f_\theta(\lambda)$. Estimating d is by solving the minimization,

$$\arg\min_{C,d} L(C, d) = \frac{1}{m}\sum_{j=1}^{m}\left\{\log(C\lambda_j^{-2d}) + \frac{I(\lambda_j)}{C\lambda_j^{-2d}}\right\}$$

where $I(\lambda_j)$ is the periodogram evaluated for a degenerated range of m harmonic frequencies $\lambda_j = \frac{2\pi j}{T}$, $j = 1, \ldots, m << [\frac{T}{2}]$, where $[.]$ represents the integer part operator, bounded by the bandwidth parameter m, which increases with the sample size T but more slowly. This bandwidth m must satisfy $\frac{1}{m} + \frac{m}{T} \to 0$ as $T \to \infty$. If $m = [\frac{T}{2}]$, this estimator is a Gaussian estimator for the parametric model $f(\lambda) = C\lambda^{-2d}$.

An estimator for d is given by

$$\hat{d}_{GSP} = \arg\min_{d}\left\{\log\left(\frac{1}{m}\sum_{j=1}^{m}\frac{I(\lambda_j)}{C\lambda_j^{-2d}}\right) - \frac{2d}{m}\sum_{j=1}^{m}\log(\lambda_j)\right\}. \qquad (10)$$

Robinson (1995a) showed that $\sqrt{m}(\hat{d}_{GSP} - d) \xrightarrow{d} N(0, 1/4)$ and is valid in the presence of some form of conditional heteroscedasticity, Robinson and Henry (1999). In general, Phillips and Shimotsu (2004) have shown that the ranges of consistency and asymptotic normality for the model type in (3) are thesame as those of the GPH estimator.

3 Long Memory Models

Several studies have dealt with models that provide useful ways of analyzing the relationships between the conditional mean and variance of a process exhibiting long memory and slow decay in its levels. In our analysis we consider an autoregressive fractional integrated moving average ($ARFIMA$) process model, Granger and Joyeux (1980), Hosking (1981), the ($FIGARCH$), Baillie et al. (1996), a combination of fractional integrated process for the mean with regular $GARCH$ process for conditional variance and the HYperbolic GARCH, ($HYGARCH$) of Davidson (2004).

The **ARFIMA(p,d,q)** model represented as

$$\Phi(L)(1-L)^d(z_t-\mu)=\Theta(L)\varepsilon_t \tag{11}$$

where $\varepsilon_t \sim i.i.d(0,\sigma_\varepsilon^2)$ extends the integration order of the conventional $ARMA$ model to a non-integer value between 0 and 1. $\Phi(L)=1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p$ and $\Theta(L)=1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q$ are the autoregressive and moving average polynomials in the lag operator L respectively. d is the long memory parameter and $(1-L)^d$ is the fractional difference filter as defined in equation 3. The $ARFIMA$ process displays persistence for $0<d<0.5$ and anti-persistence for $-0.5<d<0$. For $|d|>0.5$, the process is non-stationary as it has finite variance.

In our application, the values for p and q are chosen such that the ordered pair (p,q) minimizes the AIC criterion. We estimate the model parameters μ, ϕ, θ and d by maximum likelihood approach of Doornik and Ooms (2004) that allows for (break-) regressors in the mean and structural changes in the variance, and by non-linear least squares estimation method of Beran (1994) that is asymptotically efficient in the presence of $GARCH$ errors.

The **FIGARCH(p,δ,q)** model of Baillie et al. (1996) given as

$$\Phi(L)(1-L)^d\varepsilon_t^2=\omega+\Theta(L)\nu_t \tag{12}$$

where $\nu_t=\varepsilon_t^2-\sigma_t^2$ combine the fractional integrated process for the mean with regular $GARCH$ process for the conditional variance. The conditional variance can be represented as

$$\sigma_t^2=\frac{\omega}{1-\theta(L)}+\left[1-\frac{\phi(L)(1-L)^\delta}{1-\theta(L)}\right]\varepsilon_t^2 \tag{13}$$

with $0\leq\delta\leq 1$. The δ in $FIGARCH$ does not have the same interpretation of persistence as d in $ARFIMA$. The fractional differencing operator in the $ARFIMA$ model applies to the constant term in the mean equation while in $FIGARCH$ it does not apply to ω in the variance equation. We base our analyis on the $FIGARCH$ parametrization proposed by Chung (1999),

$$\Phi(L)(1-L)^d(\varepsilon_t^2-\sigma^2)=\Theta(L)(\varepsilon_t^2-\sigma^2) \tag{14}$$

where σ^2 is the unconditional variance of ε_t. The conditional variance is formulated as

$$\sigma_t^2 = \sigma^2 + \left[1 - \frac{\phi(L)(1-L)^\delta}{1-\theta(L)}\right](\varepsilon_t^2 - \sigma^2) \tag{15}$$

For $p = q = 1$, Chung (1999) shows that $\sigma^2 > 0$ and $0 \leq \phi_1 \leq \theta_1 \leq 1$ is a sufficient condition for positive σ_t^2. When $\delta = 0$ or 1, the $FIGARCH$ model nests the $GARCH(p,q)$ and $IGARCH$ processes respectively. The $IGARCH$ model is a short memory process having no variance and while the $FIGARCH$ has shortest memory with $\delta > 0$ closest to 1. If $\delta > 0$ the $FIGARCH$ is a non-stationary long memory process, otherwise is a stationary long memory process, Laurent and Peters (2002). The fractional difference filter is defied by,

$$\begin{aligned}
(1-L)^\delta &= \sum_{j=0}^{\infty} \frac{\Gamma(\delta+1)}{\Gamma(j+1)\Gamma(\delta-j+1)} L^j & (16)\\
&= 1 - \delta L - \frac{1}{2}\delta(1-\delta)L^2 - \frac{1}{6}\delta(1-\delta)(2-\delta)L^3 - \ldots & (17)\\
&= 1 - \sum_{j=1}^{\infty} C_j(\delta) L^j & (18)
\end{aligned}$$

such that $C_1(\delta) = \delta$, $C_2(\delta) = \frac{1}{2}\delta(1-\delta)$, etc. By construction, $\sum_{j=1}^{\infty} C_j(\delta) = 1$ for any δ, belonging to same class type models as the $IGARCH$.

The hyperbolic $GARCH$ model, **HYGARCH(p,α,d,q)** of Davidson (2004) extends the conditional variance of the $FIGARCH(p,\delta,q)$ model by introducing weights to the difference operator in equation 12 such that $(1-L)^d = \left[(1-\alpha) + \alpha(1-L)^d\right]$. The parametrization of $HYGARCH(p,\alpha,d,q)$ models is given by

$$\sigma_t^2 = \frac{\omega}{1-\theta(L)} + \left[1 - \frac{\phi(L)\left\{1+\alpha(1-L)^d\right\}}{1-\theta(L)}\right]\varepsilon_t^2 \tag{19}$$

where α are weights to $(1-L)^d$. The parameters α and d are assumed positive. The $HYGARCH(p,\alpha,d,q)$ nest $GARCH$ models (for $\alpha = 0$), $IGARCH$ (for $\alpha = d = 1$) and $FIGARCH$ (for $\alpha = 1$ or $\log\alpha = 0$). When $\alpha < 1$ ($\log\alpha < 0$) the process is stationary.

