Predictability of Stock Return Volatility from GARCH Models

Amit Goyal*
Anderson Graduate School of Management, UCLA

May 2000

Preliminary and Tentative - Comments Solicited

Abstract

This paper focuses on the performance of various GARCH models in terms of their ability of delivering volatility forecasts for stock return data. Volatility forecasts obtained from a variety of mean and variance specifications in GARCH models are compared to a proxy of actual volatility calculated using daily data. In-sample tests suggest that a regression of volatility estimates on actual volatility produces R^2s of less than 8%. An interesting by-product is evidence of significantly negative relation between unexpected volatility and stock returns. Finally, out-of-sample tests indicate that a simpler ARMA specification performs better than a GARCH-M model.

JEL classification: C22, C55

Keywords: GARCH, Volatility, Out-of-Sample

*Mailto: amit.goyal@anderson.ucla.edu. US mail address: Anderson Graduate School of Management at UCLA, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095-1481, Tel: (310)825-8160. I would like to thank Javier Gomez, Richard Roll, Pedro Santa-Clara, Walter Torous and Rossen Valkanov for helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

Since the introduction of ARCH models by Engle (1982), there has been a veritable explosion of papers analyzing models of changing volatility. A survey paper by Bollerslev, Chou, and Kroner (1992) lists more than 100 papers on this subject. Some of the more popular variants of models of changing volatility have proved to be various forms of GARCH models. In these models, the volatility process is time varying and is modeled to be dependent upon both the past volatility and past innovations. These models have been used in many applications of stock return data, interest rate data, foreign exchange data etc.

In this paper, we focus upon one aspect of GARCH models, namely, their ability to deliver volatility *forecasts*. In other words, these models are useful not only for modeling the historical process of volatility but also in giving us multi-period ahead forecasts. These forecasts are, of course, of great value in applications to stock return data (portfolio allocation, dynamic optimization, option pricing etc.). We evaluate the performance of these models in terms of their ability to give adequate forecasts. One traditional difficulty in constructing these tests is that the volatility process is inherently unobservable. We surmount this problem by using a proxy of monthly volatility calculated using daily data. Since our alternative measure of volatility is essentially model free and is estimated using higher frequency data, we have more faith in the reliability of these volatility estimates. Various specifications for the mean equation and variance equation are entertained. We perform both in-sample and out-of-sample tests on these GARCH specifications.

The overall result is that GARCH models are unable to capture entirely the variation in volatility. A regression of volatility estimates from GARCH models on (our proxy of) actual volatility produces R^2 of usually below 8%. However, on a positive note, the GARCH predictions of volatility usually (approximately 50% of the time on monthly frequency) lie within the confidence intervals of our proxy of actual volatility implying that GARCH models are not wholly inadequate measures of actual volatility. An interesting by-product of this investigation is the resolution of a puzzling feature of traditional asset pricing tests whereby the correlation between returns and volatility is usually found to be insignificantly positive in spite of strong theoretical reasons to expect a strong relation. Theoretical models such as Merton (1973) predict a *positive* correlation between *expected* volatility and stock returns. We confirm the findings of French, Schwert, and Stambaugh (1987) who find a significantly *negative* relation between *unexpected* volatility and asset returns. Finally, out-of-sample tests seem to indicate that a simpler ARMA model on our measure of volatility performs better than the GARCH model, albeit statistically insignificantly so. Robustness checks using intraday data suggest that our results are not dependent on our choice of frequency of data. In other words, simple volatility measures calculated using high frequency data are as good, if not better, proxies for actual volatility than more sophisticated measures constructed using GARCH models.

This paper focuses only on GARCH models for changing volatility. Alternative models of stochastic volatility such as models of stochastic volatility or implied volatility models from option pricing are not at debate here. In addition, various other measures of volatility based on volume, price range have been proposed in the literature. For instance, Alizadeh, Brandt, and Diebold (1999) focus on volatility obtained from price range which they find has close to Gaussian properties making it preferable to use than traditional estimators like the squared returns or the absolute returns. An interesting extension of this paper might be to run a horse race between these alternative models of changing volatility.

Our work is closely related to the recent work done by Andersen and Bollerslev (1998) and Andersen, Bollerslev, Diebold, and Ebens (1999). These papers show that the traditional tests of various volatility models which rely on ex-post squared returns as the realized volatility (see, for instance, Pagan and Schwert (1990)) are unreliable as squared (or absolute) returns are a very noisy (although an unbiased) estimate of volatility. The authors go on to recommend the use of high frequency data in volatility estimation in the same spirit as our paper. Our paper distinguishes itself from the other work in clearly demonstrating the weaknesses of various GARCH models and in its focus on testing the asset pricing implication of volatility forecasts.[1]

The rest of this paper is organized as follows. Section 2 briefly discusses the data. In-sample tests are conducted in Section 3 while out-of-sample performance is analyzed in Section 4. A few robustness checks are done in Section 5. Section 6 concludes.

[1] Note that we rely on high frequency data to calculate alternative measures of volatility only. This is different in focus from Andersen and Bollerslev (1997a) who show that GARCH models estimated on high frequency data can display different characterstics than GARCH models estimated on low frequency data.

2 Data and Method

The data used in this paper are the daily and monthly series of the CRSP value weighted returns (including dividends) from July 1962 to December 1998. During this period, the average monthly return was 1.1% with a maximum of 16.6% and a minimum of -22.5%. The standard deviation calculated from monthly returns was 4.4%. We also compute a time series of volatility over the same period. Although volatility is inherently unobservable, we proxy for it by computing a measure of volatility by using daily returns. Since July 1962, CRSP has also provided a daily series of returns. We use the intra-month variance of the returns to proxy for the variance in that month. Specifically, monthly variance is calculated using the equation

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it} r_{i-1t} \tag{1}$$

where N_t is the number of trading days in month t and r_{it} is the return on the ith day of month t. The second term accounts for the autocorrelation observed in daily returns (Autocorrelation in daily returns from 07/03/1962 to 12/31/1998 is 0.176). This correction is used by French, Schwert, and Stambaugh (1987).[2] Figure 1 shows a plot of the realized returns and our proxy for the monthly standard deviation.

[Insert Figure 1: Realized Returns and Standard Deviation]

Clearly volatility shows a substantial time variation. The mean of our proxy of standard deviation is 4.33% which is slightly less than the standard deviation of 4.37% computed using only monthly returns. This is another illustration of the well known instance of variance ratios failing to reject the null of random walk for stock returns.[3]

One immediate issue that arises is how good a proxy of actual volatility do we have.[4] One should realize that both the measures of volatility considered in this paper (viz. our measure and the GARCH measure) are *sample* estimates of *actual* volatility. We are aware of the sampling error being made in treating our measure as the benchmark *actual* volatility. We feel that our measure is a better proxy for actual volatility for four main reasons. First, it is

[2] Similar to Schwert (1989, 1990), we find that modifications to this definition, such as including the mean term and/or excluding the autocorrelation term, do not change the results.

[3] See Lo and MacKinlay (1988, 1989) for further details on variance ratio tests.

[4] I thank Prof. Richard Roll for the following discussion.

well known that estimate of second moment becomes more precise as the sampling frequency is increased. As first noted by Merton (1980), if an asset follows a geometric Brownian motion, its volatility can be estimated arbitrarily accurately if the frequency of sample is high enough. Nelson (1990a) has proved a similar property under some additional restrictions in models of changing volatility. Second, various models of changing volatility like stochastic volatility models or GARCH models are essentially filtering processes that make use of the information in the *entire* estimation period to produce volatility estimates at *one particular* point in time. If the volatility is changing over time, it seems reasonable to assume that an estimate that relies on only limited sample time would be more "true" proxy of actual volatility.[5] Third, casual observation suggests that in the months where we see widely different returns are the months where volatility is the highest. Calculating volatility using daily data captures this aspect of data reasonably well whereas this feature is lost by estimating GARCH models on monthly frequency.[6] Finally, under the assumption of Gaussian errors, the standard error of our estimate of variance is $2\sigma^4/(N_t - 1)$, where N_t is the total number of observations and σ is the actual standard deviation. Assuming a sample estimate of 5% for monthly standard deviation estimated using 21 days during the month, the 95% confidence interval for standard deviation is [3.5% 6.6%]. We will show later that not only is this range usually smaller than the deviation between the GARCH estimates and our proxy of actual volatility but also the confidence intervals of GARCH estimated of volatility are much wider. All these reasons make us comfortable in using volatility from equation (1) as a reasonable proxy for actual volatility.[7]

[5] Christoffersen and Diebold (1997) and Andersen, Bollerslev, Diebold, and Ebens (1999) make related points about our measure of volatility being more reliable as it is essentially model free.

