Can GARCH Models Capture the Long-Range Dependence in Financial Market Volatility?*

John M. Maheu
University of Toronto
Dept. of Economics
jmaheu@chass.utoronto.ca

first draft March 2002
this draft October 2002

Abstract

This paper investigates if component GARCH models introduced by Engle and Lee (1999) and Ding and Granger (1996) can capture the long-range dependence observed in measures of time-series volatility. Long-range dependence is assessed through the sample autocorrelations, two popular semiparametric estimators of the long-memory parameter, and the parametric fractionally integrated GARCH (FIGARCH) model. Monte Carlo methods are used to characterize the finite sample distributions of these statistics when data are generated from GARCH(1,1), component GARCH and FIGARCH models. For several daily financial return series we find that a two-component GARCH model captures the shape of the autocorrelation function of volatility, and is consistent with long-memory based on semiparametric and parametric estimates. Therefore, GARCH models can in some circumstances account for the long-range dependence found in financial market volatility.

JEL classification: C22,C52
Keywords: component GARCH, FIGARCH, fractional integration, long-memory

*I have benefited from discussions with Richard Baillie, Wing Chan, Christian Gouriéroux, Lynda Khalaf, Alex Maynard, Tom McCurdy, and seminar participants at University of Guelph, Laval, Toronto, and Wilfrid Laurier. I thank the Social Sciences and Humanities Research Council of Canada for financial support.

1 Introduction

Since the introduction of ARCH models by Engle (1982) there has been an explosion of research on the conditional variance process in speculative returns data. Time-varying volatility is not only important in forecasting future market movements but is central to a host of financial issues, such as portfolio diversification, risk management, and derivative pricing. Although it is common to find a significant statistical relationship between current measures of volatility and lagged values it has been very difficult to find models that adequately capture the time-series dependencies observed in the data. One linear measure of the time-series dependencies in volatility is the autocorrelation function. Ding, Granger, and Engle (1993) show the absolute value of S&P500 returns has the long-memory property in that the sample autocorrelation function (ACF) of $|r_t|$ decays very slowly and remains significant even at long lags. Evidence of long-range dependence in measures of volatility has been documented in many studies. For instance, see Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001), Andersen and Bollerslev (1997), Baillie (1996), Bollerslev and Wright (2000), Breidt, Crato, and de Lima (1998), Granger and Jeon (2001), Hwang (2001), Lobato and Savin (1998), and Ray and Tsay (2000).

Standard models of the conditional variance imply an exponential rate of decay for the autocorrelation function of squared innovations. For example, the GARCH(1,1) model posses this feature and cannot capture the degree of persistence in the sample ACF of absolute returns. Typically autocorrelations from a GARCH(1,1) model start too high and decay much faster than the data implies. This observation has been used to conclude that standard volatility models, such as the GARCH(p,q) class, are unable to account for the long-range dependence found in measures of volatility.

To parsimoniously capture long-range dependence in volatility, Baillie, Bollerslev, and Mikkelsen (1996) and Ding and Granger (1996) propose fractionally integrated autoregressive conditional heteroskedasticity (ARCH) models. The fractionally integrated GARCH (FIGARCH) model of Baillie, Bollerslev, and Mikkelsen (1996) can be viewed as a fractionally integrated ARMA (ARFIMA) specification for the squared innovations. Like the fractionally integrated class of processes $I(d)$, introduced by Granger (1980), Granger and Joyeux (1980), and Hosking (1981) the FIGARCH model avoids the sharp distinction between $I(0)$ and $I(1)$ processes by allowing d to take a value between 0 and 1. Therefore, the ACF of the volatility process can possess a rate of decay somewhere between the extremes of an exponential rate ($I(0)$) and infinite persistence ($I(1)$).

Extensions to the FIGARCH specification include Baillie, Bollerslev, and

Mikkelsen (1996) who propose an exponential version (FIEGARCH) while McCurdy and Michaud (1997) extend the parameterization to include the asymmetric power ARCH structure of Ding, Granger, and Engle (1993). The FIGARCH model has been successfully applied in several areas of empirical finance. Bollerslev and Mikkelsen (1999) investigate the economic value of FIEGARCH forecasts of volatility, while Brunetti and Gilbert (2000) extend the model to a bivariate framework, and Baillie, Cecen, and Han (2000) and Beltratti and Morana (1999) study high frequency data with the model.

In recent years there has been a great deal of research on nonlinear models that exhibit or approximate long-memory. In most cases, the model incorporates some form of regime switching between states. Some examples of this area of work are Diebold and Inoue (2001), Gourieroux and Jasiak (2001), Gourieroux and Robert (2001), Granger and Hyung (1999), Granger and Teräsvirta (1999), Liu (2000) and Mikosch and Starica (2000b).

This paper continues this line of work by considering the component GARCH (CGARCH) models introduced by Ding and Granger (1996) and Engle and Lee (1999), and their ability to capture the long range dependence in volatility. We consider a variant of the model in Engle and Lee (1999), which allows N separate components to contribute to the conditional variance. Each component allows the variance innovations to decay at a different rate. In the case of the two-component GARCH (CGARCH(2)) model, one component captures the long-run movements in volatility while the second component accounts for the noisier short-run movements. The CGARCH(2) model implies a restricted GARCH(2,2) model and therefore squared innovations follow an ARMA model (see Bollerslev (1986)). In addition, being a member of the GARCH family, all the theoretical results on GARCH models are applicable, including the recent results on stationarity (Giraitis, Kokoszka, and Leipus (2000)) and conditions for the existence of higher order moments (He and Terasvirta (1999)).

To analyze the performance of the CGARCH model we estimate and compare it to a benchmark GARCH(1,1), and FIGARCH(1,d,1) models for several daily equity return series and two foreign exchange (FX) rates. To measure the degree of long-memory that a CGARCH model can produce, we use Monte Carlo methods to characterize the finite sample distribution of several statistical measures of long-memory. For example, we simulate and compare the average sample ACF at long lags for the absolute value of return data generated from empirically realistic GARCH(1,1), CGARCH(2) and FIGARCH models. In addition we also characterize the finite sample distribution of two popular semiparametric estimates of the fractional differencing parameter when absolute returns are generated from these models.

The relationship between a two-component model of the conditional variance and the estimate of the long-memory parameter d from a FIGARCH model is explored. That is, using simulated data from a CGARCH(2) model, we consider the finite sample distribution of the fractional differencing parameter from a FIGARCH parameterization. Furthermore, this experiment is reversed, and results reported on the ability of the component GARCH structure to approximate long-memory in volatility when data is generated from a FIGARCH model.

In most cases, we find statistical evidence that a CGARCH(2) specification is favored over a GARCH(1,1) model. Conventional wisdom suggests an AR model with a long lag structure is necessary to accurately approximate a long-memory process. Although the CGARCH(2) model implies a short-memory ARMA structure in the squared innovations, we find that it provides a good fit to the sample autocorrelations of volatility from equity returns. Moreover, the CGARCH model produces data, that according to statistical measures of long-range dependence already discussed, suggests there is long-memory in volatility. Consistent with prior work, a GARCH(1,1) specification cannot account for the long-run time-series dependence in volatility.

The results are less clear for FX rates which are DEM-USD and JPY-USD returns. Overall the FIGARCH model provides the best description of the serial correlations found in DEM-USD volatility. However, none of the models can capture the autocorrelation structure in the volatility of JPY-USD rates. All models imply a decay in the ACF that is too fast. One reason the CGARCH(2) model performs worse for the FX market as compared to the equity market is the sample size is much smaller. In general, we found that a moderate to large dataset was needed to accurately identify the two components in volatility. Furthermore, one component must have a persistence parameter very close to 1 to approximate the dependence in volatility. We show a possible solution for smaller sample sizes is to impose a unit root for the trend component.

This paper is organized as follows. The next section defines long-memory and discusses its measurement. Section 3 introduces a component GARCH model of the conditional variance while Section 4 reviews the fractionally integrated GARCH model. Results are presented in Section 6 with conclusions in Section 7.

2 Defining and Measuring Long-Memory

According to Granger and Ding (1996) a series $\{y_t\}_{t=0}^{\infty}$ is said to have long-memory if it displays a slowly declining autocorrelation function (ACF) and an infinite spectrum at zero frequency. Specifically, the series y_t is said to be a stationary long-memory process if the ACF, $\rho(k)$ behaves as,

$$\rho(k) \approx c|k|^{2d-1} \text{ as } |k| \to \infty \quad (2.1)$$

where $0 < d < .5$ and c is some positive constant. The ACF in (2.1) displays a very slow rate of decay to zero as k goes to infinity and $\sum_{k=-\infty}^{\infty} |\rho(k)| = \infty$. This slow rate of decay can be contrasted with ARMA processes which have an exponential rate of decay, and satisfy the following bound,

$$|\rho(k)| \leq ba^k, \quad 0 < b < \infty \quad 0 < a < 1. \tag{2.2}$$

and consequently, $\sum_{k=-\infty}^{\infty} |\rho(k)| < \infty$. A process that satisfies (2.2) is termed short-memory. Equivalently, long-memory can be defined as a spectrum that goes to infinity at the origin. This is,

$$f(\omega) \approx c\omega^{-2d} \text{ as } \omega \to 0. \tag{2.3}$$

Additional definitions and statistical issues are dealt with in the surveys by Baillie (1996) and Beran (1994).

A simple example of long-memory is the fractionally integrated noise process, $I(d)$, with $0 < d < 1$ which is,

$$(1 - L)^d y_t = u_t, \tag{2.4}$$

where L is the lag operator, and $u_t \sim iid(0, \sigma^2)$. This model includes the traditional extremes of a stationary series, $I(0)$ and a nonstationary series $I(1)$. The fractional difference operator, $(1 - L)^d$ is well defined for a fractional d and the ACF of this process displays a hyperbolic decay consistent with (2.1). A model that incorporates the fractional difference operator is a natural starting point to capture long-memory. This is the motivation for the ARFIMA and FIGARCH class of models. However, not all long-memory processes are $I(d)$ processes. As mentioned in the introduction, nonlinear processes can also generate long-memory. Furthermore, an AR(p) model, usually with p very large, can provide an approximation to a long-memory process.

