Modeling and Forecasting the Volatility of the Nikkei 225 Realized Volatility Using the ARFIMA-GARCH Model

ISAO ISHIDA
Faculty of Economics and
Graduate School of Public Policy
University of Tokyo

TOSHIAKI WATANABE
Institute of Economic Research
Hitotsubashi University

October, 2008

Preliminary and incomplete

Abstract

In this paper, we apply the ARFIMA-GARCH model to the realized volatility and the continuous sample path variations constructed from high-frequency Nikkei 225 data. While the homoskedastic ARFIMA model performs excellently in predicting the Nikkei 225 realized volatility time series and their square-root and log transformations, the residuals of the model suggest presence of strong conditional heteroskedasticity similar to the finding of Corsi et al. (2007) for the realized S&P 500 futures volatility. An ARFIMA model augmented by a GARCH(1,1) specification for the error term largely captures this and substantially improves the fit to the data. In a multi-day forecasting setting, we also find evidence of predictable time variation in the volatility of the Nikkei 225 volatility captured by the ARFIMA-GARCH model. A battery of specification tests including the BDS, CCK, and Hong-Li tests for detecting higher-order dependence are run. The results of these tests reveal various forms of misspecification remaining in the ARFIMA-GARCH model, which suggest the model can be further improved upon.

JEL classification: C22, C53, G15.

Keywords: ARFIMA-GARCH, Volatility of realized volatility, Realized bipower variation, Jump detection, BDS test, Hong-Li test, High-frequency Nikkei 225 data

1 Introduction

Volatility plays key roles in the theory and applications of asset pricing, optimal portfolio allocation, and risk management. This fact, together with the development of econometric tools for volatility analysis and empirical evidence for the predictability of the volatility of numerous markets, spurred the phenomenal growth of the volatility literature as well as the birth of an entire financial risk management industry during the last quarter century. It is well-documented by now that time-variation in financial market volatility is to some extent predictable but stochastic, hence volatility itself is volatile. This paper empirically investigates whether the volatility of the Japanese stock market volatility is predictably time-varying. The obtained empirical results indicate that it is indeed time-varying with some predictable component.

The pace of progress at the frontier of volatility research has had at least four notable upsurges. The first followed the invention of the ARCH model by Engle (1982) and led to the development of the first set of econometric procedures for the empirical analysis of time-varying volatility (see e.g. Bollerslev et al. 1994), and our deeper understanding of the empirical properties, e.g., volatility clustering, leverage effects in volatility, and fat-tails, of many financial time series. The second wave centered around the *stochastic volatility* (SV) modeling, which capitalized on and often contributed in turn to the concurrent development in the Bayesian statistical analysis using the Markov chain Monte Carlo procedure (see e.g. Shephard 2005). This paper mainly concerns modeling and forecasting of the volatility of the *realized volatility*, and is part of the currently continuing third upsurge involving *realized volatility* measures, which was ignited by the recent availability of intraday financial data collected near or at the tick-by-tick frequency and the need to harness such high-frequency data fraught with microstructure noise and apparent short-term seasonalities before the rich information contained can be tapped into. The seminal paper by Andersen and Bollerslev (1998) defined the sum of squared intraday returns as the realized volatility (RV) for the day, and proposed to use it as a proxy for the ex post realization of the daily volatility. The squared daily return, typically used as a measure of the ex-post daily volatility in earlier volatility prediction studies, is a very noisy, albeit unbiased, proxy for the conditional variance. On the other hand, under ideal conditions in the absence of microstructure effects, the RV not only is an unbiased and much less noisy measure of the conditional variance but also converges in probability to the integrated variance over the measurement period as the sampling frequency increases to infinity if the asset price follows a diffusion process. Hence, RV may be considered an almost observable measure of volatility. A further case for the RV can be made based on the work of Hansen and Lunde (2005a) and Patton (2005), who showed how the use of noisy proxies for the ex-post volatility such as the squared daily return may lead to a choice of an inferior volatility prediction model.

In a milestone paper, Andersen et al. (2003) used a Gaussian fractional VAR model, i.e., vector-ARFIMA model with MA order zero, for directly modeling and forecasting several exchange rate RV series, building on earlier empirical investigations that had found long-range dependence and approximate normality of the daily log RV time series constructed from high-frequency intraday data of exchange rates (e.g. Andersen et al. 2001a) and stock prices (e.g. Andersen et al. 2001b). They provided compelling empirical evidence for the superiority in predictive accuracy of this direct "reduced-form" daily RV modeling approach over the daily returns approach with short- or long-memory GARCH-type models. Earlier papers including Andersen et al. (1999), Blair et al. (2001) and Martens (2001) that explored how to take advantage of intraday data within the GARCH framework also underlie the shift of the focus of the volatility literature to this approach. Ebens (1999) was also among the first to apply the ARFIMA model directly to RV time series. His ARFIMAX model for the RV of the DJIA index portfolio returns incorporated terms to capture the leverage effect, a well-documented stylized fact about equity returns. Koopman et al. (2005) conducted an extensive forecast performance comparison study of the ARFIMA model for the RV series of the S&P 100 stock index and the unobserved components RV (UC-RV) model of Barndorff-Nielsen and Shephard (2002) as well as more traditional GARCH-type and SV models based on daily returns and their implied-volatility-augmented versions, and reported that the ARFIMA model outperformed the other models although the performance of the UC-RV model was nearly as good.

Corsi (2004) proposed the heterogeneous autoregressive (HAR) model for the RV as an alternative to the ARFIMA model. The HAR-RV model employs a few predictor terms that are past daily RVs averaged over different horizons (typically a day, a week, and a month), and is capable of producing slow-decay patterns in autocorrelations exhibited by many RV series. Because of the ease in estimation and extendability of the baseline model, the HAR model has quickly become popular for modeling the dynamics of RV and other related volatility measures. Corsi et al. (2005), after finding strong conditional heteroskedasticity and non-Gaussianity in the ARFIMA and HAR residuals of the RV of the S&P 500 futures, introduced the HAR-GARCH-NIG model, which augments the basic HAR model with NIG (normal inverse Gaussian) distributed standardized innovations having a GARCH volatility dynamic structure. Subsequent papers by Bollerslev et al. (2005) and Andersen et al. (2007a) also formulated the HAR errors as a GARCH process. Note that when the time series being modeled measures volatility, its volatility is related to the volatility of volatility of the primitive price process. The volatility of volatility of an asset price process is an important determinant of the tail property of the distribution of the asset's return, precise modeling of which is crucial, for example, for managing the extreme risk of a portfolio involving the asset, or pricing and hedging of out-of-the-money options written on the asset. Furthermore, a suite of volatility options and futures at the Chicago Board of Options Exchange (CBOE) and the CBOE Futures Exchange (CFE) written on U.S. equity market volatility indices, as well as a variety of over-the-counter volatility derivatives on major indices around the world, have been traded actively in recent years. For each of these products, the underlying is itself some measure of market volatility. As far as these volatility derivative products are concerned, the volatility of volatility is a second-moment property, rather than a fourth-moment property, of each of the respective underlying processes, and hence potential payoffs to accurate volatility-of-volatility modeling may be substantial.

In this paper, we use the ARFIMA model with the GARCH(1,1) specification for the error term (ARFIMA-GARCH), to empirically investigate the dynamic behavior of the volatility of the daily RV of the Nikkei 225 index. The Nikkei 225 is the most widely watched indicator of the overall moves of the Japanese stock market, one of the largest in the world in terms of capitalization and trading volume. Shibata (2004, 2008), Shibata and Watanabe (2004), Watanabe (2005), Watanabe and Sasaki (2006), Watanabe and Yamaguchi (2006) among others studied the RV of the Nikkei 225 index or index futures, and reported empirical findings similar to those obtained for other major markets[1]. Although neither a volatility index calculated and disseminated on a real-time basis by a major financial organization nor an exchange-traded volatility derivative has been introduced for any of the Japanese equity indices, Japanese-equity-related volatility derivatives have recently been traded actively over-the-counter. Several papers applied the ARFIMA-GARCH model to lower frequency macroeconomic and financial time series (e.g. Baillie et al. 1996, Ling and Li 1997, Ooms and Doornik 1999), and a large number of papers in the RV literature employ the ARFIMA model without a conditionally heteroskedastic error specification to fit daily RV series (e.g. Oomen 2001, Giot and Laurent 2004 as well as those already referenced above). To the author's knowledge, this paper is the first to apply the ARFIMA-GARCH model to RV time series. Although several recent RV studies used the HAR-GARCH in place of the ARFIMA-GARCH model primarily for ease of estimation, it is not difficult to estimate a low-order ARFIMA or ARFIMA-GARCH(1,1) model by the conditional sum of squares (CSS) estimator either. The CSS may be considered an approximate Gaussian maximum likelihood estimator with all pre-sample innovations of the series set to zero or the unconditional mean. One advantage of the

[1] See also Takahashi et al. (2007), which applied their novel Bayesian SV approach that uses both trading-hour RV and daily returns to the TOPIX, another widely watched index representing the Japanese stock market.

ARFIMA model is that it has the fractional integration parameter d explicitly incorporated into the model, allowing one to estimate it jointly with the other parameters, a feature not shared by the HAR model, which is not formally a long-memry model.

Another notable recent development in the RV literature is the approach due to Barndorff-Nielsen and Shephard (2004, 2006a) of decomposing the RV into the contribution of continuous sample path variations and that of jumps. Extending the theory of quadratic variation of semimartingales, Barndorff-Nielsen et al. (2006a) provided an asymptotic statistical foundation for this decomposition procedure under very general conditions; See also an exposition paper by Barndorff-Nielsen et al. (2006b). Andersen et al. (2007a) used the HAR framework to study the roles of these two distinct components in RV prediction while Andersen et al. (2007b) documented improvements in the RV forecasting accuracy achieved by modeling these components separately. In light of these results, we estimate and remove the jump contributions from the daily Nikkei 225 RV using the Barndorff-Nielsen procedure modified by Andersen et al. (2007), and study the continuous sample path variations as well as the raw realized volatility. Removal of the estimated jump component did not affect our empirical results very much, but given the recent interest in the literature in this methodology, we document both results. We find strong empirical evidence of conditional heteroskedasticity in the ARFIMA errors, and some evidence of predictability of the time variation in the volatility of the Nikkei 225 realized volatility.

The pace of progress in realized volatility research is yet to quiet down, but it is worth briefly mentioning before closing the introductory section that another new sub-area of active research has already emerged, triggered by the introduction of volatility indices by several major exchanges. This strand of the literature, which may be called the fourth wave of volatility research, studies the empirical properties of volatility indices directly, rather than treating volatility as a latent process and attempting to conduct inferences about it through GARCH-type models or SV models applied to daily asset returns data. By far the most prominent of the volatility indices is the CBOE Volatility Index (VIX). The VIX index is a measure of the market's expectation of the S&P 500 index volatility over the next 30 days implied by the prices of some of the S&P 500 options traded at the CBOE. As mentioned above, the CBOE and its futures exchange (CFE) already offer volatility options and futures based on some of their volatility indices, and obviously the statistical properties of the levels of these indices are of interest on their own right. Moreover, to the extent that the stock and options markets are integrated, the VIX as well as other indices formulated in a similar fashion may be regarded as the latent volatility of the underlying index turned into an essentially observable quantity. As is the case with the RV and related measures, this observability makes it possible to explore finer structures of volatility processes. Going beyond the analytically tractable class of affine jump diffusion models, Bakshi et al. (2006), Dotsis et al. (2007), and Duan and Yeh (2007) among others estimated the constant-elasticity-of-variance(-of-variance) and other continuous-time volatility models using daily VIX data and their empirical results are broadly in line with the predictable variation of the volatility of the Nikkei 225 realized volatility that we document in this paper.

The remainder of the paper is organized as follows. Section 2 briefly reviews the results from the theory of bipower variation and jump component extraction. Section 3 describes the data and summary statistics. Section 4 reviews the ARFIMA-GARCH model and reports estimation and forecasting results. Section 5 concludes.

2 Realized variance, realized bipower variation, and jump component extraction

The starting point of the realized volatility research is the recognition of the well-known result in the theory of continuous-time stochastic processes that the volatility of a process would be completely known if we observed a continuous record of the sample path of the process. Although in reality we do not obtain a continuous record and only observe the realized sample path of the process at discrete points in time, the sum of squared increments of the process approaches the integrated variance as the return measurement intervals shrink to zero. More precisely, if the process is a continuous semimartingale, under mild regularity conditions,

$$RV_t := \sum_{j=1}^{1/\Delta} |r_{t+j\Delta,\Delta}|^2 \xrightarrow{p} \int_t^{t+1} \sigma_s^2 ds \text{ as } \Delta \downarrow 0 \tag{1}$$

where $r_{t+j\Delta,\Delta}$ is the increment over the interval $[t+(j-1)\Delta, t+j\Delta]$ (in our context, the process is the log of the Nikkei 225 index level process so that $r_{t+j\Delta,\Delta}$ is the log return), σ_t is the diffusion coefficient (instantaneous volatility) of the process, time t has a daily unit so that RV_t is the tth day realized variance[2]. We will hereafter use the term *realized volatility* (RV) to refer both to RV_t defined in (1) and more loosely to other related measures such as the realized bipower variation defined below[3]. If the process is a semimartingale with finite-activity jumps, i.e., only a finite number of jumps occurring in any finite time interval, such as Poisson jumps, then the realized variance converges to the quadratic variation, which can be decomposed

[2]Stock exchanges (the Tokyo Stock Exchange in our case) are not open 24 hours a day. Strictly speaking, time t here is used in two different ways: when we refer to the tth day RV and when we divide the trading hours $[t, t+1]$ of a particular day into $1/\Delta$ intervals. Our notation here also glosses over the existence of a lunch break.

[3]Some authors preserve the term *realized volatility* strictly for $\sqrt{RV_t}$, and call RV_t the *realized variance*.

into the integrated variance (the continuous sample path variation) and the sum of squared jump sizes:

$$RV_t \xrightarrow{p} \int_t^{t+1} \sigma_s^2 ds + \sum_{t<s\leq t+1} \kappa_s^2 \text{ as } \Delta \downarrow 0 \tag{2}$$

where κ_s is the size of the jump occurring at time s. Barndorff-Nielsen and Shephard (2004, 2006c) showed that even in the presence of jumps the *realized bipower variation*

$$BV_t := \frac{\pi}{2} \sum_{j=2}^{1/\Delta} \left| r_{t+j\Delta,\Delta} \right| \left| r_{t+(j-1)\Delta,\Delta} \right| \xrightarrow{p} \int_t^{t+1} \sigma_s^2 ds \tag{3}$$

holds under mild conditions, and proposed to use

$$RV_t - BV_t \xrightarrow{p} \sum_{t<s\leq t+1} \kappa_s^2 \tag{4}$$

or

$$J_t^* := (RV_t - BV_t)^+ \tag{5}$$

as an estimator for the sum of realized squared jumps $\sum_{t<s\leq t+1} \kappa_s^2$. J_t^* is known to take non-zero, small values very frequently due to measurement errors and due possibly to the presence of jumps of infinite-activity types. Based on the asymptotic distributional theory for these quantities developed by Barndorff-Nielsen and Shephard (2004, 2006c) and Barndorff-Nielsen et al. (2006a) and an extensive simulation study by Huang & Tauchen (2005), Andersen et al. (2007a) introduced what they call a shrinkage estimator for the jump contribution

$$J_t := I\left(Z_t > \Phi_a\right) \cdot \left(RV_t - BV_t\right) \tag{6}$$

where I is an indicator function, $Z_t := \frac{(RV_t - BV_t) RV_t^{-1}}{\sqrt{\left((\pi/2)^2 + \pi - 5\right) \max\left(1, TQ_t BV_t^{-2}\right) \Delta}}$ is asymptotically standard normally distributed, $\mu_1 := \sqrt{2/\pi}$, $\Phi_a := \Phi(a)$ is the standard normal distribution function, and the (standardized) *realized tripower variation*

$$TQ_t := \Delta^{-1} 4^{-1} \pi^{3/2} \Gamma\left(\frac{7}{6}\right)^{-3} \sum_{j=3}^{1/\Delta} \left| r_{t+j\Delta,\Delta} \right|^{4/3} \left| r_{t+(j-1)\Delta,\Delta} \right|^{4/3} \left| r_{t+(j-2)\Delta,\Delta} \right|^{4/3} \xrightarrow{p} \int_t^{t+1} \sigma_s^4 ds \text{ as } \Delta \downarrow 0 \tag{7}$$

The last convergence result holds even in the presence of jumps. a is usually set to values such as .999 so that J_t picks up only "significant" jumps. With J_t, another estimator of the continuous sample path variation:

$$C_t := RV_t - J_t \tag{8}$$

may be used in place of BV_t.

In this paper, we use the microstructure-effects-robust versions of BV_t and TQ_t, due also to Andersen et al. (2007a). In these versions, the summands are respectively $(1-2\Delta)^{-1}\left|r_{t+j\Delta,\Delta}\right|\ \left|r_{t+(j-2)\Delta,\Delta}\right|$ and $(1-4\Delta)^{-1}\left|r_{t+j\Delta,\Delta}\right|^{4/3}\left|r_{t+(j-2)\Delta,\Delta}\right|^{4/3}\left|r_{t+(j-4)\Delta,\Delta}\right|^{4/3}$, skipping an interval of length Δ in sampling short-period returns. The definitions of J_t and C_t are modified accordingly as well.

3 Data and summary statistics

3.1 Calculation of five-minute returns from minute-by-minute Nikkei 225 data and RV measures

Nihon Keizai Shinbun, Inc. (Nikkei) computes and disseminates the Nikkei 225 index once every minute during the trading hours of the Tokyo Stock Exchange (TSE) (09:00-15:00 with a 11:00-12:30 lunch break)[4]. In this paper, we use minute-by-minute Nikkei 225 index data provided directly by Nikkei and maintained by the Center for Advanced Research in Finance at the University of Tokyo. The sample period spans March 11, 1996 through August 31, 2007. From the minute-by-minute data, we construct a series of five-minute log differences multiplied by one hundred, which we call "five-minute (percentage) returns." This choice is made to strike a balance between alleviating the microstructure-related noise and increasing the precision of volatility measurement, following the standard practice of the RV literature in handling high-frequency intraday data from highly liquid markets. Andersen et al. (2000) and the other empirical papers dealing with the RV of the Nikkei 225 cash or futures referenced in the introduction also used five-minute returns. For further discussions of the miscrostructure-related issues, see Hansen and Lunde (2005a) and references therein.

The official minutely Nikkei 225 index starts at 09:01 for the morning session, and the first five-minute return in a given day that we calculate is for the 09:05-09:10 interval as in Andersen and Bollerslev (2000). Removal of the first four observations partially alleviates possible effects of the sluggish response of the Nikkei 225 to the information accumulated overnight (or over a weekend or holiday) on volatility measurement. Given the TSE trading rules, the 9:05 index value is unlikely to have fully impounded all overnight information in it when overnight domestic and overseas events move the overall Japanese stock market up or down by a large amount relative to the previous close[5], and again our choice here is an attempt for a

[4] Most of the major TSE-listed companies are dually listed on the Osaka Securities Exchange (OSE), which closes at 15:10. The OSE is the only domestic exchange where the Nikkei 225 futures and options are traded. Nikkei, however, calculates the Nikkei 225 index based on electronic data feed from the TSE with generally higher volumes in individual stock trading.

[5] The TSE trading rules limit the range of prices at which shares of each individual stock is allowed to trade. For each stock,

balance between higher precision and lower bias. For the afternoon session, the index starts at 12:31. Since the effect of the TSE trading rules hampering the Nikkei 225 index from quickly reflecting the information accrued over a lunch break is likely to be much milder, the 12:31-12:35 interval is retained and used in our calculation of the first five-minute return.

For the end price of the last five-minute return calculation of each session, we use the closing price of the session. Due to such factors as delayed arrival of individual stock price data from the TSE and the real-time nature of the Nikkei 225 calculation and dissemination, the final few observations of each session are occasionally marked by time stamps up to several minutes later than 15:00 (11:00). For sessions with such observations as well, we use the last recorded observation of the session for closing the last five-minute interval. In total, there are 53 five-minute returns for a typical trading day, 23 from the morning session and 30 from the afternoon session. We exclude sessions from half trading days including the first and last trading days of each year from our sample, retaining 2,802 trading days.