4 Empirical Analysis

The factor loadings series data are obtained from DSFM for implied volatility on the German DAX index market from January 1999 to February 2003 (data available at, http://sfb649.wiwi.hu-berlin.de/fedc). Table 1 presents descriptive statistics. Plots of sample autocorrelation functions, spectrum and periodogram (on log-log plane) are shown in Figure 3. The autocorrelation are positive and decay hyperbolically to zero as the lag increases. A linear relationship in the periodogram on log-log plane indicates the presence of self-similarities, the fluctuations in a power-law fashion. Figure 4 shows time series plots in absolute returns of three factor loadings series.

Since unit root tests are known to perform relatively poorly in distinguishing between $I(1)$ and the $I(d)$ alternatives for $d < 1$, Diebold and Rudebusch (1991), we apply model independent tests, V/S and *LobRob* for $I(0)$ against $I(d)$ alternatives. With no data driven guideline for the choice of truncation lags m, we use different values ($m = 2, 3, 5, 7, 10, 20, 50$) in the V/S test and ($m = 30, 50, 150, 200, 300$) in *LobRob* test.

Results in Tables 2 and 3 for the V/S and *LobRob* tests respectively indicate long-range dependence in all three factor loadings levels. For absolute returns, the tests indicate long memory in $|z1|$ and $|z3|$ while antipersistence could not be rejected in $|z2|$. Both tests results reject long memory for all factor loadings returns.

Table 4 shows the $\hat{d}_{GPH}$ and $\hat{d}_{GSP}$ estimates of d for the series in levels, returns and absolute returns. To evaluate the sensitivity of results for the $\hat{d}_{GPH}$ estimator, we report estimates of d for bandwidth $m = T^{\alpha}$ where $\alpha = 0.5, 0.525, 0.575, 0.60, 0.80$ and $T = 1052$ is the sample size. For the GSP estimator the bandwidth is chosen such that $m = [\frac{T}{4}], [\frac{T}{8}], [\frac{T}{16}], [\frac{T}{32}], [\frac{T}{64}]$. Results for series in levels show $0.5 \leq d < 1$; for the return series most estimates from $\hat{d}_{GPH}$ and $\hat{d}_{GSP}$ are in $-0.5 \leq d < 0$ while estimates of d for the absolute returns are within $0 \leq d < 0.5$.

To guarantee that the long memory diagnosis is not a consequence of occasional or structural break such as the 11^{th} September, 2001 terrorist attack on the World Trade Center, we use subsamples of the data to examine whether long run dependencies can be uncovered, Anderson and Bollerslev (1998). This approach is possible given that the value of d is not affected by temporal aggregation, Bollerslev and Wright (2000). Results, not presented here indicate long memory for short span of the data in levels and absolute returns. This therefore suggest that long memory is an inherent

series	mean	std. dev.	min.	max.	skew.	kurt	J.Bera (p value)
z_1	1.165	0.253	0.465	1.538	-0.820	2.698	0.000
z_2	0.005	0.035	-0.286	0.104	-0.251	6.958	0.000
z_3	0.000	0.028	-0.103	0.134	0.933	5.452	0.000

Table 1: *Summary statistics for factor loadings times series on the German DAX index market from January* 1999 *to February* 2003*, a sample of* 1039 *observations.*

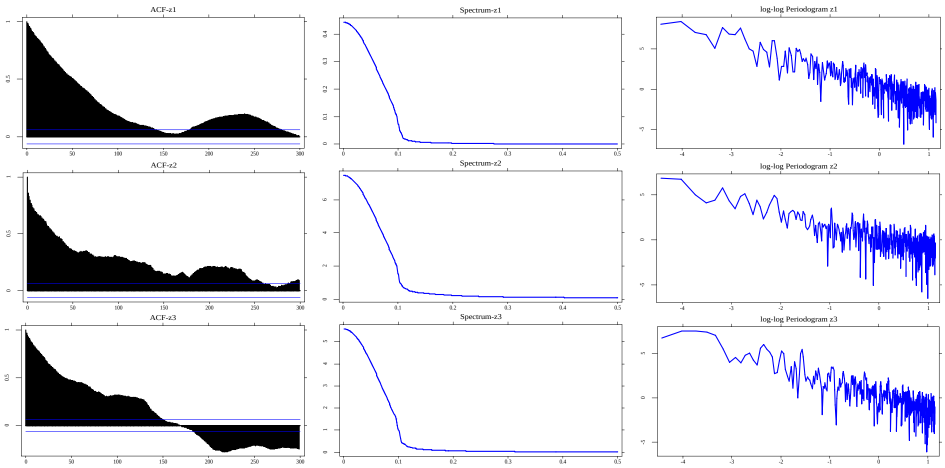

Figure 3: *Plots of sample autocorrelation functions (lag length 300), spectrum and periodogram (in the log-log plane) of the factor loadings series.*

characteristics of the factor loading series.

To summarize, our analysis suggest long-range dependence in loading levels as well as in absolute returns for $z1$ and $z3$. In general no long memory in returns was detected and evidence of antipersistence in the absolute returns for $z2$ could not be ruled out. We therefore interpret that there is quite some correlation structure in the loadings in levels as well as in absolute returns. This implies some degree of persistence and the expectation of a slow decay in impulse responses. We also observe that the long range dependence is different for each factor loading such that it could be interpreted in terms of a long term, middle long term and short term impact on the dynamics of IVS. The first factor loading, $z1$ is highly persistent and influences all options similarly, irrespective of maturity. The impact of the second factor loading, $z2$ gradually diminishes for longer maturities, while the third factor governs large volatility changes in relatively short maturities.