To assess whether a particular model can produce long memory, ideally we would like to study the population ACF to see if its behavior satisfies or approximates the above definition of long-memory. In practice, the theoretical ACF is not available and we must resort to estimating it with the usual sample quantities,

$$\hat{\rho}(k) = \frac{\frac{1}{T}\sum_{t=k+1}^{T}(y_t - \bar{y}_t)(y_{t-k} - \bar{y})}{\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y}_t)^2}, \quad \bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t. \tag{2.5}$$

Under fairly general conditions the sample ACF will be a consistent estimate of the population ACF, however, there may be cases where the sample ACF deviates substantially from theoretical quantities or the population ACF may not be defined. This could be due to finite sample issues, or because moments of the unconditional distribution do not exist. For example, Ding and Granger (1996) show the ACF of squared innovations is time-varying for an integrated GARCH (IGARCH) model and a stationary GARCH model with no fourth moment. The ACF can be approximated by a constant only for large samples and the autocorrelations display an exponential rate of decay. Similarly, Mikosch and Starica (2000a) and Davis and Mikosch (1998) show that for heavy tailed GARCH models interpreting the sample ACF of squared innovations can be problematic, particularly if the unconditional variance does not exist. For instance, if the model is an IGARCH the limiting distribution of the sample autocorrelations of the squared innovations is nondegenerate. On the other hand, if the unconditional variance exists but the 4th moment does not, the sample ACF of the absolute value of returns is consistent but converges to population values at a slower rate than the usual square-root of the sample size. Therefore, a drawback of using the sample ACF to identify long-memory is that our conclusions may be dependent on the sample size and the existence of higher-order moments. These issues are less likely to be a problem for the absolute value of returns which is used in this paper as a volatility proxy. In addition, to account for any finite sample properties when estimating the ACF we use Monte Carlo methods to provide confidence bands.

A second approach to measuring the degree of long-memory has been to use semiparametric methods. This allows one to sidestep the specific parametric form, which if misspecified, could lead to an inconsistent estimate of the long-memory parameter. We consider the two most frequently used estimators of the long memory parameter d. The first is the Geweke and Porter-Hudak (1983) (GPH) estimator, based on a log-periodogram regression. Suppose $y_0, ..., y_{T-1}$ is the dataset and define the periodogram for the first m ordinates as,

$$I_j = \frac{1}{2\pi T} \left| \sum_{t=0}^{T-1} y_t \exp(i\omega_j t) \right|^2 \tag{2.6}$$

where $\omega_j = 2\pi j/T$, $j = 1, 2, ..., m$, and m is a chosen positive integer. The estimate of d can then be derived from a linear regression of $\log I_j$ on a constant and the variable $X_j = \log |2 \sin(\omega_j/2)|$, which gives,

$$\hat{d} = -\frac{\sum_{j=1}^{m}(X_j - \bar{X}) \log I_j}{2\sum_{j=1}^{m}(X_j - \bar{X})}. \tag{2.7}$$

Robinson (1995b) provides formal proofs of consistency and asymptotic normality for the Gaussian case with $-.5 < d < .5$. The asymptotic standard error is $\pi/\sqrt{24m}$. The bandwidth parameter m must converge to infinity with the sample size, but at a slower rate than $\sqrt{T}$. Clearly, a larger choice of m reduces the asymptotic standard error, but the bias may increase. The bandwidth parameter was set to $\sqrt{T}$ in Geweke and Porter-Hudak (1983), while Hurvich, Deo, and Brodsky (1998) show the optimal rate to be $O(T^{4/5})$. Recently, Deo and Hurvich (2001) have shown that the GPH estimator is also valid for some non-Gaussian time-series, and Velasco (1999) has shown that consistency extends to $.5 \leq d < 1$, and asymptotic normality to $.5 \leq d < .75$.

The other popular semiparametric estimator is due to Robinson (1995a). The estimator is also based on the log-periodogram and solves,

$$\hat{d} = \underset{d}{\operatorname{argmin}} R(d) \tag{2.8}$$

$$R(d) = \log\left(\frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d} I_j\right) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j. \tag{2.9}$$

This estimator is asymptotically more efficient that the GPH estimator and consistency and asymptotic normality of $\hat{d}$ are available under weaker assumptions than Gaussianity. The asymptotic standard error for $\hat{d}$ is $1/(2\sqrt{m})$. Robinson and Henry (1999) have shown this estimator to be valid in the presence of some forms of conditional heteroskedasticity.

3 A Component GARCH Model

Define the information set Φ_t of daily returns to be $\{r_t, r_{t-1}, \ldots, r_1\}$. In the following we assume the conditional mean specification is,

$$r_t = \mu + \epsilon_t, \quad \epsilon_t = \sigma_t z_t \text{ and } z_t \sim i.i.d.(0,1). \tag{3.1}$$

To motivate a component GARCH model consider the well known GARCH(1,1) model introduced by Bollerslev (1986),

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3.2}$$

where σ_t^2 is measurable with respect to Φ_{t-1}. This model can be rearranged to,

$$\sigma_t^2 = \omega + \rho\sigma_{t-1}^2 + \alpha\nu_{t-1} \tag{3.3}$$

where, $\rho = \alpha + \beta$. $\nu_t = \epsilon_t^2 - \sigma_t^2$ is the innovation to the conditional variance, and is a martingale difference sequence with respect to Φ_{t-1}. This formulation emphasizes the autoregressive nature of the conditional variance in contrast to the ARMA representation of the squared innovations in Bollerslev (1986). The GARCH(1,1) specification in Equation 3.3 is similar to an AR(1) model in the conditional variance, and consequently it inherits many of the AR(1) model's features. For instance, in the case of $0 < \rho < 1$, (3.3) has an unconditional value of $\omega/(1-\rho)$, shocks (ν_t) affecting the conditional variance decay exponentially, and the speed of mean reversion in σ_t^2 is governed by ρ. In addition, ρ and α are important factors in determining the 4th and higher order moments of the unconditional distribution. The simple GARCH(1,1) model provides a limited degree of freedom to simultaneously capture the dynamics in the conditional and unconditional distributions. Additional properties of GARCH models can be found in the recent surveys by Gourieroux (1997), and Palm (1996).

A flexible alternative that decouples these dependences is the component GARCH model introduced by Engle and Lee (1999) and Ding and Granger (1996). Consider a two-component GARCH model labeled CGARCH(2),

$$\begin{aligned} \sigma_t^2 &= \sigma_{t,1}^2 + \sigma_{t,2}^2 &(3.4)\\ \sigma_{t,1}^2 &= \omega + \rho_1 \sigma_{t-1,1}^2 + \alpha_1 \nu_{t-1} &(3.5)\\ \sigma_{t,2}^2 &= \rho_2 \sigma_{t-1,2}^2 + \alpha_2 \nu_{t-1}. &(3.6) \end{aligned}$$

The extension to a N component model is obtained by adding extra components as in (3.6). In the CGARCH(2) specification two separate autoregressive components contribute to the overall conditional variance at time t. One component captures the long-run impact of an innovation while the second component captures the short-run transitory effect from a variance innovation. Each component reacts to the most recent innovation in a different way and with a different rate of decay associated with lagged variance innovations. For example, assuming $0 < \rho_i < 1, \ \ i = 1, 2$, then,

$$\begin{aligned} \sigma_{t,1}^2 &= \frac{\omega}{(1-\rho_1)} + \alpha_1(\nu_{t-1} + \rho_1 \nu_{t-2} + \rho_1^2 \nu_{t-3} + \cdots) \qquad (3.7)\\ \sigma_{t,2}^2 &= \alpha_2(\nu_{t-1} + \rho_2 \nu_{t-2} + \rho_2^2 \nu_{t-3} + \cdots). \end{aligned}$$

It is clear that the effect of a past innovation on the conditional variance is,

$$\frac{\partial \sigma_t^2}{\partial \nu_{t-k}} = \alpha_1 \rho_1^{k-1} + \alpha_2 \rho_2^{k-1}, \qquad (3.8)$$

while it is $\alpha\rho^{k-1}$ for the GARCH(1,1) model. Moreover, the expected value of the components in (3.7) is,

$$E\sigma_t^2 = E\sigma_{t,1}^2 = \frac{\omega}{(1-\rho_1)}, \text{ and } E\sigma_{t,2}^2 = 0. \qquad (3.9)$$

It is important to note that although the CGARCH(2) model relaxes the parameter restrictions for the unconditional variance and the speed of mean reversion in the GARCH(1,1) model, it still belongs to the GARCH class. In fact it is a restricted GARCH(2,2) model,

$$\begin{aligned} \sigma_t^2 &= (1-\rho_2)\omega + (\alpha_1+\alpha_2)\epsilon_{t-1}^2 - (\rho_1\alpha_2+\rho_2\alpha_1)\epsilon_{t-2}^2 \\ &\quad +(\rho_1+\rho_2-\alpha_1-\alpha_2)\sigma_{t-1}^2 - (\rho_1\rho_2-\rho_1\alpha_2-\rho_2\alpha_1)\sigma_{t-2}^2. \end{aligned} \tag{3.10}$$

As the reformulated CGARCH model (3.10) makes clear the individual components are not separately identified. For identification purposes we impose $\rho_1 > \rho_2$. As noted by Engle and Lee (1999) constraining coefficients to, $1 > \rho_1 > \rho_2 > 0$, $\rho_2 > \alpha_1 + \alpha_2$, $\omega > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$ ensures a positive conditional variance and is sufficient for covariance stationarity. Since the CGARCH model implies a GARCH model the vast body of results pertaining to GARCH models are applicable. Karanasos (2000) derives the theoretical autocorrelations for the squared innovations from a related component GARCH model while He and Terasvirta (1999) provide results on the fourth moment for the general GARCH(p,q) model.

In general, a stationary CGARCH model has an exponentially decaying autocorrelation function of ϵ_t^2. By combining several components into volatility, each having an exponential decay, provides a flexible structure to capture a slowly decaying autocorrelation function as seen in many financial time series.

4 Fractionally Integrated GARCH

The possibility of fractional integration in the conditional variance or the squared innovations has been noted recently by several authors, including Ding and Granger (1996), and Ding, Granger, and Engle (1993). Formally the Fractionally Integrated GARCH (FIGARCH) has been introduced in Baillie, Bollerslev, and Mikkelsen (1996). The FIGARCH(p,d,q) model of the conditional variance can be motivated as an ARFIMA model applied to the squared innovations,

$$(1-\phi(L))(1-L)^d\epsilon_t^2 = \omega + (1-\beta(L))\nu_t, \tag{4.1}$$

where $\beta(L) = \beta_1 L + \cdots + \beta_p L^p$, $\phi(L) = \phi_1 L + \cdots + \phi_q L^q$ and L is the lag operator, and $0 \leq d \leq 1$ is the fractional integration parameter. The roots of $(1-\beta(L))$ and $(1-\phi(L))$ are assumed to lie outside the unit circle. Recall that $\nu_t = \epsilon_t^2 - \sigma_t^2$ and rearranging (4.1) the FIGARCH(p,d,q) model can be expressed as,

$$\sigma_t^2 = \omega + \beta(L)\sigma_t^2 + (1-\beta(L)-(1-\phi(L))(1-L)^d)\epsilon_t^2. \tag{4.2}$$

The chief advantage of the FIGARCH(1,d,1) structure is that it parsimoniously decouples the long-run and short-run movements in volatility by adding 1 additional parameter to a GARCH(1,1) model. The long-run component is captured by the fractional differencing parameter d and the short-run component by the lag polynomials.