We then calculate RV_t for each day using all five-minute returns from the day. For the other associated measures such as BV_t involving lagged returns, we do not treat the morning session's last few and the afternoon session's first few as consecutive observations since there is an intervening lunch break. This entails a loss of several more five-minute return observations for these measures. $RV_t^* := RV_t + R_{n,t}^2 + R_{l,t}^2$ may be used to define the RV for the day, where $R_{n,t}$ is the previous-day-close-to-open (15:00-9:05) overnight return and $R_{l,t}$ is the lunch break (11:00-12:30) return[6]. Since the TSE is open only 4.5 hours a day, however, it would be a stretch to treat RV_t^* an observed realization of the volatility for the whole day. We therefore concentrate on RV_t and related measures calculated from five-minute returns when investigating the behavior of the volatility of volatility.

3.2 Properties of the realized volatility and related measures

Summary statistics for various returns, RV measures, and their log and square-root transformations are presented on Tables 1. In addition to the sample skewness and kurtosis, the Jarque-Bera (JB) statistic is

---

this range is initially set around the reference price based on the last execution price, and when no trade takes place within it for five minutes due to an order imbalance, it is given a shift of a predetermined size once every five minutes until a trade takes place. The TSE announces either the upper or the lower end of the range as the "special quote." The special quotes of the constituent individual stocks, if there are any, are used in place of the more stale last transaction prices in the calculation of the Nikkei 225 index.

[6] In their empirical analysis of the Nikkei 225 futures RV, Shibata and Watanabe (2004) used RV_t^* and several other measures with more sophisticated weighting schemes that extend Martens (2002).

presented for each series. It is designed to jointly measure the deviations of the sample skewness and kurtosis from their respective population values, zero and three, under normality. For an $i.i.d.$ normal series, the JB statistic is asymptotically distributed χ^2 (two degrees of freedom). The simulation work of Thomakos and Wang (2003) has shown, however, that the JB statistic grossly overrejects the null of normality if the data are a sample path of a long-memory process. Hence, it should be treated here as an informal descriptive statistic.

For checking temporal dependence, the first-order sample autocorrelation and the Ljung-Box statistics of orders 5, 10, and 22 (corresponding to roughly one week, two weeks, and a month) for no serial correlations up to their respective orders are shown for each series. Since the usual Bartlett's standard error, $T^{-1/2} = 0.019$, is biased under heteroskedasticity, the heteroskedasticity-adjusted standard error for the first-order autocorrelation and Ljung-Box statistics due to Diebold (1988) are also presented. Previous empirical studies have documented that daily volatility measures such as the daily return squared, the absolute daily return, and various daily RV measures of financial time series appear to have long-memory properties. For long memory processes, the influence of shocks does not last forever unlike in the case of integrated processes, but decays very slowly relative to the geometrically fast decay for short-memory processes. Formally, there are several definitions of long memory. A usual definition of long memory for a covariance stationary time series, which we adopt in this paper, is that

$$\sum_{k=-\infty}^{\infty} |\gamma(k)| = \infty \tag{9}$$

holds where $\gamma(k)$ is the kth order autocovariance. Another definition of long memory for a covariance stationary process is that $\gamma(k)$ decays hyperbolically, i.e.,

$$\gamma(k) \sim k^{2d-1} l(k) \tag{10}$$

as $k \to \infty$, where l is some slowly varying function and $d < 1/2$ is called the *long-memory parameter*. If the process is a covariance stationary one satisfying some regularity conditions, (10) with $d \in (0, 1/2)$ implies (9). For a review of alternative definitions and their relations to each other, see e.g. Palma (2007).

Before estimating the ARFIMA and the ARFIMA-GARCH models for the RV series, we estimate the long-memory parameter d for our various series via two popular semiparametric estimators $\widehat{d}_{GPH}$, $\widehat{d}_{Robinson}$, due respectively to Geweke-Porter-Hudak (1983) (GPH) and Robinson (1995a), with the bandwidth parameter m set at $m = T^{0.7}$ where T is the sample size. The asymptotic standard errors for $\widehat{d}_{GPH}$ and $\widehat{d}_{Robinson}$ are respectively $\pi/\sqrt{24m}$ and $1/(2\sqrt{m})$, and hence the latter is asymptotically more efficient for a given m. d is equal to the d in the fractional differencing parameter d if the process is a fractionally integrated process, reviewed in the next section, and its definition can be extended to the nonstationary region $d \geq 1/2$. The region of the true d over which $\widehat{d}_{GPH}$ and $\widehat{d}_{Robinson}$ are consistent and asymptotically normal extends beyond $1/2$. For these properties, neither $\widehat{d}_{GPH}$ nor $\widehat{d}_{Robinson}$ requires Gaussianity, and the latter does not require conditional homoskedasticity. See Robinson (1995b), Velasco (1999), Deo and Hurvich (2001), and Robinson and Henry (1999).

Summary statistics for R_t, $R_{n,t}$, $R_{am,t}$ (9:05-11:00 returns), $R_{l,t}$, and $R_{pm,t}$ (12:30-15:00 returns) are presented on the top part of Table 1. All of these log return series have means insignificantly different from zero and exhibit evidence for nonnormal unconditional distributions (in particular $R_{am,t}$) and little evidence for the presence of autocorrelations except that there is some weak evidence of autocorrelations in $R_{l,t}$ and $R_{pm,t}$.

Looking at the summary statistics for $RV_{am,t}$ (the morning RV component), $RV_{pm,t}$ (the afternoon RV component) and $RV_t = RV_{am,t} + RV_{pm,t}$, their unconditional distributions all seem to be highly nonnormal with very large positive values of sample skewness and kurtosis. The LB statistics indicate that each of the three series is highly significantly serially correlated. The values of the first-order sample autocorrelations, 0.429 ($RV_{am,t}$), 0.355 ($RV_{pm,t}$), and 0.544 (RV_t), are at medium-persistent levels and well below one, but the values of $\widehat{d}_{GPH}$ (.460 for $RV_{am,t}$, .364 for $RV_{pm,t}$, and .470 for RV_t with standard errors .040) and $\widehat{d}_{Robinson}$ (0.454 for $RV_{am,t}$, 0.376 for $RV_{pm,t}$, and 0.468 for RV_t with standard errors .0311) are significantly positive but below the stationary/nonstationary border of 1/2 for all three, indicating that autocorrelations decay slowly for these series. Note, however, that for RV_t and $RV_{am,t}$, $\widehat{d}_{GPH}$ and $\widehat{d}_{Robinson}$ are within two standard errors from 1/2.

Deviations from normality seem to be vastly reduced by the square-root transformation, but remain large. The log transformation brings down the sample skewness and kurtosis values for each series even further and close to zero (-.177 for $RV_{am,t}$, -.182 for $RV_{pm,t}$, and -.130 for RV_t) and three (3.272 for $RV_{am,t}$, 3.260 for $RV_{pm,t}$, and 3.234 for RV_t) respectively. Each of these two transformations increases the values of the first-order sample autocorrelation, LB statistics, and the two semiparametric estimates of d. For example, for RV_t, the first-order sample autocorrelation increases to .663 ($\sqrt{RV_t}$) and .713 ($\ln RV_t$), and $\widehat{d}_{GPH}$, $\widehat{d}_{Robinson}$ increase to values in excess of 1/2 (557, .524 for $\sqrt{RV_t}$ and .584, .533 for $\ln RV_t$). Other than our point estimates of d being in the nonstationary region, these results are roughly in line with earlier empirical findings about RV measures constructed from high-frequency intraday exchange rate and stock returns data, which led to the popularity of the Gaussian ARFIMA model; See e.g. Andersen et al. (2003).

Additionally, summary statistics for $RV_{NW,t}$, $RV_{HL,t}$, $RV_{NWHL,t}$ and their respective transformed versions are shown on Table 1. $RV_{NW,t} := RV_{NW,am,t} + RV_{NW,pm,t}$ where $RV_{NW,am,t}$ and $RV_{NW,pm,t}$ are respectively the Newey-West-type estimates of the morning-session and afternoon-session RVs. $RV_{HL,t}$ ($RV_{HLNW,t}$) weights the four components $R^2_{n,t}$, $RV_{am,t}$, $R^2_{l,t}$, $RV_{pm,t}$ ($R^2_{n,t}$, $RV_{NW,am,t}$, $R^2_{l,t}$, $RV_{NW,pm,t}$) more efficiently. All of these estimators are due to Hansen and Lunde (2005a, 2005b). $RV_{NW,t}$ behaves similarly to RV_t and C_t, and we do not further report their dynamic properties in this paper. We do not further investigate $RV_{HL,t}$ or $RV_{NWHL,t}$ for the same reason we do not study RV_t^*.

Turning next to the Barndorff-Nielsen decomposition of RV_t into the contribution of squared jumps and that of continuous sample path variations, we set $a = .999$ in (6). Again, we present separate results for the morning, the afternoon, and the whole daily trading hours, but focus on the whole-trading-day statistics in our brief discussion. Summary statistics of the square-root and log ($\ln(1 + J_t)$ for J_t) transformed series are also presented on Table 1. The sample mean of $I(Z_t > \Phi_a)$, which is an estimate of the unconditional jump probability over the trading hours of a day, is .159, implying a little more than one jump occurrences per week[7]. On average, the jump contribution J_t comprises about 4% $(= .045/1.087)$ of the RV over trading hours. Not surprisingly, given the results of previous studies on the S&P 500 futures and other financial time series (e.g. Andersen et al. 2007a, Andersen et al. 2007b), J_t is distinctly less persistent although the unobserved conditional jump probability series (as opposed to the realized jump series J_t) might be more persistent. And the standard deviation of J_t (.154) is not negligible relative to that of C_t (1.03). Hence, using C_t, purged of the jump component with a different dynamic behavior, may reveal a higher-resolution picture of the dependence structure of C_t. For our analysis via the ARFIMA-GARCH model, we use C_t as well as RV_t. It turns out, however, that the ARFIMA-GARCH estimation results for C_t are similar to those for RV_t.

4 Modeling and forecasting the conditional mean and the conditional variance of the RV with the ARFIMA-GARCH model

4.1 The ARFIMA-GARCH model and its estimation

The ARFIMA model, introduced by Granger and Joyeux (1980), and Hosking (1981), is a natural extension of the ARIMA model for parsimoniously modeling time series with long memory. An ARFIMA(p, d, q) process $\{Y_t\}$ may be defined as a causal solution to

$$\phi(L)(1 - L)^d (Y_t - \mu) = \psi(L)\varepsilon_t \tag{11}$$

[7] Using a less stringent value $a = .99$, the estimated jump probability becomes .296, or an average once per 3.38 trading days.

where $\mu \in \mathbb{R}$, $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\psi(L) = 1 + \psi_1 L + \cdots + \psi_q L^q$ are respectively the AR and MA operators, sharing no common roots, $(1-L)^d$ is the fractional differencing operator, and $\{\varepsilon_t\} \sim WN(0, \sigma^2)$, $0 < \sigma^2 < \infty$. If all roots of $\phi(z)$ and $\psi(z)$ lie outside of the unit disk and $d \in (-1, 1/2)$ holds, then there is a unique covariance stationary solution which is invertible[8]. It can be shown that the autocovariance function of this solution satisfies

$$\gamma(k) \sim |k|^{2d-1} c \tag{12}$$

as $|k| \to \infty$ where c is a constant (see e.g. Palma 2007, p.48). In particular, it exhibits long memory if $d > 0$. Hence, the fractional integration parameter d in (11) corresponds to the long-memory parameter d in (12). Also of note is that an ARFIMA process with $d \geq 1/2$ is nonstationary but still mean-reverting as long as $d < 1$ (See Baillie 1996, p.21). If $d \in [1/2, 1)$, we may interpret $\{Y_t\}$ as a process, starting from some finite past and satisfying (11), which becomes a stationary ARFIMA($p, d-1, q$) process after being differenced once.

Following Ding et al. (1993) that found extremely slow decay patterns in the sample autocorrelation functions of daily volatility measures such as absolute returns, features meant to capture the long-range dependence in volatility have been incorporated into GARCH-type models (e.g. the FIGARCH model) and stochastic volatility models; see e.g. Baillie et al. (1996a), Bollerslev and Mikkelsen (1996), Breidt et al. (1998), Deo and Hurvich (2003). Since the daily RV time series appears to have long-memory properties and is a series of observed quantities, the ARFIMA model is a natural modeling choice. Baillie et al. (1996b) extended the ARFIMA model to include a GARCH specification for conditional heteroskedasticity and used it to analyze the inflation rate time series from the G7 and three other high inflation countries. See also Hauser and Knust (1998, 2001) for applications of the ARFIMA model with ARCH errors. In this paper, we apply the ARFIMA-GARCH(1,1) model in which $h_t := E_{t-1}\left[\varepsilon_t^2\right]$, the conditional variance of ε_t with respect to the sigma-field $\sigma(Y_{t-1}, Y_{t-2}, \cdots)$, is given the following formulation:

$$h_t = \omega + \beta h_{t-1} + \alpha \varepsilon_{t-1}^2 \tag{13}$$

For the stationary ARFIMA model with Gaussian homoskedastic errors, Sowell's (1992) algorithm for exact maximum likelihood estimation is available. For the ARFIMA-GARCH model, however, no closed-form expression for the exact likelihood function is available. Hence, we employ the conditional sum of squares

[8]Bondon and Palma (2007) recently proved invertibility for $d > -1$, relaxing the condition $d > -1/2$ often cited in the literature.

(CSS) estimator:

$$\widehat{\theta} := \underset{\theta}{\operatorname{argmax}} L(\theta) \tag{14}$$

where $\theta := (\mu, \theta_1')'$, $\theta_1 := (\phi_1, \cdots, \phi_p, \psi_1, \cdots, \psi_q, d, \omega, \alpha, \beta)'$ and

$$L(\theta) := -\frac{1}{2} \sum_{t=1}^{T} \left(\ln h_t(\theta) + Z_t(\theta)^2 \right) \tag{15}$$

$$Z_t(\theta) := \varepsilon_t(\theta) / \sqrt{h_t(\theta)}$$

$$\varepsilon_t(\theta) := \psi(L)^{-1} \phi(L) (1-L)^d (Y_t - \mu)$$

$$h_t(\theta) = \omega + \beta h_{t-1}(\theta) + \alpha \varepsilon_{t-1}(\theta)^2 \tag{16}$$

with the presample values of $Y_t - \mu$, $t = 0, -1, -2, \cdots$ set to zero and $h_1(\theta)$ set to some initial value. If $d \geq 1/2$, Y_t does not have an unconditional mean and μ cannot be interpreted as such. For this case, $(1-L)^d (Y_t - \mu) = (1-L)^{d-1} \Delta Y_t$ but note that $\Delta Y_1 = Y_1 - Y_0 = Y_1 - \mu$ under our choice of the presample values so that μ does not disappear. We equate $h_1(\theta)$ to the value of the sample variance as is often done in the estimation of GARCH-type models. In estimation, we do not impose $d < 1/2$ (plus $\alpha + \beta < 1$ for the GARCH specification) necessary for covariance stationarity. We estimate the ARFIMA with a homoskedastic specification, i.e., $\omega = \sigma^2$, $\alpha = \beta = 0$, by the CSS estimator as well for comparability of results across homoskedastic and GARCH specifications. The objective function maximized by the CSS estimator is essentially the Gaussian maximum likelihood (without the constant $-\frac{T}{2} \ln(2\pi)$) for the AR(∞) representation of the ARFIMA-GARCH model with the conditional distribution of ε_t specified as $\mathcal{N}(0, h_t)$, conditional on the presample values and h_1. Hence, the CSS estimator for the ARFIMA-GARCH model is a long-memory analogue of what we usually refer to as the (quasi-)maximum likelihood estimator, (Q)MLE, in the short-memory ARMA-GARCH setting. Note that the exact likelihood is not available in closed form for the latter case either. Extending the results of Beran (2004) for the ARFIMA model, Ling and Li (1997) showed that the CSS estimator is $\sqrt{T}$-consistent and asymptotically normal for the ARFIMA (p, d, q)-GARCH(P, Q) model. Their results are also valid for the nonstationary case in which the true value of d is larger than 1/2. This is a desirable property particularly because d estimates are found to be near or greater than 1/2 in our semiparametric estimation as well as other studies of financial time series data, and is another justification for using the CSS estimator. Of course, one could work with a differenced series, but with estimated d near the boundary $1/2$, one cannot be confident about the appropriateness of such a procedure Ling and Li (1997) derived $\sqrt{T}$-consistency for $\widehat{\theta}_1$ assuming that μ is known. Hence, strictly speaking, their results are not applicable to $\widehat{\theta}$. However, it seems reasonable to expect that $\left(T^{1/2-d} \widehat{\mu}, T^{1/2} \widehat{\theta}_1' \right)'$ has a Gaussian limiting distribution as long as the other conditions of Ling and Li (1997) are satisfied and

$d < 1/2$ [9]. Another caveat is that the asymptotic results of Ling and Li (1997) are based on the assumption $E\left[\varepsilon_t^4\right] < \infty$. If $\{\varepsilon_t\}$ follows a GARCH(1,1) process and the standardized error sequence $\left\{Z_t := \varepsilon_t h_t^{-1/2}\right\}$ is $i.i.d.$ with $\kappa := E\left[Z_t^4\right] < \infty$, this condition amounts to $\kappa\alpha^2 + 2\alpha\beta + \beta^2 < 1$. Our estimation results indicate that a covariance staionarity condition $\alpha + \beta < 1$, let alone the more stringent condition $\kappa < \infty$, may not be satisfied by some of the series we study. However, this is not likely to invalidate estimation and inference based on Ling and Li (1997). Hansen and Lee (1994) established (local) consistency and asymptotic normality of the Gaussian QMLE for the GARCH(1,1) model when the true values of α, β may not be in the covariance stationarity region $\alpha + \beta < 1$ but are in the strict stationarity region satisfying $E\left[\ln\left(\beta + \alpha Z_t^2\right)\right] < 0$. Jensen and Rahbek (2004), focusing on the estimation of (α, β) and working with a different set of assumptions on $\{Z_t\}$ proven that the consistency and asymptotic normality results of Hansen and Lee (1994) extend to the case of conditional variance explosion $E\left[\ln\left(\beta + \alpha Z_t^2\right)\right] > 0$. Unlike in our case, however, they both assumed $Y_t = \mu + \varepsilon_t$ with either μ a constant to be estimated jointly with the other parameters (Hansen and Lee 1994) or $\mu = 0$ known (Jensen and Rahbek 2004). To our knowledge, rigorous asymptotic theory for the ARFIMA-GARCH model with $\alpha + \beta \geq 1$ is not yet available in the literature.

Given the strong evidence to be reported shortly against conditional normality of the standardized error Z_t except when Y_t is $\ln RV_t$ or $\ln C_t$, we also present the robust standard errors of Bollerselv and Wooldridge (1992) interpreting our estimator as a QMLE. For the models with the GARCH specification, the Bollerslev-Wooldridge standard errors are robust to distributional misspecification of Z_t under correct specification of the conditional mean and variance and regularity conditions. We use the BIC[10] as our model selection criteria and confine our search for the best model to a total of 64 models: The full ARFIMA(2,d,2)-GARCH(1,1) model and its 63 restricted versions (at least one of the ARFIMA parameters, $\phi_1, \phi_2, \psi_1, \psi_2, d,$ is fixed at zero and/or no conditional heteroskedasticity $(\alpha, \beta) = 0$). μ, ω are always estimated with the other parameters. Let us denote, for example, the ARFIMA(2,d,0) model with the first-order AR coefficient restricted to be zero as the ARFIMA($\{2\}$,$d, 0$). We denote the other models with second-order terms similarly.

When modeling a time series of daily currency or stock returns, the conditional mean is small relative

[9] For the case of $d \geq 1/2$, Ling and Li (1997) considered estimating the parameters in $\phi(L)(1-L)^{d-m}\left((1-L)^m Y_t - \mu\right) = \psi(L)\varepsilon_t$ with $\mu \neq 0$ unknown and estimated separately from the other paramters, where m is the smallest positive integer such that $d - m < 1/2$. If $m = 1$, this implies that $\{Y_t\}$ has a linear time trend, which is counterintuitive in our case where Y_t is a measure of volatility.

[10] $BIC = -2L^*\left(\widehat{\theta}\right) + N \ln T$ where N is the number of parameters, $L^*\left(\widehat{\theta}\right)$ is the value of the maximized log likelihood function $L^*\left(\widehat{\theta}\right) = L\left(\widehat{\theta}\right) - T\ln(2\pi)$, and $L(\theta)$ is as defined in (15).

to the variance so that setting it to zero or a constant instead of fitting a more elaborate model usually does not much affect the estimation of the volatility process. Since the RV is highly persistent, it is essential to adequately model the conditional mean part of the RV even if one's primary interest is in the conditional variance of the RV. Otherwise, misspecification in the mean may masquerade as conditional heteroskedasticity. See Diebold and Nason (1990) for a study focusing on this issue.