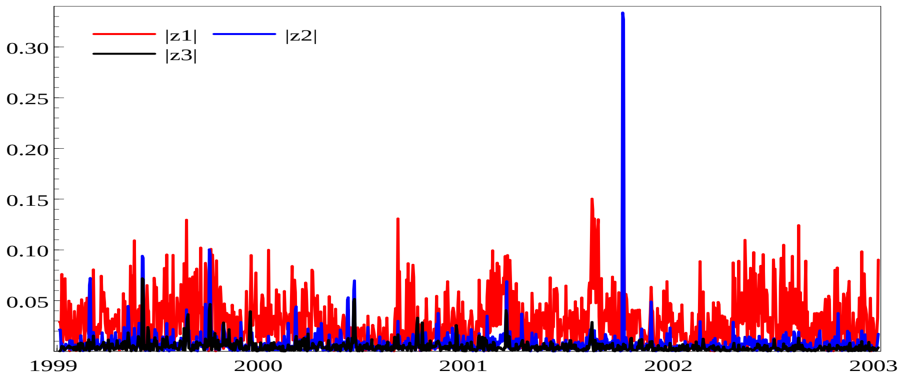

Figure 4: *Time series plots of the three factor loading series in absolute returns from* 04.01.1999 − 25.02.2003

m	2	3	5	7	10	20	50
			level				
m	2	3	5	7	10	20	50
$z1$	4.63	3.29	2.22	1.68	1.24	0.68	0.32
$z2$	3.77	2.89	1.99	1.54	1.15	0.65	0.31
$z3$	3.54	2.68	1.82	1.38	1.02	0.57	0.27
			r_t				
m	2	3	5	7	10	20	50
$z1$	0.02	0.02	0.03	0.03	0.04	0.05	0.06
$z2$	0.00	0.01	0.01	0.01	0.01	0.02	0.03
$z3$	0.01	0.01	0.02	0.02	0.03	0.03	0.04
			$\|r_t\|$				
m	2	3	5	7	10	20	50
$z1$	0.24	0.21	0.19	0.15	0.13	0.09	0.07
$z2$	0.05	0.05	0.05	0.04	0.05	0.04	0.04
$z3$	0.84	0.78	0.72	0.64	0.56	0.44	0.28

Table 2: *The rescaled variance V/S test for $I(0)$ against $I(d)$ for series in levels, return (r_t) and absolute return $(|r_t|)$. q is the truncation lag. If the evaluated statistics are over the critical value, 0.1869 for I(0), we fail to reject the alternative hypothesis that the series display long memory.*

m	30	50	150	200	300
			level		
m	30	50	150	200	300
$z1$	-6.08	-10.60	-30.07	-38.46	-53.60
$z2$	-6.50	-11.05	-28.83	-35.98	-49.20
$z3$	-5.72	-9.86	-28.33	-36.33	-50.70
			r_t		
m	30	50	150	200	300
$z1$	0.65	0.96	2.69	2.58	2.83
$z2$	2.02	1.47	5.14	5.76	7.03
$z3$	1.40	1.61	2.91	2.99	4.25
			$\|r_t\|$		
m	30	50	150	200	300
$z1$	-1.25	-1.88	-7.76	-10.04	-12.42
$z2$	0.715	0.37	-0.01	-0.52	-1.00
$z3$	-4.83	-6.81	-9.08	-11.13	-9.69

Table 3: t_{LobRob}: *Semiparametric test for I(0) of a time series against long-memory and antipersistence for factor loadings in levels, return (r_t) and absolute return ($|r_t|$). Short memory is rejected against long-memory if the test statistic is in the lower tail of the standard normal distribution. If the test statistic is in upper tail of the standard normal distribution, short memory is rejected against antipersistent.*

m	$\widehat{d}_{GPH}$: z1 level	r_t	$\|r_t\|$	m	$\widehat{d}_{GSP}$: z1 level	r_t	$\|r_t\|$
32	0.95	-0.20	0.31	16	0.77	-0.36	-0.21
38	0.90	-0.25	0.18	32	0.83	-0.20	0.28
54	0.98	-0.15	0.17	65	0.95	-0.09	0.29
65	0.96	-0.14	0.19	131	0.89	-0.17	0.31
261	0.97	-0.10	0.27	263	0.93	-0.12	0.27
m	$\widehat{d}_{GPH}$: z2 level	r_t	$\|r_t\|$	m	$\widehat{d}_{GSP}$: z2 level	r_t	$\|r_t\|$
32	0.69	-0.63	-0.09	16	0.79	-0.86	-0.38
38	0.73	-0.52	0.02	32	0.68	-0.62	-0.10
54	0.84	-0.35	-0.04	65	0.75	-0.34	-0.01
65	0.85	-0.39	-0.05	131	0.67	-0.39	-0.02
261	0.70	-0.40	-0.00	263	0.66	-0.37	0.03
m	$\widehat{d}_{GPH}$: z3 level	r_t	$\|r_t\|$	m	$\widehat{d}_{GSP}$: z3 level	r_t	$\|r_t\|$
32	0.95	-0.24	0.33	16	0.68	-0.50	0.69
38	0.85	-0.32	0.31	32	0.86	-0.27	0.38
54	0.99	-0.16	0.36	65	0.84	-0.24	0.35
65	0.90	-0.23	0.35	131	0.83	-0.22	0.21
261	0.84	-0.22	0.17	263	0.85	-0.17	0.17

Table 4: *The Log periodogram ($\hat{d}_{GPH}$) and the Gaussian semiparametric ($\hat{d}_{GSP}$) estimates of d for levels, returns and absolute returns. Bandwidth m for GPH estimator is $m = T^{\alpha}$ with $\alpha = 0.5, 0.525, 0.575, 0.60, 0.8$ and $T = 1052$ is the sample size. For the GSP estimator the bandwidth is chosen such that $m = [\frac{T}{4}], [\frac{T}{8}], [\frac{T}{16}], [\frac{T}{32}], [\frac{T}{64}]$.*

Level	$z1$	$z2$	$z3$
ARFIMA	$(5, d, 4)$	$(3, d, 3)$	$(1, d, 5)$
d	0.29 (1.78)	0.53 (4.87)	0.29 (0.74)
ϕ_1	0.59 (3.13)	-0.23 (-0.60)	0.96 (26.10)
ϕ_2	0.07 (0.31)	0.74 (4.01)	
ϕ_3	0.29 (1.84)	0.34 (0.99)	
ϕ_4	0.50 (2.71)		
ϕ_5	-0.47 (-3.92)		
θ_1	0.19 (0.94)	0.17 (0.35)	-0.51 (-1.41)
θ_2	0.04 (0.29)	-0.71 (-3.49)	-0.01 (-0.24)
θ_3	-0.17 (-1.37)	-0.24 (-0.64)	-0.09 (-2.62)
θ_4	-0.58 (-4.78)		0.03 (0.79)
θ_5			-0.08 (-2.44)
constant	-0.10	0.01	
$Ln(\ell)$	1892.33	2755.75	3585.06
AIC	-3764.66	-5453.51	-7152.12

Table 5: *ARFIMA estimation of factor loading series in levels, $z1$, $z2$ and $z3$ from 04.01.1999 to 25.02.2003. The ϕ coefficients correspond to the autoregressive part and the θ coefficients to the moving average part. t-value of the estimated parameters in brackets, $Ln(\ell)$ is the log-likelihood and (AIC) Akaike Information Criterion.*

4.1 Long Memory Models Application

For appropriate representation of the factor loadings series dynamics and the possibility of improved forecasting, we model the long memory in levels and absolute returns using the $ARFIMA$, $FIGARCH$ and $HYGARCH$ models. These models are known to describe volatility reasonably well and provide flexible structure that captures slow decaying autocorrelation functions. Estimation results with $ARFIMA$ model for series in levels and absolute returns are reported in Tables 5 and 6 respectively. Estimates of d are highly significant across all time series in levels. The t-statistic are highly significant to reject the null hypothesis ($H_0 : d = 0$) at 1% significance level. Results for absolute returns $|z2|$ does confirm earlier tests results in that antipersitence in $|z2|$ may not be rejected.