For the case of $d = 0$, the FIGARCH reduces to a standard GARCH(p,q) model ($\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$.) that is reparameterized such that $\phi_i = \alpha_i + \beta_i$. A $d = 1$ implies an IGARCH process of Engle and Bollerslev (1986). A $0 < d < 1$ is the FIGARCH model, and $(1 - L)^d$ can be defined through a Maclaurin series expansion as,

$$(1 - L)^d = 1 - dL + \frac{d(d-1)L^2}{2!} - \frac{d(d-1)(d-2)L^3}{3!} + \cdots \quad (4.3)$$

$$= \sum_{i=0}^{\infty} \frac{\Gamma(i-d)L^i}{\Gamma(i+1)\Gamma(-d)} \quad (4.4)$$

where $\Gamma(\cdot)$ is the gamma function. Baillie, Bollerslev, and Mikkelsen (1996) show this infinite expansion implies a slow hyperbolic rate of decay for coefficients of the lagged squared innovations and persistent impulse response weights. Since $(1 - L)^d$ enters the FIGARCH model in (4.2) for estimation and forecasting it must be truncated at a finite value. This gives rise to a truncation error which can be assessed by noting that if $L = 1$, then $(1 - L)^d = 0$, for $0 < d \leq 1$. In the following analysis the expansion in (4.3) is truncated at 1000 terms.

The FIGARCH process has an infinite unconditional variance and hence violates weak stationarity. However, Baillie, Bollerslev, and Mikkelsen (1996) suggest that the IGARCH results of Nelson (1990) and Bougerol and Picard (1992) can be used to show strict stationarity and ergodicity for $0 \leq d \leq 1$. Asymptotic normality and consistency of the quasi-maximum likelihood estimator (QMLE) have not been formally proven for the general FIGARCH(p,d,q) model, but the simulation results in Baillie, Bollerslev, and Mikkelsen (1996) show the QMLE to be well behaved.

5 Data

We consider daily returns from 3 equity indices and 2 foreign exchange (FX) spot rates. In all cases returns are defined as $r_t = 100 \log(p_t/p_{t-1})$, where p_t is the time t price of the asset. The equity data are the New York Stock Exchange (NYSE) composite index, 1966/1/4-2001/8/31, Standard and Poor's 500 (S&P500) Composite Index, 1928/1/4-2001/6/5, and the Dow Jones Industrial Average (DJIA), 1928/10/1-2000/1/11. FX data consists of spot rates for the German Mark/US dollar (DEM-USD), and the Japanese Yen/US dollar (JPY-USD) both from 1979/01/02 - 2001/05/16. In an effort to concentrate our analysis on the volatility dynamics a linear filter was applied to the conditional mean of returns to remove any low order autocorrelation. An AR(2) filter was used for both the S&P500, and DJIA returns, an AR(1) for the NYSE returns, and no adjustment was made to the FX returns. Subsequently, all results are for the filtered version of the data.

In analyzing long-range dependence in volatility we follow Ding, Granger, and Engle (1993), Ding and Granger (1996) and Wright (2000) in using $|r_t|$ for equity returns and $|r_t|^{.25}$ for FX returns, where the case for long-memory appears to be the strongest. Table 1 presents summary statistics for all return data. The lower part of this table reports semiparametric estimates of the long-memory parameter using the local Whittle estimator of Robinson (1995a), and the GPH estimator. Following Taqqu and Teverovsky (1996) and Andersen, Bollerslev, Diebold, and Labys (2001) we plotted estimates of d as a function of the bandwidth parameter, m and looked for a flat region that provides a reasonable tradeoff between bias and asymptotic variance. For a wide range of values, ($\sqrt{T} \leq m \leq T^{.8}$) we found $m = T^{.7}$ fit this criteria for all time-series and was used in computing Table 1 results. All estimates of d range from .3 to .36 with the exception of the Robinson estimator for the JPY-USD volatility which is .27. Evaluated by an asymptotic t-statistic, short memory ($d = 0$) is strongly rejected in all cases in favor of long-memory.

6 Results

First, we briefly summarize our empirical evidence and then proceed to a detailed discussion. Maximum likelihood estimates of GARCH(1,1), CGARCH(2) and FIGARCH(1,d,1) models assuming $z_t \sim N(0, 1)$ for our time series data are presented in Tables 2-6. We appeal to the quasi maximum likelihood estimator (QMLE) results of Bollerslev and Wooldridge (1992) which provides consistency and asymptotic normality for more general innovation distributions. Robust standard errors appear in parenthesis in the tables. At the bottom of these tables summary statistics are provided for the standardized residuals as well as the Ljung and Box (1978) (LB) statistic, $Q^2(30)$ for remaining serial correlation in the squared standardized residuals using 30 lags.

From the estimates we simulate samples from the respective models assuming normal and t-distributed innovations as well as sampling with replacement from the standardized residuals $\{\hat{\epsilon}_t/\hat{\sigma}_t\}_{t=1}^{T}$. In the case of t-distributed innovations the models are re-estimated assuming the correct distribution for z_t. In the following we present results based on sampling from the standardized residuals with additional results available from the author upon request. The size of a generated sample is set to the number of observations used in estimation.

Table 7 reports 95% confidence intervals of $\hat{d}$ from two semiparametric estimators of the long-memory parameter when the data are obtained from CGARCH(2) and FIGARCH data generating processes (DGP). Similarly, Table 8 reports 95% confidence intervals for $\hat{d}$ for the FIGARCH(1,d,1) model when samples are generated from the CGARCH(2) model. Figure 1 displays features of the CGARCH model estimated from S&P500 data. Figures 2-9 display the sample ACF of volatility from various models and examples. Panel A in these figures show the average sample ACF from the models and data while panels B and C report 95% confidence intervals for the CGARCH and FIGARCH results.

Model estimates contained in Tables 2-6 all suggest that a two component GARCH model provides a statistical improvement over a GARCH(1,1) (CGARCH(1)) model with the exception of the DEM-USD estimates. For example, for NYSE results in Table 2, the asymptotic t-statistics on ρ_2 and α_2 are both significant and the likelihood ratio (LR) test of the CGARCH(1) versus the CGARCH(2) model is 38.18 with a p-value of .512e-8 assuming standard asymptotic theory. All estimates of the CGARCH(2) models show there to be one very persistent component with ρ_1 close to 1 and a less persistent transitory component. Figure 1 plots the trend component, $\sigma^2_{t,1}$ and the transitory component, $\sigma^2_{t,2}$ of the conditional variance for S&P500 returns. In this example, the transitory component explains about 65% of the movements in the conditional variance. Panel C of this figure shows the effect of a past variance innovation on the conditional variance (Equation 3.8) for both the CGARCH and GARCH models. Clearly the two component version offers an added degree of flexibility not found in the GARCH(1,1) specification.

The FIGARCH estimates indicate the presence of long-memory, with an estimate of d being about .43 for all data series except the JPY-USD. These estimates are consistent with past studies. The long-memory parameter d, is highly significant for all series. For instance the t-statistics for $d = 0$ range from a low of 5.33 (DEM-USD) to a high of 10.02 (S&P500). Similar to other models the $Q^2(30)$ statistics suggest the FIGARCH model adequately captures the autocorrelations in the conditional variance. Note that for the DEM-USD results, the FIGARCH estimates are a local maximum as the CGARCH(1) model which is nested in the FIGARCH has a better loglikelihood value. This local optimum was robust to a range of starting values and we include it for comparison purposes.

Based on the parameter estimates reported in Tables 2 through 6 most GARCH(1,1) and CGARCH(2) models are covariance-stationary and satisfy the sufficient conditions of He and Terasvirta (1999) for a finite 4th moment. The only exception to this is the GARCH(1,1) model for S&P500 returns which violates the 4th moment condition. In the following, we simulate a return series from these models and compute the sample autocorrelations of the absolute value of returns. To account for the finite sample properties of the estimated ACF confidence bands are included.

The average sample ACF and confidence bands are calculated for $|r_t|$ in the case of equity returns and $|r_t|^{.25}$ for FX returns. These quantities are based on 20000 separate draws from the respective DGP. To minimize the effect of startup conditions on the results, the first 20000 observations in each draw from the DGP were discarded. Figures 2-6 display the average ACF for all 3 models as well as the 95% confidence intervals for the CGARCH and FIGARCH models. In each case, the GARCH(1,1) model displays the fast exponential rate of decay that is mentioned extensively in the literature as a reason why the GARCH class of models are inadequate for modeling long-range dependence. Surprisingly, the CGARCH model can produce an average ACF very similar to the FIGARCH model, which by construction has long-memory. For the NYSE, S&P500 and the DJIA returns there is very little to distinguish between the CGARCH and FIGARCH models. As measured by the confidence intervals, the CGARCH and FIGARCH specifications are both candidate models for producing the ACF observed in the NYSE data. However, the performance of both models deteriorates if we consider the S&P500 and DJIA results. For example, the average sample ACF in Figure 4 for the CGARCH and FIGARCH models is well below the data. Furthermore, based on the 95% confidence intervals both models appear to be inconsistent with the data for autocorrelations of 1800 or greater. A significant challenge for volatility models is to recreate the apparent plateau found in S&P500 and DJIA autocorrelations of volatility from 1000-2000.

For the equity series the average sample ACF from the CGARCH model closely matches the rate of decay from the FIGARCH model. However, the CGARCH specification does not perform as well for the shorter sample of FX rates. Figure 5 suggests that the FIGARCH model is doing a bit better in capturing the ACF of $|r_t|^{.25}$ for DEM-USD rates. All models are inadequate in capturing the structure for the JPY-USD rates as shown in Figure 6.

