



No. 2002/04

**Forecasting Stock Market Volatility and
the Informational Efficiency of the DAX-
index Options Market**

Holger Claessen / Stefan Mittnik



Center for Financial Studies

an der Johann Wolfgang Goethe-Universität ■ Taunusanlage 6 ■ D-60329 Frankfurt am Main

Tel: (+49)069/242941-0 ■ Fax: (+49)069/242941-77 ■ E-Mail: ifk@ifk-cfs.de ■ Internet: <http://www.ifk-cfs.de>

Forecasting Stock Market Volatility and the Informational Efficiency of the DAX-index Options Market*

Holger Claessen / Stefan Mittnik **

Abstract:

Alternative strategies for predicting stock market volatility are examined. In out-of-sample forecasting experiments implied-volatility information, derived from contemporaneously observed option prices or history-based volatility predictors, such as GARCH models, are investigated, to determine if they are more appropriate for predicting future return volatility. Employing German DAX-index return data it is found that past returns do not contain useful information beyond the volatility expectations already reflected in option prices. This supports the efficient market hypothesis for the DAX-index options market.

JEL Classification: G14, C22, C53

Keywords: Market efficiency, implied volatility, GARCH, combined forecasting

* Support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. We thank the Deutsche Börse AG for providing the data.

** Holger Claessen, Deutsche Bank AG, 31 West 52nd Street, New York, NY 10019-6160, U.S.A.

Prof. Stefan Mittnik, Ph.D., Institute of Statistics and Econometrics, University of Kiel, Olshausenstr. 40, D-24118 Kiel; phone: +49 (431)-880-2166; fax: +49 (431)-880-7605; email: mittnik@stat-econ.uni-kiel.de and Center for Financial Studies, Taunusanlage 6, D-60329 Frankfurt

1 Introduction

The expected volatility of financial markets is a key variable in financial investment decisions. For example, it is common practice to reduce asset allocation decisions to a two-dimensional decision problem by focusing solely on the expected return and risk of an asset or portfolio, with risk being related to the volatility of the returns. The volatility of returns plays also a central role in the valuation of financial derivatives such as options and futures, and can, in fact, have a greater influence on the value of derivative securities than price movements in the underlying assets. Options can be used either as part of a dynamic hedging strategy, to protect a portfolio against adverse price movements, or as a speculative asset which gains from expected price changes. To assess the fair value of an option or to hedge market risk, an investor needs to specify his or her expectations regarding future volatility.

There are basically two approaches to generate volatility forecasts. One is to extract information about the variance of future returns from their history; the second is to elicit market expectations about the future volatility from observed option prices. Assuming that an option pricing model correctly represents investors' behavior, the *implied volatility* can be derived from observed option prices and other observable variables by appropriately "inverting" the option pricing model. In informationally efficient markets the implied volatility should reflect the information contained in past returns. Hence, forecasts based on past returns should not outperform forecasts based on implied volatilities.

A number of empirical investigations support the idea of using implied volatility as a predictor for future volatility (see, for example, Latane and Rendleman, 1976; Chiras and Manaster, 1978; Gemmill, 1986; Shastri and Tandon, 1986; Scott and Tucker, 1989). However, the null hypothesis that historic returns add no information to that already contained in implied volatilities was empirically rejected by Day and Lewis (1992) in the case of the S&P100 index and by Lamoureux and Lastrapes (1993), who investigated several individual stocks. Xu and Taylor (1995) found support for this hypothesis using currency options traded on the Philadelphia stock exchange. These three studies follow the approach of Day and Lewis (1992) and consider implied volatility as an exogenous variable in the conditional variance equation of a GARCH(1,1) model. By doing so, one combines current market expectations, as reflected in option prices, with past return information captured by a

standard GARCH model. The importance of the information contained in these two different sources is then judged by the statistical significance of the corresponding parameter estimates.

In this paper we investigate a range of alternative strategies for predicting volatility in financial markets. We apply these approaches to daily returns on the DAX index, the major German stock index, and analyze their in- and out-of-sample performance. As a measure of implied volatility we will use the *Volatility DAX* (VDAX). The VDAX is a publically reported volatility measure, which has recently been introduced by the German options and futures exchange (DTB).²

The paper is organized as follows. Section 2 briefly summarizes the forecasting strategies considered. In-sample results are discussed in Section 3; out-of-sample results are presented in Section 4. Section 5 concludes.

2 Volatility Forecasting Models

Before discussing specific volatility forecasting models the question of how to approximate volatility, which is an unobservable variable, needs to be addressed. Given daily return data, the sample standard deviation over a time interval spanning the h trading days $T + 1, \dots, T + h$, i.e.,

$$V_{T+1, T+h} = \sqrt{\frac{1}{h-1} \sum_{i=1}^h (r_{T+i} - \bar{r}_{T+i})^2},$$

is commonly used as the estimate of this period's average volatility. Here, r_t denotes the asset return for trading day t ; and $\bar{r}_t = \frac{1}{h} \sum_{i=1}^h r_{t+i}$ is the average return over this period. Assuming 250 trading days per year, $V_{T+1, T+h} \sqrt{250}$ represents the annualized average volatility. Below, we will use $V_{T+1, T+h}$ as the true future average volatility over interval $[T + 1, T + h]$.

We investigate eight alternative approaches to forecasting stock market volatility. Two of them, the moving average and the random walk model, use information about past returns in a rather naive manner. We also consider a standard GARCH(1,1) model; a modified GARCH(1,1) model taking weekend and holiday effects into account; an autoregressive model for squared past returns; implied volatility (IV) information; a GARCH(1,1) model combined with IV information; and, finally, we

²Details on the construction of the VDAX will be given in the subsequent section.

consider combined forecasts, following the lines of Granger and Ramanathan (1984). The remainder of this section briefly summarizes these eight approaches.

2.1 Historical Moving Average Model

A widely used estimator for future volatility is the square root of a moving average of past squared returns. If we adjust for mean returns, the volatility forecast is given by the sample standard deviation

$$\hat{V}_{T+1,T+h} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (r_{T-i} - \bar{r}_T)^2} \quad .$$

Below we choose a window length of one calendar year, i.e., $N = 250$ trading days.

2.2 Random Walk Model

If we assume a simple random walk model for the average volatility, the forecast for the next h -day interval is given by the volatility of the past h days, i.e.,

$$\hat{V}_{T+1,T+h} = V_{T-h+1,T}.$$

2.3 Standard GARCH(1,1) Model

The autoregressive conditional heteroskedasticity (ARCH) models introduced by Engle (1982) and their generalization, the so-called GARCH models (Bollerslev, 1986) (see also Bollerslev et al., 1992, 1994) have been the most commonly employed class of time series models in the recent finance literature. These models have been very successful in describing the behavior of financial return data. Their appeal comes from the fact that they can capture both volatility clustering and unconditional return distributions with heavy tails—two stylized facts associated with financial return data. The estimation of a GARCH model involves the joint estimation of a mean and a conditional variance equation. For the forecast comparison we found a GARCH(1,1) model combined with an AR(1) model for the mean to be appropriate, i.e.,

$$r_t = \mu + \phi r_{t-1} + \epsilon_t,$$

with the conditional variance of ϵ_t , conditioned on the return information available up to time $t - 1$, being given by

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \quad (1)$$

The returns were modeled by an AR(1) model, since the coefficient ϕ has been found significant over several sub-samples of the data. The GARCH(1,1) specification (1) coincides with what has generally been regarded as an appropriate representation in the empirical literature (see, for example, Akgiray, 1989; Baillie and DeGennaro, 1990; Lamoureux and Lastrapes, 1990; Schwert and Seguin, 1990; Engle et al., 1991; West et al., 1993).