Estimation results for the $FIGARCH(1, \delta, 1)$ and $HYGARCH(1, d, 1)$ models are reported in Tables 7 and 8 respectively. We assume the Student-t distribution because it can appropriately account for leptokurticity exhibited by high frequency financial data, Pagan (1996). Under student-t distributed innovations, the $FIGARCH$ and $HYGARCH$ long memory parameter es-

Absolute returns *ARFIMA*	$\|z1\|$ $(2,d,2)$	$\|z2\|$ $(1,d,5)$	$\|z3\|$ $(1,d,2)$
d	0.30 (4.73)	-0.32 (-0.71)	0.24 (2.34)
ϕ_1	-0.79 (-4.65)	0.89 (5.89)	0.57 (4.28)
ϕ_2	0.01 (0.02)		
θ_1	0.61 (2.90)	0.06 (-0.31)	-0.40 (-2.11)
θ_2	-0.22 (-1.09)	-0.48 (-5.59)	-0.28 (-8.40)
θ_3		-0.06 (-0.61)	
θ_4		-0.03 (-0.40)	
θ_5		-0.00 (-0.07)	
const.	0.3		
$Ln(\ell)$	2381.53	2913.99	3927.21
AIC	-4753.06	-5813.98	-7846.42

Table 6: *ARFIMA estimation of factor loading series in absolute returns, $|z1|$, $|z2|$ and $|z3|$ from 04.01.1999 to 25.02.2003. The ϕ coefficients correspond to the autoregressive part and the θ coefficients to the moving average part. t-value of the estimated parameters in brackets, $Ln(\ell)$ is the log-likelihood and (AIC) Akaike Information Criterion.*

timates indicate long-memory in levels and absolute returns. Besides, the student-t distribution parameter, ν are significantly different from zero, indicating strong fat tail phenomena. Estimates, $\delta > 0$ in $FIGARCH$ models suggest non-stationary long memory characteristics in levels and absolute returns for the first and third factor loadings series, whereas the second series $(\delta < 0)$ indicate a stationary long memory behavior. We assess models fit through the log-likelihood, $Ln(\ell)$, the Akaike Information Criterion, (AIC) and the performance of the Box-Pierce (Q^2) statistic for testing remaining serial correlation in the squared standardized residuals, McLeod and Li (1983).

The $FIGARCH$ and $HYGARCH$ models perform well in describing the high persistence existing in the conditional variance. The $Q^2(24)$ statistics suggests that the $FIGARCH$ model can better capture the autocorrelations in the conditional variance for the series in levels while the $HYGARCH$ is more appropriate in the case of absolute returns. The $FIGARCH$ models report higher loglikelihood values for the series in levels while the $HYGARCH$ values are higher for absolute returns. Moreover, models in absolute returns produce better fit than those in levels, which confirms the findings of Ding et al. (1993), that absolute returns are the most appropriate indicators to represent the long memory volatility processes.

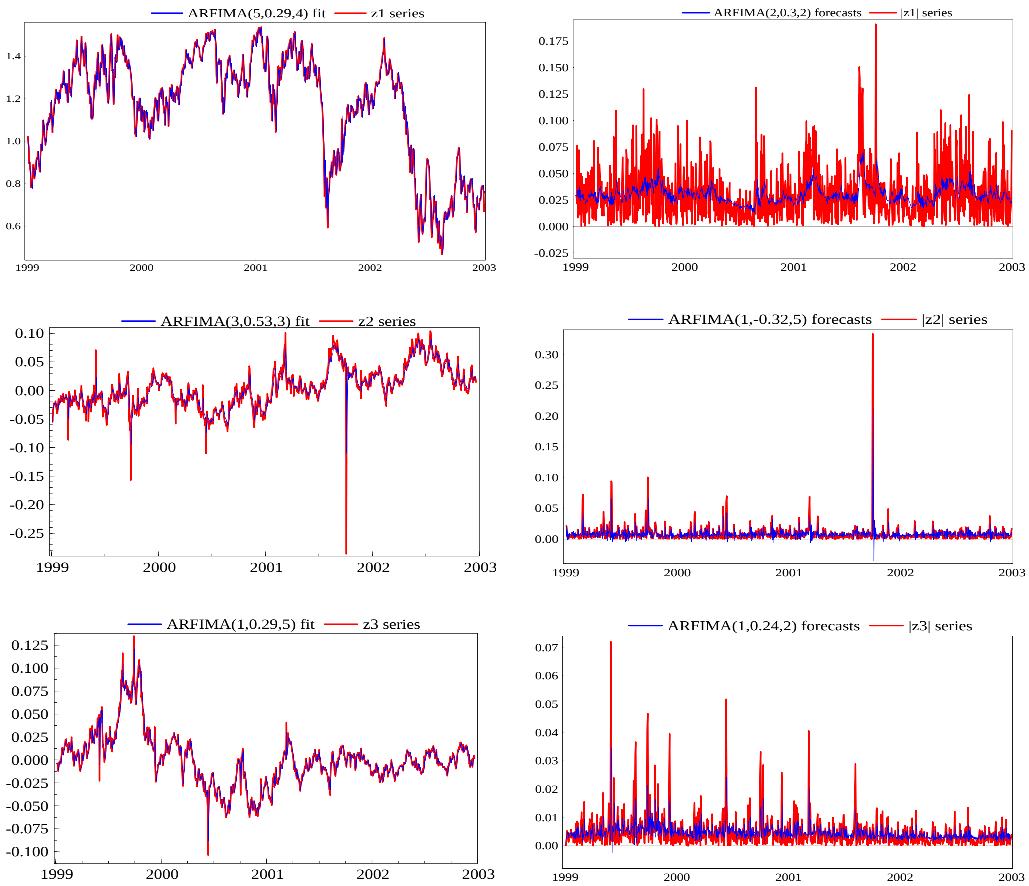

Figure 5: *Actual series (red) and in-sample fit (blue) for the estimated* $ARFIMA(p, d, q)$ *model in levels and absolute returns. Time interval from* $04.01.1999 - 25.02.2003$*, with* 1039 *observations.*

We examine the in-sample fit of the $ARFIMA$ model for the series in levels and absolute returns, Figure 5 as well as the conditional variance forecast in levels, Figure 6 and absolute returns, Figure 7 for the $FIGARCH$ and $HYGARCH$ models. Table 9 show in-sample forecast performance evaluated on the basis of the Root Mean Square Error ($RMSE$) and the Mean Absolute Prediction Error ($MAPE$). The main findings are that the $FIGARCH$ and $HYGARCH$ models show better forecast performance than the $ARFIMA$ model and seems to successfully achieve the aim of modeling the long memory behavior of volatility in a parsimonious way.