Our second piece of evidence on the CGARCH model's ability to capture long-range dependence comes from the finite sample distribution of the semiparametric estimator of the long-memory parameter listed in Table 7. This table reports 95% confidence intervals for $\hat{d}$ based on 1000 replications when data are generated from CGARCH and FIGARCH models. Results for several choices of bandwidth are presented. Overall, both models can produce estimates of long-memory typically seen in the data and in particular reported in Table 1. Based on the Robinson estimator with $m = T^{.7}$, both models bracket the estimates of d reported in Table 1 except for the JPY-USD. However, none of the models can account for the d estimates in Table 1 for both FX rates using the GPH estimator. Consistent with the

ACF estimates, the CGARCH model tends to produce a lower estimate of d for the FX returns. In some cases, for a bandwidth parameter $m = T^{.4}$ the confidence intervals of both the CGARCH and FIGARCH models contain 0. The CGARCH model is not fractionally integrated but does possess large autoregressive components, something that has been shown in Agiakloglou, Newbold, and Wohar (1993), Cheung (1993), and Geweke (1998) to cause large bias in semiparametric estimators of d.

To compliment the semiparametric estimates we also study the link between a two-component GARCH model and the distribution of $\hat{d}$ in a FIGARCH(1,d,1) model. That is, for each of the 5 DGPs of the CGARCH model we generated a sample of data and estimated a FIGARCH(1,d,1) model. 1000 replications were performed and from this 95% confidence intervals for the FIGARCH parameters are listed in Table 8. For the smaller datasets, the length of the confidence intervals is quite large, however for the 2 large datasets (S&P500 and DJIA) the confidence intervals are smaller and bracket values of d typically found in other studies. In addition, the average d reported in the last column of this table is remarkably close to the FIGARCH estimates of d in Tables 3 and 4. Figure 7 provides more information on the finite sample distribution of $\hat{d}$ with a density estimate in the case of the S&P500 CGARCH(2) DGP. The FIGARCH model in this experiment is misspecified, however these results underscore the tight link between a true long-memory volatility model and a two-component variance structure. To reverse this argument one might ask how well the CGARCH(2) can capture the dependence in a true long-memory volatility model? Figure 8 plots the average ACF of the CGARCH(2) model when data is obtained from one draw of the S&P500 FIGARCH DGP. The CGARCH specification provides a very good approximation to the serial correlations in this dataset.

Our results suggest that the CGARCH(2) specification captures the long-range dependence in volatility at least as well as a FIGARCH model for equity data. This does not appear to extend to the shorter FX data series. One possibility is that the persistence of the two components is not accurately identified for these shorter series. Indeed the asymptotic standard error for ρ_1 is larger for the JPY-USD data than for the equity data. A practical solution to using smaller datasets is to impose $\rho_1 = 1$. Figure 9 shows the ACF when a $\rho_1 = 1$ is imposed. Comparing this to Figure 6 shows a remarkable improvement.

In this paper we have confined ourselves to a two component version of the model. It is possible that a 3 component GARCH model may improve matters. We found that a third component was very difficult to identify for the shorter datasets. However, for the two longest equity series a 3 component GARCH model added very little to the CGARCH(2) results. With larger datasets constructed from high frequency intraday data it may be possible to identify more than 2 components in the conditional variance process. Indeed it would be interesting to see if our results continue to hold for data of a higher frequency.

In summary the results from the simulated autocorrelations, estimates of long-memory from semiparametric and parametric methods show that the CGARCH(2) specification can capture long-memory as defined by these statistics. Of course the CGARCH specification is not a long-memory model as defined in Section 2. Nevertheless, whatever the true long-range structure in financial market data, the CGARCH model can provide a good approximation to the unknown process under certain conditions. These conditions are: one component that has a persistence parameter very close to 1, and a moderate to large dataset to accurately identify the two components in volatility.

7 Conclusion

It has been suggested that combining several autoregressive components in volatility could provide a good description of the long-range dependence found in financial market volatility. This paper investigates this and finds a two component GARCH model can capture the long-range dependence in equity volatility at least as well as a FIGARCH model. The component model does not perform as well for the shorter sample of foreign exchange rates. Our results suggest that moderate to large datasets are needed for the component model to accurately capture long-range structure in volatility. A possible solution for smaller sample sizes is to impose a unit root for the trend component.

Model Summary

CGARCH(1)

$$
\begin{aligned}
r_t &= \mu + \epsilon_t, \quad \epsilon_t = \sigma_t z_t, \quad z_t \sim i.i.d.(0,1) \\
\sigma_t^2 &= \sigma_{t,1}^2 \\
\sigma_{t,1}^2 &= \omega + \rho_1 \sigma_{t-1,1}^2 + \alpha_1 \nu_{t-1} \\
\nu_{t-1} &= \epsilon_{t-1}^2 - \sigma_{t-1}^2
\end{aligned}
$$

CGARCH(2)

$$
\begin{aligned}
r_t &= \mu + \epsilon_t, \quad \epsilon_t = \sigma_t z_t, \quad z_t \sim i.i.d.(0,1) \\
\sigma_t^2 &= \sum_{i=1}^{2} \sigma_{t,i}^2 \\
\sigma_{t,1}^2 &= \omega + \rho_1 \sigma_{t-1,1}^2 + \alpha_1 \nu_{t-1} \\
\sigma_{t,2}^2 &= \qquad \rho_2 \sigma_{t-1,2}^2 + \alpha_2 \nu_{t-1}, \\
\nu_{t-1} &= \epsilon_{t-1}^2 - \sigma_{t-1}^2
\end{aligned}
$$

FIGARCH(1,d,1)

$$
\begin{aligned}
r_t &= \mu + \epsilon_t, \quad \epsilon_t = \sigma_t z_t, \quad z_t \sim i.i.d.(0,1) \\
\sigma_t^2 &= \omega + \beta \sigma_{t-1}^2 + (1 - \beta - (1-\phi)(1-L)^d)\epsilon_t^2
\end{aligned}
$$

Table 1: Descriptive Statistics: Daily Returns

	NYSE	S&P500	DJIA	DEM-USD	JPY-USD
Mean	-.773e-9	.802e-9	.124e-8	.003	-0.008
Stdev	.874	1.133	1.104	.676	.695
Skewness	-1.543	-0.368	-0.652	-0.166	-0.516
Kurtosis	42.309	26.117	31.941	4.979	6.703
Min	-20.589	-22.722	-25.487	-4.144	-5.710
Max	8.121	16.306	14.117	3.218	3.358
Observations	8979	19600	18745	5644	5644
Long-Memory Estimates	$\|r_t\|$	$\|r_t\|$	$\|r_t\|$	$\|r_t\|^{.25}$	$\|r_t\|^{.25}$
Robinson	.333	.368	.381	.304	.268
GPH	.348	.327	.366	.323	.301

This table reports summary statistics for daily returns, and semiparametric estimates of the long-memory parameter for $|r_t|$ and $|r_t|^{.25}$. Asymptotic standard errors for the Robinson and GPH estimator are $\pi/\sqrt{24m}$, and $1/(2\sqrt{m})$, respectively. m is the bandwidth parameter and set to $T^{.7}$, where T is the sample size for both semiparametric estimators.

Table 2: Model Estimates, NYSE Composite

	CGARCH(1)	CGARCH(2)		FIGARCH(1,d,1)
μ	.013 (.008)	.014 (.007)	μ	.015 (.008)
ω	.010 (.003)	.828e-3 (.863e-3)	ω	.022 (.008)
ρ_1	.988 (.004)	.999 (.001)	β	.610 (.081)
α_1	.081 (.019)	.013 (.006)	ϕ	.296 (.062)
ρ_2		.974 (.008)		
α_2		.076 (.017)		
			d	.424 (.077)
lgl	-10364.825	-10345.736		-10344.596
Residual Statistics, $\hat{\epsilon}_t/\hat{\sigma}_t$				
Skewness	-.475	-.424		-.409
Kurtosis	7.308	6.563		6.450
$Q^2(30)$	26.341 [.658]	26.300 [.660]		22.570 [.832]

This table reports ML estimates for GARCH(1,1), CGARCH(2) and FIGARCH(1,d,1) models. Bollerslev and Wooldridge (1992) robust standard errors appear in parenthesis and p-values appear in square brackets. $Q^2(30)$ is the LB statistic for remaining serial correlation in the squared standardized residuals using 30 lags.

Table 3: Model Estimates, S&P500

	CGARCH(1)	CGARCH(2)		FIGARCH(1,d,1)
μ	.029 (.008)	.029 (.007)	μ	.029 (.007)
ω	.013 (.003)	.001 (.626e-3)	ω	.028 (.007)
ρ_1	.995 (.004)	.999 (.772e-3)	β	.585 (.051)
α_1	.113 (.021)	.022 (.008)	ϕ	.283 (.040)
ρ_2		.947 (.017)		
α_2		.100 (.016)		
			d	.448 (.045)
lgl	-25196.981	-25097.770		-25096.242
Residual Statistics, $\hat{\epsilon}_t/\hat{\sigma}_t$				
Skewness	-.672	-.642		-.633
Kurtosis	11.174	9.964		9.986
$Q^2(30)$	22.500 [.835]	14.826 [.991]		12.040 [.998]

See notes to Table 2.

Table 4: Model Estimates, DJIA

	CGARCH(1)	CGARCH(2)		FIGARCH(1,d,1)
μ	.020 (.006)	.023 (.006)	μ	.022 (.006)
ω	.009 (.002)	.001 (.729e-3)	ω	.025 (.005)
ρ_1	.993 (.002)	.998 (.871e-3)	β	.609 (.043)
α_1	.079 (.012)	.017 (.007)	ϕ	.288 (.036)
ρ_2		.954 (.014)		
α_2		.075 (.011)		
			d	.429 (.044)
lgl	-23433.827	-23363.477		-23356.511
Residual Statistics, $\hat{\epsilon}_t/\hat{\sigma}_t$				
Skewness	-.583	-.5511		-.5481
Kurtosis	8.490	8.145		8.101
$Q^2(30)$	30.223 [.454]	16.350 [.979]		14.376 [.993]

See notes to Table 2.

Table 5: Model Estimates, DEM-USD

	CGARCH(1)	CGARCH(2)		FIGARCH(1,d,1)
μ	.007 (.008)	.007 (.008)	μ	.007 (.008)
ω	.010 (.003)	.004 (.002)	ω	.014 (.006)
ρ_1	.979 (.007)	.993 (.004)	β	.672 (.062)
α_1	.065 (.008)	.019 (.013)	ϕ	.304 (.043)
ρ_2		.965 (.015)		
α_2		.049 (.015)		
			d	.437 (.082)
lgl	-5532.828	-5531.356		-5533.389
Residual Statistics, $\hat{\epsilon}_t/\hat{\sigma}_t$				
Skewness	-.1519	-.155		-.159
Kurtosis	4.442	4.466		4.536
$Q^2(30)$	18.419 [.951]	19.345 [.932]		17.107 [.971]

See notes to Table 2.