Instead of assuming normally distributed innovations we use the generalized error distribution (GED) (Taylor, 1994) for maximizing the likelihood function, from which the normal distribution is a special case. The density function of the GED takes the form

$$f(\epsilon_t; \nu) = \frac{\nu \exp(-\frac{1}{2} |\epsilon_t/\lambda|^\nu)}{\lambda 2^{1+1/\nu} \Gamma(1/\nu)}, \quad \lambda = \left(\frac{2^{-2/\nu} \Gamma(1/\nu)}{\Gamma(3/\nu)} \right)^{1/2},$$

with $\nu > 0$. Here, ν plays the role of a tail-thickness parameter and determines the shape of the distribution. It is estimated simultaneously with all other model parameters. For $\nu = 2$ we have $\epsilon_t \sim N(0, 1)$, while for $\nu < 2$ the distribution has thicker tails than the normal. The reason for considering the GED is the fact that, although the unconditional distribution of ϵ_t in a GARCH model with conditional normal errors has fatter tails than the normal distribution, for many financial time series the standardized residuals $\hat{\epsilon}_t/\hat{\sigma}_t$ still appear to be leptokurtic. Therefore, assuming a leptokurtic unconditional distribution for ϵ_t seems more appropriate.

The daily s -step-ahead variance forecast from a GARCH(1,1) model can be derived by recursion

$$\hat{\sigma}_{t+s}^2 = \begin{cases} \hat{\omega} + \hat{\alpha} \hat{\epsilon}_t^2 + \hat{\beta} \sigma_t^2, & s = 1, \\ \hat{\omega} + (\hat{\alpha} + \hat{\beta}) \hat{\sigma}_{t+s-1}^2, & s > 1, \end{cases} \quad (2)$$

so that the forecast for the average future volatility for the next h -day period is given by

$$\hat{V}_{T+1, T+h} = \sqrt{\frac{1}{N_T} \sum_{j=1}^{N_{T,h}} \hat{\sigma}_{T+j}^2}, \quad (3)$$

with $N_{T,h}$ denoting the number of trading days in that h -day period.

2.4 GARCH(1,1) Model with Weekend and Holiday Effects

The standard GARCH specification (1) ignores the fact that days where there is no trading, such as weekends and holidays, may have an effect on volatility. A GARCH model allowing for such effects has been proposed in Engle et al. (1993) and Noh et al. (1994). They incorporate a variable, d_t , which reflects the number of calendar days elapsed since the most recent trading day, and a parameter, $\delta \geq 0$, as exponent of d_t , determining the degree with which non-trading days affect volatility.³ For successive trading days during the week we have $d_t = 1$; for Mondays we have $d_t = 3$, in general. To illustrate the empirical relevance of this phenomenon we computed the return variance for trading days for which $d_t = 1$ and for trading days following weekends or holidays, i.e., $d_t > 1$. In our sample, for trading days with $d_t > 1$ (216 observations) the return variance exceeded that for trading days with $d_t = 1$ (767 observations) by 75.7%.

To include this weekend and holiday effect Engle et al. (1993) modify the conditional variance equation of the GARCH(1,1) model so that:

$$\sigma_t^2 = d_t^\delta [\omega + d_{t-1}^{-\delta} (\alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2)] \quad (4)$$

Dividing $\alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$, in (4), by d_{t-1}^δ scales past volatility information so that it corresponds to that of a one-day non-trading period. The multiplication by d_t^δ inflates the one-day volatility according to the length of the non-trading period between days t and $t - 1$ and the average variance rate during this period. The resulting GARCH(1,1) forecasting recursion is of the form

$$\hat{\sigma}_{t+s}^2 = \begin{cases} d_{t+1}^\delta [\hat{\omega} + d_t^{-\delta} (\hat{\alpha} \hat{\epsilon}_t^2 + \hat{\beta} \sigma_t^2)], & s = 1, \\ d_{t+s}^\delta (\hat{\omega} + d_{t+s-1}^{-\delta} (\hat{\alpha} + \hat{\beta}) \hat{\sigma}_{t+s-1}^2)], & s > 1, \end{cases}$$

while relationship (3) still applies.

2.5 Autoregressive Model

Bollerslev (1986) showed that a GARCH process for ϵ_t can be expressed in terms of an ARMA model for ϵ_t^2 . Using a long autoregression for the squared residuals in (1)

³Engle et al. (1993) interpret parameter δ as average *speed of the variance rate* since the previous close. For $\delta = 1$, the variance rate would not decrease on non-trading days. Therefore, we expect δ to be less than 1.

as an approximation for the ARMA representation of ϵ_t^2 , one may expect that an OLS estimation of such an autoregression performs more or less like an estimated GARCH model. An autoregression of order p or, in short, an AR(p) model for ϵ_t^2 implies predictions

$$\hat{\epsilon}_{T+s}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\epsilon}_{T+s-1}^2 + \dots + \hat{\alpha}_p \hat{\epsilon}_{T+s-p}^2,$$

for future ϵ_t^2 's, with $\hat{\epsilon}_{T+j}^2 = \epsilon_{T+j}^2$ for $j \leq 0$. Then, the forecast for the average volatility over the next h days is

$$\hat{V}_{T+1, T+h} = \sqrt{\frac{1}{N_t} \sum_{j=1}^{N_t} \hat{\epsilon}_{T+j}^2}.$$

2.6 Implied Volatilities (IVs)

If financial markets use information efficiently, past information should not help to predict future volatility. Ideally, the volatility implied by observed option prices reflects the volatility expectation of the market participants and should, under the assumption of rationality, yield an efficient predictor. A potential problem associated with IVs is that one assumes an underlying option pricing model. For example, the widely used Black–Scholes model requires constant volatility. However, it is well known that the IV derived from the Black–Scholes model for an at–the–money option yields an approximately unbiased estimator of the average volatility over the remaining life of the option, if volatility is stochastic and uncorrelated with aggregate consumption (see Hull and White, 1987; Feinstein, 1989). This is due to the fact that for at–the–money options the Black–Scholes model is almost linear in average volatility. Therefore, focusing on at–the–money or near–at–the–money options not only helps to minimize the bias induced by thinly traded options, but should also minimize the specification error due to assuming the validity of the Black–Scholes model.⁴

The Frankfurt Stock Exchange has published daily an implied volatility index, called VDAX, since May 12, 1994; it has been reconstructed backward until January 1, 1992. The VDAX combines volatilities implied by call and put DAX–index

⁴However, any specification error arising from the inappropriateness of the Black–Scholes model makes it less likely that IVs are the best predictors, even when markets are efficient. Therefore, the results for the IV predictor in the following forecast comparison can be interpreted as conservative estimates of the information contained in IVs.

options for eight different strike prices and two different times to maturity. First, for each individual time to maturity a subindex of implied volatilities for the different strike prices is determined by a nonlinear least squares procedure.⁵ The two times to maturity are chosen such that they encompass a duration of 45 days. Linear interpolation between the two time indices gives an implied volatility of a synthetic DAX index option with a remaining life of exactly 45 days.