[6] Section 5 shows that our results are not dependent on our choice of monthly versus daily frequency.

[7] As suggested by one of the referees, another heuristic way of checking the reasonableness of volatility obtained from equation (1) is to conduct the following Monte Carlo experiment. Assume that the true underlying process for daily returns is a GARCH process. The parameters for the daily process are calibrated to match US data. Using this simulated data, two estimates of variance are obtained. One, according to equation (1) and second, based on GARCH estimates on monthly data. The next issue is that of a benchmark. While aggregation results for GARCH models are difficult to obtain, we consider a simple benchmark. Each month's *actual* volatility is taken to be the sum of daily volatilities from the generated data. The exercise is done for 438 months with 20 days in each month. For each iteration, we then run a OLS regression of actual volatility on the two estimates of volatility. The exercise is repeated 5000 times. The average R^2 of the regression of actual volatility on estimates from equation (1) is 87% while the average R^2 of the regression of actual volatility on GARCH estimates is only 28%. This lends further credence to our belief that volatility estimates from higher frequency data are more reliable than the more sophisticated GARCH estimates.

3 In-Sample Tests

In this section we evaluate various GARCH models in terms of their in-sample performance. Subsection 3.1 carries out tests for simpler GARCH model with different mean specifications while subsection 3.2 carries out these tests for other volatility specifications. Recognizing recent evidence on stock return predictability, we subject these models to further robustness checks by introducing instrumental variables in subsection 3.3.

3.1 Different Mean Specifications

We begin by evaluating the simpler GARCH models first introduced by Bollerslev (1986). Volatility, h_t, is assumed to evolve according to the equation

$$h_t = \gamma_1 + \gamma_2\, h_{t-1} + \gamma_3\, \epsilon_{t-1}^2 \tag{2}$$

We entertain various specifications for the mean equation.[8] All these specifications are nested in the equation

$$r_t = \mu + \delta\, h_t^p + \epsilon_t \tag{3}$$

The traditional GARCH-M model sets $p = 1$ which corresponds to expected returns being proportional to variance of the returns. Specification involving the standard deviation ($p = 0.5$) has also been found to be statistically significant. Exclusion of μ corresponds to an exact form of asset pricing restriction in the spirit of Merton (1973). The results of the estimation are presented in Table 1.

[Insert Table 1: Estimation of GARCH Models with Different Mean Specifications]

Panel A of Table 1 presents the results of Quasi-Maximum-Likelihood Estimation. Two types of t-statistics are reported. The numbers in (parenthesis) are calculated using the outer product of the scores of the likelihood function. To account for departures from normality, another variance-covariance matrix is estimated along the lines of Bollerslev and Wooldridge

[8] ARCH-M models were first introduced by Engle, Lilien, and Robins (1987).

(1992) and these robust t-statistics are reported in [brackets].[9] The results are, for the most part, as expected. The volatility process is highly persistent with γ_2 close to 0.87 (corresponding to a half life of approximately 5 months) and an extremely significant t-statistic. Sum of γ_2 and γ_3 is close (but not equal) to 1.0 which suggests that the volatility process might be integrated. Such a specification, however, is not entertained in this paper.[10] Somewhat disconcerting, although not surprising in the light of previous research, is the fact that the δ coefficient is usually insignificant whenever estimated along with μ. On the basis of just the estimation results of Panel A, there seems to be little to choose between the variance specification ($p = 1$) and standard deviation specification ($p = 0.5$).[11]

Some diagnostic information on the estimation is presented in Panel B of the same table. The first six columns give descriptive statistics on the standardized residuals $\epsilon_t/\sqrt{h_t}$. Although the standard deviation of these residuals is close to 1.0 (as it should be), other descriptive statistics illustrate some of the deficiencies of the GARCH model. In all cases, the mean of the residuals is negative (although statistically insignificantly different from 0.0). The residuals display statistically significant negative skewness and excess kurtosis.[12] Bollerslev (1987) has proposed a model with t-distributed error terms which is useful in accounting for excess conditional kurtosis but introduces additional complexity of estimating the degrees of freedom for the t-distribution. The mean of the estimated volatility process, however, compares favorably with the unconditional variance (19.27) of the returns. Also the Ljung-Box statistic for the standardized residuals is always less than the 95% critical value of 21.03 suggesting that there is not much autocorrelation in the returns. Similarly, the Ljung-Box statistic for the squared residuals is not significant suggesting that GARCH(1,1) model is an adequate description of the volatility process and no higher lags are needed to capture the autocorrelation. The likelihood values are similar across models once we include the term involving δ. As we are unable to establish the superiority of standard deviation model over the variance model, we choose to

[9] It is well known that the likelihood function for models of changing volatility is close to flat near the optimum. In this situation, analytical derivatives prove to be superior to numerical derivatives. At the same time, one of the main advantages of Bollerslev and Wooldridge (1992) standard errors is that it requires calculation of only first derivatives. See Appendix B for details on the construction of these standard errors.

[10] See Nelson (1990b) for stationarity properties of GARCH(1,1) processes.

[11] Models M6 and M7 in Table 1 do seem to suggest that p is closer to 1.0 than to 0.5. However, the standard error associated with the estimated p is large to preclude any statistically meaningful conclusions.

[12] In large samples of normally distributed data, estimators of skewness and kurtosis have means of 0 and 3 and variances of 6/T and 24/T respectively. T=438 for our sample.

work with the more traditional variance model in out-of-sample tests done in Section 4.

One of the by products of this estimation is a time series of the one period ahead forecasts of the volatility. As mentioned in the introduction, one of the objectives of the various GARCH models is to provide good estimates of volatility which can then be used for a variety of purposes including portfolio allocation, performance measurement, option valuation etc. One of the problems with measuring the accuracy of these forecasts (and indeed one of the reasons why GARCH models are so popular) is that volatility is inherently unobservable. We bypass this problem of unobservability by using our proxy of volatility (equation (1)).

[Insert Table 2: Performance of GARCH Models with Different Mean Specification]

We evaluate the performance of the GARCH models in Table 2. This table reports two sets of regressions. The first regression is simply to check the ability of predicted volatility from GARCH models (denoted by h_t) to forecast the actual volatility (σ_t^2). Specifically the following regression is estimated

$$\sigma_t^2 = a + b\, h_t + u_t \tag{4}$$

A good forecast should have the properties : $a = 0$, $b = 1$ and a high R^2. This equation is estimated using the usual OLS procedure and White's (1980) heteroskedasticity consistent t-statistics are given below the coefficient estimates in brackets. We see that the GARCH models satisfy two of the desirable properties viz. a is insignificantly different from 0 and b is insignificantly different from 1. However, the R^2 of the regression never rises above 5%. It is worth exploring the cause of this. In Figure 2, we plot the actual versus the predicted standard deviation from the simplest GARCH-M model (model M2 in Table 1). It is clear from the figure that while $\sqrt{h_t}$ does a good job of tracking σ_t, it is much less variable and is unable to capture entirely the variation in actual volatility. But we advocate caution in interpreting this figure. As we emphasized in Section 2, both our measure of volatility (σ_t^2) and the GARCH volatility (h_t) are subject to estimation error. The confidence intervals of these measure are plotted in Figure 3.[13] From this figure, we see that the GARCH estimate

[13]Confidence intervals of σ_t^2 is estimated using the asymptotically approximate standard error of $2\sigma_t^2/N_t$, where N_t is the number of observations (days in the month) used to construct σ_t^2. Note that this is an approximation as we ignore the correction of autocorrelation term in the standard error.

lies within the confidence intervals in 50% of the observations. At the same time, the GARCH estimate (h_t) is also subject to sampling error because the parameters of the volatility equation are estimated with error.[14] As is evident from the figure, the sampling error in these parameter estimates makes the confidence interval for the GARCH volatility rather wide. The conclusion we would draw from these two figures is twofold. First, both σ_t^2 and h_t are beset by problems of sampling error. However, the confidence intervals of σ_t^2 are smaller. Second, although casual empiricism would advocate the use of σ_t^2, on statistical grounds, it is difficult to gauge the (un)reasonableness of GARCH models.