	$z1$	$z2$	$z3$	$\|z1\|$	$\|z2\|$	$\|z3\|$
μ	1.329	-0.002	-0.003	0.024	0.005	0.003
	(11.190)	(-0.257)	(-3.733)	(25.460)	(22.660)	(21.240)
ω	25.011	5.156	3.534	8.264	5.482	0.590
	(2.368)	(2.566)	(0.538)	(2.336)	(3.363)	(4.203)
δ	0.460	-0.031	0.716	0.291	-0.040	0.172
	(4.280)	(-1.364)	(18.330)	(2.693)	(-1.344)	(2.181)
ϕ_1	0.161	0.742	-0.865	0.085	0.823	0.309
	(1.580)	(10.100)	(-14.110)	(0.636)	(14.870)	(1.931)
β_1	0.510	-0.012	-0.261	0.248	-0.031	-0.032
	(3.214)	(-1.111)	(-3.965)	(1.586)	(-1.830)	(-0.415)
ν	10.870	2.560	30.965	4.919	2.243	2.520
	(3.185)	(11.320)	(1.349)	(6.016)	(22.710)	(12.660)
$Ln(\ell)$	1938.359	3302.076	3250.264	2424.146	3786.055	4378.002
AIC	-3.717	-6.404	-6.243	-4.654	-7.274	-8.413
$Q(24)$	58.793	19.167	4095.820	164.692	11.872	36.460
	[0.000]	[0.691]	[0.000]	[0.000]	[0.972]	[0.036]
$Q^2(24)$	15.226	0.097	7.319	11.789	0.092	2.156
	[0.852]	[1.000]	[0.998]	[0.961]	[1.000]	[1.000]

Table 7: *$FIGARCH$ estimation of the factor loading series in levels and absolute returns with t statistics in parentheses. Significance is at 5% level. Estimation is with the Student distribution with ν degrees of freedom. $Ln(\ell)$ is the value of the maximized likelihood. $Q(24)$ and $Q^2(24)$ are the Box-Pierce statistic for remaining serial correlation in the standardized and squared standardized residuals respectively, using 24 lags with p-values in square brackets. The critical value at significant level of 5% is 36.4 .*

	$z1$	$z2$	$z3$	$\|z1\|$	$\|z2\|$	$\|z3\|$
μ	1.304	-0.001	-0.003	0.024	0.004	0.003
	(11.520)	(0.703)	(-2.061)	(26.290)	(20.230)	(19.900)
ω	0.000	0.184	0.156	0.790	0.344	0.069
	(-0.203)	(5.055)	(3.344)	(1.885)	(2.267)	(2.120)
d	0.082	0.980	0.919	0.975	0.985	0.001
	(1.164)	(32.630)	(13.220)	(7.073)	(46.840)	(11.610)
ϕ_1	-0.861	0.456	-0.893	-0.024	0.421	0.271
	(-20.720)	(9.161)	(-18.660)	(-0.256)	(2.820)	(1.663)
β_1	-0.897	0.321	-0.107	0.730	0.548	-0.050
	(-26.300)	(0.037)	(-1.791)	(5.715)	(4.431)	(-0.776)
ν	11.065	3.106	24.200	5.175	2.414	2.668
	(3.059)	(7.918)	(1.413)	(5.538)	(9.835)	(8.027)
$\log(\alpha)$	0.836	0.083	-1.661	-0.120	-0.504	4.953
	(1.003)	(1.586)	(-7.471)	(-1.699)	(-2.553)	(10.200)
$Ln(\ell)$	1934.661	3301.990	3256.190	2428.212	3785.100	4378.439
AIC	-3.708	-4.562	-6.252	-4.660	-7.270	-8.412
$Q(24)$	59.270	7315.110	4096.210	153.997	10.412	37.739
	[0.000]	[0.000]	[0.000]	[0.000]	[0.988]	[0.027]
$Q^2(24)$	19.704	13.140	4.499	11.364	0.084	2.042
	[0.610]	[0.930]	[0.999]	[0.940]	[1.000]	[1.000]

Table 8: *$HYGARCH$ estimation of the factor loading series in levels and absolute returns with t statistics in parentheses. Significance is at 5% level. Estimation is with the Student-t distribution with ν degrees of freedom.* $\log(\alpha)$ *is the log of weight α, to the difference operator$(1-L)^d$. $Ln(\ell)$ is the value of the maximized likelihood. $Q(24)$ and $Q^2(24)$ are the Box-Pierce statistic for remaining serial correlation in the standardized and squared standardized residuals respectively, using 24 lags with p-values in square brackets.*

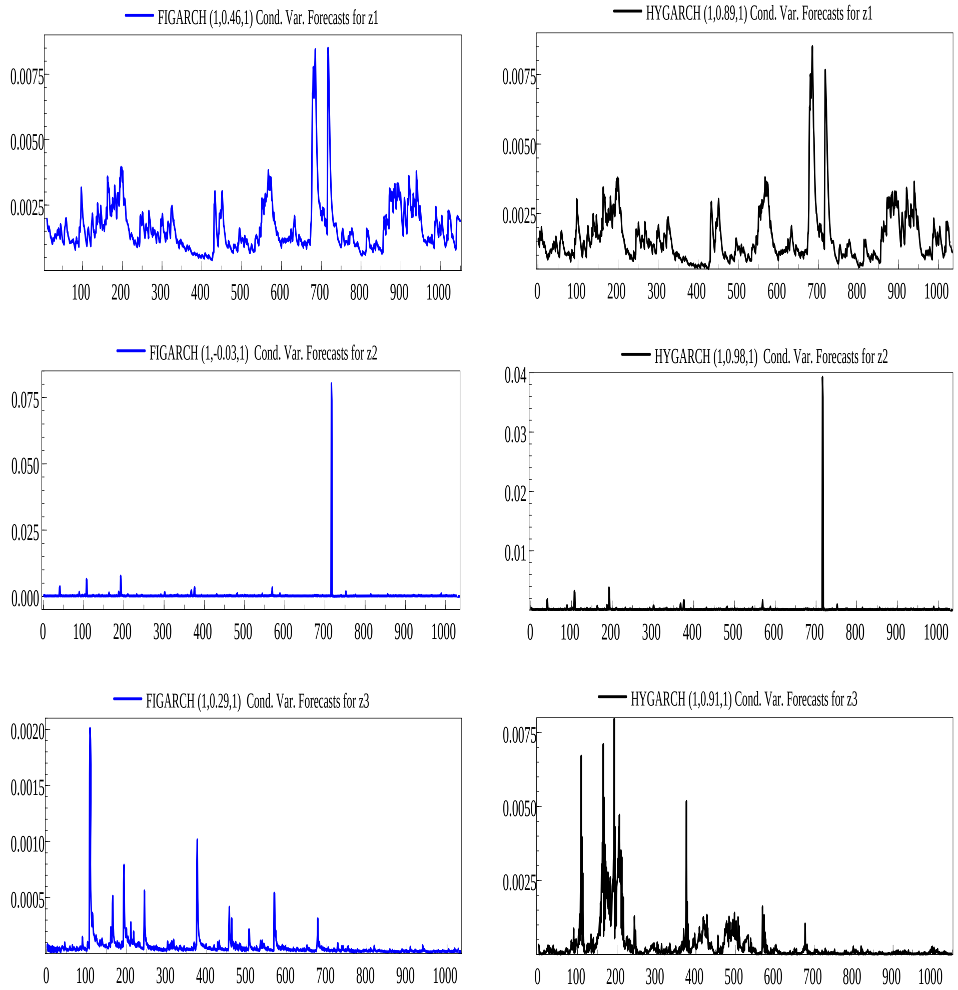

Figure 6: *Left-Right panels: $FIGARCH$ and $HYGARCH$ conditional variance forecast of factor loading in levels. Time interval from* 04.01.1999 – 25.02.2003, *with* 1039 *observations.*