Instead of employing the VDAX as a predictor for future volatility, one could use IVs derived from the shortest maturity index option of each day as an estimate of the next period's average volatility.⁶ We computed daily IVs from transaction options data recorded at the German options exchange DTB. Only those options that were less than 10% in- or out-of-the-money as well as options with a remaining lifetime of more than 10 trading days are used. Thus, the maturities of the remaining options are varying between 11 and 32 calendar days. Independently of the particular IV variant used, the resulting volatility forecast is given by

$$V_{T+1,T+h} = IV_T.$$

2.7 GARCH(1,1)–IV Model

The GARCH(1,1)–IV model combines the information contained in both past returns and option prices by adding an IV variable as an exogenous variable to the GARCH model (see Day and Lewis, 1992; Lamoureux and Lastrapes, 1993; and Xu and Taylor, 1995). Here, we adopt the specification

$$\begin{aligned} r_t &= \mu + \phi r_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{GED}(0, \sigma_t^2), \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma IV_{t-1}^2/250. \end{aligned} \quad (5)$$

The resulting variance forecast equations correspond to those of the pure GARCH case, but with $\gamma IV_t^2/250$ added to forecasting recursion (2).

⁵The computation of each subindex is based on IVs of options with strike prices lying in a ± 100 index point interval around the actual DAX level. For our sample period, this corresponds to an exclusion of all options which were more than 4% – 7% in- or out-of-the-money.

⁶Clearly, IVs derived from longer maturity options can also serve as indicators for future volatility. But since the longest forecast horizon considered in our forecast comparison is one month, short-term options are expected to be more appropriate for our purposes.

Model (1) can be interpreted as the special case of model (5) where $\gamma = 0$. In the context of testing the efficient market hypothesis the special case of (5) given by

$$\sigma_t^2 = \omega + \gamma IV_{t-1}^2/250 \quad (6)$$

is of interest. Here, the conditional variance equation contains no explicit GARCH terms. If the options market is indeed informationally efficient, the constraints $\alpha = \beta = 0$ implied by (6) should not be rejected.

A problem with this approach is the maturity mismatch of the GARCH and the IV forecasts. Whereas the GARCH model predicts the conditional variance for the next period (here, the next day), the IV variable represents market expectations of the *average* daily volatility over the option's remaining lifetime. Xu and Taylor (1994) proposed a Kalman filtering approach for modeling volatility expectation as an unobservable factor. In their forecasting comparison for currency options data Xu and Taylor (1995) found no evidence suggesting that the choice of the IV predictor (short maturity or next period's volatility estimate) affects the predictive power of the mixed GARCH–IV model.

2.8 GARCH(1,1)–IV Model with Weekend and Holiday Effects

* The GARCH–IV model with weekend and holiday effects simply combines models (4) and (5), so that the conditional variance equation takes the form

$$\sigma_t^2 = d_t^\delta [\omega + d_{t-1}^{-\delta} (\alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma IV_{t-1}^2/250)]. \quad (7)$$

The forecast recursions for the conditional variances are obtained as for model (4), but now with term $IV_t^2/250$ being included.

2.9 Combined Forecasts

An alternative way of combining forecasts is simply to compute linear combinations of forecasts generated by alternative models. To do so, a vector of weights, indicating to what extent an individual forecast enters the combined forecast, needs to be specified. An intuitive way of combining forecasts is to average the individual forecasts. A more elaborate technique is the regression–based approach proposed by Granger and Ramanathan (1984). This requires forecasts of volatility realizations

of each of the individual model. When the series of past realized volatility is then regressed on these ex-post forecasts, the estimated regression coefficients can be used as weights for a linear combination. Granger and Ramanathan (1984) showed that under a mean-squared-error criterion an unrestricted regression including a constant term should be preferred. The inclusion of a constant term should lead to an unbiased combined forecast, even if individual forecasts are biased.

3 In-sample Results

Table 1 displays some sample statistics for the 982 daily DAX index returns from February 3, 1992 through December 29, 1995. Assuming 250 trading days per year the standard deviation for the daily returns amounts to an annualized standard deviation of 14.6%. The high volatility of the daily returns relative to the mean return causes the deviation of the mean return from zero to be statistically insignificant. The highly significant excess kurtosis and the marginally significant negative skewness of the daily DAX returns match stylized facts associated with financial return data and indicate that the normal distribution is an inappropriate assumption for the DAX returns. This is also supported by the Jarque-Bera test, which tests the normal hypothesis by means of the sample skewness and kurtosis. Another stylized fact, shared by many other asset return series, is the conditional heteroskedasticity of the returns as is reflected by both the Ljung-Box statistics and Engle's Lagrange multiplier ARCH tests. The squared returns and, even more so, the absolute returns display highly significant autocorrelations. Though the null of no ARCH effects cannot be rejected at a significance level of 10% for a lag-length $p = 8$, it can, for example, be rejected at any reasonable level for $p = 16$.

In summary, the DAX index return series seems best described by an unconditional leptokurtic distribution and possesses significant conditional heteroskedasticity. This makes a (G)ARCH model with GED disturbances a natural candidate for forecasting DAX index return volatility.

Table 2 compares the in-sample performance of the GARCH models described in the previous section.⁷ The GARCH models (1) and (4) indicate a high persistence

⁷In the following we report only estimates of the conditional variance equation and not the mean equation, since it is of limited interest here.

of volatility shocks with the persistence measure, $\alpha + \beta$, equal to .9802 and .9795 with standard errors .0282 and .0255, respectively. This suggests that an integrated GARCH (IGARCH) model could be appropriate. Reestimation of the IGARCH restriction $\alpha + \beta = 1$ leads only to a small decrease of the maximum likelihood value. A standard likelihood–ratio test against the unrestricted model (1) would reject the IGARCH hypothesis at the 5% level, but not at the 1% level.

Point estimates of about 1.5 for the tail–thickness parameter, ν , with the standard errors around 0.1 clearly indicate the rejection of the normal assumption (i.e., $\nu = 2$). A likelihood–ratio test of the restriction $\nu = 2$ against the unrestricted model (1) clearly supports this conclusion with a test value of 26.40, which is significant even at the 0.5% level.

Figure 1 compares the kernel–estimated empirical density of the DAX returns with the estimated normal density and the GED density for $\nu = 1.5$ and illustrates that the GED represents a more appropriate distributional assumption than the normal density.