[Insert Figure 2: Actual and Predicted Standard Deviation]

[Insert Figure 3: Confidence Intervals of Volatility Estimates]

One of the puzzling features of the GARCH-M models is that the empirically estimated relation between returns and volatility (the δ parameter in Equation (3)) is insignificant (see Scruggs (1998)). At first sight, this is surprising because the ICAPM model of Merton (1973) predicts a positive correlation between these two variables (See Appendix A for a brief overview of Merton's asset pricing model). However, one should be careful in interpreting the theoretical evidence. Merton's model predicts a positive relation between expected volatility and expected

[14] Constructing confidence intervals for GARCH volatility estimates is a non-trivial exercise. This is easily seen in the context of the simplest GARCH volatility equation

$$h_t = \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2$$

Our estimation gives us estimates $\hat{\gamma}_1$, $\hat{\gamma}_2$ and $\hat{\gamma}_3$. These are, of course, asymptotically normally distributed with a covariance matrix given by the inverse of the information matrix from the likelihood function. The issue is that h_t is not observed and has to be computed recursively. This implies that h_t for all $t > 1$ has a non-standard distribution and, therefore, it becomes hard to figure out the appropriate size of the interval. In this paper, we adopt a slightly ad-hoc approach. In particular, the upper tail of the confidence interval is constructed using the upper limits of the confidence intervals of parameters ($\gamma' s$) and using the same series recusrively in the volatility equation. In other words, the upper tail of the forecast (h_t^u) is constructed as

$$h_t^u = \hat{\gamma}_1^u + \hat{\gamma}_2^u h_{t-1}^u + \hat{\gamma}_3^u \epsilon_{t-1}^2$$

where the superscipt u denotes the upper limit of the confidence interval for the γ parameters. The delta method is equivalent to the above exercise because the equation for h_t is linear in γ's and therefore the whole covariance matrix of γ parameters is not required.

stock returns. However, GARCH models estimate an essentially predictive regression as h_t is a predetermined variable. In fact as French, Schwert, and Stambaugh (1987) point out, a stronger test of the asset pricing relation can be carried out by including the unexpected change in volatility in the return equation. This is done via the equation

$$r_t = c + d\,\hat{\sigma}_t^2 + e\,(\sigma_t^2 - \hat{\sigma}_t^2) + u_{2,t} \tag{5}$$

where $\hat{\sigma}_t^2$ is the (expected) actual volatility obtained from equation (4).

The expected sign of coefficient e is negative. The intuition is simple. If there is a negative shock in period $t-1$, this raises upward the predicted volatility h_t. At the same time, this raises the expected discount rate for future periods as discount rate is positively related to the risk. In absence of any correlation with the cash flows, an increase in the discount rate would reduce the stock price. This induces a negative relation between current returns and the unexpected change in volatility. Equation (5) is estimated using weighted least squares (WLS) as the innovations are highly heteroskedastic. We use the actual standard deviation σ_t as weights for estimation purposes. The results are completely in accord with the intuition. The coefficient e is significantly negative with t-statistic close to -2.7.

The lessons from Table 2 are twofold. One, the estimates of volatility obtained from GARCH models are inadequate in capturing the entire variation in the actual volatility of returns although they seem to lie within the confidence intervals for our other measure of volatility. Second asset pricing tests embedded in GARCH models are misspecified as they deal with predetermined rather than contemporaneous variables. Once cognizance is taken of this fact, we get the expected statistically significant *negative* relation between *unexpected* stock volatility and stock returns.

3.2 Different Volatility Specifications

In this subsection we go on to explore other specifications of volatility process. We restrict our attention to three of the most popular specifications. The first one is exponential GARCH

(EGARCH) introduced by Nelson (1991) which parameterizes the volatility process as

$$\begin{aligned}\log\ h_t &= \gamma_1 + \gamma_2 \log\ h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \mathrm{E}\left(\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} \right) \right] \\ &= \gamma_1 + \gamma_2 \log\ h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right] \qquad (6)\end{aligned}$$

This specification has two main advantages. First, it allows h_t to respond asymmetrically to positive and negative shocks ϵ_{t-1}. Second, because of the exponential specification, there are no non-negativity constraints of the γ parameters.

The second specification we explore is the asymmetric GARCH (AGARCH) model of Engle and Ng (1993).[15] The volatility equation is

$$h_t = \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 (\epsilon_{t-1} + \gamma_4)^2 \qquad (7)$$

The parameter γ_4 is typically negative and thus AGARCH model also allows for asymmetric response of volatility to positive and negative shocks.

The final specification is due to Glosten, Jagannathan, and Runkle (1993). This model (which we name as GJRGARCH) specifies the volatility process as

$$h_t = \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 + \gamma_4 S_{t-1}^- \epsilon_{t-1}^2 \qquad (8)$$

where $S_{t-1}^- = 1$ if $\epsilon_{t-1} < 0$ and 0 otherwise. Engle and Ng (1993) find that this is the best parametric model in explaining Japanese stock return data.

In the remaining part of this section, we explore the properties of the four models given in equations (2), (6), (7) and (8). The estimation results for these four models of volatility and two different specifications for return equation are given in Table 3. The first two rows of Table 3 repeat the first two rows of Table 1. As before, t-statistics calculated from the outer product of the score are given in (parenthesis) while Bollerslev and Wooldridge (1992) t-statistics are given in [brackets].

[Insert Table 3: Estimation of GARCH Models with Different Volatility Specifications]

[15] Engle and Ng (1993) credit a 1990 paper by Robert F. Engle as having introduced the AGARCH model. However, the citation is unfortunately missing from the reference list.

In Table 3, γ_3 is significantly negative and γ_4 is significantly positive for EGARCH models. Moreover the likelihood value is higher than that of GARCH (see the last column in Panel B) suggesting that there is indeed an asymmetric response of shocks to volatility and the EGARCH model does a good job of capturing this asymmetry. The coefficients of the AGARCH model are also as expected, with γ_4 significantly negative. However, the estimates from GJRGARCH model are surprising. The coefficient γ_3 is close to zero suggesting that *only* the negative innovations have an impact on volatility. While the possibility of this being a local (rather than a global) maximum cannot be completely ruled out, nevertheless this hugely asymmetric behavior where only the negative shocks effect the volatility is interesting. The implications of this finding are left for a future study.

A noteworthy feature of the estimation is the absence of a significant relation between return and volatility (t-statistic of δ is always below 1.64). Further diagnostics of this estimation are presented in Panel B. We see again that the standardized residuals are negatively skewed and display excess kurtosis. On the basis of the log likelihood value, the EGARCH model seems to be superior to all other models.

[Insert Table 4: Performance of GARCH Models with Different Volatility Specification]

In Table 4, we repeat the exercise of evaluating these different models in terms of their ability to deliver forecasts of volatility. As before, the volatility estimates from the various GARCH models are much less variable than the actual volatility but do a reasonable job of tracking the actual volatility (coefficient b is close to 1.0). In asset pricing tests involving the predicted and actual volatility, coefficient e is significantly negative. On the whole, different GARCH specifications of the volatility process seem unable to produce good forecasts of actual volatility.

3.3 Instrumental Variables

In recent years, there has been a huge literature documenting evidence in favor of predictability of asset returns.[16] Of the various instruments that have been proved to have some forecasting

[16]For a partial list see Fama and French (1988, 1989), Ferson (1989), Ferson and Harvey (1991), Harvey (1989), and Keim and Stambaugh (1986).

ability, we choose the three most significant: the dividend yield, the term premium and the default premium. The dividend yield ("Dvy") is the monthly dividend yield on the CRSP value weighted index. The term premium ("Term") is defined as the difference between 10 year Treasury bond yield and 3 month Treasury bill yield. The default premium ("Def") is defined as the difference in yields between BAA and AAA rated corporate bonds. The source for "Term" and "Def" is Citibase. Summary statistics on these instruments are given in Table 5.