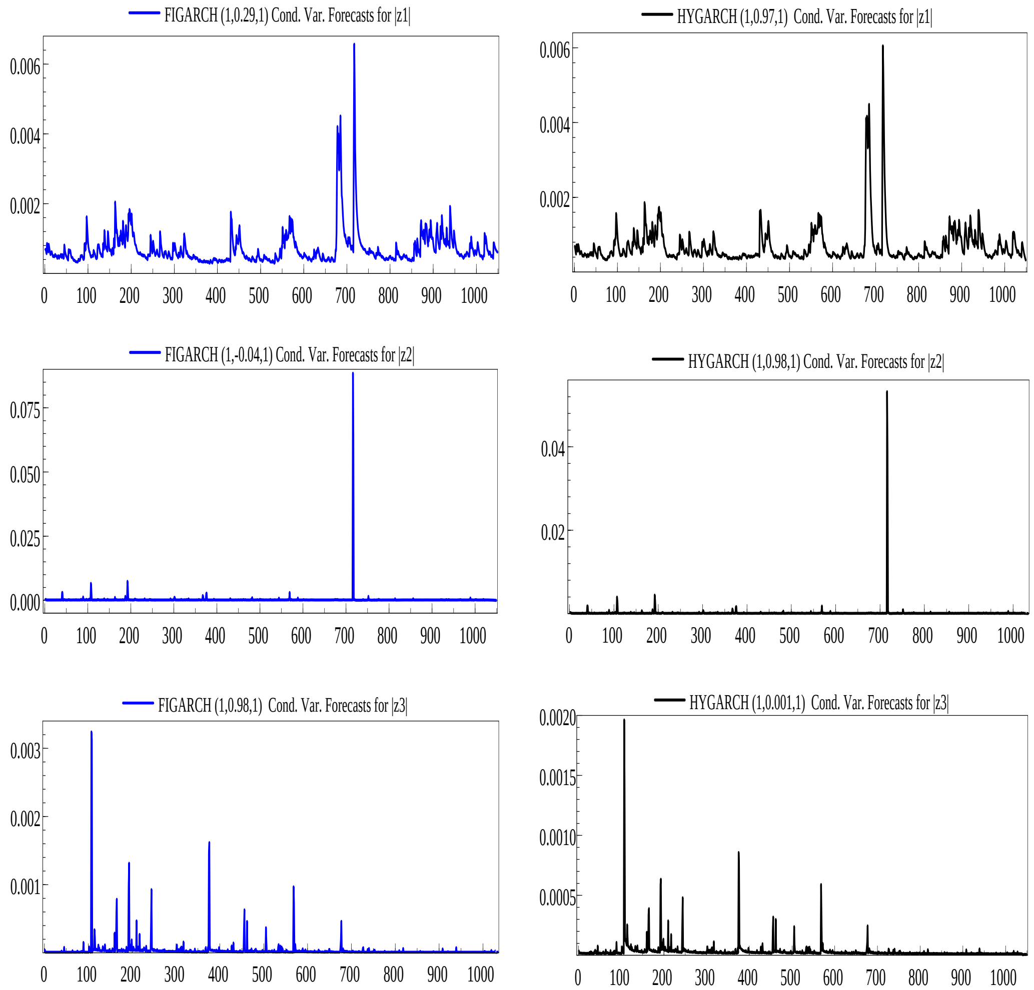

Figure 7: *Left-Right panels: $FIGARCH$ and $HYGARCH$ conditional variance forecast of absolute returns. Time interval from* 04.01.1999 − 25.02.2003*, with* 1039 *observations.*

	$ARFIMA$	$FIGARCH$	$HYGARCH$
$RMSE$			
$z1$	0.026	0.171	0.171
$z2$	0.006	0.001	0.001
$z3$	0.005	0.001	0.001
$\|z1\|$	0.019	0.001	0.001
$\|z2\|$	0.007	0.001	0.001
$\|z3\|$	0.002	0.001	0.001
$MAPE$			
$z1$	3.180	0.991	0.991
$z2$	27.600	1.166	0.860
$z3$	68.720	2.837	9.413
$\|z1\|$	52.657	22.910	16.380
$\|z2\|$	70.577	1.892	2.409
$\|z3\|$	41.706	9.134	10.39

Table 9: *In-sample performance of the five-step ahead forecast of the estimated $ARFIMA$, $FIGARCH$ and $HYGARCH$ models for the factor loading series in levels and absolute returns. The measures of forecast accuracy are the Root Mean Square Error ($RMSE$) and the Mean Absolute Prediction Error ($MAPE$).*

5 Conclusion

We present an empirical investigation of long memory dynamics in the factors of Implied Volatility Strings. The factor loadings series are obtained by applying a Dynamic Semiparametric Factor Model ($DSFM$) for implied volatility strings on the German DAX index market. Long range dependence in the factor loadings series is tested using the rescaled variance V/S and the semiparametric *LobRob* tests. We estimated the degree of long memory based on the log-periodogram GPH regression estimator and the GSP estimator based on the Whittle approximate maximum likelihood estimate. Results are indicative of long-range dependence in the factor loading series in levels and absolute returns. The factors can be interpreted in terms of a long term, middle long term and short term impact on the dynamics of IVS. The first factor loading, $z1$ is highly persistent and influences all options similarly, irrespective of maturity. The impact of the second factor loading, $z2$ gradually diminishes for longer maturities and the third factor governs large volatility changes in relatively short maturities. Such dependence or persistence has importance implications for short-term trading and long range investment strategies. As a consequence, hedging strategies of a long position should take into consideration the long-memory effects in a short position in a call option. This would certainly provide more secure protection against negative effects of long-range persistence in volatility. On the other hand, better results could be obtained for models that price and hedge derivative securities when there is prior information on long-memory volatility in terms of expectation on the potential level of volatility and the rate at which volatility changes.