Model (5) includes the VDAX as an exogenous variable in the conditional variance equation. We observe a sharp drop in the GARCH coefficient–sum, $\alpha + \beta$, from 0.9802 to 0.1459 and a relatively high estimate of 0.6545 for the IV coefficient, γ . However, all parameters in the variance equation turn out to be insignificant, indicating a serious identification problem when combining the two information sources. A likelihood–ratio test of model (5) against model (1), constraining the IV parameter to zero, gives a χ^2 –value of 17.46, which is significant even at the 0.5% level. On the other hand, testing model (5) against model (6), which constrains the GARCH parameters α and β to zero, yields a clearly insignificant χ^2 –value of 0.08. This indicates that the IVs derived from the DAX index options reflect the information contained in past DAX return and supports the efficient market hypothesis for the DAX options market.

This finding differs from the results reported by Day and Lewis (1992) for the conditional volatility of the S&P100 index returns and those of Lamoureux and Lastrapes (1993) for individual stock returns. These studies suggest that both IVs and GARCH components have incremental explanatory power for conditional variances. Using the same methodology, Xu and Taylor (1995) found evidence in support of the efficient market hypothesis for currency options data from the Philadelphia stock exchange. For each of the currencies under investigation they could not reject the

hypothesis that the GARCH component contains no additional information.

Model (4) suggests significant weekend and holiday volatility effects with a t -value of 4.58 for adjustment parameter δ . This conclusion is supported again by a likelihood-ratio test of the restricted model (1) against model (4) significant at the 0.5% level. For a typical weekend (i.e., $d_t = 3$) we have $d_t^\delta = 1.695$, which means that the model estimates Monday variances to exceed the return variance of any other pair of adjacent trading days (i.e., $d_t = 1$) by 69.5%. This is consistent with the 75.7% increase computed from the raw return data.

4 Out-of-sample Comparison

We now consider the out-of-sample predictive power of the alternative forecasting models described in Section 2. Of particular interest is the question of whether or not the sufficiency of the IV information for the return history found in the in-sample results carries over to out-of-sample comparisons. Below, we evaluate the forecasting performance for one-, two- and four-week horizons. For each horizon, we constructed sequences of nonoverlapping forecasting intervals.

Both the GARCH forecasts and those of the autoregressive model for squared returns are based on rolling-sample estimates given by 250 trading days prior to the forecasting period, for each new forecasting period. The models are reestimated for each period; i.e., the forecast anchored in period T , $\hat{V}_{T+1, T+h}$, is based on models estimated from sample $r_t, r_{t-1}, \dots, r_{T-249}$. In accordance with this, we chose a length of 250 trading days for the rolling standard deviation of past returns.

The forecasts for the one-, two- and four-week intervals start on January 4, 1993.⁸ We recorded the forecasts of each model until December 30, 1993, and then perform regressions, to derive the first combined forecast. With each new prediction the weights for combining forecast were updated. By doing so, we were left with 104 one-week, 52 two-week and 26 four-week forecasts for the two-year period from January 3, 1994 through December 29, 1995.

⁸Since the sample period starts at February 3, 1992 the first 4 (2, 1) forecasts are based on samples with less than 250 observations.

4.1 Some General Results

In the previous section we presented in-sample fits for alternative GARCH(1,1) model estimated from the entire data set. Looking now at the estimation results for the 156 weekly one-year rolling samples for the years 1993–1995 allows us to take a closer look at the appropriateness of the competing specifications over different sub-periods. Parameters of particular interest, plotted in Figure 2, are the persistence measure $\alpha + \beta$, parameter δ capturing the weekend/holiday volatility effect, and the tail-thickness parameter, ν . The first plot in Figure 2 shows that shocks to volatility are highly persistent for a large part of the sample. For a long subperiod the estimates favor an IGARCH process. This is consistent with the fact that the IGARCH hypothesis was not clearly rejected for the full sample. The other two plots in Figure 2 exhibit some variation in the behavior of $\hat{\delta}$ and $\hat{\nu}$.

As an alternative way of comparing the different GARCH specifications Figure 3 displays twice the difference between the log-likelihood values of three non-restricted models and their restricted counterparts. This can be viewed as rolling likelihood-ratio tests. The horizontal lines in these three plots represent the 95% critical value when testing the restricted against the unrestricted model. The first plot in Figure 3 shows that the hypothesis of an insignificant weekend/holiday volatility effect can be rejected for half of the sample (for 78 of 156 subsamples). The fact that the informational content in IVs outweighs that contained in the GARCH components σ_{t-1}^2 and ϵ_{t-1}^2 is reflected by the second and third plots of Figure 3. Whereas the hypothesis $\gamma = 0$ for the GARCH model (5) could be rejected for 101 of the 156 subsamples, the hypothesis of insignificant GARCH components (i.e., $\alpha = \beta = 0$) could not be rejected for a single subsample.

4.2 Out-of-sample Forecasting Results

The out-of-sample forecasting results are summarized in Tables 3, 4 and 5. We employ three criteria to compare the volatility forecasts of the models: the mean prediction error (ME); the mean squared prediction error (MSE); and the proportion of correctly predicted directions, i.e., of the volatility movements, (D).

The main conclusion to be drawn from Tables 3, 4 and 5 is that in terms of the MSE a combination of GARCH forecasts together with an IV variable seem to perform best. For all three forecast horizons, GARCH specification (5) with IV

as exogenous variable has the smallest MSE within the group of individual (i.e., uncombined) forecasting models. The GARCH–IV model makes use of IVs of short maturity options rather than the VDAX, because the former outperformed the latter for all horizons. Compared to the pure GARCH model the mixed version leads to a remarkable improvement of the forecasting accuracy.⁹ This is even more remarkable in view of the poor performance of the purely IV–based forecasts which tend to overstate forward volatility. For each forecast horizon the IV variables display an average absolute bias of about 3%. However, despite this bias, which could be corrected by embedding the IV variable into the GARCH model, IVs clearly contain relevant information about future volatility movements.

The dominance of the GARCH–IV forecasting model is mainly due to the inclusion of the IV information. This is indicated by the performance of the GARCH(0,0)–IV model, which models the conditional variance simply as a linear transformation of the squared IVs without lagged σ_{t-1}^2 and ϵ_{t-1}^2 . The results show that the omission of the GARCH components causes only a marginal loss in predictive power, whereas the exclusion of the IV information in a standard GARCH(1,1) model entails a significant drop in forecasting accuracy. This is in accordance with the in–sample results.

Estimating an AR(1) model for the squared IVs the estimate for the AR coefficient is 0.9731, indicating a high persistence in volatility expectations. This matches very closely the estimate of 0.9802 for the persistence measure, $\alpha + \beta$, in the estimated GARCH model (1b).

Incorporating weekend and holiday effects into the standard GARCH(1,1) model reduces, as shown in Tables 3–5, the ME and MSE of the GARCH forecasts. Therefore, only these GARCH forecasts are considered in the combinations of forecasts.