Recently, the HAR model proposed by Corsi (2003) has also been frequently used in the RV literature because of its flexibility in terms of accommodating predetermined variables other than pure lagged dependent variables and ease of estimation; See e.g. Andersen et al. (2007), Bollerslev et al. (2005), Forsberg and Ghysels (2006), and Shibata (2008). The HAR model is a high-order AR model which restricts the parameters in an intuitively appealing way, and can also be extended to have conditionally non-Gaussian, conditional heteroskedastic error specifications; See Corsi et al. (2007). Although it can roughly match slowly decaying autocorrelation patterns exhibited by observations of many RV time series, the HAR model is formally not a long-memory process. Of course, lack of the long memory property in the formally defined sense per se is not a shortcoming of the HAR model. Long memory being an asymptotic property of $\gamma(k)$ as $|k| \to \infty$, the concept is strictly speaking not applicable to a finite segment of a time series anyway, and a primary objective of using long-memory-type models is to give a parsimonious approximation to the DGP that generates RV data with slowly decaying sample autocorrelations. For this purpose, both ARFIMA and HAR models seem to do an admirable job. Nevertheless, it is still of interest to estimate the long-memory parameter d in a simple unified framework provided by the ARFIMA model.

4.2 ARFIMA-GARCH estimation results

Many empirical studies of time series data set aside a hold-out sample for out-of-sample performance evaluation. In our case, however, we already know, based on the results reported in the literature, that the ARFIMA model would perform well as a time series model for the conditional mean for our entire sample, and we are not running a horse race of a myriad of models in out-of-sample forecasting. Rather, our objective is simply to get a first handle on the time-series behavior of the volatility of the Nikkei 225 RV using the ARFIMA-GARCH model. We do search for the best order within the ARFIMA(2,d,2)-GARCH(1,1) class, but it is done in an attempt to adequately filter out the conditional mean component before the volatility of the RV can be analyzed. Remaining serial correlations in the ARFIMA residuals may lead to spurious predictability of the volatility of the RV. Hence, we first use the entire sample (2802 daily observations) available to us to estimate the full ARFIMA(2,d,2)-GARCH(1,1) model and 63 restricted versions. For order selection, we employ the BIC.

In applying the ARFIMA-GARCH model, we work with transformed series $\ln RV_t$, $\ln C_t$ and $\sqrt{RV_t}$, $\sqrt{C_t}$ as well as the original series RV_t, C_t. It should be emphasized that, although the results for all six series are presented and discussed together and some comparisons in terms of model adequacy are made across the two series RV_t and C_t and across different transformations, each series is being investigated on its own right, and that the BIC values and other specification test statistics and forecast performance measures are not directly comparable across the six series. It is beyond the scope of this article to address the issue of which series should be used as the LHS variable of the ARFIMA-GARCH model in a forecast exercise with a particular loss function[11] although our results contain information relevant to this issue.

Table 3 summarizes the parameter estimation results. For brevity, only the results for the best short-memory model, the best homoskedastic model and its GARCH(1,1) counterpart, and the best overall model and its homoskedastic counterpart (if the best overall model is not a homoskedastic one). Table 4 presents summary statistics and some specification test statistics for the residuals of these models.

4.2.1 Estimation results for the ARFIMA model with the constant error variance specification

Our estimation results for the ARFIMA model with the constant error variance specification and various order restrictions are similar to those previously reported in the literature for the RVs of the Nikkei 225 cash index and index futures and other financial time series. The order selected by the BIC is the ARFIMA(0,d,1) for all series except $\ln RV_t$. For $\ln RV_t$, the ARFIMA(2,d,0) is selected. So the effect of filtering out the jump component is not so large as to influence the ARFIMA order selection for our sample except when log transformation is applied. The difference in the BIC values between the ARFIMA(2,d,0) and the ARFIMA(0,d,1) is rather small (Estimation details of the latter are not reported). Note that for all series the BIC selects an order (p,d,q) with either $p > 0$ or $q > 0$, but not both, and d not restricted to be zero.

The estimated values of d in the BIC-selected homoskedastic ARFIMA models are 0.496 (RV_t), 0.519 $(\ln RV_t)$, 0.539 $(\ln RV_t)$, 0.494 (C_t), 0.516 $\left(\sqrt{C_t}\right)$, 0.509 $(\ln C_t)$. At usual levels, all of these are significantly higher than zero, but none are significantly different from the nonstationarity boundary of 1/2. Hence it is hard to conclude from $\widehat{d}$ whether these series are nonstationary.

On Table 4, the first-order sample serial correlations (the Bartlett's standard errors $T^{-1/2} = 0.019$, and the heteroskedasticity-adjusted standard errors given in parentheses) and the LB portmanteau statistics (the first lines) for testing the null of no serial correlations of orders up to 5, 10, and 22 (roughly corresponding

[11] Suppose, for example, that our goal is to minimize the mean squared errors in forecasting RV_t one-step-ahead. Since $\exp\left(E_{t-1}\left[\ln RV_t\right]\right) \neq E_{t-1}\left[RV_t\right]$ by Jensen's inequality, we need to specifiy more than the first two condtional moments of $\ln RV_t$ to produce an optimal forecast for RV_t if we are to work with $\{\ln RV_t\}$ instead of directly with $\{RV_t\}$.

to a week, two weeks, and a month) are presented on Table 4. The sample autocorrelations are smaller than 0.01 in absolute magnitude and are hence insignificant at usual levels even using $T^{-1/2} = 0.019$, which is much smaller than the heteroskedasticity-adjusted versions. The values of the LB statistics for the residuals of the models of the raw series RV_t and C_t are apparently large enough, whether or not the degrees of freedom is reduced by the number of estimated parameters in obtaining the χ^2 critical values, to indicate that serial correlations are not adequately filtered out by the selected ARFIMA models although for the log and square-root transformed series they are not significant. However, in the presence of conditional heteroskedasticity, the Bartlett's standard error is not a consistent estimator of the standard deviations of the sample serial correlations, and consequently the LB statistic is not asymptotically distributed $\chi^2(k)$ under the null of no serial correlations of orders up to k. In particular, if the squared series is positively autocorrelated, the Bartlett's standard error overestimates the estimation precision of sample autocorrelations and as a result the LB statistics overreject the null. Note that under the null of conditional homoskedasticity, the squared ARFIMA errors are serially uncorrelated, but that under correct specification of the conditional mean only, the squared residuals may be serially correlated while $\left(\widehat{\mu}, \widehat{\phi}_1, \cdots, \widehat{\phi}_p, \widehat{\psi}_1, \cdots, \widehat{\psi}_q\right)$ may still consistently estimate the true values. Therefore, we also report Diebold's (1988) heteroskedasticity-adjusted LB statistics (on the second lines), according to which there is no strong evidence of remaining serial correlations in the residuals of any of the best homoskedastic models. The reductions in the LB statistics due to the adjustment are particularly substantial for the residuals of the models of the raw series RV_t and C_t.

Although the estimates of the long-memory parameter d, when not restricted to be zero, are significantly greater than zero and the BIC favors the long-memory models, the heteroskedasticity-adjusted LB statistics for the residuals of the BIC-selected short-memory models are only marginally significant for the raw series RV_t and C_t and insignificant for the four transformed series. For each series, the best short-memory ARMA model achieves this by mobilizing more AR and MA terms than the best homoskedastic long-memory model, but yet are not quite as successful[12].

The very large values of sample skewness and kurtosis of the residuals from the raw series RV_t and C_t shown on Table 4 suggest that the unconditional distributions of the error term are highly nonnormally distributed. The square-root transformation of RV_t and C_t vastly reduces the values of these measures of nonnormality, but still far above those of a normal distribution. The log transformation bring the residual distributions closer to normality, but the sample kurtosis values are still nearly 4.

[12]To save space, only the results for the BIC-selected models among the short-memory models with or without the GARCH specification are presented. The same is true, however, when the best models are selected from among the homoskedastic models.

In sum, it appears that parsimonious ARFIMA models are reasonably successful in removing serial correlations although some evidence of remaining serial correlations is found for the residuals of the models of the raw series RV_t and C_t.

The discrepancies between the values of the unadjusted LB statistics and those of the heteroskedasticity-adjusted ones already indicate the presence of conditional heteroskedasticity in the form of serial correlations in ε_t^2. Next we directly test the null of no serial correlations in ε_t^2. McLeod and Li (1983) showed that, under the null of homoskedasticity, the asymptotic distribution of the sample autocorrelations in the squared residuals $\widehat{\varepsilon}_t^2$ is multivariate standard normal, not affected by the replacement of the unobservable errors ε_t of an ARMA model by the residuals $\widehat{\varepsilon}_t$. Hence, parameter estimation does not impact the asymptotic null distribution of the LB statistic, called the McLeod-Li statistic in this context, which is $\chi^2(k)$ using sample serial correlations of k different orders in the squared residuals. The McLeod-Li statistics for the squared residuals from the selected models are all very large, cleanly rejecting the null of no serial correlations of orders up to 5, 10, and 22 at usual levels. With this evidence, we turn to modeling conditional heteroskedasticity with the ARFIMA model augmented by the GARCH(1,1) specification for the error term.

4.2.2 Estimation results for the ARFIMA-GARCH(1,1) model

We next compare all 64 versions of the ARFIMA(2,d,2)-GARCH(1,1) model including those without the constant error variance restriction $\alpha = \beta = 0$. For each of the six series, the BIC selects a model with the GARCH(1,1) specification, more specifically, the ARFIMA(2,d,0)-GARCH(1,1) for RV_t, $\sqrt{RV_t}$, the ARFIMA(0,d,1)-GARCH(1,1) for $\sqrt{C_t}$, $\ln C_t$, and the ARFIMA($\{2\}$,d,1)-GARCH(1,1) for $\ln RV_t$, C_t. The estimation results for the selected model are shown at the bottom of the section of Table 3 for each series, together with the results for the homoskedastic version of it. As expected from the LB statistics for the squared residuals from the conditionally homoskedastic models, the improvement in the log likelihood value achieved by giving the GARCH(1,1) specification for the error process is substantial for each series. Also as expected, the addition of the GARCH(1,1) specification to the homoskedastic model of the same ARFIMA order does not alter the ARFIMA parameter estimates very much in most cases, and consequently the distributions of the residuals remain similarly nonnormal (Table 4). In spite of this, for RV_t, $\sqrt{RV_t}$, $\ln RV_t$, C_t, the ARFIMA order changes when the GARCH(1,1) specification is allowed, which is not surprising since the differences in the BIC values amongst the highest BIC-ranking homoskedastic versions are small.

As for the estimation of the GARCH equation, we first discuss the results for the raw series RV_t, C_t. The point estimates of the volatility-of-the-RV persistence measure $\alpha + \beta$ are in excess of the covariance stationarity threshold of one although not significantly so. The Bollerslev-Wooldridge standard errors of the GARCH parameters are rather large, in fact several times larger than the Hessian-based ones and renders even $\widehat{\alpha} = 0.263$ insignificantly different from zero. The values of the sample skewness and kurtosis and the JB normality test statistic of the standardized residuals $\widehat{Z}_t := \varepsilon_t\left(\widehat{\theta}\right)/h_t^{-1/2}\left(\widehat{\theta}\right)$ (shown on Table 4) indicate very high degrees of conditional nonnormality of the error terms of the models for the raw series RV_t and C_t[13]. This at least partially explains the large discrepancies between the Bollerslev-Wooldridge standard errors and the Hessian-based standard errors and suggest that substantial efficiency gains may be obtained by adopting a non-Gaussian QMLE, but non-Gaussian QMLE requires additional conditions for ensuring consistency under distributional misspecification; See Newey and Steigerwald (1997). As a check of the GARCH(1,1) specification, we calculate the LB statistics for no correlations in the squared standardized residuals (Table 4). Unlike in the case of the McLeod-Li test statistics for no serial correlations in the squared raw residuals, parameter estimation affects the asymptotic distribution in this case. Bollerslev and Mikkelson (1996) suggest a heuristic adjustment of reducing the degrees of freedom of the χ^2 distribution by the number of estimated parameters. The values of the LB statistics are greatly reduced by standardization although they are still high enough to reject the null whether the degrees-of-freedom adjustment is applied or not (when the number of lag orders is 5, this adjustment obviously cannot be applied)[14] except for the statistic for lag orders up to 22 and $Y_t = RV_t$.

As for the four transformed series, both α and β are estimated to be significantly above zero, but the estimates of $\alpha + \beta$ do not exceed one, although for the log transformed series, it is very close to one. For the log transformed series, α estimates are rather small, which together with $\alpha + \beta$ estimated to be nearly one implies a slowly varying conditional mean process as shown on Figures 2a and 2b. Again, the values of the LB statistics for the squared standardized residuals are substantially reduced from the versions calculated for the raw squared residuals, but are still large enough to reject the null of no serial correlations. For $\sqrt{RV_t}$ and $\sqrt{C_t}$, the degrees of nonnormality in the residuals are hugely reduced but not to near normal levels. For $\ln RV_t$ and $\ln C_t$, the raw residuals are not highly nonnormal to begin with, and the effects of standardization are much smaller than in the case of the raw or square-root transformed series. Consistent

[13]Kulperger and Yu (2005) proved that the asymptotic distribution of the Jarque-Bera-type moment-based distributional test statistic based on $\widehat{Z}_t$ rather than Z_t is $\chi^2(2)$, unaffected by parameter estimation if the conditional variance is correctly specified and $\{Z_t\}$ is an i.i.d. sequence. For testing normality, these conditions imply that the Gaussian QMLE is the MLE. One of their additional assumptions is that the conditional mean of the observed variable Y_t is zero, which is not satisfied by our ARFIMA-GARCH case.

[14]Li and Mak (1994) proposed a more elaborate statistic that correctly accounts for this effect.

with near normality, the discrepancies between the Bollerlsev-Wooldridge standard errors and the Hessian-based ones are not very large for these log-transformed variables. However, the values of the JB statistic still reject the null of normality easily.

All in all, the addition of the GARCH(1,1) specification for the ARFIMA error term helps to capture much of the serial correlations in the squared residuals though not completely. We may augment the ARFIMA model by other GARCH-type specifications that have proven successful in improving the GARCH(1,1) fit in more traditional volatility (rather than the volatility-of-the-RV) prediction contexts, but such attempts are beyond the scope of the current paper.

4.3 Further specification tests

The LB statistics are for testing the null of no remaining serial correlations in the residuals and the squared residuals, and are not designed to detect more general forms of serial dependence. Hence, we run a battery of other tests capable of detecting nonlinear dependence. In this paper, we only attempt to model the conditional mean by the homoskedastic ARFIMA model and the first two conditional moments by the ARFIMA-GARCH model. By estimating them using the CSS, which is interpreted as the Gaussian QMLE, we do not take a stand on neither the higher-order dependence structure nor the shape of the conditional distribution (beyond zero mean and unit variance) of $\{Z_t\}$. Nevertheless, it is of interest for a variety of reasons to test the independence and (standard) normality of $\{Z_t\}$ jointly or one at a time assuming the other. The difficulty is that we observe $\left\{\widehat{Z}_t\right\}$ but not $\{Z_t\}$, and the influence of parameter estimation may or may not vanish in the distributions of the test statistics even asymptotically as we have discussed above in the contexts of the LB and JB portmanteau statistics. Unfortunately, for all of the diagnostic statistics that we employ, asymptotic theory assumes some form of mixing for $\{Y_t\}$ and/or $\sqrt{T}$-consistency of the parameter estimator for establishing invariance in the presence of parameter estimation (the "nuisance-parameter-free" property) or justifying the adjustment when invariance does not hold whereas an ARFIMA process with $d > 0$ is not mixing and $\widehat{\mu}$ cannot be expected to be $\sqrt{T}$-consistent. Hence, strictly speaking, the diagnostic tests in this subsection as well as the more traditional ones used in the previous subsections remain somewhat informal.

The values of the statistics are summarized on Table 5. The first columns show the values of the BDS nonlinearity test statistics due to Brock et al. (1996), which test the null of $\{Z_t\} \sim i.i.d.$ and has power to detect higher-order dependence. All different pairs of two consecutive segments of a fixed length k are taken from an observed time series and the number of cases in which the distance (using a particular measure) between the two segments in a pair is shorter than a preset value ϵ is counted. After standardization by its consistent estimator and the asymptotic standard deviation, this number becomes the BDS statistic with (k, ϵ), which is asymptotically distributed $\mathcal{N}(0,1)$ under the null of $\{Z_t\} \sim i.i.d.$ De Lima (1996) showed that under some conditions (including a mixing condition on $\{Y_t\}$ and $\sqrt{T}$-consistency of the estimator) the BDS applied to $\left\{\ln \widehat{Z}_t^2\right\}$, where $\widehat{Z}_t$ is the standardized residual of an ARCH model, is nuisance-parameter free (Note that if $\{Z_t\} \sim i.i.d.$, then $\left\{\ln Z_t^2\right\} \sim i.i.d$ holds as well). Caporal et al. (2005) investigated by Monte Carlo simulations the finite-sample size properties of the BSD test statistics under GARCH(1,1) DGPs with various combinations of true parameter values (estimated by the Gaussian QMLE) and distributions of Z_t not necessarily satisfying De Lima's (1996) sufficiency conditions and reported that they are well-behaved for $T \geq 1000$. We apply the BDS test to the log standardized residuals, and following Chen and Kuan (2005), we set ϵ equal to 0.75 times the sample standard deviation of the series under investigation (in our case $\left\{\ln \widehat{Z}_t^2\right\}$), and calculate the sample standard deviation based on the 1000 bootstrap resamples from the empirical distribution of $\widehat{Z}_t$ and use it for normalization. The results for $k = 2, \cdots, 5$ are presented (on the first lines). For RV_t and C_t, the null of $\{Z_t\} \sim i.i.d.$ is strongly rejected using the log squared residuals from the best homoskedastic model, but not rejected using those from the best overall model, which has the GARCH(1,1) specification, except when $k = 5$. For $\sqrt{RV_t}$ and $\sqrt{C_t}$, the null is strongly rejected for the best homoskedastic model except when $k = 5$ (but the rejection is not as strong as in the case of the log transformed variables), and not rejected for any $k = 2, \cdots, 5$. Finally for $\ln RV_t$ and $\ln C_t$, rejection never occurs, indicating lack of power in light of serial correlations in the squared residuals detected by the LB statistics. We also present the BDS statistics calculated with $\left\{\widehat{Z}_t\right\}$, which violates one of De Lima's (1996) sufficient conditions for invariance, (on the second lines) because the impact of parameter estimation on the variance of the statistic is accounted for by the bootstrap although we still presume that the asymptotic normality of the BSD statistics extend to this case, and use the standard normal critical values for inference. The pattern is similar to the case of the log squared residuals, but the values of the BDS statistic are much larger, rejecting the null in more cases.

We next turn to a test of time reversibility due to Chen et al. (2000) (hereafter the CCK statistic). $\{Z_t\} \sim i.i.d.$ implies time reversibility, which in turn implies that the unconditional distribution of $Z_t - Z_{t-k}$ is symmetric around the origin. The CCK statistic tests this symmetry and is calculated as $\xi_k / \widehat{\sigma}_k$, where $\xi_k := (T-k)^{-1/2} \sum_{t=k+1}^{T} \xi(Z_k)$ and $\widehat{\sigma}_k$ is a consistent estimator for the standard deviation of ξ_k. There are a variety of functions that can be chosen as $\xi(\cdot)$. Following Chen and Kuan (2005), we use $\xi(x) := \gamma x / \left(1 + \gamma^2 x^2\right)$, replace the true errors with the residuals $\left\{\widehat{Z}_t\right\}$, and calculate $\widehat{\sigma}_k$ by bootstrap similarly to the case of the BDS statistics[15], and present the results for $k = 1, 2, 3$ and $\gamma = 0.5, 1$. A general tendency

[15] Chen and Kuan (2005) showed that, unlike in the case of the BDS statistic, the impact of parameter estimation on the CCK

again is that the values of the CCK statistics decrease when square-root transformation is applied to Y_t and turn insignificantly or marginally negative when log-transformation is applied. The CCK statistics appear to be less sensitive to the addition of the GARCH specification to the model than the BDS statistics are, particularly for $\ln RV_t$ and $\ln C_t$.