For an appropriate representation of the series dynamics and the possibility of improved forecasting, we model the long memory in volatility via the class of flexible processes, the $ARFIMA$, $FIGARCH$ and $HYGARCH$ models. Our results indicate that these models appear to capture the slow decaying autocorrelation function and therefore are applicable in mimicking the dynamics of the factor loadings. In comparison, models in absolute returns have better performance, confirming the findings of Ding et al. (1993), that absolute returns are the most appropriate indicator to represent the long memory volatility processes. It would be interesting to find out if there are persistent time scales that are of local importance or influence the factor loading time series. Therefore, a possible extensions for future research would include studies on spectral analysis or wavelet transform to identify such persistent time scales. In addition, the resulting long range dependence and evidence of fat tail phenomenon also provides a natural extension to investigate long memory value-at-risk. This would be useful to regulators, derivative market participants and practitioners whose interest is to reasonably forecast stock market movements.

References

Anderson, T. and Bollerslev, T., (1998): Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Long Run Dependencies. *Journal of Finance*, **53**: 219–265.

Baillie, R. T., (1996): Long memory processes and fractional integration in economics. *Journal of Econometrics*, **73**: 5–59.

Baillie, R. T., Bollerslev, T., Mikkelson, H., (1996): Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **14**: 3–30.

Beran, J., (1994): Statistics for long memory processes. *Chapman & Hall*, New York.

Bhardwaj, G. and Swanson, N.R., (2004): An Empirical Investigation of the Usefulness of ARFIMA Models For Predicting Macroeconomic and Financial Times series. *Working paper*, Rutgers University.

Bollerslev, T., (1987): A conditionally heteroskedastic time series model for speculative prices and rates of returns. *Review of Economics and Statistics*, **69**: 542-547.

Bollerslev, T. and Jubinski, D., (1999): Equity trading volume and volatility: Latent information arrivals and common long-run dependencies. *Journal of Business and Economic Statistics*, **17**: 9 - 21.

Bollerslev, T. and Mikkelsen, H., (1996): Modelling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics*, **73**: 151 - 184.

Bollerslev, T., and Wright, J.H., (2000): Semiparametric estimation of Long-memory volatility dependencies: The role of high-frquency data. *Journal of Econometrics*, **Vol.98**, No.1, pp.81-106.

Borak, S., Härdle, W. and Fengler, M., (2005): DSFM fitting of Implied Volatility Surfaces, Proceedings 5th International Conference on Intelligent System Design and Applications, IEEE Computer Society Number P2286, Library of Congress Number 2005930524.

Borak, S., Härdle, W., Mammen, E. and Park, B. U. (2007): Time series modeling with semiparametric factor dynamics. *SFB 649 Discussion Paper 2007-023*, Humboldt-Universität zu Berlin.

Brown, R.L., Durbin, J., and Evans, J.M., (1975): Techniques for testing the constancy of regression relationship over time. *Journal of the Royal Statistical Society B*, **37**: 149 192.

Cheung, Y. W., and Lai, K. S., (1995): A search for long memory in international stock market returns. *Journal of International Money and Finance*, **14**, 597-615.

Chung, C. F., (1999): Estimating the Fractionally Integrated GARCH Model. *working paper*, National Taïwan Univerity.

Davidson, J., (2004): Moments and memory properties of linear conditional heteroskedasticity models, and a new model. *Journal of Business and Economic Statistics*, **22**, 16-29.

Deo, R. S. and Hurvich, C. M., (2001): On the Log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models. *Econometric Theory*, **17**(686-710).

Diebold, F.X. and Inoue, A., (2001): Long memory and Regime Switching. *Journal of Econometrics*, **105**, 131–159.

Diebold, F.X. and Rudebusch, G.D., (1991): On the Power of the Dickey-Fuller Test Against Fractional Alternatives. *Economic letters*, **35**, 155–160.

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

Ding, Z., Granger, C.W.J. and Engle, R.F, (1993): A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**, 83 - 106.

Doornik, J.A., (2006a): An Introduction to OxMetrics 4. A software System for Data Analysis and Forecasting. *Timberlake Consultant Ltd.*, first edn.

Doornik, J.A., Ooms, M., (2004): Inference and forecasting for ARFIMA models with an Application to US and UK inflation. *Studies in Nonlinear Dynamics and Econometrics 8*, No. **2**, Article 14.

Elliott, G., Rothenberg, T. J and Stock, J. H., (1996): Efficient tests for an autoregressive unit root. *Econometrica* , **64**, 813–836.

Fengler, M., Härdle, W. and Mammen, E., (2007): A Semiparametric Factor Model for Implied Volatility Surface Dynamics. *Journal of Financial Econometrics*, **Vol.5**, No.2,000-000.

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

Giacomini, E., Handel, M. and Härdle, W., (2006): Time Dependent Relative Risk Aversion. *SFB 649, Discussion paper 2006-020*, Humboldt-Universität zu Berlin.

Giraitis, L., Kokoszka, P. and Leipus, R., (2000): Stationary ARCH models: Dependence structure and central limit theorem. *Econometric Theory* **16**, 3-22.

Giraitis, L., Kokoszka, P. and Leipus, R., (2001): Testing for long memory in the presence of a general trend. *Journal of Applied Probability*, **38**, 1033–1054.

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G., (1999): Semiparametric Estimation of the Intensity of Long-Memory in Conditional Heteroskedasticity. *G.R.E.Q.A.M.*, 99a24, Universite Aix-Marseille III.

Ghysels, E., Santa-Clara, P. and Valkanov (2004): Predicting Volatility: Getting the Most out of Return Data Sampled at Different Frequencies. *Scientific Series*, 2004s–19.

Granger, C.W.J., (1980): Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, **14**: 261–279.

Granger, C.W.J. and Ding, Z., (1995): Some properties of absolute returns: An alternative measure of risk. *Annales d'Économie et de Statistique*, **40**, 67 - 91.

Granger, C.W.J. and Ding, Z., (1996): Varieties of Long-Memory Models. *Journal of Econometrics*, **73**, 61–77.

Granger, C. W. J. and Joyeux, R., (1995): An introduction to long memory times series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–39.

Hall, P., Härdle, W., Kleinow, T. and Schmidt, P., (2000): Semiparametric Bootstrap Approach to Hypothesis Tests and Confidence Intervals for the Hurst Coefficient. *Statistical Inference for Stochastic Processes*, **3**, 263-276.

Herzberg, M., Sibbertsen, P., (2004): Pricing of options under different volatility models. *Technical Report/ Universität Dortmund, SFB 475, Komplexitätsreduktion in Multivariaten Datenstrukturen*, 2004, **62**.

Hillebrand, E., (2005): Neglecting parameter changes in GARCH models. forthcoming. *Journal of Econometrics*, **129**.

Hosking, J.R.M., (1981): Fractional Differencing. *Biometrika*, **68**(1), 165-176.

Hurvich, C., Doe, R. and Brodsky, J., (1998): The Mean Square Error of Geweke and Porter-Hudak's estimator of the Memory Parameter of a Long-Memory time series. *Journal of Time Series Analysis*, **19**, 19-46.