[Insert Table 5: Summary Statistics of Instruments]

As shown in this table, all three instrumental variable are highly persistent though not so highly cross-correlated. The Ljung-Box statistic is extremely significant at any conventional level of significance. Moreover, the Augmented Dickey-Fuller statistic is strongly consistent with the presence of a unit root in the dividend yield series and close to unit root in the other two variables. To convert these series into stationary series, the procedure employed by Lamont (1998) is used. Specifically, the variables are stochastically detrended by subtracting a prior 12 month moving average. As the last three columns of Panel A show, this is sufficient to remove the evidence of unit roots although the detrended variables still remain highly autocorrelated.[17]

[Insert Table 6: OLS Regression of Stock Returns on Instrumental Variables]

Table 6 presents some preliminary OLS regressions of stock returns on the stochastically detrended instruments. The adjusted R^2 is quite meagre (less than 2%). This might either be because these instruments have low predictive power at monthly frequency or it might be because these variable have lost their forecasting ability in the 1990's. The residuals from the regression show negative skewness and excess kurtosis. So we also present two sets of t-statistics. One in parenthesis are normal OLS t-statistics and the one in brackets are Newey-West's t-statistics correcting for autocorrelation and heteroskedasticity. From these results, it appears that the only variable with significant predictive ability is the default premium. We continue to use all the three instrumental variables in GARCH estimations to follow.

[17] An unfortunate side effect of this transformation is that the time series of "Def" is close to zero. It is therefore multiplied by 100 in all the following regressions to bring it to the same scale as the other variables.

In our next set of GARCH estimations, we introduce these instrumental variables in the mean equation.[18] The results of this estimation are presented in Table 7. The performance evaluation is done in Table 8. We limit our discussion of these tables just to note that all the results are virtually unchanged. The risk premium δ is insignificant, predicted volatility h_t is unable to capture the variation of actual volatility σ_t^2 and the re-estimated asset pricing tests show the expected negative relation between the stock returns and the unpredictable component of volatility.

[Insert Table 7: Estimation of GARCH Models with Instrumental Variables]

[Insert Table 8: Performance of Various GARCH Models with Instrumental Variables]

4 Out-of-Sample Tests

Until this point, the focus of our tests has been in-sample performance. However, a market participant does not have the benefit of foresight. He can make forecasts conditional only upon the historical information. In a related setting, out-of-sample performance of mean predictability has been found to be poor even if the models exhibited good in-sample performance.[19] In this section, we evaluate GARCH models in terms of their out-of-sample performance.

For brevity, we concentrate only on the GARCH-M model.[20] We can report that results for other GARCH models are qualitatively the same. We construct out-of-sample forecasts by using the following procedure.[21] We allow an initial period of estimation before constructing

[18] An alternative approach would be to make the risk premium parameter δ dependent on these instrumental variables. For further details on this approach, see Chou, Engle, and Kane (1992), De Santis and Gerard (1997), Dumas and Solnik (1995), Ferson and Harvey (1999), and Harvey (1991).

[19] See Bossaerts and Hillion (1999), Goyal and Welch (1999), and Pesaran and Timmerman (1995).

[20] An exponentially weighted moving average of squared returns (integrated GARCH) is quite popular amongst practitioners, most notably RiskMetrics, for forecasting. The forecasts obtained from this model are however, very similar in properties to those obtained from the GARCH model and are therefore not explored in great detail in this paper.

[21] Our focus is only on one step ahead forecasts of conditional variance. Moreover, we realize that each forecast is subject to uncertainty and has an associated forecast interval. We ignore the variance of these forecasts and instead choose to work with a time series of one step ahead forecasts. See Baillie and Bollerslev (1992) for analytical expressions of s-step ahead forecasts from GARCH and ARMA models.

any forecast. Thus our first forecast is for period January 1975 even though the sample period starts in July 1962. For each forecast period u, we run a regression for $t = 1..u-1$ of the following GARCH-M model

$$\begin{aligned} r_t &= \mu + \delta\, h_t + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned} \tag{9}$$

The forecast for period u is then given by

$$\hat{h}_u = \gamma_1 + \gamma_2 h_{u-1} + \gamma_3 \epsilon_{u-1}^2 \tag{10}$$

In other words, our forecasts are obtained by recursively estimating the GARCH model by expanding the estimation window one period at each iteration.[22] This gives us a time series u =January 1975 ... December 1998 of out-of-sample forecasts from GARCH model. The proxy for actual volatility is again taken to be the monthly standard deviation calculated from daily data using equation (1).

To benchmark this forecast against an alternative, we choose a simple ARMA model[23] on the realized volatility.

$$(1-\theta_1 L)\sigma_t^2 = \phi_0 + (1+\phi_1 L + \phi_2 L^2) u_t \tag{11}$$

where L is the lag operator.

Fitting the entire data from July 1962 to December 1998, we find that there is little to choose between ARMA(1,2) and ARMA(2,1) models on the basis of Akaike and Schwarz criteria and therefore decide for the ARMA(1,2) since a similar specification has been previously used by French, Schwert, and Stambaugh (1987). The sample estimates for the entire period (with White's heteroskedasticity consistent t-statistics in parenthesis) are

$$\begin{array}{ccc} (1-\ 0.765\ L)\,\sigma_t^2 & = & 4.409 + (1\ -0.625\ L + 0.011\ L^2)\,u_t \\ (9.256) & & \qquad\qquad (-6.077) \quad\ (0.180) \end{array} \tag{12}$$

Note that it appears that the AR roots and the MA roots are close to being equal and to canceling each other. If this were truly the case, then we know that the coefficients cannot be

[22] Note that the standard error of the estimated parameters is ignored in these calculations.

[23] Note that the ARMA model has statistical properties which are very close to those obtained from (1). In this case, of course, (1) is estimated using daily data whereas the ARMA model is estimated on monthly data. It is unclear how a daily ARMA would aggregate to monthly frequency and this issue is left for future exploration.

identified and a white noise will do as well. However, a Wald test for the sum of AR and MA coefficients being equal to zero is overwhelmingly rejected (χ^2 statistic of 9.98 with a p-value of less than 0.1%).

Repeating the procedure of recursive estimation, for $t = 1..u-1$, we estimate the following model

$$(1 - \theta_1 L)\sigma_t^2 = \phi_0 + (1 + \phi_1 L + \phi_2 L^2)u_t \tag{13}$$

The forecast for period u is then given by

$$\hat{\sigma}_u^2 = \phi_0 + \theta_1 \sigma_{u-1}^2 + (\phi_1 L + \phi_2 L^2)u_u \tag{14}$$

[Insert Figure 4: Out-of-Sample Forecasts of Standard Deviation]

Figure 4 shows a plot of the actual standard deviation (σ_t) versus the GARCH forecasts ($\sqrt{\hat{h}_t}$) and the ARMA forecasts ($\hat{\sigma}_t$). It is easy of see that the ARMA forecast seems to be doing a better job at forecasting the variation in actual volatility. Further evidence is presented in Table 9. GARCH forecasts errors have a higher mean and a higher standard deviation than those of ARMA forecast errors. Root mean squared error (RMSE) of the GARCH forecast at 15.60% is worse than that of the ARMA forecast at 15.22%. To assess the statistical significance, we use a simple jackknife. We randomly shuffle each of the 287 forecast errors from the two models and compute a new RMSE difference to simulate a draw from the null of no difference. This procedure is repeated 100,000 times and the location of actually observed RMSE difference within this null distribution is reported. The actually observed RMSE difference of 0.38% lies at the 80th percentile (i.e. at the 20% level on a one-sided test) implying that although the GARCH forecast does perform worse than the ARMA forecast, there is no statistical difference between their performance.