For testing the correctness of the specification of the joint distribution of $\{Y_t\}$ in its entirety, we may use the nuisance-parameter-free Hong-Li statistics. In our Gaussian ARFIMA-GARCH case, it is equivalent to jointly testing the independence and (standard) normality of $\{Z_t\}$ (as opposed to the JB statistic that tests the normality under the maintained hypothesis of independence). The Hong-Li statistics are based on the observation that, under the null of correct model specification in its entirety (as opposed to just the first two conditional moments), the probability integral transformed series $\{U_t\}$ implied by the model is a sequence of $i.i.d.$ uniform $[0,1]$ random variables, and in particular the joint density of (U_t, U_{t-k}) should be $f(u_1, u_2) = 1$ over $[0,1] \times [0,1]$. Hong and Li (2005) showed that, under the null, a properly normalized measure of the distance[16] (call it the Hong-Li statistic of order k, $Q_{HL}(k)$) between $f(u_1, u_2) = 1$ and $\widehat{f}(u_1, u_2)$, a nonparametric estimate of $f(u_1, u_2)$ constructed using $\left\{\widehat{U}_t\right\}$, is asymptotically distributed $\mathcal{N}(0,1)$. Furthermore, they showed that the asymptotic distribution of $(Q_{HL}(1), \cdots, Q_{HL}(K))$ is standard multivariate normal and hence that of a portmanteau statistic $W_{HL}(K) := \sqrt{K}\sum_{k=1}^{K} Q_{HL}(K)$ is $\mathcal{N}(0,1)$ under the null. Noting that negative values of $\rho_{HL}(k)$ occur only under the null if the sample size is sufficiently large, they suggest using the upper-tailed critical values for individual $\rho_{HL}(k)$'s and $Q_{HL}(K)$. For this reason, the portmanteau statistic is a scaled sum rather than a scaled sum of squares that would yield a $\chi^2(K)$ statistic. For the choice of the kernel function and the bandwidth parameter involved in nonparametric density estimation, we follow Hong and Li (2005). $Q_{HL}(1)$, $W_{HL}(5)$, $W_{HL}(10)$, $W_{HL}(22)$ are shown on Table 4. The overall pattern across variables and models is similar to the case of the JB statistics for the standardized residuals, and as expected, the null of $i.i.d.$ normality is very strongly rejected for RV_t, C_t, $\sqrt{RV_t}$, and $\sqrt{C_t}$. However, it is of note that, while all Hong-Li statistics still cleanly reject the null for the best model for $\ln RV_t$, the statistics values are much reduced for $\ln C_t$ and $Q_{HL}(1)$ only marginally rejects the null for $\ln C_t$ at the 5% level (Note that the upper-tailed 5% critical value is 1.645). This seemingly contradicts the strong rejection of the normality of Z_t for $\ln C_t$ by the JB statistic. However, recall that the

---

statistic is of the same stochastic order as the CCK statistic calculated from the true $\{Z_t\}$ provided that the parameter estimator is $\sqrt{T}$-consistent. Hence, correcting the asymptotic variance based on the assumption of no parameter estimation errors is crucial here, and boostrapped $\widehat{\sigma}_k$ serves the purpose. It might turn out to be the case, however, that using the ARFIMA-GARCH standardized residuals, the effect of the error in estimating μ asymptotically dominates the other terms.

[16] Their expression of the statistic has a typo, which is corrected in Egorov et al (2006, Footnote 11).

JB statistic tests the normality under the maintained hypothesis of $\{Z_t\} \sim i.i.d.$ whereas the Hong-Li statistics in our context test the independence and standard normality jointly. If the $i.i.d.$ assumption does not hold as indicated by the LB statistics for the squared standardized residuals, the JB statistic may overreject normality.

Hong and Li (2005) proposed another set of nuisance-parameter-free statistics, which they call the "separate inference" statistics, designed to detect possible sources of misspecification when $Q_{HL}(k)$ or $W_{HL}(K)$ reject the null of correct specification. To obtain the Hong-Li separate inference statistic $M(m,l)$ for given m and l, we calculate cross-correlations of all orders $j \geq 1$ (up to a truncation point) in $\widehat{U}_t^m$ and $\widehat{U}_{t-j}^j$ and take a weighted average of the squares of them, which after normalization is asymptotically distributed $\mathcal{N}(0,1)$ under the null of correct specification of the entire joint density of $\{Y_t\}$. $M(m,l)$ for $(m,l) = (1,1), (2,2), (3,3), (4,4), (1,2)$, and $(2,1)$ are shown on Table 5. $M(m,m)$ for $m = 1,2,3,4$ are meant to detect autocorrelations in level, volatility, skewness, and kurtosis respectively, and $M(1,2)$ and $M(2,1)$ ARCH-in-mean and leverage effects. Note that, in our context, leverage means that the volatility of the RV responds asymmetrically to positive and negative shocks in the RV, and hence is different from the phenomenon of asymmetric reactions of equity market volatility to positive and negative returns[17]. These $M(m,l)$ statistics are revealing. Substantial reductions in the values of $M(2,2)$ in all cases are as expected. But it is rather surprising, for example, to find the very high value of $M(4,4)$, 28.74, for the best model of $\ln C_t$ with nearly normal $\widehat{Z}_t$ relative to mere 0.74 for the best model of C_t with highly positively skewed and leptokurtic $\widehat{Z}_t$. Failure of the best model for $\ln C_t$ at these $M(m,l)$ tests should be kept in mind in predicting the entire conditional density for risk management or choosing the variable to directly target given a loss function.

4.4 RV prediction

For one-step prediction of the RV measures, we use

$$\widehat{Y}_{t+1|t} := \widehat{\mu} + \sum_{s=1}^{t} \widehat{\pi}_s \left(Y_{t+1-s} - \widehat{\mu}\right) \tag{17}$$

where $\widehat{\mu}$ is the CSS estimate of the unconditional mean and $\widehat{\pi}_s$ are the coefficients in the AR(∞) expansion of the ARFIMA model, implied by the ARFIMA parameter estimates. While the Durbin-Levinson algorithm may be applied to calculate the best linear one-step predictor based on the finite past, the formula (17) is

[17] As mentioned in the introduction, some studies incorporated a specification for the latter effect in time series models for the RV by using daily returns data.

more in line with our CSS estimator. We also evaluate the performance of the selected model for each of the RV series in forecasting the k-day average RV, $k^{-1}\sum_{s=1}^{k} Y_{t+s}$. For this, we use $k^{-1}\sum_{s=1}^{k}\widehat{Y}_{t+s|t}$ as our k-day forecast where

$$\widehat{Y}_{t+k|t} := \widehat{\mu} + \sum_{s=1}^{k-1}\widehat{\pi}_s\left(\widehat{Y}_{t+s|t} - \widehat{\mu}\right) + \sum_{s=0}^{t-1}\widehat{\pi}_{k+s}\left(Y_{t-s} - \widehat{\mu}\right) \tag{18}$$

For evaluating predictive accuracy, we mainly look at R^2 from the Mincer-Zarnowiz regression of the realization of the target variable on the prediction (plus an intercept).

The results for $k = 1, 5, 10, 22$ are summarized on Table 5. Again, the results for RV_t, $\sqrt{RV_t}$, $\ln RV_t$ and their respective jump-free versions C_t, $\sqrt{C_t}$, $\ln C_t$ are similar. In conformance with the previously reported results for the RVs of the Nikkei 225 index and other financial time series, the log transformed series $\ln RV_t$ and $\ln C_t$ appear to be most predictable, followed by $\sqrt{RV_t}$ and $\sqrt{C_t}$. For example, R^2 is nearly 60% when the target is the one-step-ahead $\ln RV_t$ or $\ln C_t$. Although R^2 tapers off as the horizon increases, predictability of the average daily RV over the next month is still substantial (nearly 30% for $\ln RV_t$ or $\ln C_t$). Excellent performance of the ARFIMA-GARCH model can be visually confirmed by the time series plot (Figure 1) of the ARFIMA-GARCH fit together with the target series.

4.5 Prediction of the variance of the RV

We next evaluate the performance of the ARFIMA-GARCH model selected for each of the six series in predicting the volatility of the RV. For one-step-ahead forecasting, we use $\widehat{h}_{t+1}$ as the volatility-of-the-RV forecast and $\widehat{\varepsilon}_{t+1}^2 = \left(Y_{t+1} - \widehat{Y}_{t+1}\right)^2$ as the target, whether the object of our interest is $Var_t\ (Y_{t+1}) = E_t\left[\varepsilon_{t+1}^2\right]$ or some other measure of the volatility of the RV. Our ARFIMA-GARCH estimation results indicate that the variability of $\widehat{\varepsilon}_{t+1}^2$ is much higher than that of $E_t\left[\varepsilon_{t+1}^2\right]$, which would lead to an apparently low R^2 value of the Mincer-Zarnowitz regression even if the time-variation in $E_t\left[\varepsilon_{t+1}^2\right]$ could be well approximated by the ARFIMA-GARCH model. This parallels the solution by Andersen and Bollerslev (1998) of a puzzle regarding the low R^2 phenomenon prevalent in earlier volatility prediction studies based on data sampled at daily or lower frequencies. The introduction of the RV measures in empirical finance has rendered financial volatility nearly observable, but here we are back in the unobservability territory. What makes the problem even more complicated in our context is that the RV is highly predictable unlike daily asset returns and that $Var\left(E_t\left[Y_{t+1}\right]\right)/Var\left(Y_{t+1}\right)$ seems to be large and hence the effect of the deviation of $\widehat{Y}_{t+1}$ from the true $E_t\left[Y_{t+1}\right]$ due to possible model misspecification, parameter estimation errors and other factors on the Mincer-Zarnowitz R^2 is potentially more serious for volatility-of-volatility prediction than for equity or currency return volatility prediction.

For forecasting the multi-day volatility of the RV, we use $k^{-1}\sum_{s=1}^{k}\widehat{h}_{t+s|t}$, where $\widehat{h}_{t+s|t}$ is the s-period-ahead single-day conditional variance of the RV implied by the ARFIMA-GARCH, as our forecast, and the daily squared errors averaged over the horizon,

$$\upsilon_t\left(k\right) := k^{-1}\sum_{s=1}^{k}\left(Y_{t+s}-\widehat{Y}_{t+s|t}\right)^2 \tag{19}$$

as a proxy for the target.

The right half of Table 5 summarizes the forecasting performance evaluation results for horizons 1, 5, 10, and 22 days. As expected, R^2's are in fact low, but not negligible. For brevity, we do not report the details of the results of our multi-day-ahead single-day volatility-of-the-RV forecasting exercise, but the predictability does not increase as we attempt to forecast further into the future. In spite of this, R^2 values are higher for multi-day average forecasts than for a one-day ahead forecast. This may happen because our target proxy $\upsilon_t\left(k\right)$, being squared forecast errors aggregated over forecast horizons of a week to several weeks, is a sort of realized volatility (of the RV). Although the sampling frequency is much coarser here, a noise reduction effect similar to those observed in the daily volatility forecasting studies using RV measures constructed from intraday high frequency asset returns may lead to a higher R^2 in Mincer-Zarnowitz regressions with longer-horizon targets and predictors. For $\sqrt{RV_t}$ and $\sqrt{C_t}$, the R^2 values are respectively 3.63% and 3.55% (one-day-ahead, single-day), 8.23% and 7.60% (5 days), 6.87% and 6.50% (10 days), 5.44% and 5.27% (22 days). Presumably, the effect of decreasing predictability outweighs the noise reduction as the horizon increases beyond a week or so. For $\ln RV_t$ and $\ln C_t$, the R^2 values are respectively 1.70% and 3.31% (one-day-ahead, single-day), 3.68% and 7.29% (5 days), 4.65% and 9.30% (10 days), 6.15% and 12.50% (22 days). It is not clear why the squared errors appear to be more predictable for $\ln C_t$ than for $\ln RV_t$ in spite of the similar behavior of C_t and RV_t observed in other respects.

5 Concluding remarks

In this paper, we investigated the volatility of the daily Nikkei 225 realized volatility. Although much of the recent advances in volatility research has been due to the recognition that high-frequency intraday data make daily volatility essentially observable in the form of the realized volatility and related measures, we are back to the condition of unobservability when we move one order higher in terms of the moments from volatility to volatility of volatility. This makes evaluation of models such as the ARFIMA-GARCH for forecasting the volatility of the RV a difficult task. Nevertheless, the GARCH specification for the conditional variance of the RV added to the ARFIMA specification for the conditional mean of the RV seems to capture the persistent time variation in the conditional variance of the Nikkei 225 RV regardless of whether we use the raw, square-root- or log-transformed series. Filtering out the jump component from the RV using a version of the Barndorff-Nielsen procedure does not seem to have much impact on the estimation of the ARFIMA-GARCH model or the degree of predictability in the RV. However, the volatility of the RV seems to be more predictable for the jump-free RV than the usual RV at the forecast horizon of about a month.

Acknowledgments

The authors are grateful to Yoshihiro Yajima and Yuji Sakurai for helpful discussions.

References

[1] Andersen, T.G., and T. Bollerslev, 1998, Answering the skeptics: Yes, standard volaitlity models do provide accurate forecasts, *International Economic Review*, 39, 885-905.

[2] Andersen, T.G., T. Bollerslev, and J. Cai, 2000, Intraday and interday volatility in the Japanese stock market, *Journal of International Financial Markets, Institutions and Money*, 10, 107-130.

[3] Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics*, 61, 43-75.

[4] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2001, The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, 96, 42-55.

[5] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2003, Modeling and forecasting realized volatility, *Econometrica*, 71, 579-625.

[6] Andersen, T.G., T. Bollerslev, and F.X. Diebold, 2007, Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility, *Review of Economics and Statistics*, 89, 701-720.

[7] Andersen, T.G., T. Bollerslev, and X. Huang, 2007, A reduced form framework for modeling volatility of speculative prices based on realized variation measures, Working paper, Northwestern University and Duke University.

[8] Andersen, T.G., T. Bollerslev, and S. Lange, 1999, Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon, *Journal of Empirical Finance*, 6, 457-477.

[9] Baillie, R.T., 1996, Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, 73, 5-59.

[10] Baillie, R.T., C. Chung, and M.A. Tieslau, 1996, Analysing inflation by the fractionally integrated ARFIMA-GARCH model, *Journal of Applied Econometrics*, 11, 23-40.

[11] Baillie, R.T., T. Bollerslev, and H.O. Mikkelsen, 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 74, 3-30.

[12] Bakshi, G., Nengjiu Ju, and Hui Ou-Yang, 2006, Estimation of continuous-time models with an application to equity volatility dynamics, *Journal of Applied Econometrics*, 82, 227-249.

[13] Barndorff-Nielsen, O.E., S.E. Graversen, J. Jacod, M. Podolskij, and N. Shephard, 2006a, A central limit theorem for realised power and bipower variation of continuous semimartingales, in Yu Kabanov, Robert Lipseter, and Jordan Stoyanov, eds., *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, Springer Verlag: Berlin.

[14] Barndorff-Nielsen, O.E., S.E. Graversen, J. Jacod, and N. Shephard, 2006b, Limit theorems for bipower variation in financial econometrics, *Econometric Theory*, 22, 677-719.

[15] Barndorff-Nielsen, O.E., and N. Shephard, 2002, Econometric analysis of realised volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, Series B*, 64, 253-280.

[16] Barndorff-Nielsen, O.E., and N. Shephard, 2004, Power and bipower variation with stochastic volatility and jumps, *Journal of Financial Econometrics*, 2, 1-37.

[17] Barndorff-Nielsen, O.E., and N. Shephard, 2006c, Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics*, 4, 217-252.

[18] Beran, J., 1994, Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models, *Journal of the Royal Statistical Society B*, 57, 659–672.

[19] Blair, B.J., S. Poon, and S.J. Taylor, 2001, Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency returns, *Journal of Econometrics*, 105, 5-26.

[20] Bollerslev, T., R.F. Engle, and D.B. Nelson, 1994, ARCH models, in Daniel McFadden and Robert F. Engle, eds., *Handbook of Econometrics Vol.4*, Elsevier: Amsterdam.

[21] Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen, 2005, A discrete-time model for daily S&P 500 returns and realized variations: Jumps and leverage effects.

[22] Bollerslev, T., and H.O. Mikkelsen, 1996, Modeling and pricing long memory in stock market volatility, *Journal of Econometrics*, 73, 151-184.

[23] Bollerslev, T., and J. Wooldrige, 1992, Quasi-maximum likelihood estimation and inference for dynamic models with time-varying covariances, *Econometric Reviews*, 11, 143-172.

[24] Bondon, P., and W. Palma, 2007, A class of antipersistent processes, *Journal of Time Series Analysis*, 28, 261-273.

[25] Bos, Charles S., Philip Hans Franses, and Marius Ooms, 1999, Long memory and level shifts: Reanalyzing inflation rates, *Empirical Economics*, 24, 427-449.

[26] Brock, W.A., W.D. Dechert, J.A. Scheinkman, and B. LeBaron, 1987, A test for independence based on the correlation dimension, *Econometric Reviews*, 15, 197-235.

[27] Caporale, G.M., C. Ntantamis, T. Pantelidis, and N. Pittis, 2005, The BDS test as a test for the adequacy of a GARCH(1,1) specification: A Monte Carlo study, *Journal of Financial Econometrics*, 3, 282-309.

[28] Chen, Y., R.Y. Chou, and C. Kuan, 2000, Testing time reversibility without moment restrictions, *Journal of Econometrics*, 95, 199-218.

[29] Chen, Y., and C. Kuan, 2002, Time irreversibility and EGARCH effects in US stock index returns, *Journal of Applied Econometric*s, 17, 565-578.

[30] Corsi, F., 1994, A simple long memory model of realized volatility, Unpublished manuscript, University of Southern Switzerland.

[31] Breidt, F.J., N. Crato, P. de Lima, 1998, On detection and estimation in stochastic volatility, *Journal of Econometrics*, 83, 325-348.

[32] 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.

[33] Diebold, F.X., 1988, *Empirical modeling of exchange rate dynamics*, Springer.

[34] Diebold, F.X., and J.A. Nason, 1990, Nonparametric exchange rate prediction? *Journal of International Economics*, 28, 315-332.

[35] 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.

[36] Dotsis, G., D. Psychoyios, and G. Skiadopoulos, An empirical comparison of continuous-time models of implied volatility indices, *Journal of Banking and Finance*, 31, 3584-3603.

[37] Duan, J., and C. Yeh, 2007, Jump and volatility risk premiums implied by VIX, Working paper, University of Tronto and National Taiwan University.

[38] Ebens, H., 1999, Realized stock volatility, Working paper, John Hopkins University.

[39] Egorov, A., Y. Hong, and H. Li, 2006, Validating forecasts of the joint probability density of bond yields: Can affine models beat random walk?, *Journal of Econometrics* 135, 255-284.

[40] Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica*, 50, 987-1008.

[41] Giot, P., and Sébastien L., Modelling daily value-at-risk using realized volatility and ARCH type models, *Journal of Empirical Finance*, 11, 379-398.

[42] Geweke, J., and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, *Journal of Time Series Analysis* 4, 221-238.

[43] Giacomini, R., and Halbert W., 2006, Tests of conditional predictive ability, *Econometric*a 74, 1545-1578.

[44] Granger, C.W.L., and R. Joyeux, 1980, An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-29.

[45] Forsberg, L., and Eric G.s, 2007, Why do absolute returns predict volatility so well?, *Journal of Financial Econometrics*, 5, 31-67.

[46] Hansen, P.R., and A. Lunde, 2005, Realized variance and market microstructure noise, *Journal of Business and Economic Statistics*, 24, 127-161.

[47] Hansen, P.R., and A. Lunde, 2005, A realized variance for the whole day based on intermittent high-frequency data, *Journal of Financial Econometrics*, 3, 525-554.

[48] Hansen, P.R., and A. Lunde, 2006, Consistent ranking of volatility models, *Journal of Econometrics*, 131, 97-121.

[49] Hauser, M.A., and R.M. Kunst, 1998, Fractionally integrated models with ARCH errors: With an application to the Swiss 1-month Euromarket interest rate, *Review of Quantitative Finance and Accounting* 10, 95-113.

[50] Hauser, M.A., and R.M. Kunst, 2001, Forecasting high-frequency financial data with the ARFIMA-ARCH model, *Journal of Forecasting*, 20, 501-518.

[51] Hosking, J.R., 1981, Fractional differencing, *Biometrika*, 68, 165-176.