The results for the other individual forecasting methods can be summarized as follows. The RW forecasts, though on average with the lowest bias, exhibit by far the highest MSE. The MA model can substantially improve forecast accuracy, but shows—apart from the two IV–based forecasts 4 and 5—the highest ME. This, together with the results for the RW model, indicates that processing information in

⁹The forecast equations for the mixed GARCH model differ from the forecast equations for the pure GARCH model by adding $\gamma IV_t^2/250$ to the recursion (2). We also computed IV forecasts by an ARIMA approach and substituted $IV_t^2/250$ by $\hat{IV}_{t+s}^2/250$ for $s > 1$, but this did not improve the volatility forecasts.

past returns in a naive fashion yields inaccurate volatility forecasts. A 15th-order autoregression for the squared returns leads to a marginal reduction of both the ME and MSE.¹⁰ Specifying the conditional heteroskedasticity by either an AR(15) model for squared returns or a standard GARCH(1,1) model does not give rise to a dramatically different forecast performance. Although, the GARCH forecasts consistently outperform the autoregressive forecasts. Finally, the forecasts of a GARCH(1,1) model with a conditional normal distribution, which are not reported here, are very close to those derived from the GED-GARCH model. This suggests that distributional assumptions are of secondary importance in volatility forecasting, even though the explanatory power of the GARCH model is strongly affected by this assumption.

Turning now to the results for the combined forecasts (see Tables 3–5) we find that combining forecasts does not automatically improve accuracy. Combining forecasts seems to introduce a substantial amount of noise. A combination, which includes the MA, RW, AR(15), IV and GARCH(1,1) models (denoted by c, 1, 2, 3, 5, 7, in Tables 3–5), compares reasonable well with GARCH-IV forecasts. However, a much better performance, in terms of the MSE, can be obtained by simply combining IV and GARCH (c, 5, 7) forecasts. This way of combining the two sources of information seems to be even more efficient than combining them within the GARCH framework. For all three horizons, the combination of IV and GARCH forecasts outperforms the GARCH-IV forecasts in terms of the ME and MSE.

The insignificance of the informational content in the GARCH forecasts, given the IV information, is again confirmed by excluding the GARCH forecast from the combined model. Only minor reductions, if any, occur in the forecast accuracy relative to the linear combination of the GARCH and IV forecast. In fact, small improvements can be observed when omitting the GARCH model. The results for the combined approach in Tables 3–5 indicate that an overall reduction of the mean bias of the forecasts can indeed be achieved by this method. All combined forecasts yield a smaller ME than each of the individual forecasting techniques. An exception is the RW method. This is illustrated in Figures 4 and 5, displaying the forecasting results for the two-week horizon.

¹⁰Order 15 was chosen, because the autocorrelation coefficient for the squared returns at that lag was highly significant ($\rho_{15} = 0.1225$ with a t -value of 3.70). We also tried an autoregressive model for the absolute returns, as in Schwert (1989). We do not report these results, because they are inferior to those of an autoregressive model for squared returns.

An obvious fact revealed in Figure 4 is the relative smoothness of all time series forecasts relative to the IV forecasts. Market expectations, which appear to become obsolete with the arrival of new information, seem to be better in explaining the frequent changes in volatility than the smoothed GARCH-based forecasts. The two top plots in Figure 5 show that the IVs' tendency to overstate the true volatility can be corrected by either embedding an IV variable within a conditional-variance GARCH equation, or by combining forecasts in the sense of Granger and Ramanathan (1984). The differences between the IV forecasts and the bias-corrected IV forecasts, which are almost all positive, are displayed in Figure 5. The differences between the combined and the IV forecasts are quite stable; whereas the differences between the GARCH(0,0)-IV and IV forecasts fluctuate considerably. This is caused by the smoothing effects of the GARCH model. Figure 5 also shows the change in volatility forecasts, if we rely only on the bias-corrected IV information instead of using a combination of GARCH and IV information. In all cases, this difference is smaller than 2%; and in 102 of 104 cases it is smaller than 1%, indicating the insignificant incremental information content in GARCH forecasts given IV information.

5 Summary

In this paper we have investigated the suitability of several forecasting techniques for the volatility of the returns of the German DAX index and examined whether or not the German DAX-index options market is informational efficient. We have focused on the problem of whether or not implied volatility information, derived either from observed option prices or from time series models such as GARCH models, are useful in predicting future return volatility. By combining both sources of information, our in-sample fitting and out-of-sample forecasting results give strong support for the hypothesis that historic returns contain no information beyond the market's volatility expectation that is reflected in DAX-index option prices.

However, the hypothesis that implied volatility is an unbiased estimate for one-, two- or four-week ahead realized volatility must be rejected. It tends to overstate the actual volatility of DAX-index returns. By including the implied volatility information into the GARCH equation or by combining the implied volatility forecast with a constant term, following the lines of Granger and Ramanathan (1984), one

can correct for this mean bias.

In summary, we conclude that implied volatility is a biased but highly informative predictor for future volatility. Moreover, implied volatilities are informationally efficient relative to other historic volatility information sources. Altogether, our findings support the efficient market hypothesis for the DAX-index options market.

References

- Agiakloglou, C. and Newbold, P. (1992), Empirical Evidence on Dickey–Fuller–Type Tests, *Journal of Time Series Analysis*, 13, 471–483.
- Akgiray, V. (1989), Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts, *Journal of Business*, 62, 55–80.
- Baillie, R.T. and DeGennaro, R.P. (1990), Stock Returns and Volatility, *Journal of Financial and Quantitative Analysis*, 25, 203–214.
- Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992), ARCH Modeling in Finance, *Journal of Econometrics*, 52, 5–59.
- Bollerslev, T., Engle, R.F. and Nelson, D.B. (1994), ARCH Models, in Engle, R. and McFadden, D., editors, *Handbook of Econometrics*, volume 4, chapter 49, Elsevier Science B.V., Amsterdam, The Netherlands.
- Chiras, D.P. and Manaster, S. (1978), The Information Content of Option Prices and a Test of Market Efficiency, *Journal of Financial Economics*, 6, 213–234.
- Day, T.E. and Lewis, C.M. (1992), Stock Market Volatility and the Information Content of Stock Index Options, *Journal of Econometrics*, 52, 267–287.
- Engle, R.F. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the U.K. Inflation, *Econometrica*, 50, 987–1008.
- Engle, R.F., Hong, T., Kane, A. and Noh, J. (1991), Arbitrage Valuation of Variance Forecasts, *Advanced Futures and Options Research*, 6, 393–415.
- Engle, R.F., Kane, A., and Noh, J. (1996), Index–Option Pricing with Stochastic Volatility and the Value of Accurate Variance Forecasts, *Review of Derivatives Research*, 1, 139–157.
- Feinstein, S.P. (1989), A Theoretical and Empirical Investigation of the Black–Scholes Implied Volatility, Dissertation, Yale University, New Haven, CT.