[Insert Table 9: Out-of-Sample Performance of GARCH and ARMA Models]

Our out-of-sample exercise has concentrated only on comparing RMSE of forecasts from alternative models. As has been noted earlier, estimates of volatility are crucial inputs in portfolio optimization and Value-at-Risk measures. In this paper we are focusing only on second moments and abstract away from the issue of forecasting the first moments. It is interesting to see, however, whether these differences in volatility forecasts lead to economically significant utility levels.[24] A simple calculation (details available upon request) shows that if we assume CARA utility then the relative utility of two investors with different forecasts for volatility is given by

$$\frac{Eu_1}{Eu_2} = \exp\left[-\frac{1}{2}(\bar{r} - r_f)^2 \left(\frac{1}{\bar{\sigma}_1^2} - \frac{1}{\bar{\sigma}_2^2}\right)\right] \tag{15}$$

where $\bar{r}$ is the average return on stocks, r_f is the riskfree rate, and $\bar{\sigma}_1^2$ and $\bar{\sigma}_2^2$ are the averages of the two forecasts of volatility. Using the same forecasts as in Table 9, the ratio of expected utility of an investor forecasting using ARMA model to that of the investor forecasting using GARCH is obtained as 0.6835,[25] which roughly means that expected utility using ARMA is 32% higher than that using GARCH.

5 Robustness Checks

In this section, we perform a few robustness checks. Unreported results show that there is no significant difference between using the CRSP value weighted returns or the CRSP equal weighted returns or returns on the S&P500 index. A potentially more damaging issue is the frequency of data used. It is conceivable that volatility estimated from daily data is more precise than GARCH volatility estimated from monthly data simply because of the higher frequency of daily data. While in some respects, this is precisely one of the points of this paper viz. *simple* volatility estimates from high frequency data are more reliable than the more *complicated* volatility estimates from GARCH, we would still like to confirm that the results in this paper are not artifacts of the large frequency difference between monthly and daily data.

We use the intraday exchange rate data for conducting this robustness exercise. Specifically, the data used is the US Dollar - Deutsch Mark exchange rate data for year 1996 at half hour intervals. This data is obtained from Olson Associates and has been extensively used previously (see Andersen and Bollerslev (1997b)). Panel A of Table 10 presents some summary statistics on the data used. Half-hourly returns are indistinguishable from 0 and are highly correlated at

[24] See West, Edison, and Cho (1993) for detailed utility based comparisons of models of exchange rate volatility.

[25] Note that since utilities are measured as negative numbers, a ratio of less than 1.0 is favorable for the numerator investor.

lags upto 4.[26] We then calculate the daily return by compounding the 48 half-hourly returns for that day. Daily variance is calculated using equation similar to Equation (1) modified to account for significant correlation at higher lags in the exchange rate data. Specifically, we use the following equation to estimate the daily volatility[27]

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it} r_{i-1t} + 2\sum_{i=3}^{N_t} r_{it} r_{i-2t} \tag{16}$$

Panel B of Table 10 shows the results of estimating a GARCH model on daily returns of the exchange rate data. The results are more or less as expected although the ARCH coefficient of 0.045 is not significant.

[Insert Table 10: GARCH Estimation of Exchange Rate Data]

In Figure 5, we plot the actual daily volatility calculated using Equation (16) and the volatility process estimated from GARCH model of Panel B of Table 10. We see that Figure 5 looks very much like Figure 2 which leads us to belive that the inability of GARCH models to predict volatility is not an artefact of the frequency of the data. Moreover, since the daily volatility is computed using 48 intra-day observations in Equation (16), confidence intervals of σ_t^2 are much smaller (not plotted in Figure 5). In this case, we find that the GARCH estimate of volatility lies within these confidence intervals in only 11% of the cases.

[Insert Figure 5: Actual and Predicted Standard Deviation from GARCH Estimation of Exchange Rate Data]

6 Conclusion

In this paper, we have focused on only one aspect of GARCH models, namely their ability to deliver one period ahead forecasts of volatility. We have compared these forecasts to a proxy of actual volatility calculated using daily stock returns and have found that the volatility series

[26]Under the assumption of iid samples, the standard error of autocorrelation coefficients is $1/\sqrt{T}$.

[27]Table 10 shows that autocorrelations of intraday returns are significant at lags upto 4. However, we use only two lags in estimating the daily volatility. Moreover, we make no adjustment for the mean return. These modifications, however, produce no substantial changes in the results.

obtained from GARCH models is too smooth to capture the entire variation in actual volatility. We have demonstrated that this result is not an artifact of our choice of monthly/daily frequency but is quite independent of the frequency chosen. However, one cannot wholly reject the GARCH models in favor of our measure of volatility as the GARCH volatility frequently lies within the confidence interval of our other measure.

Alternative estimates of volatility, such as stochastic volatility estimates or implied volatility estimates, are not at debate in this paper. For example, a recent paper by Christensen and Prabhala (1998) suggests that implied volatility embedded in option prices has better forecasting ability than has been previously assumed in literature. One extension of this paper could be to analyze how well does implied volatility predict the actual volatility.

An obvious issue at this stage is that if GARCH forecasts are not adequate, what do we use as an alternative? Preliminary evidence presented in the section on out-of-sample tests indicates that simpler ARMA specifications do a good job of predicting future volatility.

Appendices

A Asset Pricing

Assume that we have N assets with excess returns $\tilde{r}$. Let the first two moments of these assets be given by

$$\mu = E[\tilde{r}], \qquad \Sigma = \text{var}[\tilde{r}] \tag{A.1}$$

Assuming a representative agent economy, it follows from Merton' intertemporal consumption-investment model (see Merton (1973)), that the optimal holdings (α) in these risky assets are given by

$$\alpha = -\frac{J'}{WJ''}\Sigma^{-1}\mu \tag{A.2}$$

where the subscripts refer to the partial derivatives of the $J(W)$ value function and W is the level of wealth. Rearranging the terms in this equation, we get

$$\mu = A\Sigma\alpha \tag{A.3}$$

where $A = -\frac{WJ''}{J'}$ is the market price of risk. Premultiplying the above equation by α', we get

$$\mu_m = A\sigma_m^2 \tag{A.4}$$

where μ_m and σ_m^2 are respectively the expected return and the variance on the market.

The above was a model with time invariant parameters. Introducing time varying moments imposes no additional difficulty and we write the full model as

$$\begin{aligned} \tilde{r}_{m,t+1} &= \mu_{m,t} + \tilde{\epsilon}_{t+1} \\ &= A\sigma_{m,t}^2 + \tilde{\epsilon}_{t+1} \end{aligned} \tag{A.5}$$

where $\mu_{m,t} = E_t[\tilde{r}_{m,t+1}]$, $\sigma_{m,t}^2 = \text{var}_t[\tilde{r}_{m,t+1}]$ and $\tilde{\epsilon}_{t+1}|\mathcal{F}_t \sim N(0, \sigma_{m,t}^2)$.

Thus Merton's model predicts a positive relation between expected volatility and returns.

B Bollerslev and Wooldridge Standard Errors

All the models in the paper can be nested in the general form of

$$r_t = \mu_t(\theta) + \epsilon_t \tag{B.1}$$

$$\epsilon_t \sim N(0, h_t(\theta)) \tag{B.2}$$

where θ is the vector of unknown parameters. The log likelihood function apart from a constant is given by

$$\begin{aligned} l_t(\theta) &= \frac{1}{2}\log h_t + \frac{1}{2}\frac{\epsilon_t^2}{h_t} \equiv \frac{1}{2}\log h_t + \frac{1}{2}\frac{(r_t - \mu_t)^2}{h_t} \\ \mathcal{L}_T(\theta) &= \sum_{t=1}^{T} l_t(\theta) \end{aligned} \tag{B.3}$$

The score of the likelihood function is

$$\begin{aligned} s_t(\theta) &= \frac{\partial \mu_t}{\partial \theta}\frac{\epsilon_t}{h_t} + \frac{1}{2}\frac{\partial h_t}{\partial \theta}\left[\frac{\epsilon_t^2}{h_t^2} - \frac{1}{h_t}\right] \\ \mathcal{S}_T(\theta) &= \sum_{t=1}^{T} s_t(\theta) \end{aligned} \tag{B.4}$$

The maximum likelihood estimates $\hat{\theta}$ are obtained from the equation $S_T(\hat{\theta}) = 0$. Bollerslev and Wooldridge (1992) show that the variance of $\hat{\theta}$ is given by

$$\text{var}(\hat{\theta}) = A_T(\hat{\theta})^{-1} B_T(\hat{\theta}) A_T(\hat{\theta})^{-1} \tag{B.5}$$

where

$$\begin{aligned} a_t(\theta) &= \frac{\partial \mu_t}{\partial \theta}\frac{\partial \mu_t}{\partial \theta'}\frac{1}{h_t} + \frac{1}{2}\frac{\partial h_t}{\partial \theta}\frac{\partial h_t}{\partial \theta'}\frac{1}{h_t^2} \\ A_T(\theta) &= \sum_{t=1}^{T} a_t(\theta) \\ B_T(\theta) &= \sum_{t=1}^{T} s_t(\theta) s_t'(\theta) \end{aligned} \tag{B.6}$$

The rest of this section provides analytical expressions for the derivatives of the mean equation and the variance equation. Note that these expressions are in the form of difference equations. In numerical computation, the initial condition is usually set to the unconditional sample variance.