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

Lobato, I. and P.M. Robinson (1998): A Nonparametric Test for I(0). *Review of Economic Studies*, **65**(3), 475495.

Lobato, I. and N.E. Savin (1998): Real and Spurious Long-Memory Properties of Stock Market Data. *Journal of Business and Economic Statistics*, **16**, 261-283.

MacKinnon, J. G., (1991): Critical values for cointegration tests. *in* C. W. J. Granger R. F. Engle (eds), *Long-Run Economic Relationships*, Oxford University Press, Oxford.

McLeod, A. I. and Li, W. K., (1983): Diagnostic checking ARMA time series models using squared residual autocorrelations. *Journal of Time Series Analysis*, **4**, 269-273.

Mandelbrot, B. B., (1971): When a price be arbitraged efficiently? a limit to the validity of the random walk and Martingale models. *Review of Economics and Statistics*, **53**, 225-236.

Mikosch, T., Starica, C., (2004): Nonstationarities in financial time series, the long range dependence and the IGARCH effects. *Review of Economics and Statistics* **86**(1), 378–390.

Newey, W.K. and West, K.D., (1994): A Simple Positive Definite, Heteroskedasticity and Autocorelation Consistent Covariance Matrix. *Econometrica*, **55**, 703–705.

Newey, W.K. and West, K.D., (1994): Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, **61**, 631–653.

Pagan, A., (1996): The econometrics of financial markets. *Journal of Empirical Finance*, **3**, 15–102.

Phillips, P.C.B and Shimotsu, K., (2004): Local Whittle estimation in nonstationary and unit root cases. *Annals of Statistics*, **32**, 656–692.

Robinson, P.M., (1995a): Gaussian Semiparametric Estimation of Long-Range Dependence. *Annals of Statistics*, **23**, 1630-1661.

Robinson, P.M., (1995b): Log-Periodogram regression of time Series with Long Range Dependence. *Annals of Statistics*, **23**(3), 1048-1072.

Robinson, P.M., (2003): Long memory time series. In Robinson, P.M. (Ed.), Times series with Long Memory. *Oxford University Press*, Oxford.

Robinson, P.M., Henry, M., (1999): Long and Short Memory Conditional Heteroskedasticity in Estimating the Memory Parameter in Levels. *Economic Theory*, **15**, 299-336.

Rydén, T., Teräsvirta, T. and Ȧsbrink, S (1998): Stylized facts of daily return series and hidden Markov model. *Journal of Applied Econometrics*, **13**: 217-244.

Sibbertsen, P., (2004): Long-Memory in Volatilities of German Stock Returns. *Empirical Economics* , **29**, 477 - 488.

Tiao, G. C. (2001): Univariate Autoregressive Moving-Average Models. In Pena, D., Tiao, G.C. and Tsay, R.S. (eds.), A Course in Time Series Analysis. *ew York (Wiley), Chapter 3.*

Velasco, C., (1999): Gaussian Semiparametric Estimation of Non-Stationary Time Series. *Journal of Time Series Analysis*, **20**: 87 127(41).

Velasco, C., (1999): Non-Stationary Log-periodogram Regresssion. *Journal of Econometrics*, **91**: 325 - 371.

Wilson, P. J., and Okunev, J., (1999): Long-Term Dependencies and Long Run Non-Periodic co-Cycles: Real Estate and Stock Markets. *Journal of Real Estate Research*, **18**, 257 - 278.

SFB 649 Discussion Paper Series 2007

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Trade Liberalisation, Process and Product Innovation, and Relative Skill Demand" by Sebastian Braun, January 2007.
002 "Robust Risk Management. Accounting for Nonstationarity and Heavy Tails" by Ying Chen and Vladimir Spokoiny, January 2007.
003 "Explaining Asset Prices with External Habits and Wage Rigidities in a DSGE Model." by Harald Uhlig, January 2007.
004 "Volatility and Causality in Asia Pacific Financial Markets" by Enzo Weber, January 2007.
005 "Quantile Sieve Estimates For Time Series" by Jürgen Franke, Jean-Pierre Stockis and Joseph Tadjuidje, February 2007.
006 "Real Origins of the Great Depression: Monopolistic Competition, Union Power, and the American Business Cycle in the 1920s" by Monique Ebell and Albrecht Ritschl, February 2007.
007 "Rules, Discretion or Reputation? Monetary Policies and the Efficiency of Financial Markets in Germany, 14th to 16th Centuries" by Oliver Volckart, February 2007.
008 "Sectoral Transformation, Turbulence, and Labour Market Dynamics in Germany" by Ronald Bachmann and Michael C. Burda, February 2007.
009 "Union Wage Compression in a Right-to-Manage Model" by Thorsten Vogel, February 2007.
010 "On $\sigma-$additive robust representation of convex risk measures for unbounded financial positions in the presence of uncertainty about the market model" by Volker Krätschmer, March 2007.
011 "Media Coverage and Macroeconomic Information Processing" by Alexandra Niessen, March 2007.
012 "Are Correlations Constant Over Time? Application of the CC-TRIG_t-test to Return Series from Different Asset Classes." by Matthias Fischer, March 2007.
013 "Uncertain Paternity, Mating Market Failure, and the Institution of Marriage" by Dirk Bethmann and Michael Kvasnicka, March 2007.
014 "What Happened to the Transatlantic Capital Market Relations?" by Enzo Weber, March 2007.
015 "Who Leads Financial Markets?" by Enzo Weber, April 2007.
016 "Fiscal Policy Rules in Practice" by Andreas Thams, April 2007.
017 "Empirical Pricing Kernels and Investor Preferences" by Kai Detlefsen, Wolfgang Härdle and Rouslan Moro, April 2007.
018 "Simultaneous Causality in International Trade" by Enzo Weber, April 2007.
019 "Regional and Outward Economic Integration in South-East Asia" by Enzo Weber, April 2007.
020 "Computational Statistics and Data Visualization" by Antony Unwin, Chun-houh Chen and Wolfgang Härdle, April 2007.
021 "Ideology Without Ideologists" by Lydia Mechtenberg, April 2007.
022 "A Generalized ARFIMA Process with Markov-Switching Fractional Differencing Parameter" by Wen-Jen Tsay and Wolfgang Härdle, April 2007.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

023 "Time Series Modelling with Semiparametric Factor Dynamics" by Szymon Borak, Wolfgang Härdle, Enno Mammen and Byeong U. Park, April 2007.

024 "From Animal Baits to Investors' Preference: Estimating and Demixing of the Weight Function in Semiparametric Models for Biased Samples" by Ya'acov Ritov and Wolfgang Härdle, May 2007.

025 "Statistics of Risk Aversion" by Enzo Giacomini and Wolfgang Härdle, May 2007.

026 "Robust Optimal Control for a Consumption-Investment Problem" by Alexander Schied, May 2007.

027 "Long Memory Persistence in the Factor of Implied Volatility Dynamics" by Wolfgang Härdle and Julius Mungo, May 2007.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".