- Gemmill, G.T. (1986), The Forecasting Performance of Stock Options on the London Traded Options Market, *Journal of Business Finance and Accounting*, 13, 535–546.
- Granger, C.W.J. and Ramanathan, R. (1984), Improved Methods of Combining Forecasts, *Journal of Forecasting*, 3, 197–204.
- Hull, J. and White, A. (1987), The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance*, 42, 281–300.
- Lamoureux, C.G. and Lastrapes, W.D. (1993), Forecasting Stock–Return Variance: Toward an Understanding of Stochastic Implied Volatilities, *The Review of Financial Studies*, 6, 293–326.
- Latane, H. and Rendleman, R.J. (1976), Standard Deviation of Stock Price Ratios Implied by Option Premia, *Journal of Finance*, 31, 369–382.
- Noh, J., Engle, R.F. and Kane, A. (1994), Forecasting Volatility and Option Prices of the S&P 500 Index, Discussion paper 93–32(R), University of California, San Diego.
- Schwert, G.W. (1989), Why does Stock Market Volatility Change over Time?, *Journal of Finance*, 44, 1115–1154.
- Schwert, G.W. and Seguin, P.J. (1990), Heteroskedasticity in Stock Returns, *Journal of Finance*, 45, 1129–1155.
- Scott, E. and Tucker, A.L. (1989), Predicting Currency Return Volatility, *Journal of Banking and Finance*, 13, 838–851.
- Shastri, K. and Tandon, K. (1986), An Empirical Test of a Valuation Model for American Options on Futures Contracts, *Journal of Financial and Quantitative Analysis*, 10, 377–392.
- Taylor, S.J. (1994), Modeling Stochastic Volatility: A Review and Comparative Study, *Mathematical Finance*, 4, 183–204.
- West, K.D., Edison, D.H.J., and Cho, D. (1993), A Utility Based Comparison of Some Models of Exchange Rate Volatility, *Journal of International Economics*, 35, 23–45.

Xu, X. and Taylor, S.J. (1994), The Term Structure of Volatility Implied by Exchange Options, *Journal of Financial and Quantitative Analysis*, 29, 57–74.

Xu, X. and Taylor, S.J. (1995), Conditional Volatility and Informational Efficiency of the PHXL Currency Options Market, *Journal of Banking and Finance*, 19, 803-821.

Table 1: Summary statistics for daily returns on the DAX index^a

mean	standard deviation	skewness	excess kurtosis	Jarque and Bera normality test	
.0002951 (.0002944)	.009225	-.1115 (0.0782)	1.5485 (0.1563)	100.15 1.000	
Ljung-Box Test				ARCH-Test	
h	r_t	r_t^2	$ r_t $	p	TR^2
16	26.73 .955	56.15 1.000	92.55 1.000	8	13.09 .891
32	37.16 0.757	89.56 1.000	151.63 1.000	12	28.57 .995

^aStandard errors are given in parentheses. Marginal significance levels are given without parentheses.

Table 2: GARCH(1,1) Estimation Results

Conditional variance equations:

$$\begin{aligned}
 (1) \quad \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\
 (5) \quad \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma \text{VDAX}_{t-1}^2/250 \\
 (6) \quad \sigma_t^2 &= \omega + \gamma \text{VDAX}_{t-1}^2/250 \\
 (4) \quad \sigma_t^2 &= d_t^\delta [\omega + d_{t-1}^{-\delta} (\alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2)] \\
 (7) \quad \sigma_t^2 &= d_t^\delta [\omega + d_{t-1}^{-\delta} (\alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma \text{VDAX}_{t-1}^2/250)]
 \end{aligned}$$

with $\epsilon_t \sim GED(\nu)$

Model Restriction	a (1) $\nu = 2$	b (1)	c (1) $\alpha + \beta = 1$	Case d (5)	e (6)	f (4)	g (7)
$\omega \times 10^4$	0.0247 (0.0029)	0.0177 (0.0025)	0.0055 (0.0021)	0.0231 (0.0293)	0.0325 (0.0363)	0.0158 (0.0024)	0.0186 (0.0130)
α	0.0513 (0.0095)	0.0493 (0.0084)	0.0568 (0.0181)	0.0104 (0.0470)		0.0493 (0.0090)	0.0355 (0.0318)
β	0.9208 (0.0249)	0.9309 (0.0318)	0.9432 (0.0564)	0.1355 (0.1457)		0.9302 (0.0281)	0.5484 (0.0638)
γ				0.6545 (0.7515)	0.7604 (0.1347)		0.2948 (0.2010)
δ						0.4805 (0.1050)	0.5980 (0.3295)
ν	2 —	1.4945 (0.0971)	1.4480 (0.0841)	1.5340 (0.1070)	1.5319 (0.0895)	1.5674 (0.0924)	1.5957 (0.0959)
log-lik	-1291.41	-1278.21	-1280.74	-1269.48	-1269.52	-1267.06	-1259.09
χ^2	26.40 b	17.46; 22.30 d; f	5.06 b		0.08 d		20.78 d

^aStandard deviations are given in parentheses.

χ^2 -values correspond to LR-tests against the case given in the subsequent row. The χ^2 -statistics are distributed as $\chi^2(1)$ for models (1) and $\chi^2(2)$ for model (6). Critical values: 3.84 (6.63) for $\chi^2(1)$ and $\alpha = .05$ (.01) and 5.99 (9.21) for $\chi^2(2)$ and $\alpha = .05$ (.01).

Table 3: Results for one-week forecasts

		01/03/94 – 12/29/95			01/04/94 – 12/30/94			01/02/95 – 12/29/95		
		ME	MSE	D (%)	ME	MSE	D (%)	ME	MSE	D (%)
1	MA	1.698	3.538	68.9	0.572	3.980	66.7	2.825	3.095	70.6
2	RW	0.185	5.655	0.9	0.127	7.387	0	0.243	3.923	1.9
3	AR(15)	1.564	3.344	69.9	0.884	3.931	64.7	2.244	2.756	74.5
4	VDAX	3.526	4.062	66.0	4.359	5.206	58.8	2.693	2.918	72.5
5	IV	3.354	3.871	68.0	4.284	4.994	62.7	2.423	2.748	72.5
6	G11	1.804	3.253	72.8	1.693	3.932	70.6	1.915	2.574	74.5
7	G11(δ)	1.330	3.101	71.8	1.210	3.777	68.6	1.450	2.425	74.5
8	G00-IV	1.244	2.932	71.8	1.889	3.656	66.7	0.600	2.208	76.5
9	G11-IV	1.302	2.908	71.8	2.099	3.672	66.7	0.506	2.145	76.5
10	G11(δ)-IV	1.312	2.934	70.9	2.108	3.729	66.7	0.515	2.138	74.5
combined:										
	c,1,2,3,5,7	-0.205	2.912	74.8	-0.878	3.639	70.6	0.466	2.184	78.4
	c,5,7	0.465	2.792	72.8	0.823	3.431	66.7	0.107	2.152	78.4
	c,5	0.503	2.776	71.8	0.903	3.425	66.7	0.104	2.126	76.5
	c,7	0.348	3.017	68.9	-0.307	3.747	64.7	1.003	2.287	72.5

The ME (MSE) entries must be multiplied by 10^{-2} (10^{-3}).