Table 6: Model Estimates, JPY-USD

	CGARCH(1)	CGARCH(2)		FIGARCH(1,d,1)
μ	.183e-3 (.008)	.004 (.008)	μ	.004 (.009)
ω	.013 (.005)	.006 (.003)	ω	.037 (.016)
ρ_1	.973 (.010)	.987 (.006)	β	.561 (.125)
α_1	.057 (.013)	.031 (.010)	ϕ	.412 (.119)
ρ_2		.702 (.147)		
α_2		.071 (.023)		
			d	.266 (.048)
lgl	-5672.058	-5658.643		-5663.066
Residual Statistics, $\hat{\epsilon}_t/\hat{\sigma}_t$				
Skewness	-.485	-.482		-.487
Kurtosis	5.716	5.591		5.660
$Q^2(30)$	36.279 [.199]	31.880 [.373]		34.816 [.249]

See notes to Table 2.

Table 7: 95% Confidence Intervals of Long Memory Estimates from Simulated Models

	Robinson			GPH		
NYSE $\|r_t\|$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$
CGARCH(2)	(.10,.60)	(.21,.57)	(.32,.48)	(.07,.68)	(.19,.60)	(.31,.49)
FIGARCH(1,d,1)	(.18,.66)	(.26,.55)	(.27,.41)	(.16,.71)	(.25,.59)	(.26,.42)
S&P500 $\|r_t\|$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$
CGARCH(2)	(.33,.84)	(.30,.60)	(.34,.50)	(.28,.87)	(.29,.64)	(.34,.52)
FIGARCH(1,d,1)	(.24,.67)	(.31,.59)	(.33,.45)	(.20,.71)	(.28,.62)	(.32,.46)
DJIA $\|r_t\|$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$
CGARCH(2)	(.24,.60)	(.26,.56)	(.32,.48)	(.22,.73)	(.24,.60)	(.31,.50)
FIGARCH(1,d,1)	(.23,.64)	(.29,.56)	(.31,.42)	(.21,.68)	(.28,.60)	(.30,.43)
DEM-USD $\|r_t\|^{.25}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$
CGARCH(2)	(.002,.48)	(.16,.48)	(.18,.31)	(-.02,.52)	(.13,.50)	(.17,.32)
FIGARCH(1,d,1)	(.15,.64)	(.23,.53)	(.19,.32)	(.10,.70)	(.19,.56)	(.17,.31)
JPY-USD $\|r_t\|^{.25}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$	$m = T^{.4}$	$m = \sqrt{T}$	$m = T^{.7}$
CGARCH(2)	(.02.50)	(.16,.45)	(.12,.24)	(-.05.52)	(.12,.46)	(.11,.25)
FIGARCH(1,d,1)	(.01.45)	(.09,.36)	(.11,.22)	(-.05.51)	(.04.39)	(.10,.23)

This table reports 95% confidence intervals for various semiparametric estimators of the long-memory parameter for data simulated from models previously estimated in Tables 2-6. Data from a particular model was simulated by sampling from the standardized innovations with 1000 replications. m is the bandwidth parameter that determines the cutoff of periodogram ordinates.

Table 8: 95% Confidence Intervals for FIGARCH(1,d,1) Estimates when Data is Generated from a CGARCH(2) Model

Parameter	NYSE	S&P500	DJIA	DEM-USD	JPY-USD
ω	(.008, .039)	(.021,.053)	(.019,.048)	(.003,.0419)	(.002,.097)
β	(.468,.890)	(.391,.662)	(.420,.678)	(.392,.932)	(.094,.981)
ϕ	(.041,.320)	(.098,.290)	(.144,.319)	(.002,.377)	(-.071,.991)
d	(.355,.940)	(.365,.570)	(.328,.526)	(.276,1.009)	(.066,.369)
average d	.515	.458	.421	.434	.215

This table reports 95% confidence intervals for FIGARCH model estimates when data is generated from the respective CGARCH(2) DGP.
FIGARCH(1,d,1) model,

$$\begin{aligned} r_t &= \mu + \epsilon_t, \;\; \epsilon_t = \sigma_t z_t, \;\; z_t \sim i.i.d.(0,1) \\ \sigma_t^2 &= \omega + \beta\sigma_{t-1}^2 + (1 - \beta - (1-\phi)(1-L)^d)\epsilon_t^2. \end{aligned}$$