GARCH

$$
\begin{aligned}
r_t &= x_t'\beta + \delta\, h_t + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \beta \\ \partial h_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ \epsilon_{t-1}^2 \\ -2\gamma_3 \epsilon_{t-1} x_{t-1} \\ -2\gamma_3 \epsilon_{t-1} h_{t-1} \end{bmatrix} + (\gamma_2 - 2\gamma_3 \delta \epsilon_{t-1}) \frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \beta \\ \partial \mu_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ x_t \\ h_t \end{bmatrix} + \delta \frac{\partial h_t}{\partial \theta}
\end{aligned}
\tag{B.7}
$$

GARCH-M

$$
\begin{aligned}
r_t &= x_t'\beta + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ \epsilon_{t-1}^2 \\ -2\gamma_3 \epsilon_{t-1} x_{t-1} \end{bmatrix} + \gamma_2 \frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ x_t \end{bmatrix}
\end{aligned}
\tag{B.8}
$$

EGARCH

$$
\begin{aligned}
r_t &= x_t'\beta + \epsilon_t \\
\log\ h_t &= \gamma_1 + \gamma_2 \log\ h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{h_t}\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \gamma_4 \\ \partial h_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 1 \\ \log h_{t-1} \\ \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \\ \frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \\ -\frac{\gamma_3 x_{t-1}}{\sqrt{h_{t-1}}} - \frac{\gamma_4 x_{t-1}}{\sqrt{h_{t-1}}} \frac{\epsilon_{t-1}}{|\epsilon_{t-1}|} \end{bmatrix} + \left(\frac{\gamma_2}{h_{t-1}} - \frac{\gamma_3 \epsilon_{t-1} + \gamma_4 |\epsilon_{t-1}|}{2h_{t-1}\sqrt{h_{t-1}}} \right) \frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \gamma_4 \\ \partial \mu_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \end{bmatrix}
\end{aligned}
\tag{B.9}
$$

EGARCH-M

$$
\begin{aligned}
r_t &= x_t'\beta + \delta\, h_t + \epsilon_t \\
\log\ h_t &= \gamma_1 + \gamma_2 \log\ h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right]
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{h_t}\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \gamma_4 \\ \partial h_t/\partial \beta \\ \partial h_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 1 \\ \log h_{t-1} \\ \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} \\ \frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \\ -\frac{\gamma_3 x_{t-1}}{\sqrt{h_{t-1}}} - \frac{\gamma_4 x_{t-1}}{\sqrt{h_{t-1}}}\frac{\epsilon_{t-1}}{|\epsilon_{t-1}|} \\ -\frac{\gamma_3 h_{t-1}}{\sqrt{h_{t-1}}} - \frac{\gamma_4 h_{t-1}}{\sqrt{h_{t-1}}}\frac{\epsilon_{t-1}}{|\epsilon_{t-1}|} \end{bmatrix} + \\
&\left(\frac{\gamma_2}{h_{t-1}} - \frac{\gamma_3 \epsilon_{t-1} + \gamma_4 |\epsilon_{t-1}|}{2h_{t-1}\sqrt{h_{t-1}}} - \frac{\delta}{\sqrt{h_{t-1}}}\left(\gamma_3 + \gamma_4 \frac{\epsilon_{t-1}}{|\epsilon_{t-1}|}\right)\right)\frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \gamma_4 \\ \partial \mu_t/\partial \beta \\ \partial \mu_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \\ h_t \end{bmatrix} + \delta \frac{\partial h_t}{\partial \theta} \qquad \text{(B.10)}
\end{aligned}
$$

AGARCH

$$
\begin{aligned}
r_t &= x_t'\beta + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \left(\epsilon_{t-1} + \gamma_4\right)^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \gamma_4 \\ \partial h_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ (\epsilon_{t-1} + \gamma_4)^2 \\ 2\gamma_3(\epsilon_{t-1} + \gamma_4) \\ -2\gamma_3(\epsilon_{t-1} + \gamma_4)x_{t-1} \end{bmatrix} + \gamma_2 \frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \gamma_4 \\ \partial \mu_t/\partial \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \end{bmatrix} \qquad \text{(B.11)}
\end{aligned}
$$

AGARCH-M

$$
\begin{aligned}
r_t &= x_t'\beta + \delta\, h_t + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \left(\epsilon_{t-1} + \gamma_4\right)^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial \gamma_1 \\ \partial h_t/\partial \gamma_2 \\ \partial h_t/\partial \gamma_3 \\ \partial h_t/\partial \gamma_4 \\ \partial h_t/\partial \beta \\ \partial h_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ (\epsilon_{t-1} + \gamma_4)^2 \\ 2\gamma_3(\epsilon_{t-1} + \gamma_4) \\ -2\gamma_3(\epsilon_{t-1} + \gamma_4)x_{t-1} \\ -2\gamma_3(\epsilon_{t-1} + \gamma_4)h_{t-1} \end{bmatrix} + \left(\gamma_2 - 2\gamma_3\delta(\epsilon_{t-1} + \gamma_4)\right)\frac{\partial h_{t-1}}{\partial \theta} \\
\frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial \gamma_1 \\ \partial \mu_t/\partial \gamma_2 \\ \partial \mu_t/\partial \gamma_3 \\ \partial \mu_t/\partial \gamma_4 \\ \partial \mu_t/\partial \beta \\ \partial \mu_t/\partial \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \\ h_t \end{bmatrix} + \delta \frac{\partial h_t}{\partial \theta} \qquad \text{(B.12)}
\end{aligned}
$$

GJRGARCH

$$\begin{aligned} r_t &= x_t'\beta + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 + \gamma_4 S_{t-1}^- \epsilon_{t-1}^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial\gamma_1 \\ \partial h_t/\partial\gamma_2 \\ \partial h_t/\partial\gamma_3 \\ \partial h_t/\partial\gamma_4 \\ \partial h_t/\partial\beta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ \epsilon_{t-1}^2 \\ S_{t-1}^-\epsilon_{t-1}^2 \\ -2\epsilon_{t-1}x_{t-1}(\gamma_3 + S_{t-1}^-\gamma_4) \end{bmatrix} + \gamma_2 \frac{\partial h_{t-1}}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial\gamma_1 \\ \partial \mu_t/\partial\gamma_2 \\ \partial \mu_t/\partial\gamma_3 \\ \partial \mu_t/\partial\gamma_4 \\ \partial \mu_t/\partial\beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \end{bmatrix} \end{aligned} \tag{B.13}$$

GJRGARCH-M

$$\begin{aligned} r_t &= x_t'\beta + \delta\, h_t + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 + \gamma_4 S_{t-1}^- \epsilon_{t-1}^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial\gamma_1 \\ \partial h_t/\partial\gamma_2 \\ \partial h_t/\partial\gamma_3 \\ \partial h_t/\partial\gamma_4 \\ \partial h_t/\partial\beta \\ \partial h_t/\partial\delta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ \epsilon_{t-1}^2 \\ S_{t-1}^-\epsilon_{t-1}^2 \\ -2\epsilon_{t-1}x_{t-1}(\gamma_3 + S_{t-1}^-\gamma_4) \\ -2\epsilon_{t-1}h_{t-1}(\gamma_3 + S_{t-1}^-\gamma_4) \end{bmatrix} + (\gamma_2 - 2\delta\epsilon_{t-1}(\gamma_3 + \gamma_4 S_{t-1}^-)) \frac{\partial h_{t-1}}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial\gamma_1 \\ \partial \mu_t/\partial\gamma_2 \\ \partial \mu_t/\partial\gamma_3 \\ \partial \mu_t/\partial\gamma_4 \\ \partial \mu_t/\partial\beta \\ \partial \mu_t/\partial\delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x_t \\ h_t \end{bmatrix} + \delta \frac{\partial h_t}{\partial \theta} \end{aligned} \tag{B.14}$$