The results for the best individual forecasts for a criterion are printed in bold-face.

Table 4: Results for two-week forecasts

		01/03/94 – 12/29/95			01/04/94 – 12/30/94			01/02/95 – 12/29/95		
		ME	MSE	D (%)	ME	MSE	D (%)	ME	MSE	D (%)
1	MA	1.076	2.517	62.7	-0.133	2.722	64.0	2.285	2.312	64.0
2	RW	0.092	3.803	0	0.207	4.466	0	-0.024	3.141	0
3	AR(15)	0.975	2.433	58.8	0.166	2.601	64.0	1.785	2.265	56.0
4	VDAX	2.899	2.681	64.7	3.599	3.081	64.0	2.199	2.282	68.0
5	IV	2.731	2.415	66.7	3.622	2.814	64.0	1.839	2.016	72.0
6	G11	1.239	2.323	64.7	1.091	2.601	60.0	1.386	2.044	72.0
7	G11(δ)	0.787	2.210	68.6	0.684	2.489	60.0	0.893	1.931	80.0
8	G00-IV	0.667	1.836	74.5	1.233	1.988	72.0	0.100	1.683	80.0
9	G11-IV	0.712	1.829	78.4	1.443	2.007	76.0	-0.018	1.651	84.0
10	G11(δ)-IV	0.711	1.848	76.5	1.457	2.044	68.0	-0.034	1.652	88.0
combined:										
	c,1,2,3,5,7	-0.428	1.967	76.5	-1.127	2.272	76.0	0.270	1.653	80.0
	c,5,7	0.313	1.744	76.5	0.867	1.821	72.0	-0.240	1.667	84.0
	c,5	0.366	1.740	76.5	0.936	1.830	72.0	-0.204	1.650	84.0
	c,7	0.451	2.191	72.5	-0.002	2.512	68.0	0.904	1.869	80.0

The ME (MSE) entries must be multiplied by 10^{-2} (10^{-3}).

The results for the best individual forecasts for a criterion are printed in bold-face.

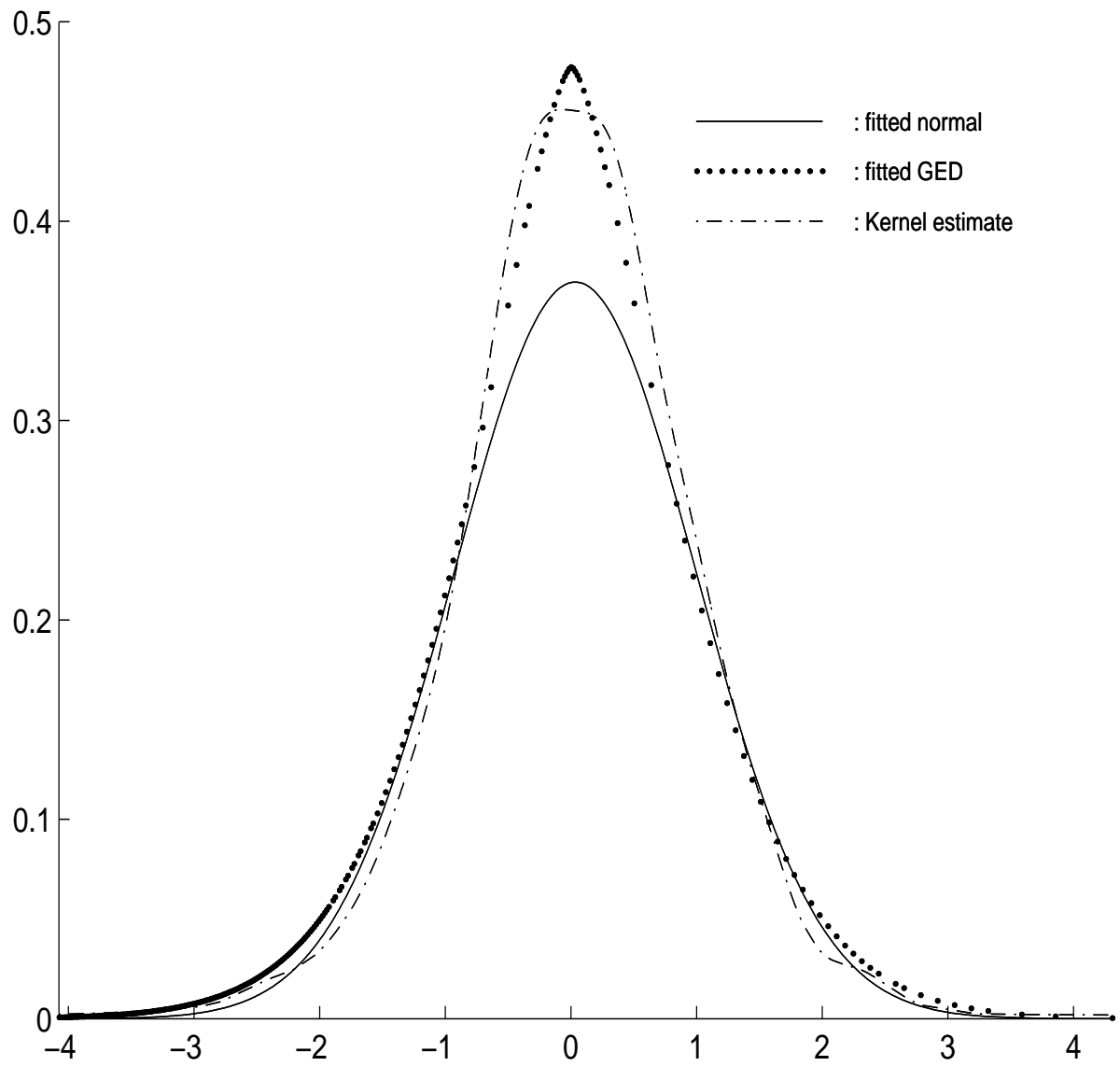


Figure 1: Comparison of Estimated and Empirical Kernel Densities

Table 5: Results for four-week forecasts

		01/03/94 – 12/29/95			01/04/94 – 12/30/94			01/02/95 – 12/29/95		
		ME	MSE	D (%)	ME	MSE	D (%)	ME	MSE	D (%)
1	MA	1.046	1.770	52.0	-0.148	1.972	66.7	2.241	1.569	41.7
2	RW	0.303	2.427	0	0.314	3.324	0	0.293	1.531	0
3	AR(15)	1.022	1.666	52.0	0.095	1.919	66.7	1.949	1.413	41.7
4	VDAX	3.070	2.161	44.0	3.901	2.456	50.0	2.239	1.866	41.7
5	IV	3.029	2.038	44.0	4.114	2.418	50.0	1.943	1.657	41.7
6	G11	1.137	1.662	56.0	1.057	1.888	66.7	1.217	1.436	50.0
7	G11(δ)	0.712	1.591	56.0	0.687	1.884	66.7	0.738	1.298	50.0
8	G00-IV	0.759	1.251	52.0	1.465	1.294	58.3	0.053	1.208	50.0
9	G11-IV	0.822	1.262	52.0	1.716	1.323	50.0	-0.072	1.202	58.3
10	G11(δ)-IV	0.624	1.554	48.0	1.290	1.889	50.0	-0.041	1.218	50.0
combined:										
	c,1,2,3,5,7	-0.204	1.589	52.0	-0.978	2.001	50.0	0.570	1.178	50.0
	c,5,7	0.605	1.089	64.0	1.166	1.027	66.7	0.045	1.152	66.7
	c,5	0.583	1.141	64.0	1.140	1.115	66.7	0.026	1.166	66.7
	c,7	0.345	1.560	44.0	-0.360	1.882	50.0	1.050	1.237	41.7

The ME (MSE) entries must be multiplied by 10^{-2} (10^{-3}).