[52] Hong, Y., and H. Li, 2005, Nonparametrci specification testing for continuous-time models with applications to term structure of interest rates, *Review of Financial Studies*, 18, 37-84.

[53] Hong, Y., H. Li, and F. Zhao, 2007, Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates, *Journal of Econometrics*, 141, 736-776.

[54] Huang, X., and G. Tauchen, 2005, The relative contribution of jumps to total price variation, *Journal of Financial Econometrics*, 3, 456-499.

[55] Koopman, S.J., B. Jungbacker, and E. Hol, 2005, Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements, *Journal of Empirical Finance*, 12, 445-475.

[56] Kulperger, R., and H. Yu, 2005, High moment partial sum processes of residuals in GARCH models and their applications, *Annals of Statistics*, 33, 2395-2422.

[57] Ling, S., and W.K. Li, 1997, On fractionally integrated autoregressive moving-average time series models with conditional heteroskedasticity, *Journal of the American Statistical Association*, 92, 1184-1194.

[58] Maheu, J, 2005, Can GARCH models capture long-range dependence? *Studies in Nonlinear Dynamics & Econometrics*, 9.

[59] Martens, M., 2001, Forecasting daily exchange rate volatility using intraday returns, *Journal of International Money and Finance*, 20, 1-23.

[60] Martens, M., 2002, Measuring and forecasting S&P 500 index-futures volatility using high-frequency data, *Journal of Futures Markets*, 22, 497-518.

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

[62] Newey, W., and D. Steigerwald, 1997, Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroskedasticity models, *Econometrica*, 65, 587-599.

[63] Oomen, R.C.A., 2001, Using high frequency stock market index data to calculate, model & forecast realized return variance, Working paper, European University Institute.

[64] Ooms, M., and J. Doornik, , Inference and forecasting for fractional autoregressive integrated moving average models, with an application to US and UK inflation, Econometric Institute Report 9947/A, Erasmus University.

[65] Palma, Wilfredo, 2007, *Long-memory Time Series: Theory and Methods*, Wiley: Hoboken, New Jersey.

[66] Patton, Andrew, 2006, Volatility forecast comparison using imperfect volatility proxies, Working paper, London School of Economics.

[67] Jensen, S.T., and A. Rahbek, 2004, Asymptotic inference for nonstationary GARCH, *Econometric Theory*, 20, 1203-1126.

[68] Robinson, P.M., 1995a, Gaussian semiparametric estimation of long range dependence, *Annals of Statistics*, 23, 1630-1661.

[69] Robinson, P.M., 1995b, Log-periodogram regression of time series with long range dependence, *Annals of Statistics*, 23, 1048-1072.

[70] Robinson, P.M., and M. Henry, 1999, Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels, *Econometric Theory*, 15, 299-336.

[71] Shephard, N., 2005, *Stochastic volatility*: *Selected Readings*, Oxford University Press: Oxford.

[72] Shibata, M., 2004, Realized volatility wo mochiita GARCH model no yosokuryoku hikaku, Chapter 2, Ph.D. Dissertation, Tokyo Metropolitan Universtiy (in Japanese).

[73] Shibata, M., 2008, Kohindo data ni yoru volatility no suitei: Realized volatility no survey to nihon no kabuka shisu oyobi shisu sakimono no jissho bunseki, *Kin-yu Kenkyu*, 27 (1), 1-54 (in Japanese).

[74] Shibta, M., and T. Watanabe, 2004, Realized volatility wo mochiita Nikkei 225 sakimono kakaku no volatility yoshoku, Osaka Securities Exchange Sakimono Option Report 16-12 (in Japanese).

[75] Takahashi, M., Y. Omori, and T. Watanabe, 2007, Estimating stochastic volatility models using daily returns and realized volatility simultaneously, Center for Advanced Research in Finance Working Paper F-108, University of Tokyo.

[76] Watanabe, T., 2005, Nikkei 225 sakimono kakaku no realized volatility, Osaka Securities Exchange Sakimono Option Report 17-5 (in Japanese)

[77] Watanabe, T., and K. Sasaki, 2006, ARCH-gata model to "realized volatility" ni yoru volatility yosoku to value-at-risk, *Kin-yu Kenkyu*, 25 (2), 39-79 (in Japanese).

[78] Watanabe, T., and K. Yamaguchi, 2006, Nikkei 225 no "realized volatility" to implied volatility, Osaka Securities Exchange Sakimono Option Report 18-12 (in Japanese).

Table 1: Summary statistics: Nikkei 225 daily returns, realized volatility and related measures

Series	*Mean*	*Std. Dev.*	*Skew.*	*Kurt.*	*Min.*	*Max.*	*Jarque-Bera*	ρ_1	S.E.(ρ_1)	*LB*(5)	*DLB*(5)	*LB*(10)	*DLB*(10)	*LB*(22)	*DLB*(22)	Long-memory *d* *GPH*	Long-memory *d* *Robinson*
$\boldsymbol{R}$	-0.007	1.423	-0.064	4.962	-7.234	7.660	451.13	-0.041	(0.022)	**19.207**	13.251	**23.805**	16.598	**37.588**	26.732	-0.018	-0.011
$\boldsymbol{R_n}$	0.039	0.631	0.009	3.480	-3.186	3.849	26.89	0.036	(0.020)	**12.112**	10.265	**14.545**	12.334	**33.345**	29.016	0.054	0.044
$\boldsymbol{R_{am}}$	-0.025	0.802	-0.242	7.700	-7.394	5.078	2,606.19	-0.033	(0.022)	**9.678**	6.342	**11.876**	7.938	**23.167**	15.856	-0.048	-0.031
$\boldsymbol{R_l}$	-0.026	0.164	0.015	4.905	-0.736	0.786	423.62	0.026	(0.024)	**16.083**	10.712	**22.007**	15.017	**57.635**	40.153	0.071	0.113
$\boldsymbol{R_{pm}}$	0.005	0.774	0.103	5.818	-3.675	4.881	932.38	-0.099	(0.025)	**53.107**	33.121	**60.194**	37.744	**80.960**	53.633	-0.107	-0.060
$\boldsymbol{RV_{am}}$	0.573	0.622	5.970	78.350	0.011	12.748	679,502.99	0.429	(0.060)	**2,253.238**	223.283	**3,635.044**	442.419	**5,801.812**	809.331	0.460	0.454
$RV_{am}^{1/2}$	0.691	0.309	1.587	9.207	0.103	3.570	5,674.29	0.569	(0.083)	**3,862.612**	214.711	**6,603.898**	406.563	**11,267.971**	803.268	0.532	0.495
$\ln RV_{am}$	-0.925	0.871	-0.177	3.272	-4.551	2.545	23.19	0.616	(0.043)	**4,532.985**	949.938	**7,921.581**	1,761.857	**14,268.434**	3,406.661	0.525	0.491
$\boldsymbol{RV_{pm}}$	0.514	0.595	12.096	335.024	0.015	18.941	12,938,877.08	0.355	(0.059)	**1,298.273**	168.714	**1,981.101**	390.580	**2,873.448**	822.065	0.364	0.376
$RV_{pm}^{1/2}$	0.655	0.292	1.773	13.810	0.122	4.352	15,111.33	0.535	(0.081)	**3,128.317**	196.911	**5,208.527**	373.510	**8,450.013**	734.922	0.504	0.472
$\ln RV_{pm}$	-1.028	0.866	-0.182	3.260	-4.202	2.941	23.35	0.600	(0.047)	**4,070.792**	710.029	**7,035.733**	1,290.913	**12,361.541**	2,405.026	0.524	0.477
$\boldsymbol{RV} = RV_{am} + RV_{pm}$	1.087	1.030	5.336	67.840	0.065	20.493	504,141.92	0.540	(0.071)	**3,237.586**	250.406	**4,970.638**	529.470	**7,586.328**	1,004.412	0.470	0.468
$RV^{1/2}$	0.968	0.388	1.459	8.520	0.255	4.527	4,551.03	0.663	(0.097)	**4,997.214**	211.504	**8,342.521**	402.196	**13,887.091**	793.872	0.557	0.524
$\ln RV$	-0.214	0.778	-0.130	3.234	-2.733	3.020	14.25	0.713	(0.030)	**5,927.105**	2,427.793	**10,322.146**	4,474.032	**18,365.815**	8,625.484	0.584	0.533
$\boldsymbol{R_n^2}$	0.399	0.628	8.937	168.634	0.000	14.812	3,240,305.54	0.062	(0.017)	**88.876**	81.982	**159.721**	133.651	**241.634**	221.775	0.197	0.202
$\|R_n\|$	0.508	0.376	1.192	7.069	0.000	3.849	2,596.51	0.097	(0.028)	**192.320**	73.951	**330.598**	131.412	**515.651**	222.424	0.261	0.250
$\ln (R_n^2)$	-2.149	2.195	-1.540	6.627	-16.151	2.695	2,643.12	0.050	(0.020)	**58.136**	44.332	**104.587**	79.669	**175.383**	133.484	0.209	0.190
$\boldsymbol{R_l^2}$	0.028	0.054	4.178	26.776	0.000	0.618	74,150.20	0.145	(0.033)	**241.331**	93.607	**396.055**	161.320	**793.223**	308.027	0.278	0.255
$\|R_l\|$	0.121	0.114	1.687	6.476	0.000	0.786	2,739.04	0.175	(0.035)	**397.477**	122.080	**713.096**	233.969	**1,358.050**	459.849	0.288	0.302
$\ln (R_l^2)$	-5.279	2.438	-1.139	5.284	-19.501	-0.481	1,213.66	0.096	(0.030)	**157.867**	53.404	**320.240**	106.402	**605.525**	213.052	0.210	0.258
$\boldsymbol{RV+R_n^2+R_l^2}$	1.514	1.321	4.040	37.141	0.082	21.151	143,704.07	0.461	(0.059)	**2,475.549**	257.636	**3,928.348**	521.401	**6,076.282**	980.450	0.451	0.458
$(RV+R_n^2+R_l^2)^{1/2}$	1.149	0.440	1.283	6.990	0.286	4.599	2,626.47	0.563	(0.090)	**3701.681**	179.179	**6248.175**	337.833	**10273.433**	651.265	0.542	0.502
$\ln (RV+R_n^2+R_l^2)$	0.139	0.752	-0.174	3.327	-2.506	3.052	26.58	0.603	(0.026)	**4291.539**	2415.713	**7482.022**	4,479.544	**12918.003**	8,479.105	0.571	0.505
$\boldsymbol{RV_{NW}}$	1.153	1.420	5.879	61.994	0.041	21.909	422,464.73	0.467	(0.077)	**2,409.258**	160.784	**3,494.761**	357.302	**4,982.016**	710.337	0.427	0.445
$RV_{NW}^{1/2}$	0.968	0.466	1.931	10.572	0.201	4.681	8,436.53	0.563	(0.087)	**3,652.271**	194.090	**6,029.797**	378.271	**9,857.148**	759.299	0.490	0.504
$\ln RV_{NW}$	-0.267	0.894	-0.006	3.301	-3.204	3.087	10.57	0.606	(0.028)	**4,327.317**	2,030.833	**7,633.914**	3,797.857	**13,702.610**	7,391.286	0.527	0.507
$\boldsymbol{RV_{HL}}$	1.514	1.289	3.637	31.792	0.083	20.565	102,960.04	0.503	(0.063)	**2,835.818**	272.087	**4,492.889**	553.123	**7,071.670**	1,057.351	0.459	0.464
$RV_{HL}^{1/2}$	1.150	0.438	1.184	6.272	0.288	4.535	1,904.83	0.596	(0.092)	**4,050.849**	189.043	**6,821.396**	355.864	**11,373.832**	690.552	0.554	0.505
$\ln RV_{HL}$	0.141	0.753	-0.186	3.261	-2.487	3.024	24.14	0.634	(0.026)	**4,679.843**	2,634.268	**8,132.032**	4,865.625	**14,158.498**	9,273.210	0.588	0.509
$\boldsymbol{RV_{HLNW}}$	1.580	1.512	3.405	23.229	0.051	17.863	53,190.72	0.412	(0.059)	**1,833.103**	220.902	**2,876.402**	452.062	**4,335.164**	867.892	0.439	0.438
$RV_{HLNW}^{1/2}$	1.156	0.494	1.233	5.956	0.226	4.226	1,730.24	0.476	(0.077)	**2,578.344**	172.929	**4,344.404**	326.894	**7,017.121**	622.922	0.505	0.475
$\ln RV_{HLNW}$	0.114	0.849	-0.217	3.232	-2.978	2.883	28.29	0.495	(0.024)	**2,891.338**	1,889.880	**5,051.168**	3,470.476	**8,568.755**	6,304.374	0.526	0.469

Table 1: Summary Statistics (Continued)

Series	*Mean*	*Std. Dev.*	*Skew.*	*Kurt.*	*Min.*	*Max.*	*Jarque-Bera*	ρ_1	S.E.(ρ_1)	*LB*(5)	*DLB*(5)	*LB*(10)	*DLB*(10)	*LB*(22)	*DLB*(22)	Long-memory d	
																GPH	*Robinson*
$\boldsymbol{I_{am}}$	0.054	0.227	3.936	16.492	0.000	1.000	28,485.68	0.082	(0.030)	**40.003**	18.276	**66.501**	30.686	**171.483**	77.806	0.102	0.148
$\boldsymbol{I_{pm}}$	0.105	0.306	2.585	7.680	0.000	1.000	5,676.54	0.063	(0.024)	**66.480**	37.339	**113.491**	65.899	**191.848**	114.649	0.137	0.189
$\boldsymbol{I}$	0.159	0.390	2.325	7.663	0.000	2.000	5,063.29	0.111	(0.030)	**154.403**	70.207	**270.145**	124.120	**515.364**	245.012	0.214	0.231
$\boldsymbol{J_{am}}$	0.018	0.111	12.378	254.404	0.000	3.126	7,450,609.97	0.024	(0.022)	**23.634**	4.842	**24.062**	5.409	**35.036**	45.902	0.115	0.094
$J_{am}^{1/2}$	0.029	0.133	5.330	36.177	0.000	1.768	141,774.19	0.049	(0.026)	**19.234**	8.781	**21.624**	11.187	**50.315**	27.145	0.108	0.102
ln (1+J_{am})	0.015	0.077	7.480	78.774	0.000	1.417	696,469.20	0.035	(0.027)	**17.319**	5.526	**17.777**	5.924	**36.053**	21.733	0.094	0.088
$\boldsymbol{J_{pm}}$	0.027	0.108	6.936	72.084	0.000	1.819	579,666.63	-0.009	(0.008)	**22.952**	11.622	**34.800**	17.749	**78.481**	37.759	0.108	0.137
$J_{pm}^{1/2}$	0.049	0.156	3.512	16.148	0.000	1.349	25,944.59	0.017	(0.017)	**39.216**	20.935	**65.200**	36.154	**124.117**	68.731	0.100	0.164
ln (1+J_{pm})	0.022	0.083	5.197	37.851	0.000	1.036	154,421.82	-0.003	(0.011)	**28.390**	13.384	**44.811**	21.804	**97.205**	45.817	0.125	0.149
$\boldsymbol{J}$	0.045	0.154	6.750	82.923	0.000	3.126	767,042.11	0.099	(0.087)	**82.202**	15.450	**95.866**	22.705	**150.284**	53.468	0.152	0.156
$J^{1/2}$	0.076	0.198	2.834	11.464	0.000	1.768	12,114.74	0.063	(0.030)	**80.413**	36.245	**114.470**	59.428	**226.672**	124.609	0.177	0.184
ln (1+J)	0.037	0.112	4.232	26.939	0.000	1.417	75,267.00	0.067	(0.046)	**71.425**	21.879	**90.850**	33.287	**172.347**	75.446	0.165	0.168
$\boldsymbol{C_{am}}$	0.554	0.617	6.094	81.196	0.011	12.748	731,228.51	0.429	(0.060)	**2,210.926**	215.305	**3,526.428**	431.947	**5,639.629**	789.466	0.426	0.446
$C_{am}^{1/2}$	0.677	0.310	1.581	9.214	0.103	3.570	5,675.75	0.567	(0.082)	**3,749.278**	216.737	**6,386.549**	411.409	**10,979.110**	818.777	0.535	0.485
ln C_{am}	-0.977	0.900	-0.225	3.272	-4.551	2.545	32.28	0.609	(0.043)	**4,322.297**	887.450	**7,584.786**	1,642.283	**13,818.262**	3,195.502	0.540	0.482
$\boldsymbol{C_{pm}}$	0.488	0.592	12.327	343.571	0.012	18.941	13,612,657.54	0.349	(0.059)	**1,247.790**	188.381	**1,891.862**	415.115	**2,714.858**	825.181	0.368	0.373
$C_{pm}^{1/2}$	0.633	0.295	1.779	13.664	0.110	4.352	14,753.15	0.531	(0.078)	**3,089.449**	208.804	**5,137.232**	396.071	**8,194.508**	766.741	0.519	0.480
ln C_{pm}	-1.115	0.910	-0.210	3.215	-4.413	2.941	26.08	0.596	(0.048)	**4,006.274**	678.345	**6,938.593**	1,234.260	**12,007.228**	2,271.501	0.515	0.486
$\boldsymbol{C}$	1.042	1.021	5.426	70.160	0.056	20.493	540,350.80	0.537	(0.070)	**3,177.375**	250.504	**4,859.086**	527.560	**7,371.936**	995.301	0.477	0.464
$C^{1/2}$	0.943	0.391	1.449	8.415	0.237	4.527	4,403.37	0.658	(0.094)	**4,880.797**	218.977	**8,129.868**	416.252	**13,491.335**	821.386	0.549	0.518
ln C	-0.278	0.811	-0.161	3.183	-2.879	3.020	16.02	0.703	(0.031)	**5,698.077**	2,265.783	**9,909.682**	4,165.279	**17,598.452**	7,995.639	0.577	0.523

The sample period is from March 11, 1996, through August 31, 2007 (2802 observations). ρ_1 and S.E.(ρ1) in parentheses are respectively the first-order sample autocorrelations and heteroskedasticity-robust standard errors (the usual Bartlett's standard errors are T-1/2 = 0.018). LB(k) are the Ljeung-Box statistics of orders up to k and DLB(k) are their heteroskedasticity adjusted versions. The 5% critical values for χ2(k) are 5.991 (k=2), 11.070 (5), 18.307 (10), 33.924 (22). The standard errors of the GPH and Robinson estimators for the long-memory parameter d are 0.034 and 0.031 respectively.