GARCH-M(Standard Deviation)

$$\begin{aligned} r_t &= \mu + \delta\sqrt{h_t} + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial h_t}{\partial \theta} &\equiv \begin{bmatrix} \partial h_t/\partial\gamma_1 \\ \partial h_t/\partial\gamma_2 \\ \partial h_t/\partial\gamma_3 \\ \partial h_t/\partial\mu \\ \partial h_t/\partial\delta \end{bmatrix} = \begin{bmatrix} 1 \\ h_{t-1} \\ \epsilon_{t-1}^2 \\ -2\gamma_3\epsilon_{t-1} \\ -2\gamma_3\epsilon_{t-1}\sqrt{h_{t-1}} \end{bmatrix} + (\gamma_2 - \gamma_3\delta\frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}}) \frac{\partial h_{t-1}}{\partial \theta} \\ \frac{\partial \mu_t}{\partial \theta} &\equiv \begin{bmatrix} \partial \mu_t/\partial\gamma_1 \\ \partial \mu_t/\partial\gamma_2 \\ \partial \mu_t/\partial\gamma_3 \\ \partial \mu_t/\partial\mu \\ \partial \mu_t/\partial\delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \sqrt{h_t} \end{bmatrix} + \frac{\delta}{2\sqrt{h_t}} \frac{\partial h_t}{\partial \theta} \end{aligned} \tag{B.15}$$

References

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold, 1999, Range-Based Estimation of Stochastic Volatility Models, Working paper.

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

———, 1997b, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance* 4, 115–158.

———, 1998, Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* 39, 885–905.

———, Francis X. Diebold, and Heiko Ebens, 1999, The Distribution of Stock Return Volatility, Working Paper.

Baillie, Richard T., and Tim Bollerslev, 1992, Prediction in Dynamic Models with Time-Dependent Conditional Variances, *Journal of Econometrics* 52, 91–113.

Bollerslev, Tim, 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307–327.

———, 1987, A Conditionally Heteroskedastic Time Series Model For Speculative Prices and Rates of Return, *Review of Economics and Statistics* 69, 542–547.

———, Ray C. Chou, and Kenneth F. Kroner, 1992, ARCH Modeling in Finance: A Review of Theory and Empirical Evidence, *Journal of Econometrics* 52, 5–59.

Bollerslev, Tim, and Jeffrey M. Wooldridge, 1992, Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances, *Econometric Reviews* 11, 143–172.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?, *Review of Financial Studies* 12, 405–428.

Chou, Ray, Robert F. Engle, and Alex Kane, 1992, Measuring Risk Aversion from Excess Returns on a Stock Index, *Journal of Econometrics* 52, 201–224.

Christensen, B.J., and N.R. Prabhala, 1998, The relation between implied and realized volatility, *Journal of Financial Economics* 50, 125–150.

Christoffersen, Peter F., and Francis X. Diebold, 1997, How Relevant is Volatility Forecasting for Financial Risk Management, Working paper.

De Santis, Giorgio, and Bruno Gerard, 1997, International Asset Pricing and Portfolio Diversification with Time-Varying Risk, *Journal of Finance* 52, 1881–1912.

Dumas, Bernard, and Bruno Solnik, 1995, The World Price of Foreign Exchange Risk, *Journal of Finance* 50, 445–479.

Engle, Robert F., 1982, Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987–1007.

——— , David M. Lilien, and Russell P. Robins, 1987, Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model, *Econometrica* 55, 391–407.

Engle, Robert F., and Victor K. Ng, 1993, Measuring and Testing the Impact of News on Volatility, *Journal of Finance* 48, 1749–1778.

Fama, Eugene F., and Kenneth R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3–25.

——— , 1989, Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics* 25, 23–49.

Ferson, Wayne E., 1989, Changes in Expected Security Returns, Risk, and the level of Interest Rates, *Journal of Finance* 44, 1191–1217.

——— , and Campbell R. Harvey, 1991, The Variation of Economic Risk Premiums, *Journal of Political Economy* 99, 385–415.

——— , 1999, Conditioning Variables and the Cross-Section of Stock Returns, *Journal of Finance* 54, 1325–1360.

French, Kenneth R., William Schwert, and Robert F. Stambaugh, 1987, Expected Stock Returns and Volatility, *Journal of Financial Economics* 19, 3–29.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle, 1993, On the Relation Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks, *Journal of Finance* 48, 1791–1801.

Goyal, Amit, and Ivo Welch, 1999, Predicting the Equity Premium, Working Paper, Anderson Graduate School of Management at UCLA.

Harvey, Campbell R., 1989, Time-Varying Conditional Covariances on Tests of Asset Pricing Models, *Journal of Financial Economics* 24, 289–317.

——— , 1991, World Price of Covariance Risk, *Journal of Finance* 46, 111–157.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics* 17, 357–390.

Lamont, Owen, 1998, Earnings and Expected Returns, *Journal of Finance* 53, 1563–1587.

Lo, Andrew W., and Craig A. MacKinlay, 1988, Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test, *Review of Financial Studies* 1, 41–66.

——— , 1989, The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation, *Journal of Econometrics* 40, 203–238.

Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867–887.

——— , 1980, On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8, 323–361.

Nelson, Daniel B., 1990a, ARCH Models as Diffusion Approximations, *Journal of Econometrics* 45, 7–38.

——— , 1990b, Stationarity and Persistence in the GARCH(1,1) Model, *Econometric Theory* 6, 318–334.

——— , 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347–370.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Pagan, Adrian R., and William G. Schwert, 1990, Alternative Models for Conditional Stock Volatility, *Journal of Econometrics* 45, 267–290.

Pesaran, Hashem M., and Allan Timmerman, 1995, Predictability of Stock Returns: Robustness and Economic Significance, *Journal of Finance* 50, 1201–1228.

Schwert, William G., 1989, Why Does Stock Market Volatility Change Over Time?, *Journal of Finance* 44, 1115–1153.

——— , 1990, Stock Volatility and the Crash of 87, *Review of Financial Studies* 3, 77–102.

Scruggs, John T., 1998, Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach, *Journal of Finance* 52, 575–603.

West, Kenneth D., Hali J. Edison, and Dongchul Cho, 1993, A utility-based comparison of some models of exchange rate volatility, *Journal of International Economics* 35, 23–45.

White, Halbert, 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica* 48, 817–838.

Figure 1: Realized Returns and Standard Deviation

This figure plots realized returns and standard deviation. The return series (r_t) is the CRSP value weighted monthly returns (including dividends). Sample period is 1962.07 to 1998.12 (T=438). Standard deviation is calculated using daily data by the formula

$$\sigma_t = \sqrt{\sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it} r_{i-1t}}$$