The results for the best individual forecasts for a criterion are printed in bold-face.

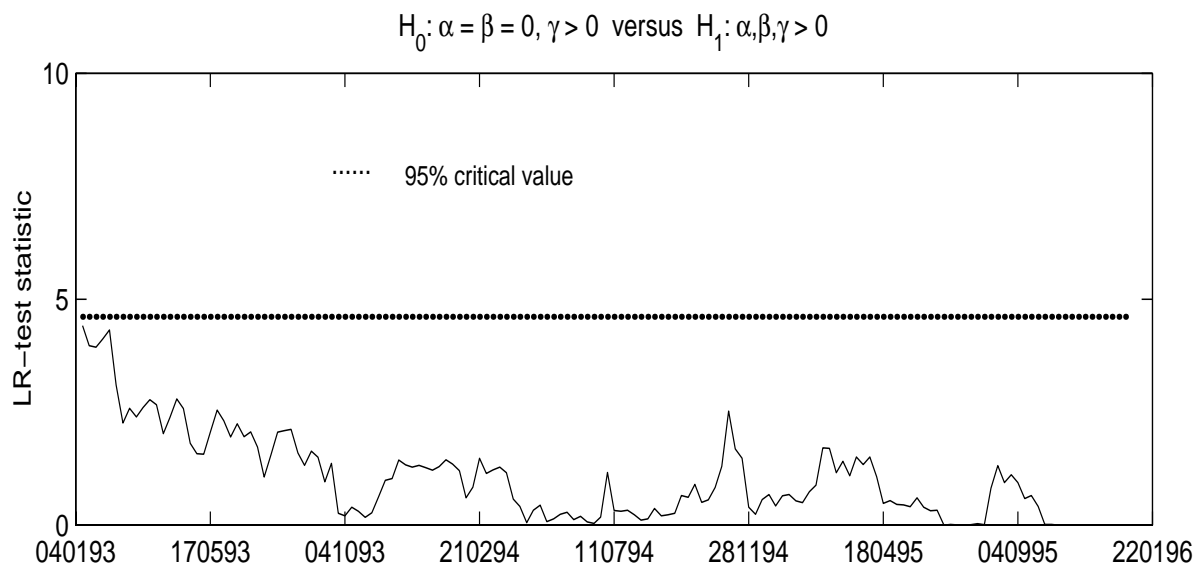
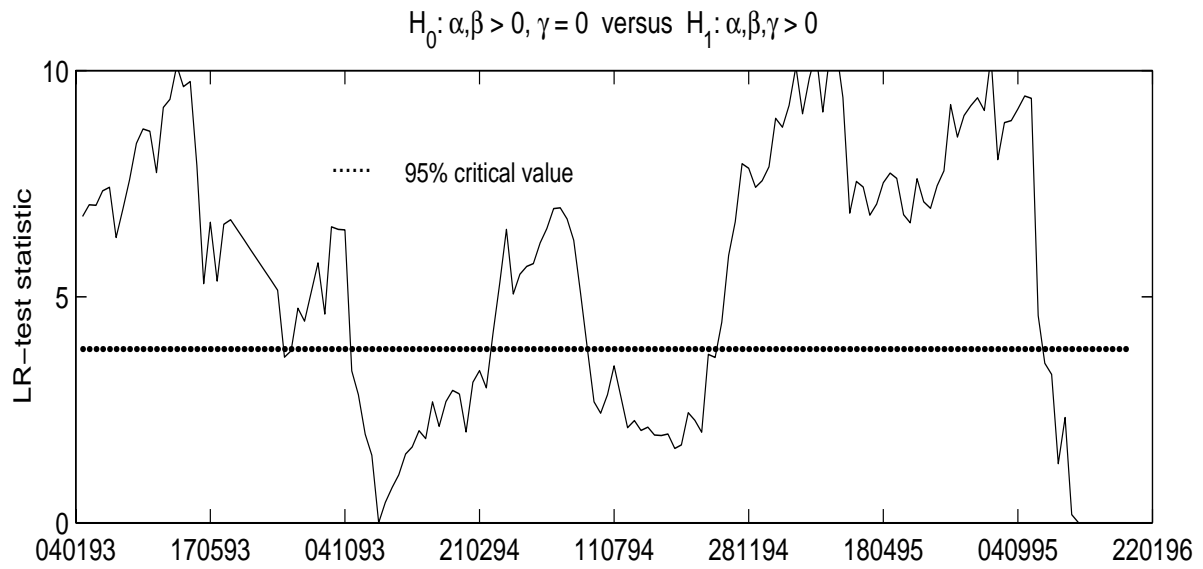


Figure 2: Rolling Likelihood-ratio Tests

CFS Working Paper Series:

No.	Author(s)	Title
2001/06	Bernd Kaltenhäuser	Explaining the Dollar-Euro Rate: Do Stock Market Returns Matter?
2001/07	Guenter Beck Axel A. Weber	How wide are European borders? – New Evidence on the Integration Effects of Monetary Unions –
2001/08	Yunus Aksoy Tomasz Piskorski	Domestic Money and US Output and Inflation
2001/09	Elke Hahn	Core Inflation in the Euro Area: Evidence from the Structural VAR Approach
2001/10	Olaf Ehrhardt Eric Nowak	Private Benefits and Minority Shareholder Expropriation – Empirical Evidence from IPOs of German Family-Owned Firms
2001/11	Daniel Gross	Country-Specific and Global Shocks in the Business Cycle
2001/12	Daniel Gross	Trade Flows and the International Business Cycle
2002/01	Stefan Feinendegen Daniel Schmidt Mark Wahrenburg	Die Vertragsbeziehung zwischen Investoren und Venture Capital-Fonds: Eine empirische Untersuchung des europäischen Venture Capital-Marktes
2002/02	Issam Hallak	Price Discrimination on Syndicated Loans and the Number of Lenders: Empirical Evidence from the Sovereign Debt Syndication
2002/03	Raimond Maurer Christian Schlag	Money-Back Guarantees in Individual Pension Accounts: Evidence from the German Pension Reform
2002/04	Holger Claessen Stefan Mittnik	Forecasting Stock Market Volatility and the Informational Efficiency of the DAX-index Options Market

Copies of working papers are available at the Center for Financial Studies or can be downloaded (<http://www.ifk-cfs.de>).