Figure 1: S&P500 Estimates

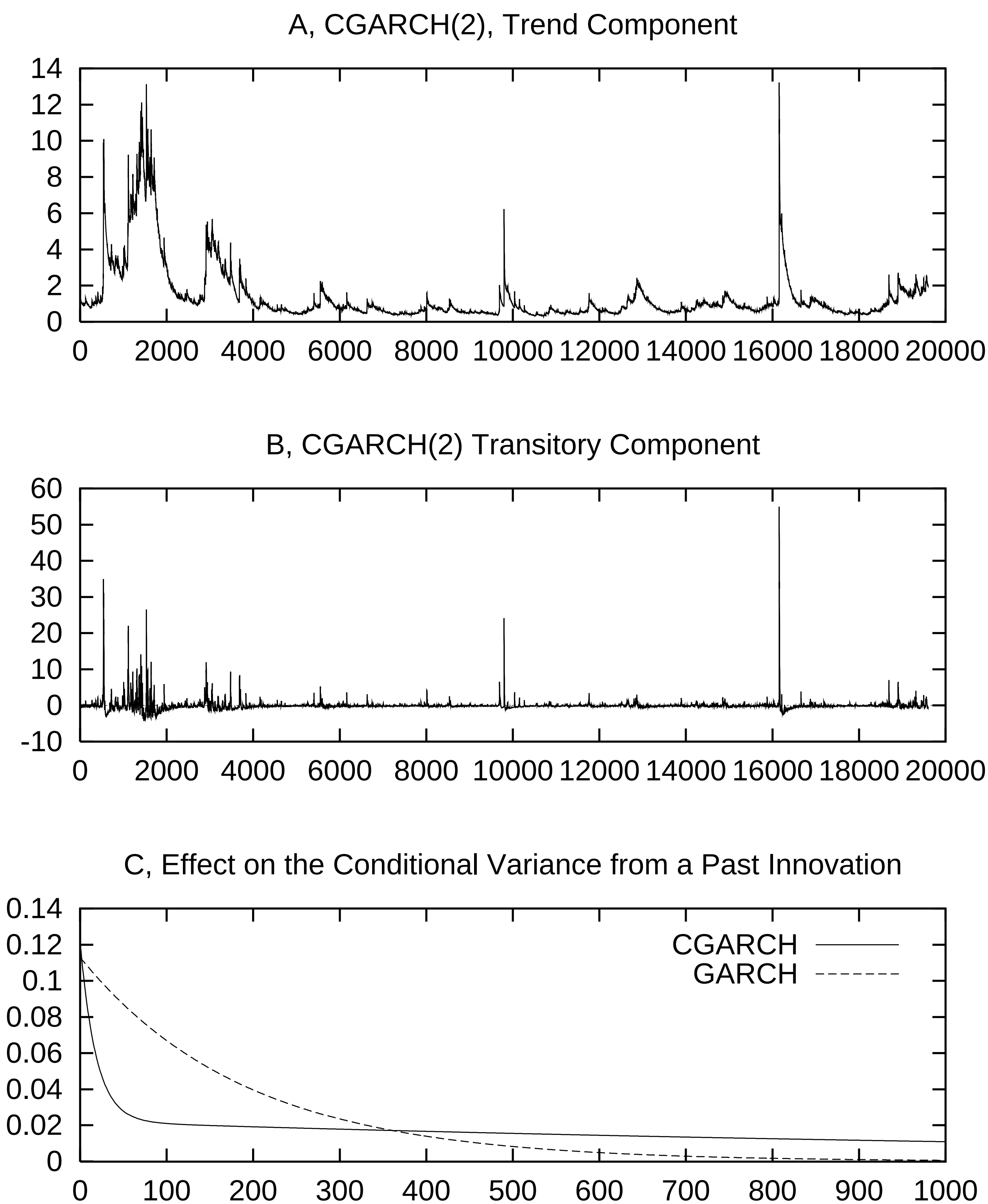


Figure 2: NYSE, Average Sample ACF of $|r_t|$, models and data

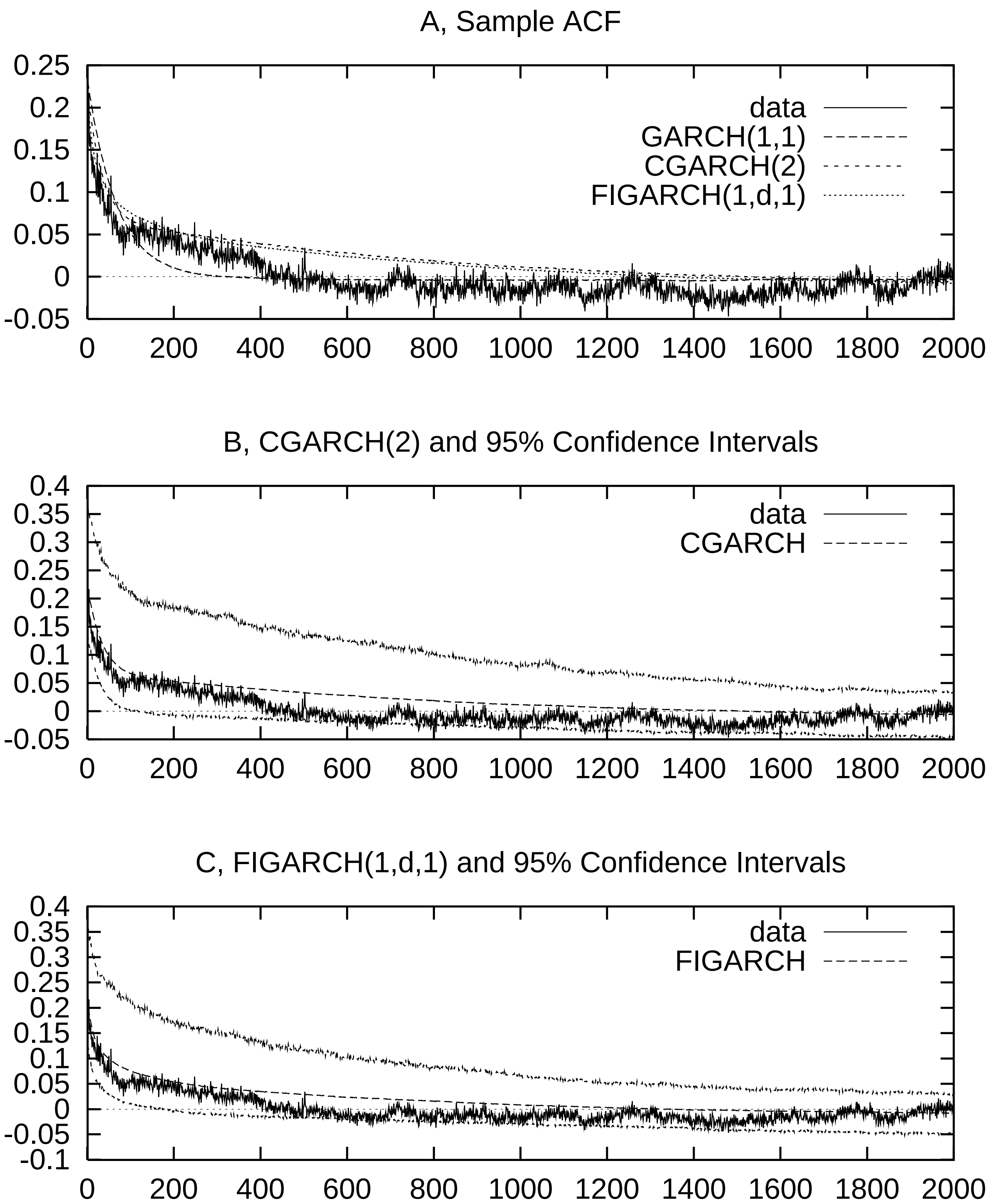


Figure 3: S&P500, Average Sample ACF of $|r_t|$, models and data

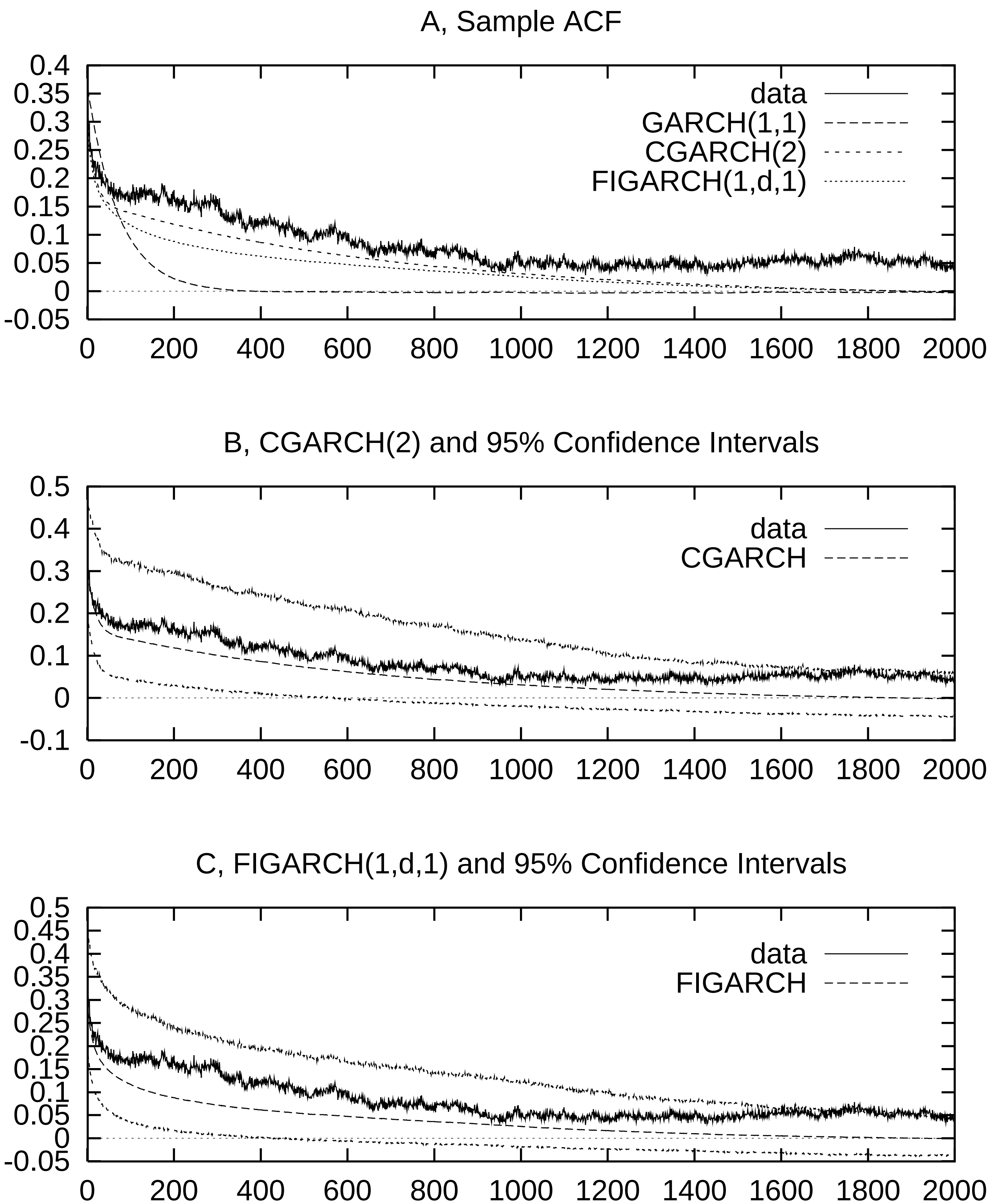


Figure 4: DJIA, Average Sample ACF of $|r_t|$, models and data

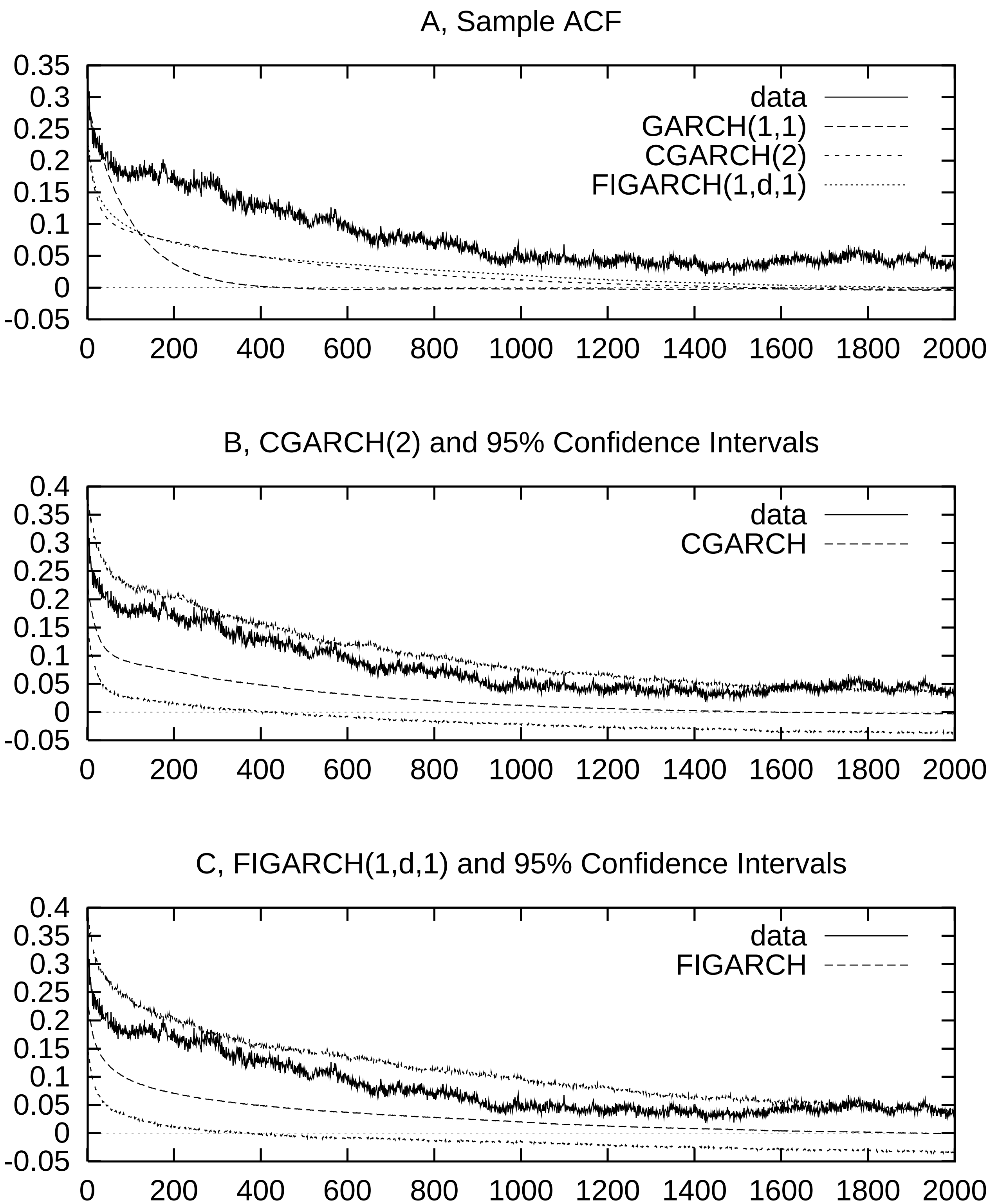