Table 2: ARFIMA-GARCH estimation results

Var.	Model	ARFIMA parameters						GARCH paramters			LL
		μ	d	φ_1	φ_2	ψ_1	ψ_2	σ^2, ω	β	α	BIC
	Best short-memory	**0.7415**		**0.3592**	**0.5887**		**-0.4579**	**0.0136**	**0.7873**	**0.2630**	**-2,410.93**
	(2,0,{2})-G	(0.0915)		(0.0252)	(0.0282)		(0.0290)	(0.0024)	(0.0260)	(0.0397)	**4,877.42**
		(0.1667)		(0.0375)	(0.0506)		(0.0438)	(0.0127)	(0.1291)	(0.1893)	
	Best homoskedastic	**0.8843**	**0.4963**			**-0.2262**		**0.6391**			**-3,348.67**
	(0,d,1)	(0.2118)	(0.0320)			(0.0389)		(0.0171)			**6,729.09**
		(0.0936)	(0.0865)			(0.1068)		(0.1352)			
	GARCH ver. of	**0.5751**	**0.5604**			**-0.2105**		**0.0106**	**0.8238**	**0.2119**	**-2,416.28**
RV	best homosked.	(0.2495)	(0.0478)			(0.0605)		(0.0015)	(0.0141)	(0.0219)	**4,880.19**
	(0,d,1)-G	(0.1917)	(0.0859)			(0.0998)		(0.0068)	(0.0577)	(0.0819)	
	Homoskedastic ver.	**0.8765**	**0.5085**	**-0.2345**	**-0.0866**			**0.6381**			**-3,346.53**
	of best overall model	(0.3450)	(0.0282)	(0.0330)	(0.0254)			(0.0170)			**6,732.75**
	(2,d,0)	(0.2259)	(0.0822)	(0.0928)	(0.0684)			(0.1344)			
	Best overall	**0.6332**	**0.6218**	**-0.2554**	**-0.1725**			**0.0115**	**0.8079**	**0.2340**	**-2,406.79**
	(2,d,0)-G	(0.5415)	(0.0384)	(0.0430)	(0.0330)			(0.0017)	(0.0184)	(0.0283)	**4,869.15**
		(0.6374)	(0.0950)	(0.0901)	(0.0774)			(0.0085)	(0.0879)	(0.1299)	
	Best short-memory	**0.8838**		**0.9741**		**-0.6178**	**-0.0949**	**0.0033**	**0.8561**	**0.0990**	**-17.30**
	(1,0,2)-G	(0.0465)		(0.0055)		(0.0223)	(0.0232)	(0.0007)	(0.0209)	(0.0142)	**90.17**
		(0.0471)		(0.0063)		(0.0231)	(0.0263)	(0.0016)	(0.0436)	(0.0263)	
	Best homoskedastic	**0.8818**	**0.5193**			**-0.1898**		**0.0700**			**-250.25**
	(0,d,1)	(0.1215)	(0.0333)			(0.0436)		(0.0019)			**532.26**
		(0.1011)	(0.0607)			(0.0819)		(0.0049)			
	GARCH ver. of	**0.7579**	**0.4979**			**-0.1545**		**0.0031**	**0.8590**	**0.0988**	**-14.18**
$RV^{1/2}$	best homosked.	(0.1019)	(0.0304)			(0.0388)		(0.0007)	(0.0200)	(0.0140)	**75.99**
	(0,d,1)-G	(0.0922)	(0.0323)			(0.0414)		(0.0015)	(0.0435)	(0.0270)	
	Homoskedastic ver.	**0.8790**	**0.5427**	**-0.2061**	**-0.0911**			**0.0698**			**-246.75**
	of best overall model	(0.1191)	(0.0260)	(0.0306)	(0.0237)			(0.0019)			**533.18**
	(2,d,0)	(0.0922)	(0.0419)	(0.0460)	(0.0330)			(0.0049)			
	Best overall	**0.7721**	**0.5346**	**-0.1831**	**-0.0988**			**0.0031**	**0.8590**	**0.0991**	**-9.68**
	(2,d,0)-G	(0.1210)	(0.0302)	(0.0355)	(0.0271)			(0.0007)	(0.0202)	(0.0143)	**74.93**
		(0.1150)	(0.0335)	(0.0382)	(0.0288)			(0.0015)	(0.0435)	(0.0277)	
	Best short-memory	**-0.1962**		**1.1414**	**-0.1597**	**-0.7730**		**0.0006**	**0.9874**	**0.0104**	**-1,970.04**
	memory	(0.0957)		(0.0323)	(0.0299)	(0.0239)		(0.0005)	(0.0045)	(0.0032)	**3,995.65**
	(2,0,1)-G	(0.0859)		(0.0375)	(0.0345)	(0.0286)		(0.0007)	(0.0056)	(0.0039)	
	Best homoskedastic	**-0.3558**	**0.5388**	**-0.1826**	**-0.0913**			**0.2434**			**-1,996.23**
		(0.2239)	(0.0257)	(0.0309)	(0.0239)			(0.0065)			**4,032.14**
	(2,d,0)	(0.2146)	(0.0282)	(0.0349)	(0.0249)			(0.0078)			
	GARCH ver. of	**-0.3739**	**0.5316**	**-0.1688**	**-0.0849**			**0.0006**	**0.9870**	**0.0106**	**-1,963.06**
lnRV	best homosked.	(0.1529)	(0.0257)	(0.0308)	(0.0239)			(0.0005)	(0.0048)	(0.0034)	**3,981.68**
	(2,d,0)-G	(0.0953)	(0.0263)	(0.0317)	(0.0233)			(0.0007)	(0.0061)	(0.0042)	
	Homoskedastic ver.	**-0.3569**	**0.5553**		**-0.0638**	**-0.1993**		**0.2434**			**-1,996.23**
	of best overall model	(0.1610)	(0.0314)		(0.0210)	(0.0368)		(0.0065)			**4,032.14**
	({2},d,1)	(0.1080)	(0.0348)		(0.0214)	(0.0417)		(0.0078)			
	Best overall	**-0.3740**	**0.5470**		**-0.0623**	**-0.1846**		**0.0006**	**0.9871**	**0.0105**	**-1,962.95**
	({2},d,1)-G	(0.2582)	(0.0324)		(0.0217)	(0.0380)		(0.0005)	(0.0048)	(0.0034)	**3,981.46**
		(0.2639)	(0.0351)		(0.0214)	(0.0414)		(0.0007)	(0.0060)	(0.0041)	

Table 2 (Cont.): ARFIMA-GARCH estimation results

Var.	Model	ARFIMA parameters						GARCH paramters			LL
		μ	d	φ_1	φ_2	ψ_1	ψ_2	σ^2, ω	β	α	BIC
	Best short-memory	**0.7155**		**0.3404**	**0.6116**		**-0.4749**	**0.0134**	**0.7857**	**0.2595**	**-2,365.42**
	memory	(0.0932)		(0.0245)	(0.0267)		(0.0296)	(0.0023)	(0.0223)	(0.0320)	**4,786.40**
	(2,0,{2})-G	(0.1343)		(0.0361)	(0.0425)		(0.0396)	(0.0110)	(0.0932)	(0.1325)	
	Best homoskedastic	**0.7893**	**0.4941**			**-0.2216**		**0.6329**			**-3,335.09**
	(0,d,1)	(0.3326)	(0.0322)			(0.0394)		(0.0169)			**6,701.93**
		(0.2169)	(0.0869)			(0.1086)		(0.1367)			
	GARCH ver. of	**0.4635**	**0.5661**			**-0.2375**		**0.0120**	**0.7993**	**0.2423**	**-2,366.79**
C	best homosked.	(0.2665)	(0.0500)			(0.0626)		(0.0018)	(0.0184)	(0.0274)	**4,781.21**
	(0,d,1)-G	(0.2163)	(0.1022)			(0.1301)		(0.0084)	(0.0730)	(0.1083)	
	Homoskedastic ver.	**0.7704**	**0.5163**		**-0.0321**	**-0.2371**		**0.6324**			**-3,333.97**
	of best overall model	(0.7140)	(0.0342)		(0.0211)	(0.0383)		(0.0168)			**6,707.64**
	({2},d,1)	(0.8760)	(0.0978)		(0.0504)	(0.1010)		(0.1344)			
	Best overall	**0.5185**	**0.6251**		**-0.0966**	**-0.2753**		**0.0120**	**0.7966**	**0.2457**	**-2,360.72**
	({2},d,1)-G	(0.3330)	(0.0467)		(0.0267)	(0.0527)		(0.0019)	(0.0191)	(0.0284)	**4,777.00**
		(0.2393)	(0.0843)		(0.0450)	(0.0879)		(0.0084)	(0.0732)	(0.1099)	
	Best short-memory	**0.6985**		**1.8988**	**-0.8990**	**-1.5908**	**0.5965**	**0.0029**	**0.8630**	**0.0982**	**-47.29**
	memory	(0.0872)		(0.0175)	(0.0174)	(0.0298)	(0.0287)	(0.0007)	(0.0198)	(0.0138)	**158.08**
	(2,0,2)-G	(0.0661)		(0.0188)	(0.0188)	(0.0354)	(0.0342)	(0.0015)	(0.0415)	(0.0271)	
	Best homoskedastic	**0.8151**	**0.5160**			**-0.1847**		**0.0723**			**-295.10**
	(0,d,1)	(0.1152)	(0.0305)			(0.0393)		(0.0019)			**621.94**
		(0.0811)	(0.0515)			(0.0681)		(0.0050)			
	GARCH ver. of	**0.7145**	**0.4990**			**-0.1633**		**0.0027**	**0.8714**	**0.0926**	**-50.53**
$C^{1/2}$	best homosked.	(0.1177)	(0.0297)			(0.0373)		(0.0006)	(0.0179)	(0.0126)	**148.68**
	(0,d,1)-G	(0.1107)	(0.0306)			(0.0397)		(0.0013)	(0.0370)	(0.0248)	
	Homoskedastic ver.	**0.8151**	**0.5160**			**-0.1847**		**0.0723**			**-295.10**
	of best overall model	(0.1152)	(0.0305)			(0.0393)		(0.0019)			**621.94**
	(0,d,1)	(0.0811)	(0.0515)			(0.0681)		(0.0050)			
	Best overall	**0.7145**	**0.4990**			**-0.1633**		**0.0027**	**0.8714**	**0.0926**	**-50.53**
	(0,d,1)-G	(0.1177)	(0.0297)			(0.0373)		(0.0006)	(0.0179)	(0.0126)	**148.68**
		(0.1107)	(0.0306)			(0.0397)		(0.0013)	(0.0370)	(0.0248)	
	Best short-memory	**-0.2637**		**1.1508**	**-0.1685**	**-0.7792**		**0.0002**	**0.9894**	**0.0102**	**-2,100.31**
	memory	(0.1042)		(0.0322)	(0.0298)	(0.0240)		(0.0002)	(0.0032)	(0.0028)	**4,256.19**
	(2,0,1)-G	(0.0945)		(0.0368)	(0.0337)	(0.0289)		(0.0003)	(0.0040)	(0.0036)	
	Best homoskedastic	**-0.5282**	**0.5089**			**-0.1641**		**0.2731**			**-2,157.44**
	(0,d,1)	(0.3443)	(0.0283)			(0.0374)		(0.0073)			**4,346.63**
		(0.4589)	(0.0330)			(0.0465)		(0.0088)			
	GARCH ver. of	**-0.5293**	**0.5030**			**-0.1452**		**0.0002**	**0.9889**	**0.0105**	**-2,095.63**
ln$C$	best homosked.	(0.2520)	(0.0273)			(0.0362)		(0.0002)	(0.0034)	(0.0030)	**4,238.89**
	(0,d,1)-G	(0.1958)	(0.0286)			(0.0398)		(0.0003)	(0.0046)	(0.0039)	
	Homoskedastic ver.	**-0.5282**	**0.5089**			**-0.1641**		**0.2731**			**-2,157.44**
	of best overall model	(0.3443)	(0.0283)			(0.0374)		(0.0073)			**4,346.63**
	(0,d,1)	(0.4589)	(0.0330)			(0.0465)		(0.0088)			
	Best overall	**-0.5293**	**0.5030**			**-0.1452**		**0.0002**	**0.9889**	**0.0105**	**-2,095.63**
	(0,d,1)-G	(0.2520)	(0.0273)			(0.0362)		(0.0002)	(0.0034)	(0.0030)	**4,238.89**
		(0.1958)	(0.0286)			(0.0398)		(0.0003)	(0.0046)	(0.0039)	

The sample period is from March 11, 1996, through August 31, 2007 (2802 observations). For each of the six series, the parameter estimates for the best short-memory model (top), the best homoskedastic model and its GARCH version with the same ARFIMA order, the homoskedastic version of the best model, and the best model (bottom), selected by the BIC from the 64 restricted versions of the ARFIMA(2,d,2)-GARCH(1,1) model, are shown with the Hessian-based (on the first line beneath the parameter estimates) and the Bollerslev-Wooldridge (on the second line) standard errors in parentheses, the log likelihood (LL) and BIC values. For $C^{1/2}$ and ln$C$, the best homooskedastic model and the best model have a common ARFIMA order. The resulting two duplicates (the GARCH version of the best homoskedastic model and the homoskedastic version of the best model) are not omitted from Table 2. The ARFIMA(2,d,1)-GARCH(1,1) model with the first-order AR coefficient restricted to be zero, for example, is denoted as ({2},d,1)-*G*.

Table 3: Residual Diagnostic Statistics

Var.	Model	Residuals ε_t mean	std.	skew.	kurt.	JB	ρ_1	LB(5)	LB(10)	LB(22)	Squared residuals ε_t^2 ρ_1	LB(5)	LB(10)	LB(22)	Standardized residuals Z_t mean	std.	skew.	kurt.	JB	Squared standardizd residuals Z_t^2 ρ_1	LB(5)	LB(10)	LB(22)
RV	Short-mem.	0.033	0.806	6.306	117.688	1,554,221.518	**-0.084**	**53.176**	**82.729**	**126.167**	0.126	118.963	120.793	133.099	0.075	0.997	3.594	31.503	100,878.194	-0.001	19.607	20.903	24.063
	(2,0,{2})-*G*						(0.075)	7.023	19.617	33.220													
	Homosked.	0.001	0.800	7.059	126.740	1,810,892.767	**0.004**	**18.174**	**28.554**	**63.534**	0.070	80.392	81.722	91.537	0.001	1.000	7.059	126.740	1,810,892.767				
	(0,*d*,1)						(0.059)	2.476	7.480	19.087													
	+ GARCH	0.005	0.803	6.556	122.018	1,673,875.665	**-0.071**	**37.032**	**49.867**	**88.438**	0.114	105.383	106.727	117.021	0.023	0.999	3.719	33.868	117,699.427	0.000	22.683	23.826	26.627
	(0,*d*,1)-*G*						(0.073)	4.885	11.312	24.346													
	Homosk..ver.of best	0.001	0.799	6.996	125.325	1,769,826.414	**-0.001**	**13.178**	**25.098**	**60.567**	0.072	82.462	83.810	94.211	0.001	1.000	6.996	125.325	1,769,826.414				
	(2,*d*,0)						(0.060)	1.473	7.115	18.701													
	Best overall	0.003	0.805	6.212	115.545	1,496,806.010	**-0.095**	**52.404**	**71.944**	**114.655**	0.128	123.181	124.726	138.032	0.027	0.999	3.629	32.263	106,124.409	0.000	20.808	22.141	25.063
	(2,*d*,0)-*G*						(0.075)	6.665	15.709	29.049													
$RV^{1/2}$	Short-mem.	0.008	0.265	1.685	14.567	16,946.021	**-0.015**	**7.272**	**22.272**	**41.796**	0.138	220.138	251.703	304.449	0.030	1.000	1.385	8.082	3,910.744	0.010	26.629	29.218	33.834
	(1,0,2)-*G*						(0.032)	3.882	13.162	25.380													
	Homosked.	0.000	0.265	1.761	14.910	18,008.312	**0.006**	**9.366**	**17.903**	**33.033**	0.128	203.585	230.174	275.851	0.000	1.000	1.761	14.910	18,008.312				
	(0,*d*,1)						(0.032)	4.726	10.283	19.831													
	+ GARCH	0.003	0.265	1.773	14.973	18,205.261	**-0.007**	**10.348**	**19.734**	**35.480**	0.131	202.744	228.330	274.450	0.006	1.000	1.400	8.313	4,211.003	0.012	28.392	31.175	35.567
	(0,*d*,1)-*G*						(0.032)	4.970	11.098	20.941													
	Homosk..ver.of best	0.000	0.264	1.729	14.685	17,336.279	**0.000**	**1.697**	**10.961**	**26.835**	0.133	209.458	236.776	286.246	0.000	1.000	1.729	14.685	17,336.279				
	(2,*d*,0)						(0.032)	0.577	6.592	16.590													
	Best overall	0.002	0.264	1.727	14.661	17,267.676	**-0.015**	**3.450**	**13.500**	**30.073**	0.137	211.853	238.638	290.040	0.004	1.000	1.388	8.239	4,103.537	0.013	27.560	30.650	35.000
	(2,*d*,0)-*G*						(0.032)	1.396	7.936	18.270													
ln RV	Short-mem.	-0.002	0.495	0.268	3.905	129.249	**-0.008**	**3.446**	**11.024**	**21.219**	0.082	51.156	78.289	88.733	-0.004	0.995	0.304	3.707	101.444	0.038	23.449	27.811	33.160
	(2,0,1)-*G*						(0.021)	2.943	9.552	19.226													
	Homosked.	-0.001	0.493	0.278	3.916	134.075	**-0.001**	**0.318**	**4.702**	**13.178**	0.094	57.204	82.510	92.203	-0.002	1.000	0.278	3.916	134.075				
	(2,*d*,0)						(0.021)	0.270	4.211	12.326													
	+ GARCH	-0.001	0.493	0.280	3.913	133.913	**-0.007**	**0.541**	**5.062**	**13.636**	0.094	57.203	82.859	92.823	-0.004	0.995	0.303	3.735	105.983	0.047	25.559	30.920	36.709
	(2,*d*,0)-*G*						(0.021)	0.447	4.521	12.732													
	Homosk..ver.of best	-0.001	0.493	0.281	3.922	136.115	**0.000**	**0.168**	**4.497**	**12.902**	0.093	56.288	81.684	91.640	-0.002	1.000	0.281	3.922	136.115				
	({2},*d*,1)						(0.021)	0.145	4.007	12.044													
	Best overall	-0.001	0.493	0.282	3.919	135.664	**-0.007**	**0.468**	**4.876**	**13.355**	0.093	56.378	82.082	92.258	-0.004	0.995	0.307	3.741	107.926	0.046	25.143	30.420	36.284
	({2},*d*,1)-*G*						(0.021)	0.393	4.337	12.448													
C	Short-mem.	0.030	0.801	6.472	122.420	1,684,532.807	**-0.061**	**32.707**	**61.594**	**109.406**	0.112	83.031	85.197	98.084	0.071	0.997	3.295	25.344	63,358.309	-0.002	31.719	33.070	36.102
	(2,0,{2})-*G*						(0.072)	5.340	17.063	30.983													
	Homosked.	0.003	0.796	7.167	131.645	1,956,131.830	**0.003**	**13.625**	**22.272**	**62.813**	0.070	52.801	54.189	64.745	0.004	1.000	7.167	131.645	1,956,131.830				
	(0,*d*,1)						(0.060)	2.237	6.465	19.133													
	+ GARCH	0.007	0.798	6.740	127.021	1,816,978.170	**-0.049**	**23.326**	**34.366**	**77.775**	0.101	70.518	72.028	83.001	0.026	0.999	3.403	27.339	74,569.180	-0.002	32.466	33.766	36.753
	(0,*d*,1)-*G*						(0.070)	3.848	9.357	23.242													
	Homosk..ver.of best	0.003	0.795	7.098	130.267	1,914,529.203	**-0.003**	**9.912**	**20.051**	**61.243**	0.074	55.928	57.387	68.599	0.004	1.000	7.098	130.267	1,914,529.203				
	({2},*d*,1)						(0.061)	1.619	6.422	19.014													
	Best overall	0.004	0.799	6.483	122.088	1,675,364.830	**-0.071**	**24.334**	**40.853**	**88.084**	0.117	86.038	87.770	101.463	0.027	0.999	3.360	26.619	70,401.910	-0.002	32.707	34.055	36.937
	({2},*d*,1)-*G*						(0.073)	3.563	11.070	25.013													
$C^{1/2}$	Short-mem.	0.009	0.269	1.726	14.521	16,888.778	**0.028**	**6.392**	**9.984**	**31.365**	0.135	182.378	217.176	273.101	0.028	1.000	1.310	7.685	3,363.966	0.013	28.953	32.971	37.029
	(2,0,2)-*G*						(0.032)	3.167	5.574	18.568													
	Homosked.	0.001	0.269	1.712	14.590	17,050.511	**0.005**	**7.183**	**11.215**	**28.938**	0.137	179.276	214.392	268.211	0.004	1.000	1.712	14.590	17,050.511				
	(0,*d*,1)						(0.032)	3.859	6.519	17.750													
	Best overall	0.004	0.269	1.730	14.660	17,270.139	**0.001**	**7.538**	**12.025**	**30.246**	0.138	177.706	211.891	265.939	0.009	1.000	1.332	7.801	3,518.869	0.011	34.676	38.197	42.318
	(0,*d*,1)-*G*						(0.032)	3.882	6.803	18.233													
ln C	Short-mem.	-0.001	0.524	0.196	3.897	111.742	**-0.016**	**5.457**	**10.085**	**19.664**	0.099	68.566	133.160	171.932	-0.003	0.987	0.250	3.636	76.444	0.033	24.887	30.485	38.184
	(2,0,1)-*G*						(0.021)	4.502	8.199	17.053													
	Homosked.	0.002	0.523	0.219	3.908	118.640	**0.006**	**5.708**	**8.300**	**16.803**	0.112	78.223	140.330	175.775	0.004	1.000	0.219	3.908	118.640				
	(0,*d*,1)						(0.022)	5.056	7.264	15.227													
	Best overall	0.002	0.523	0.221	3.908	119.116	**-0.007**	**6.063**	**8.751**	**17.454**	0.112	77.339	140.100	176.198	0.000	0.988	0.250	3.666	80.957	0.041	26.919	33.700	41.156
	(0,*d*,1)-*G*						(0.022)	5.367	7.664	15.808													

JB, ρ_1, and $LB(k)$ stand respectively for the Jarque-Bera statistic for nonnormality, first-order sample autocorrelation, and the Ljung-Box statistic for no serial correlations of orders up to k (LB for the squared residuals ε_t^2 are also called the McLeod-Li statistics). For the residuals ε_t, the heteroskedasticity-consistent standard errors for ρ_1 are given in parentheses (the usual standard error is $T^{-1/2} = 0.018$) and both the usual LB (upper lines) and the heteroskedasticity-adjusted LB (lower lines) are shown. The 5% critical values for $\chi^2(k)$ are 5.991 (k=2), 7.815 (3), 9.488 (4), 11.070 (5), 12.592 (6), 14.067 (7), 15.507 (8), 16.919 (9), 18.307 (10), 19.675 (11), 21.026 (12), 22.262 (13), 23.685 (14), 24.996 (15), 26.296 (16), 27.587 (17), 28.869 (18), 30.144 (19), 31.410 (20), 32.671 (21), 33.924 (22).

Table 4: Further specification tests

Var.	Model	BDS		Line 1: $\ln Z_t^2$ Line 2: Z_t		CCK $\gamma = 0.5$			CCK $\gamma = 1.0$			Hong-Li				Hong-Li separate inference					
		$k=2$	3	4	5	$k=1$	2	3	$k=1$	2	3	$Q_{HL}(1)$	$W_{HL}(5)$	$W_{HL}(10)$	$W_{HL}(20)$	$(m,l) = (1,1)$	(1,2) ARCH-m	(2,1) Leverage	(2,2)	(3,3)	(4,4)
RV	Short-mem.	**-0.002**	**-0.456**	**-0.593**	**-3.241**	4.716	5.451	1.834	2.002	3.158	3.359	139.297	300.546	418.937	613.858	-0.216	1.038	0.229	1.198	-1.363	-2.867
	$(2,0,\{2\})$-G	2.408	2.557	3.411	4.695																
	Homosked.	**12.646**	**12.832**	**13.249**	**-3.274**	5.577	4.750	4.683	4.469	5.177	4.230	340.963	730.094	1,002.689	1,435.300	5.420	13.893	25.255	14.293	-1.095	-2.973
	$(0,d,1)$	18.630	21.912	25.045	28.791																
	+ GARCH	**0.820**	**0.627**	**0.431**	**-2.134**	4.768	5.350	3.181	3.084	2.157	2.408	140.927	301.625	418.049	611.230	-0.213	1.416	0.311	1.790	-1.088	-2.777
	$(0,d,1)$-G	3.782	4.491	5.406	6.766																
	Homosk.. ver. of best	**12.913**	**13.312**	**13.865**	**-3.360**	5.570	4.859	4.260	4.112	5.708	4.664	341.694	728.647	1,001.012	1,431.934	4.608	15.167	25.546	14.788	-1.051	-2.959
	$(2,d,0)$	18.488	21.731	24.901	28.667																
	Best overall	**1.176**	**0.913**	**0.688**	**-2.915**	5.047	5.735	1.370	1.594	3.602	3.690	134.742	287.149	399.346	585.305	-1.223	0.622	0.475	1.454	-1.190	-2.769
	$(2,d,0)$-G	3.363	3.613	4.375	5.659																
$RV^{1/2}$	Short-mem.	**-0.540**	**-0.796**	**-0.779**	**-2.506**	2.344	2.598	0.322	-0.036	1.196	1.195	36.745	79.042	110.621	161.434	-0.880	0.517	2.622	3.203	5.156	4.152
	$(1,0,2)$-G	1.226	1.341	1.930	2.942																
	Homosked.	**2.822**	**2.822**	**2.672**	**-1.855**	2.576	2.379	1.748	1.378	2.039	1.664	56.780	123.072	167.576	239.940	1.627	2.240	23.649	48.013	14.567	7.456
	$(0,d,1)$	9.674	12.136	14.470	17.515																
	+ GARCH	**0.382**	**-0.017**	**-0.177**	**-1.761**	1.996	2.260	0.939	0.468	0.842	0.932	36.962	80.971	112.648	164.097	-0.021	1.011	2.521	3.580	5.668	4.303
	$(0,d,1)$-G	1.904	2.136	2.663	3.627																
	Homosk.. ver. of best	**2.706**	**2.640**	**2.549**	**-2.570**	2.729	2.633	1.368	1.164	2.438	1.998	56.167	120.665	164.562	236.100	0.147	2.924	23.733	49.772	14.781	7.845
	$(2,d,0)$	9.388	11.733	14.029	17.040																
	Best overall	**0.480**	**0.025**	**-0.130**	**-2.235**	2.249	2.617	0.287	-0.014	1.275	1.294	37.116	78.697	109.086	159.376	-1.723	0.408	2.728	3.495	5.239	4.127
	$(2,d,0)$-G	1.911	2.054	2.505	3.432																
$\ln RV$	Short-mem.	**0.001**	**-0.023**	**0.007**	**-0.838**	-1.434	-0.667	-2.216	-2.240	-0.673	-0.708	2.815	6.414	8.762	11.689	-0.522	3.217	0.339	3.148	7.241	25.083
	$(2,0,1)$-G	1.677	1.914	2.292	3.199																
	Homosked.	**0.363**	**0.112**	**-0.075**	**-1.691**	-1.262	-0.603	-2.092	-2.083	-0.653	-0.633	4.618	9.634	13.701	18.604	-1.449	5.065	1.187	14.761	5.425	14.710
	$(2,d,0)$	3.512	4.171	4.577	5.629																
	+ GARCH	**0.302**	**0.017**	**-0.240**	**-1.587**	-1.479	-0.635	-2.171	-2.176	-0.755	-0.705	3.677	6.522	8.524	11.296	-1.754	3.674	0.207	3.904	6.812	23.126
	$(2,d,0)$-G	2.157	2.564	2.811	3.620																
	Homosk.. ver. of best	**0.200**	**0.043**	**-0.098**	**-1.379**	-1.254	-0.615	-2.134	-2.129	-0.636	-0.654	4.686	9.729	13.761	18.607	-1.478	5.049	1.253	14.583	5.440	14.461
	$(\{2\},d,1)$	3.407	4.008	4.461	5.478																
	Best overall	**0.202**	**-0.043**	**-0.210**	**-1.363**	-1.471	-0.648	-2.207	-2.215	-0.751	-0.726	3.579	6.414	8.387	11.142	-1.836	3.754	0.261	3.799	6.632	22.747
	$(\{2\},d,1)$-G	2.110	2.492	2.747	3.545																
C	Short-mem.	**1.296**	**1.085**	**0.881**	**-1.896**	4.547	5.571	1.925	2.155	3.095	3.478	135.709	296.375	412.433	602.941	-0.188	0.690	0.117	3.569	3.368	0.896
	$(2,0,\{2\})$-G	1.511	2.004	3.174	4.439																
	Homosked.	**12.201**	**13.217**	**14.048**	**-2.592**	5.179	4.923	4.328	4.281	4.953	4.424	342.874	740.090	1,011.522	1,441.198	4.639	13.205	22.458	8.789	-2.623	-3.358
	$(0,d,1)$	18.880	22.370	25.656	29.452																
	+ GARCH	**1.947**	**2.557**	**3.709**	**5.024**	4.290	5.353	3.055	3.098	2.163	2.664	137.079	297.039	414.158	605.267	-0.619	0.854	0.138	3.689	3.008	0.419
	$(0,d,1)$-G	0.913	1.159	1.304	-1.792																
	Homosk.. ver. of best	**12.651**	**13.742**	**14.481**	**-2.846**	5.242	5.071	3.835	3.833	5.168	4.649	341.469	735.879	1,006.099	1,433.161	4.110	14.516	22.997	9.529	-2.598	-3.352
	$(\{2\},d,1)$	18.891	22.310	25.595	29.404																
	Best overall	**1.079**	**1.352**	**1.333**	**-2.360**	4.895	5.854	1.373	1.621	2.758	3.228	134.695	288.186	402.269	589.806	-1.831	0.251	0.233	3.754	3.202	0.737
	$(\{2\},d,1)$-G	1.861	2.245	3.274	4.589																
$C^{1/2}$	Short-mem.	**0.550**	**0.259**	**0.026**	**-1.182**	1.694	2.362	0.894	0.500	1.425	1.725	30.906	68.276	95.601	139.663	-0.358	-0.240	1.551	3.768	7.181	8.302
	$(2,0,2)$-G	0.818	0.912	1.846	2.922																
	Homosked.	**2.712**	**2.868**	**2.915**	**-1.253**	2.498	2.800	1.340	1.038	2.215	2.206	52.669	116.046	158.849	226.107	0.624	1.522	23.293	45.588	5.783	0.825
	$(0,d,1)$	9.168	11.606	14.520	17.888																
	Best overall	**0.351**	**0.288**	**0.094**	**-0.889**	2.097	2.663	0.828	0.334	1.102	1.423	33.284	73.196	102.143	149.058	-1.010	0.334	1.985	4.840	10.672	13.526
	$(0,d,1)$-G	0.967	1.171	2.278	3.456																
$\ln C$	Short-mem.	**-0.249**	**-0.338**	**-0.515**	**-0.154**	-1.229	-0.422	-2.242	-2.538	-0.341	-0.084	0.489	2.636	3.312	4.350	-0.494	2.941	-0.128	3.731	8.468	27.987
	$(2,0,1)$-G	0.802	0.874	1.373	2.197																
	Homosked.	**0.248**	**0.342**	**0.382**	**-0.153**	-1.531	-0.737	-2.347	-2.481	-0.202	-0.094	3.677	8.651	11.500	15.658	-0.755	5.292	1.415	28.678	6.514	21.892
	$(0,d,1)$	3.291	4.015	5.097	6.344																
	Best overall	**0.376**	**0.391**	**0.451**	**-0.184**	-1.536	-0.754	-2.099	-2.407	-0.430	-0.215	1.725	3.017	3.681	5.178	-0.846	2.767	-0.144	4.459	10.056	28.736
	$(0,d,1)$-G	1.155	1.192	1.675	2.476																

Table 5: In-sample forecast performance evaluation

Variable / Selected Model	Horizon : Day (s)	RV prediction RMSE	Mincer-Zarnowits Regression Int.	Coef.	R^2	χ^2	Volatility-of-the RV prediction RMSE	Mincer-Zarnowits Regression Int.	Coef.	R^2	χ^2
	1	0.8048	0.1245	0.8878	0.3959	8.325	7.2603	0.3977	0.2943	0.0201	103.968
			(0.0459)	(0.0485)		(0.0156)		(0.0987)	(0.0832)		(0.0000)
RV	5	0.9054	0.2291	0.7938	0.2448	17.159	4.2408	0.4836	0.2842	0.0687	177.385
(2,d,0)-G			(0.0554)	(0.0560)		(0.0002)		(0.0906)	(0.0612)		(0.0000)
	10	0.9491	0.2956	0.7345	0.1765	31.805	3.9477	0.6027	0.2240	0.0598	421.560
			(0.0527)	(0.0521)		(0.0000)		(0.1161)	(0.0389)		(0.0000)
	22	1.0100	0.4598	0.5856	0.0885	34.744	3.6034	0.7313	0.1650	0.0562	824.011
			(0.0837)	(0.0711)		(0.0000)		(0.1392)	(0.0281)		(0.0000)
	1	0.2643	0.0162	0.9856	0.5347	1.130	0.2548	0.0190	0.7189	0.0363	4.216
			(0.0201)	(0.0228)		(0.5684)		(0.0095)	(0.1700)		(0.1215)
$RV^{1/2}$	5	0.3041	0.0480	0.9551	0.3856	2.959	0.1628	0.0276	0.7678	0.0823	15.915
(2,d,0)-G			(0.0316)	(0.0351)		(0.2278)		(0.0082)	(0.1349)		(0.0003)
	10	0.3244	0.0822	0.9215	0.3033	4.279	0.1557	0.0395	0.7313	0.0687	21.515
			(0.0424)	(0.0462)		(0.1177)		(0.0088)	(0.1116)		(0.0000)
	22	0.3506	0.1751	0.8277	0.1942	6.364	0.1421	0.0524	0.7261	0.0544	18.512
			(0.0694)	(0.0701)		(0.0415)		(0.0140)	(0.1528)		(0.0001)
	1	0.4934	-0.0022	0.9931	0.5982	0.192	0.4122	0.0088	0.9528	0.0170	0.131
			(0.0099)	(0.0159)		(0.9084)		(0.0443)	(0.1914)		(0.9366)
	5	0.5675	-0.0053	0.9818	0.4691	0.394	0.2846	0.0485	0.9710	0.0368	25.042
lnRV			(0.0160)	(0.0297)		(0.8214)		(0.0440)	(0.1870)		(0.0000)
({2},d,1)-G	10	0.6065	-0.0092	0.9675	0.3947	0.589	0.2720	0.0670	1.0255	0.0465	50.321
			(0.0218)	(0.0438)		(0.7450)		(0.0493)	(0.2056)		(0.0000)
	22	0.6567	-0.0203	0.9222	0.2929	1.513	0.2766	0.0820	1.1473	0.0615	77.823
			(0.0292)	(0.0694)		(0.4694)		(0.0588)	(0.2421)		(0.0000)
	1	0.7986	0.1063	0.9014	0.3933	7.074	7.3971	0.4192	0.2631	0.0159	156.044
			(0.0433)	(0.0484)		(0.0291)		(0.1091)	(0.0676)		(0.0000)
C	5	0.9006	0.2224	0.7930	0.2397	17.408	4.3643	0.5073	0.2544	0.0547	277.402
({2},d,1)-G			(0.0534)	(0.0565)		(0.0002)		(0.0976)	(0.0488)		(0.0000)
	10	0.9450	0.2934	0.7273	0.1701	31.865	4.0076	0.6141	0.2062	0.0517	590.623
			(0.0523)	(0.0534)		(0.0000)		(0.1195)	(0.0329)		(0.0000)
	22	1.0052	0.4571	0.5727	0.0835	36.492	3.6497	0.7351	0.1543	0.0503	856.786
			(0.0816)	(0.0715)		(0.0000)		(0.1407)	(0.0277)		(0.0000)
	1	0.2689	0.0086	0.9951	0.5269	1.042	0.2641	0.0197	0.7190	0.0355	4.728
			(0.0198)	(0.0230)		(0.5939)		(0.0093)	(0.1630)		(0.0940)
$C^{1/2}$	5	0.3090	0.0271	0.9800	0.3764	1.997	0.1706	0.0306	0.7422	0.0760	17.726
(0,d,1)-G			(0.0319)	(0.0364)		(0.3685)		(0.0081)	(0.1272)		(0.0001)
	10	0.3294	0.0592	0.9490	0.2932	3.004	0.1634	0.0425	0.7081	0.0650	20.623
			(0.0437)	(0.0489)		(0.2227)		(0.0096)	(0.1148)		(0.0000)
	22	0.3548	0.1491	0.8573	0.1866	4.487	0.1498	0.0561	0.6952	0.0527	18.422
			(0.0723)	(0.0750)		(0.1061)		(0.0149)	(0.1611)		(0.0001)
	1	0.5226	0.0015	0.9979	0.5843	0.063	0.4585	0.0176	0.9062	0.0331	1.083
			(0.0106)	(0.0163)		(0.9692)		(0.0305)	(0.1204)		(0.5819)
ln$C$	5	0.5996	0.0070	1.0080	0.4534	0.170	0.3142	0.0584	0.9345	0.0729	21.464
(0,d,1)-G			(0.0175)	(0.0311)		(0.9184)		(0.0315)	(0.1198)		(0.0000)
	10	0.6386	0.0072	1.0031	0.3808	0.094	0.2987	0.0761	0.9912	0.0930	43.871
			(0.0240)	(0.0462)		(0.9542)		(0.0369)	(0.1354)		(0.0000)
	22	0.6885	-0.0011	0.9656	0.2813	0.224	0.3006	0.0897	1.1124	0.1250	69.770
			(0.0319)	(0.0750)		(0.8939)		(0.0468)	(0.1730)		(0.0000)

In volatility prediction, the target variable is RV, $RV^{1/2}$, lnRV, C, $C^{1/2}$, or ln$C$ *(one-day-ahead, or multi-days-ahead single-day)*. In volatility-of-RV prediction, the proxy for the target is the squared prediction errors (*one-day-ahead* or *multi-days average*) from the selected model for each of the six series. Newey-West standard errors for the OLS estimates of the regression intercept and the coefficient are given in parentheses. The χ^2 test statistic with p-value in parentheses is for testing the joint hypothesis of the intercept being zero and the coefficient being one.

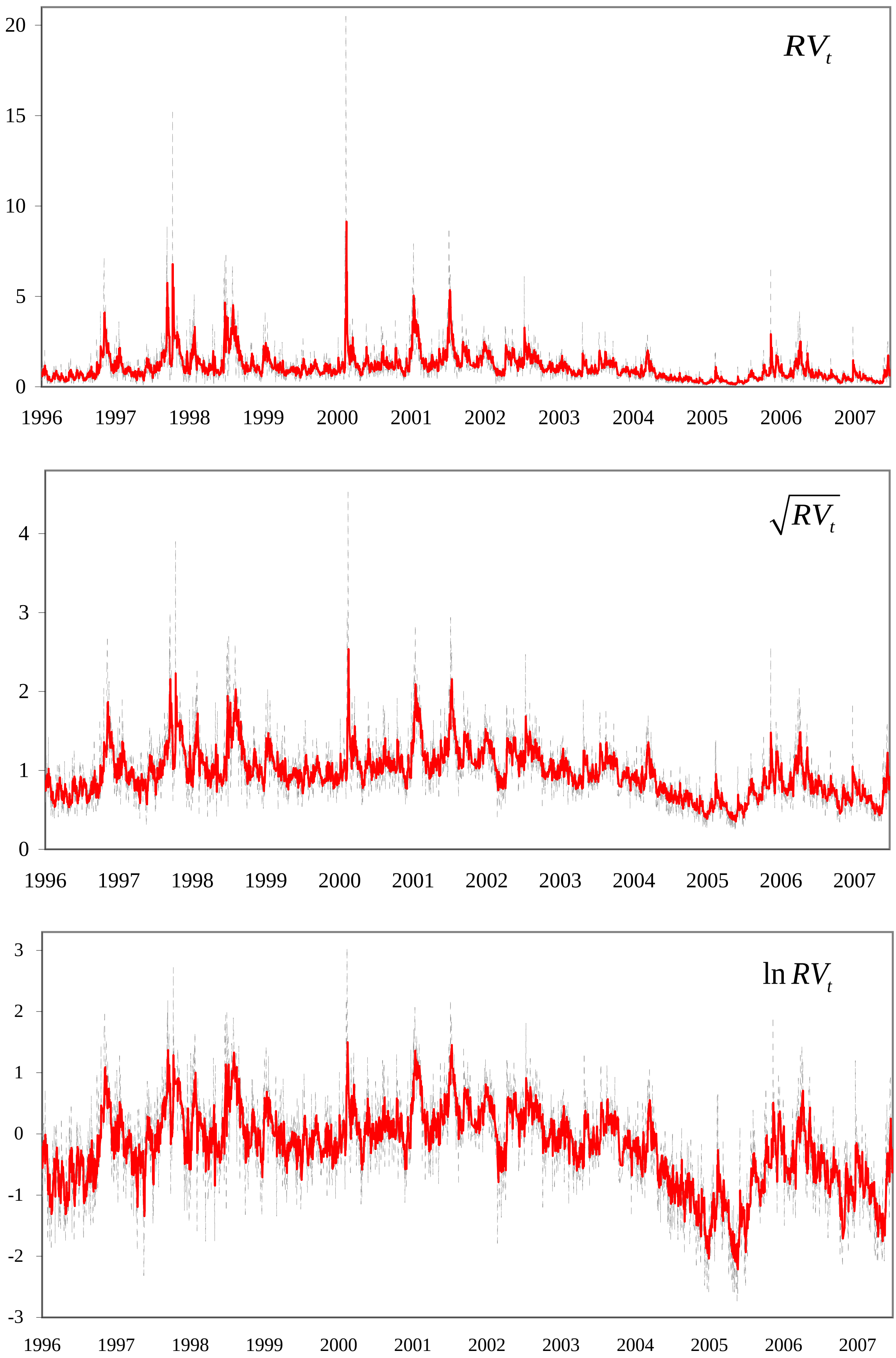

Figure 1a: The figure plots time series of daily Nikkei 225 RV, square-root RV, and log RV (the dotted lines), along with the corresponding in-sample one-day-ahead ARFIMA-GARCH forecasts (the solid lines).