Figure 5: DEM-USD, Average Sample ACF of $|r_t|^{.25}$, models and data

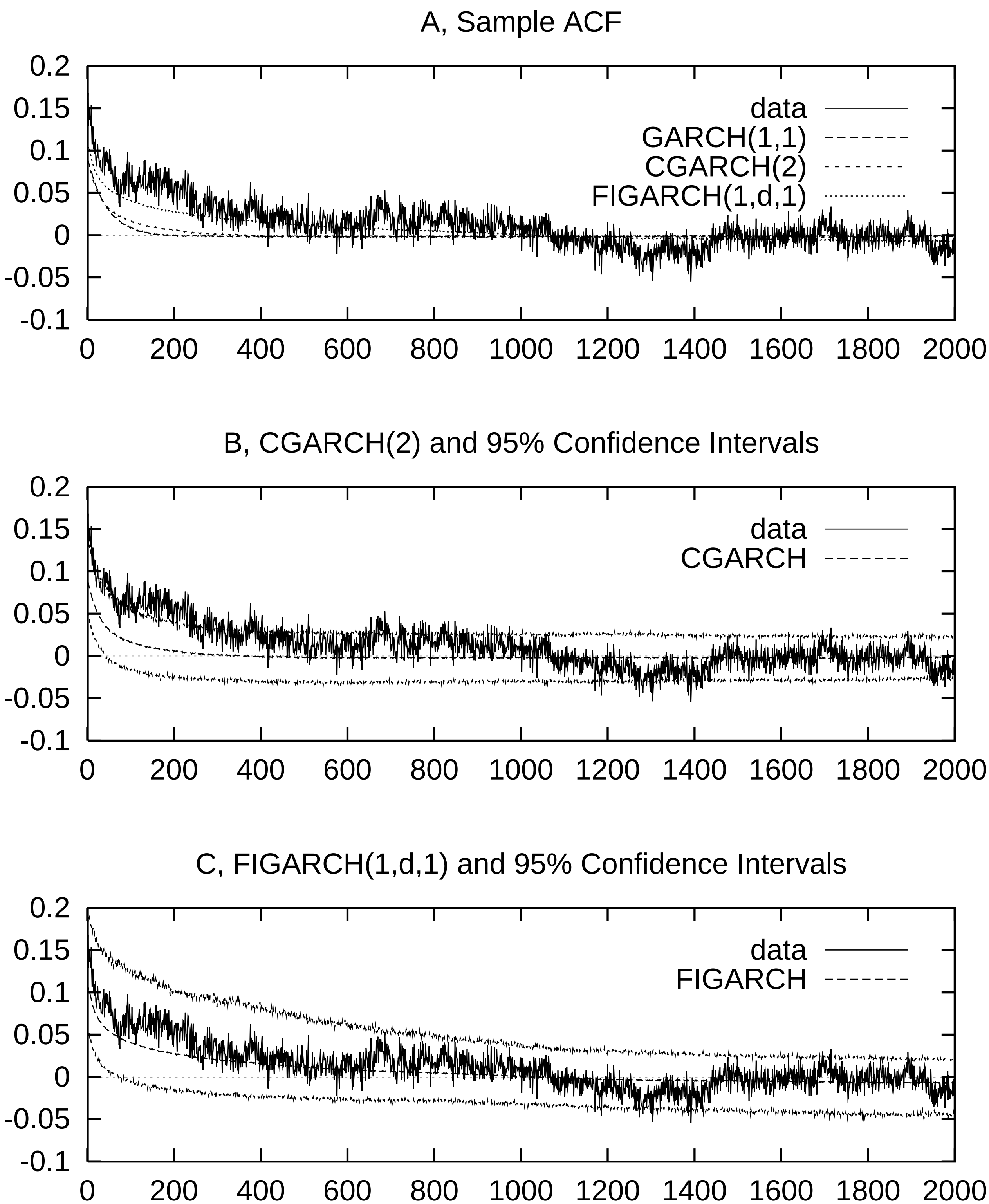


Figure 6: JPY-USD, Average Sample ACF of $|r_t|^{.25}$, models and data

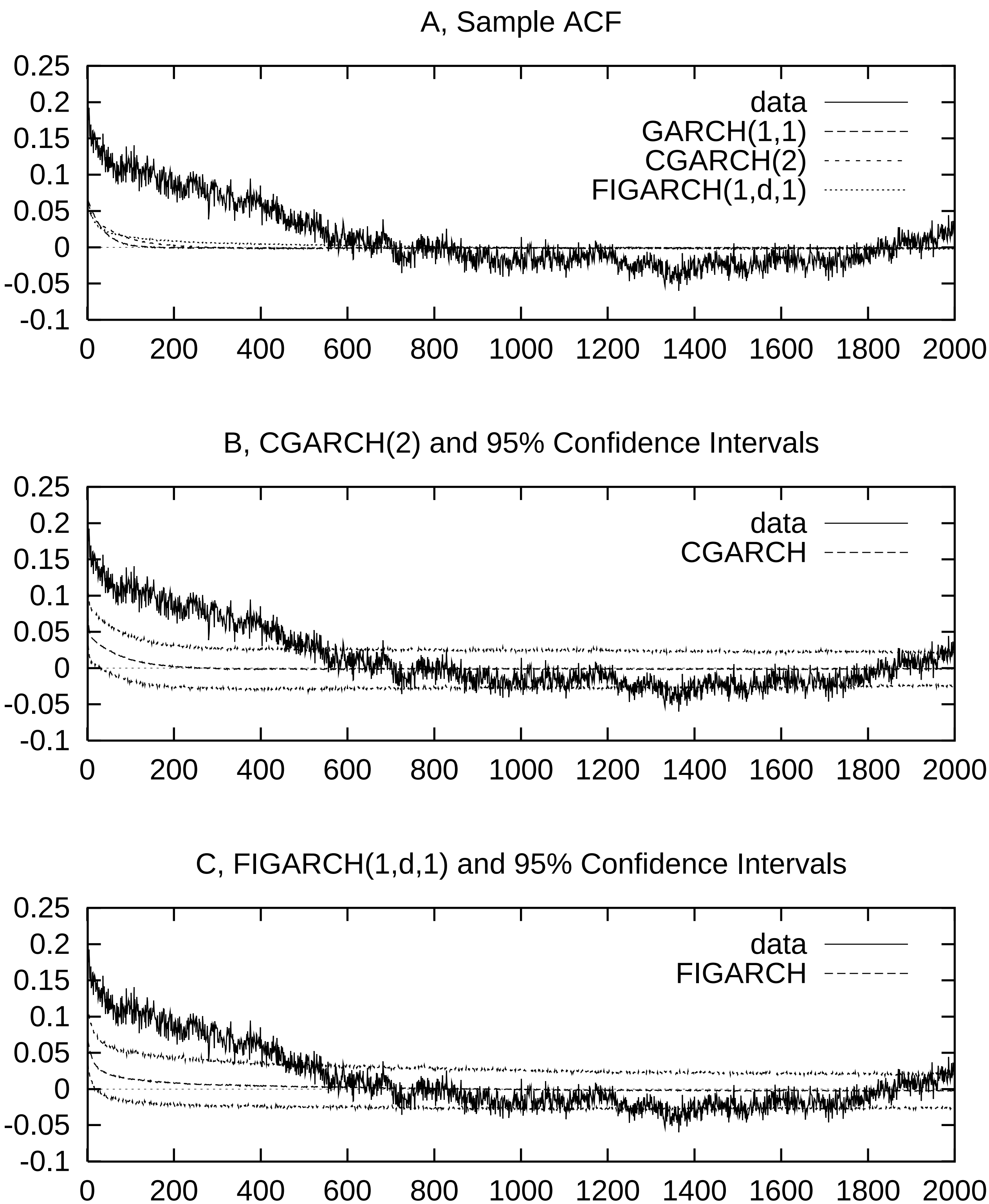


Figure 7: Density Estimate of $\hat{d}$, FIGARCH(1,d,1). Data generated from a CGARCH(2) model.

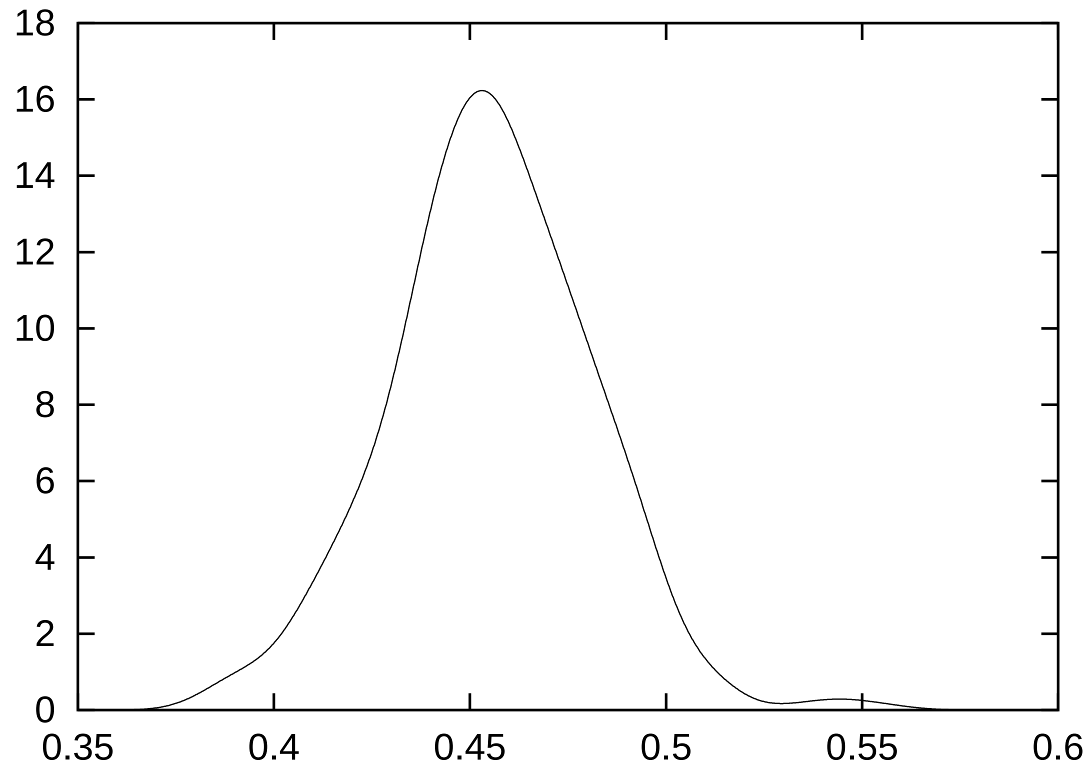


Figure 8: Average Sample ACF of $|r_t|$, data generated from a FI-GARCH(1,d,1) model