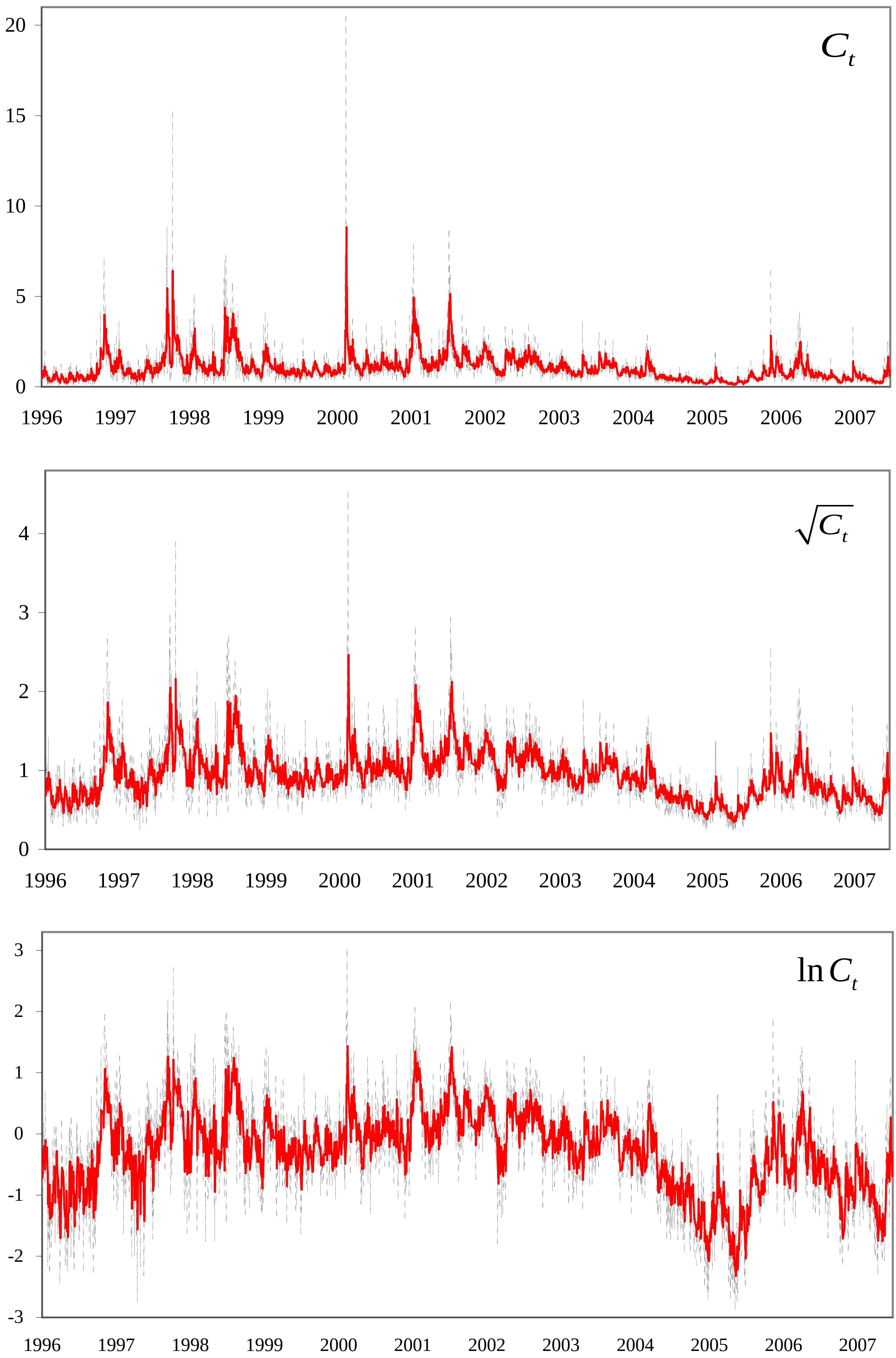

Figure 1b: The figure plots time series of daily Nikkei 225 C, $C^{1/2}$, and $\ln C$ (the dotted lines) where C is the continuous sample path component of RV, along with the corresponding in-sample one-day-ahead ARFIMA-GARCH forecasts (the solid lines).

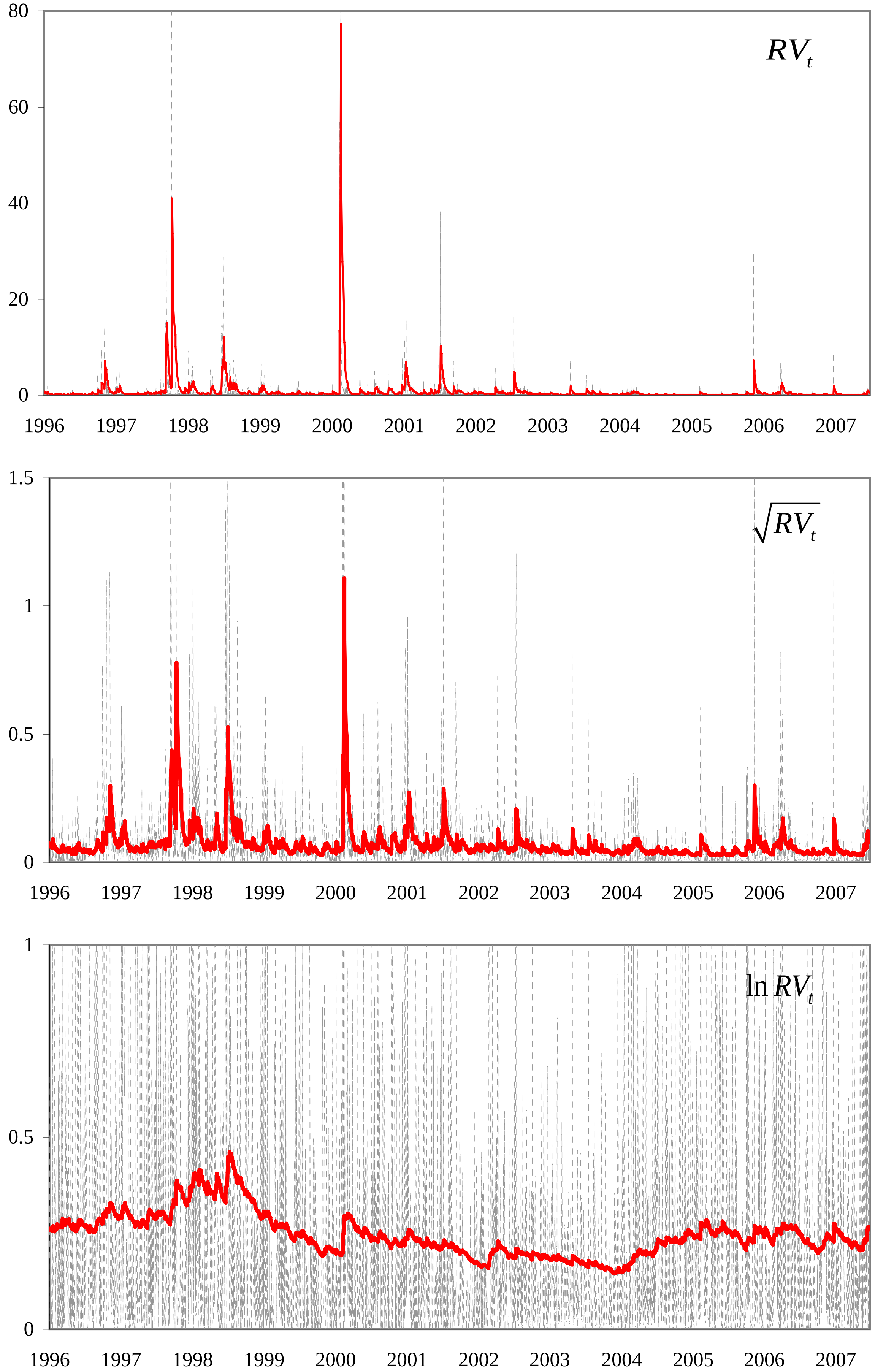

Figure 2a: The figure plots time series of squared residuals (dotted lines) and one-day-ahead GARCH conditional variance estimate (solid lines) from the ARFIMA-GARCH model for daily Nikkei 225 RV, $RV^{1/2}$, and $\ln RV$.

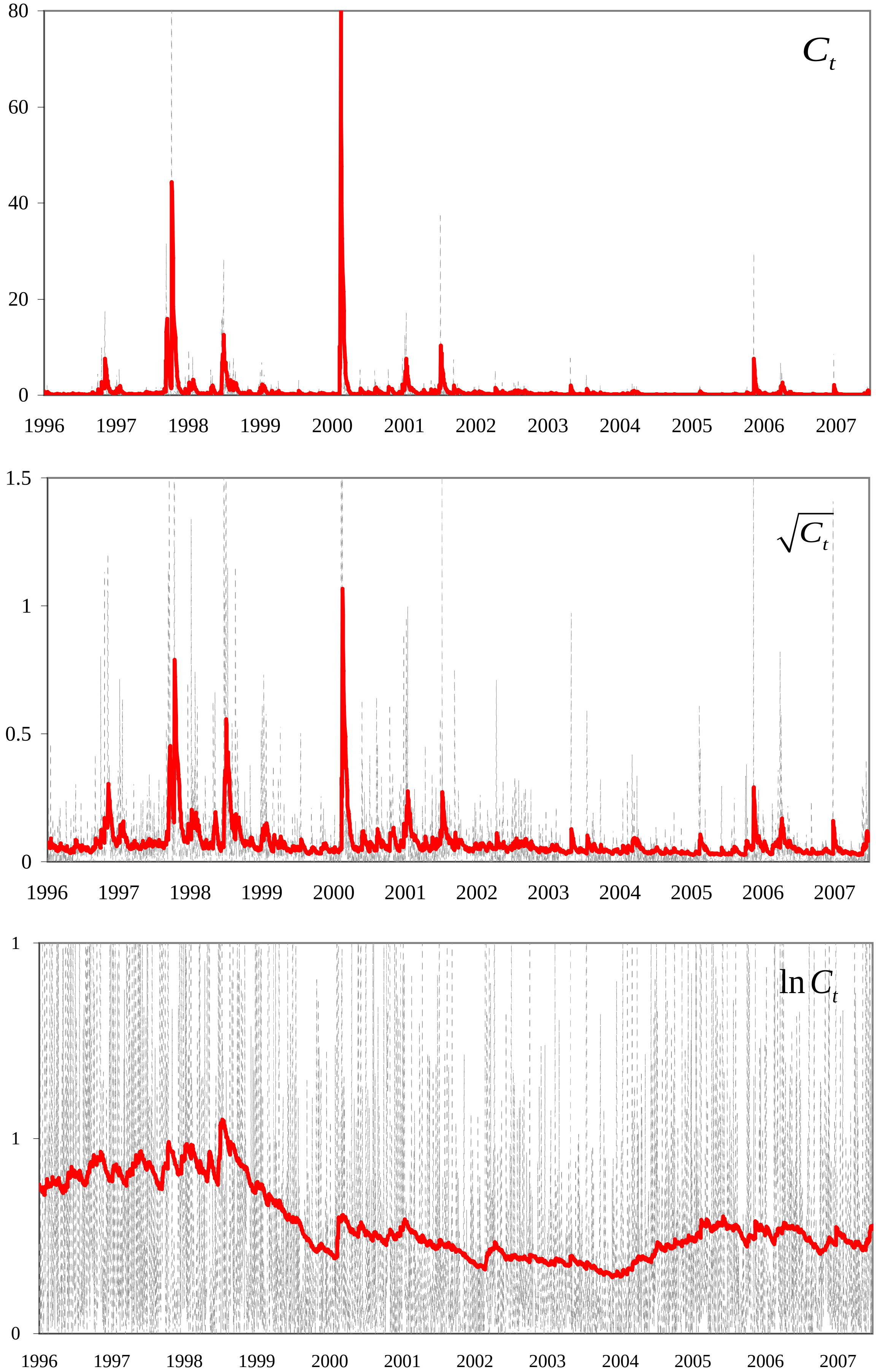

Figure 2b: The figure plots time series of squared residuals (dotted lines) and one-day-ahead GARCH conditional variance estimate (solid lines) from the ARFIMA-GARCH model for daily Nikkei 225 C, $C^{1/2}$, and $\ln C$ (the dotted lines) where C is the continuous sample path component of RV.

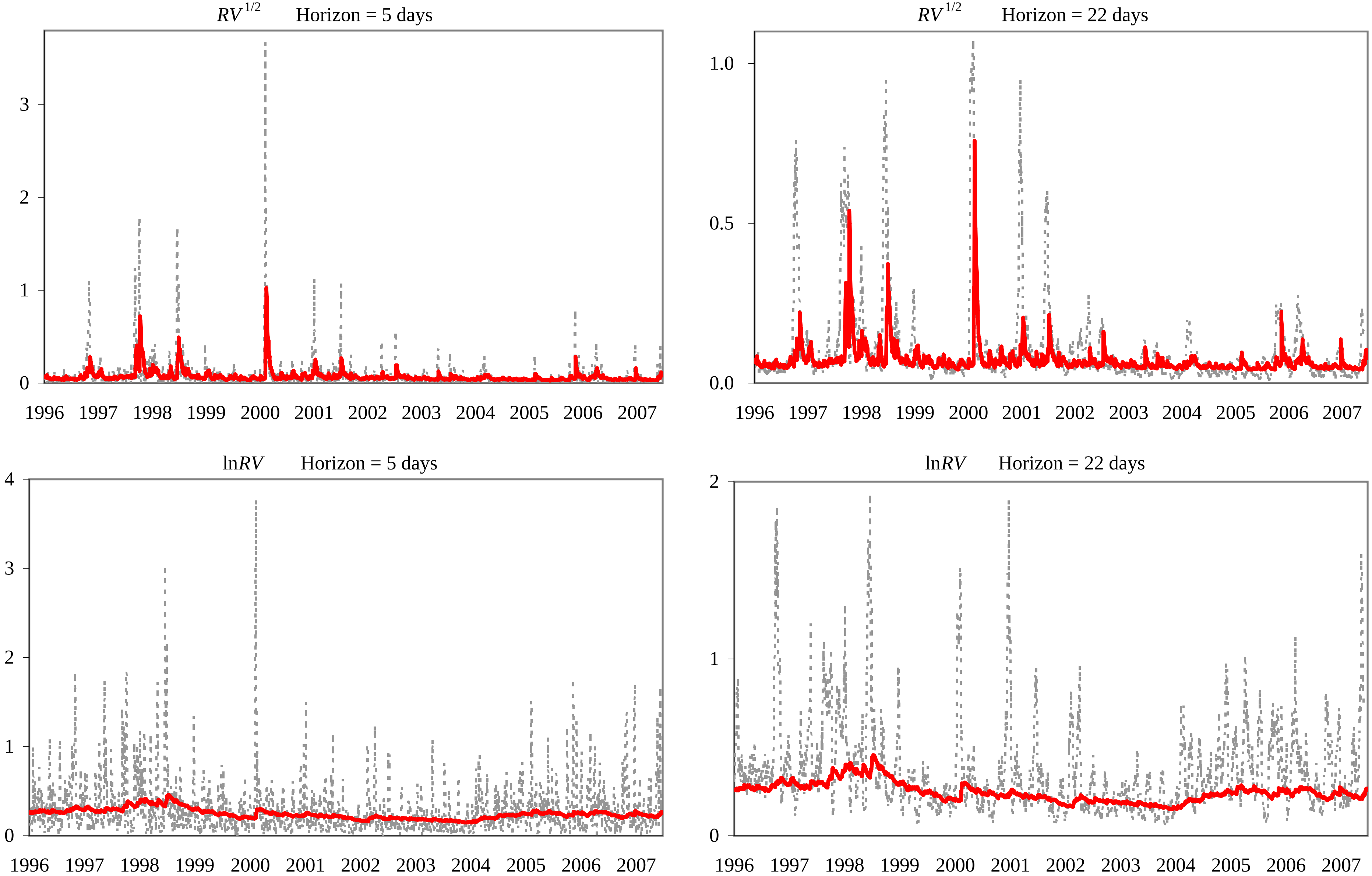

Figure 3a: The figure plots time series of squared residuals (dotted lines) and five-day-average and 22-day-average conditional variance estimates (solid lines) from the ARFIMA-GARCH model for daily Nikkei 225 $RV^{1/2}$ and lnRV.

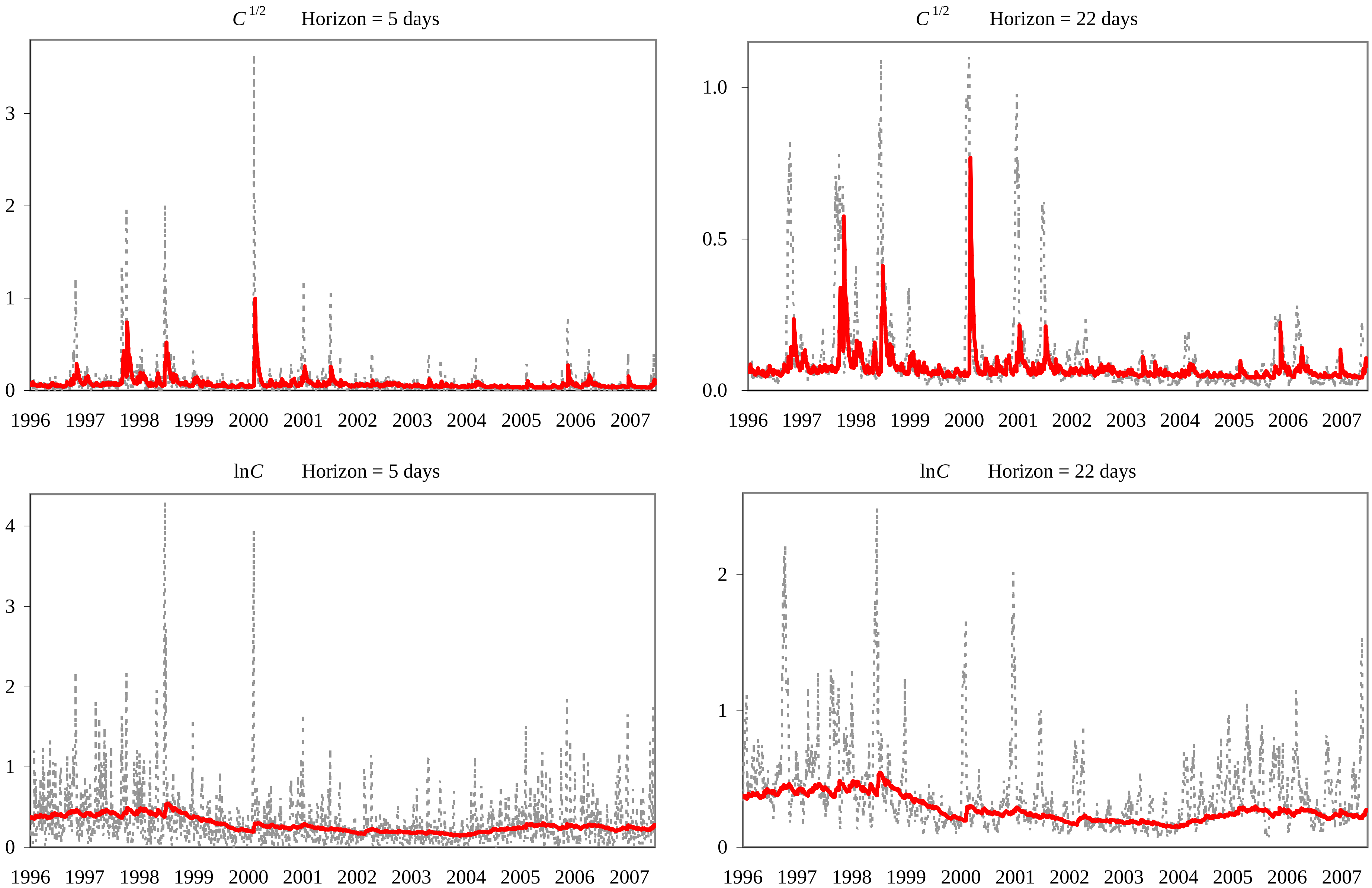

Figure 3b: The figure plots time series of squared residuals (dotted lines) and five-day- average and 22-day-average conditional variance estimates (solid lines) from the ARFIMA-GARCH model for daily Nikkei 225 $C^{1/2}$ and $\ln C$.