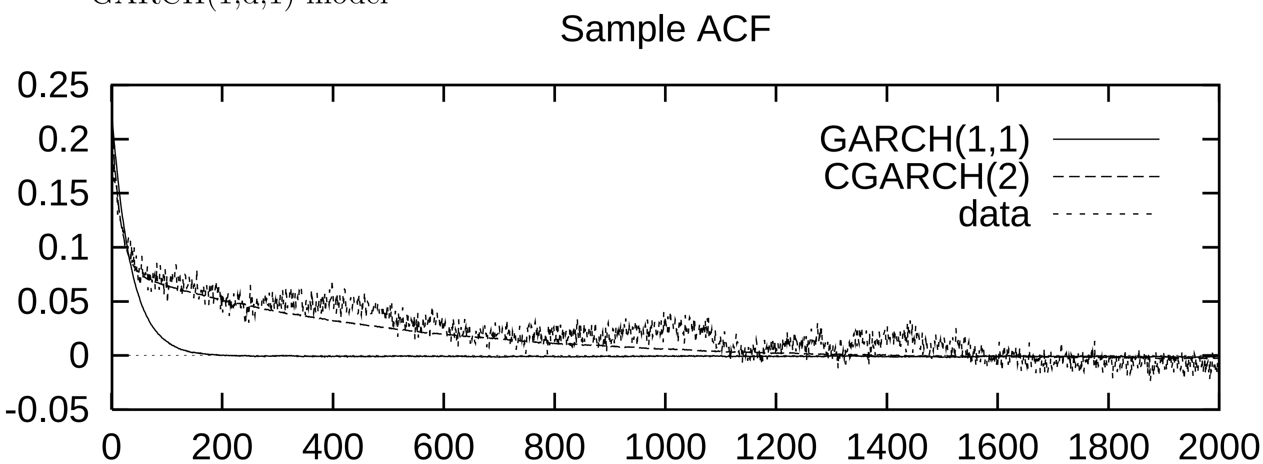
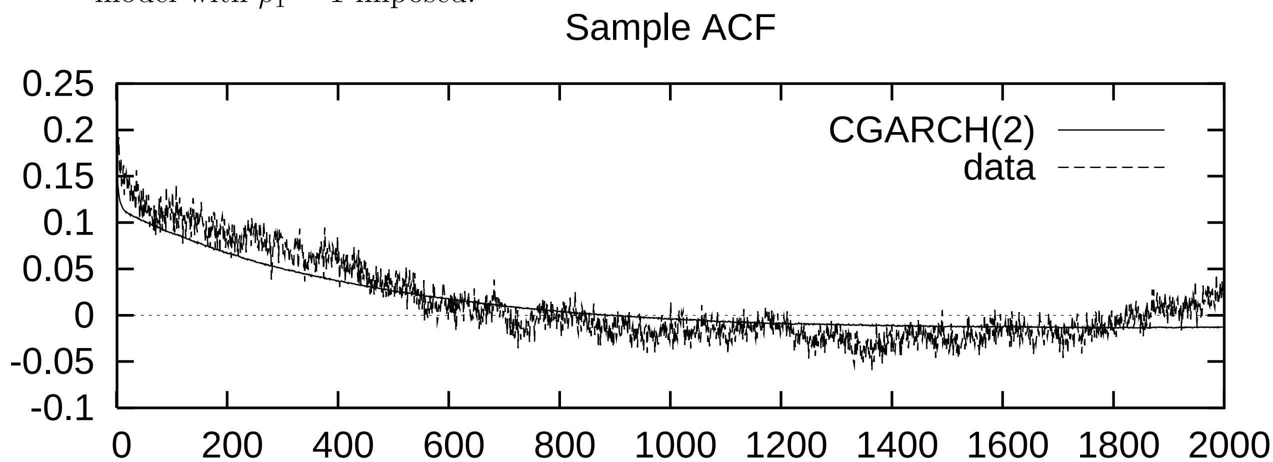


Figure 9: JPY-USD, Average Sample ACF of $|r_t|^{.25}$, data and CGARCH(2) model with $\rho_1 = 1$ imposed.



References

Agiakloglou, C., P. Newbold, and M. Wohar (1993): "Bias in an Estimator of the Fractional Difference Parameter," *Journal of Time Series Analysis*, 14(3), 235–246.

Andersen, T. G., and T. Bollerslev (1997): "Heterogenous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975–1005.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

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

Baillie, R. (1996): "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73(5), 5–59.

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

Baillie, R. T., A. A. Cecen, and Y. Han (2000): "High Frequency Deutsche Mark-US Dollar Returns: FIGARCH Representations and Non Linearities," *Multinational Finance Journal*, 4(3-4), 247–67.

Beltratti, A., and C. Morana (1999): "Computing Value at Risk with High Frequency Data," *Journal of Empirical Finance*, 6(5), 431–55.

Beran, J. (1994): *Statistics for Long-Memory Processes*, Monographs on Statistics and Applied Probability 61. Chapman and Hall.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 309–328.

Bollerslev, T., and H. O. Mikkelsen (1999): "Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics," *Journal of Econometrics*, 92, 75–99.

Bollerslev, T., and J. M. Wooldridge (1992): "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances," *Econometric Reviews*, 11, 143–172.

Bollerslev, T., and J. H. Wright (2000): "Semiparametric Estimation of Long-memory volatility dependencies: The Role of High-Frequency Data," *Journal of Econometrics*, 98, 81–106.

Bougerol, P., and N. Picard (1992): "Stationarity of GARCH Processes and of some Nonnegative Time Series," *Journal of Econometrics*, 52(115-128).

Breidt, F. J., N. Crato, and P. de Lima (1998): "The Detection and Estimation of Long Memory in Stochastic Volatility," *Journal of Econometrics*, 83, 325–348.

Brunetti, C., and C. L. Gilbert (2000): "Bivariate FIGARCH and Fractional Cointegration," *Journal of Empirical Finance*, 7, 509–530.

Cheung, Y. (1993): "Tests for Fractional Integration: A Monte Carlo Investigation," *Journal of Time Series Analysis*, 14(4), 331–345.

Davis, R. A., and T. Mikosch (1998): "The Sample Autocorrelations of Heavy Tailed Processes with Applications to ARCH," *The Annals of Statistics*.

Deo, R. S., and C. M. Hurvich (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 A. Inoue (2001): "Long memory and regime switching," *Journal of Econometrics*, 105(1), 131–159.

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

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.

Engle, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the UK inflation," *Econometrica*, 50, 987–1008.

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

Engle, R. F., and G. G. L. Lee (1999): "A Long-Run and Short-Run Component Model of Stock Return Volatility," in *Cointegration, Causality, and Forecasting*, ed. by Engle, and White. Oxford University Press.

Geweke, J. (1998): "Comment on Real and Spurious Long-Memory Properties of Stock Market Data," *Journal of Business & Economic Statistics*, 16(3), 269–271.

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.

Giraitis, L., P. Kokoszka, and R. Leipus (2000): "Stationary ARCH Models: Dependence Structure and Central Limit Theorem," *Econometric Theory*, 16, 3–22.

Gourieroux, C. (1997): *ARCH Models and Financial Applications*. Springer Series in Statistics.

Gourieroux, C., and J. Jasiak (2001): "Memory and Infrequent Breaks," *Economics Letters*, 70, 29–41.

Gourieroux, C., and C. Y. Robert (2001): "Tails and Extremal Behaviour of Stochastic Unit Root Models," manuscript, Dept. of Economics, University of Toronto.

Granger, C. W., and Y. Jeon (2001): "The Distributional Properties of Shocks to a Fractional I(d) Process Having a Marginal Exponential Distribution," *Applied Financial Economics*, 11(5), 469–74.

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 Z. Ding (1996): "Varieties of Long Memory Models," *Journal of Econometrics*, 73, 61–77.

Granger, C. W. J., and N. Hyung (1999): "Occasional Structural Breaks and Long Memory," Discussion Paper 99-14, Dept. of Economics, University of California, San Diego.

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, 15–39.

Granger, C. W. J., and T. Teräsvirta (1999): "A Simple Nonlinear Time Series Model with Missleading Linear Proporties," *Economics Letters*, 62, 161–165.

He, C., and T. Terasvirta (1999): "Fourth Moment Structure of the GARCH(p,q) Process," *Econometric Theory*, 15, 824–846.

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

Hurvich, C. M., R. S. Deo, and J. Brodsky (1998): "The Mean Squared Error of Geweke and Porter-Hudak's Estimator of the Memory Parameter of a Long-Memory Time Series," *Journal of Time Series Analysis*, 19(1), 19–46.

Hwang, Y. (2001): "Asymmetric Long Memory GARCH in Exchange Return," *Economics Letters*, 73(1), 1–5.

Karanasos, M. (2000): "Modelling Volatility Persistence: Some New Results on the Component GARCH Model," Discussion Paper 00-11, Dept. of Economics, University of York.

Liu, M. (2000): "Modeling Long Memory in Stock Market Volatility," *Journal of Econometrics*, 99, 139–171.

Ljung, L., and G. Box (1978): "On a Measure of Lack of Fit in Time-Series Models," *Biometrika*, 67, 297–303.

Lobato, I. N., and N. E. Savin (1998): "Real and Spurious Long-Memory Proporties of Stock Market Data," *Journal of Business & Economic Statistics*, 16(3), 261–268.

McCurdy, T. H., and P. K. Michaud (1997): "Capturing Long Memory in the Volatility of Equity Returns: a Fractionally Integrated Asymmetric Power ARCH Model," manuscript, Rotman School of Managment, University of Toronto.

Mikosch, T., and C. Starica (2000a): "Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process," *The Annals of Statistics*.

——— (2000b): "Long Range Dependence Effects and ARCH Modeling," University of Copenhagen, Laboratory of Actuarial Mathematics.

Nelson, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–39.

Palm, F. C. (1996): *GARCH Models of Volatility* chap. 7, Handbook of Statistics vol. 14. Elsevier Science.

Ray, B. K., and R. S. Tsay (2000): "Long-Range Dependence in Daily Stock Volatilities," *Journal of Business & Economic Statistics*, 18(2), 254–262.

Robinson, P. M. (1995a): "Gaussian Semiparametric Estimation of Long Range Dependence," *The Annals of Statistics*, 23(5), 1630–1661.

——— (1995b): "Log-Periodogram regression of time Series with Long Range Dependence," *The Annals of Statistics*, 23(3), 1048–1072.

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

Taqqu, M. S., and V. Teverovsky (1996): "Semi-Parametric Graphical Estimation Techniques for Long-Memory Data," in *Time Series Analysis in Memory of E. J. Hannan*, ed. by P. Robinson, and M. Rosenblatt. Springer-Verlag.

Velasco, C. (1999): "Non-stationary Log-periodogram Regression," *Journal of Econometrics*, 91, 325–371.

Wright, J. H. (2000): "Log-Periodogram Estimation of Long Memory Volatility Dependencies with Conditionally Heavy Tailed Returns," Board of Governors of the Federal Reserve System, International Discussion Paper, 685.