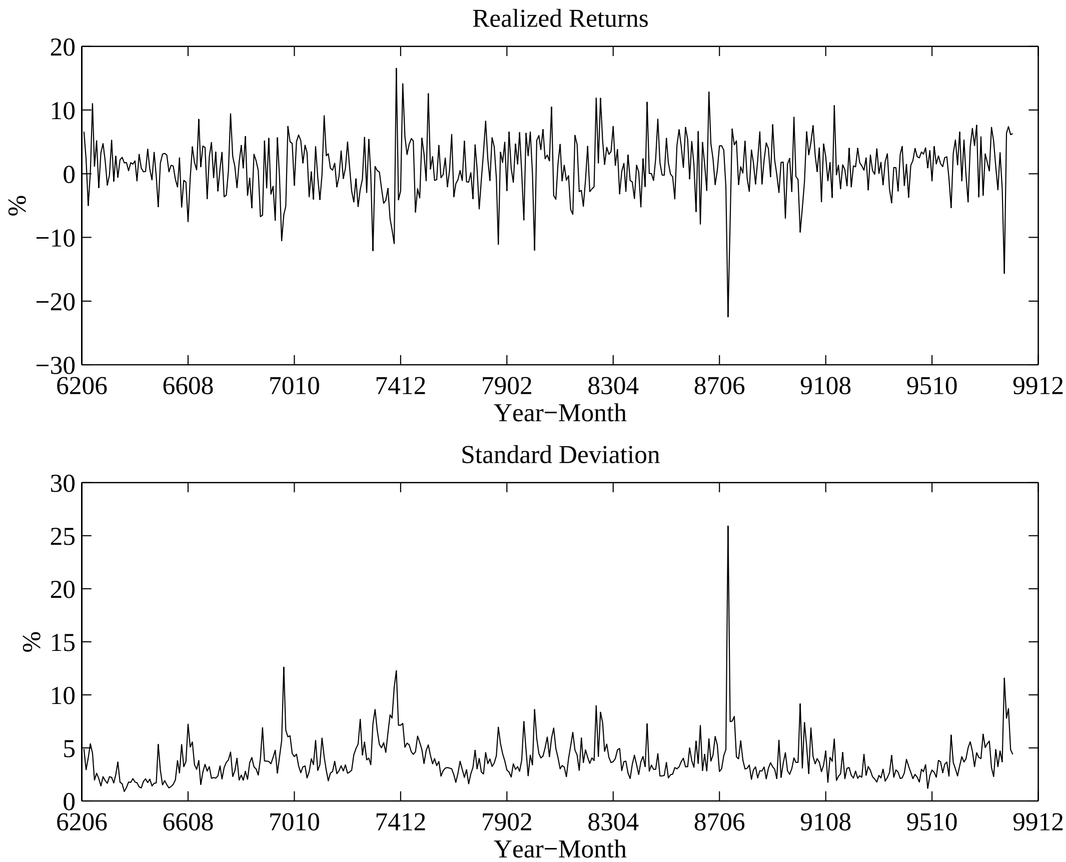


Table 1: Estimation of GARCH Models with Different Mean Specifications

This table presents results for the estimation of following GARCH models with different mean equation. The return series (r_t) is the CRSP value weighted monthly returns (including dividends). Sample period is 1962.07 to 1998.12 (T=438). First row in each estimation is the coefficient, second row (in parenthesis) is the t-statistic computed from the outer product of the scores and the third row [in brackets] is the Bollerslev and Wooldridge (1992) corrected t-statistic. Panel B presents various summary statistics from GARCH estimation. LB-1 is the Ljung-Box(12) statistic for the standardized residuals ($\epsilon_t/\sqrt{h_t}$) while LB-2 is the Ljung-Box(12) statistic for the square of standardized residuals (ϵ_t^2/h_t). logL is the log likelihood value.

$$
\begin{aligned}
r_t &= \mu + \delta\, h_t^p + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2
\end{aligned}
$$

Panel A: Estimation

	γ_1	γ_2	γ_3	μ	δ	p
M1	0.950	0.873	0.083	1.052		1
	(1.826)	(19.730)	(2.498)	(5.280)		
	[1.608]	[17.164]	[3.288]	[5.018]		
M2	1.061	0.872	0.076	-0.085	0.066	1
	(1.750)	(19.258)	(2.479)	(-0.122)	(1.692)	
	[1.786]	[16.952]	[3.106]	[-0.138]	[1.877]	
M3	1.029	0.872	0.078		0.062	1
	(1.750)	(19.962)	(2.646)		(6.034)	
	[1.802]	[17.130]	[3.186]		[4.066]	
M4	1.121	0.869	0.075	-1.506	0.624	0.5
	(1.744)	(18.511)	(2.431)	(-0.991)	(1.710)	
	[1.809]	[16.100]	[3.003]	[-1.123]	[1.899]	
M5	0.890	0.873	0.086		0.266	0.5
	(1.752)	(20.646)	(2.572)		(5.947)	
	[1.697]	[18.154]	[3.414]		[4.593]	
M6	1.048	0.873	0.076	0.061	0.042	1.112
	(1.750)	(19.314)	(2.422)	(0.027)	(0.125)	(0.560)
	[1.734]	[16.665]	[3.115]	[0.033]	[0.153]	[0.698]
M7	1.056	0.872	0.076		0.050	1.069
	(1.750)	(19.245)	(2.438)		(0.646)	(2.067)
	[1.802]	[16.984]	[3.104]		[0.702]	[2.369]

Panel B: Diagnostics from GARCH Estimation

	Mean	SDev	Min	Max	Skew	Kurt	Mean	SDev	LB-1	LB-2	logL
	Standardized Residuals ($\epsilon_t/\sqrt{h_t}$)						Volatility (h_t)				
M1	-0.01	1.00	-5.41	2.61	-0.65	5.48	19.67	8.44	8.73	5.55	-1258.19
M2	-0.03	1.00	-5.52	2.68	-0.65	5.60	19.18	7.65	8.81	6.00	-1255.14
M3	-0.03	1.00	-5.52	2.68	-0.65	5.60	19.23	7.83	8.75	5.97	-1255.14
M4	-0.03	1.00	-5.55	2.68	-0.66	5.64	19.14	7.50	8.92	5.93	-1255.22
M5	-0.03	1.00	-5.47	2.64	-0.66	5.55	19.57	8.73	8.66	5.73	-1255.81
M6	-0.03	1.00	-5.52	2.68	-0.65	5.59	19.19	7.67	8.78	6.01	-1255.13
M7	-0.03	1.00	-5.52	2.68	-0.65	5.60	19.18	7.64	8.79	6.01	-1255.13

Table 2: Performance of GARCH Models with Different Mean Specification

h_t is the predicted volatility as predicted by various models in Table 1. σ_t^2 is the actual estimate of volatility calculated using the daily returns. The following two regressions are estimated.

$$\sigma_t^2 = a + b\,h_t + u_t$$
$$r_t = c + d\,\hat{\sigma}_t^2 + e\,(\sigma_t^2 - \hat{\sigma}_t^2) + v_t$$

The first equation is estimated using OLS. Second equation uses the predicted values ($\hat{\sigma}_t^2$) from the first equation and is estimated using WLS where the weights are σ_t. White's (1980) heteroskedasticity consistent t-statistics are in brackets below the coefficient estimates. Adjusted R^2 is in percent.

	a	$b^\star$	c	d	e	R^2
M1	0.951	0.904				4.180
	[0.334]	[-0.751]				
			0.527	0.030	-0.042	6.277
			[1.396]	[1.344]	[-2.711]	
M2	-1.046	1.031				4.483
	[-0.327]	[0.208]				
			0.526	0.030	-0.042	6.335
			[1.411]	[1.361]	[-2.757]	
M3	-0.630	1.007				4.481
	[-0.200]	[0.046]				
			0.529	0.030	-0.042	6.331
			[1.420]	[1.355]	[-2.756]	
M4	-1.556	1.060				4.560
	[-0.471]	[0.386]				
			0.521	0.030	-0.043	6.360
			[1.398]	[1.375]	[-2.775]	
M5	1.178	0.897				4.415
	[0.411]	[-0.800]				
			0.542	0.029	-0.042	6.297
			[1.458]	[1.327]	[-2.741]	
M6	-0.955	1.026				4.461
	[-0.300]	[0.174]				
			0.528	0.030	-0.042	6.328
			[1.415]	[1.356]	[-2.752]	
M7	-1.035	1.031				4.469
	[-0.324]	[0.204]				
			0.527	0.030	-0.042	6.331
			[1.412]	[1.359]	[-2.754]	

$\star$ t-statistic of b is for null of $b = 1$

Figure 2: Actual and Predicted Standard Deviation

This figure plots the actual and predicted standard deviation from 1962.07 to 1998.12. Actual standard deviation is calculated using daily data and the formula

$$\sigma_t = \sqrt{\sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it} r_{i-1t}}$$

The predicted standard deviation is the volatility predicted using Model M2 of Table 1. In other words it is the square root of h_t estimated from

$$\begin{aligned} r_t &= \mu + \delta\, h_t + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned}$$

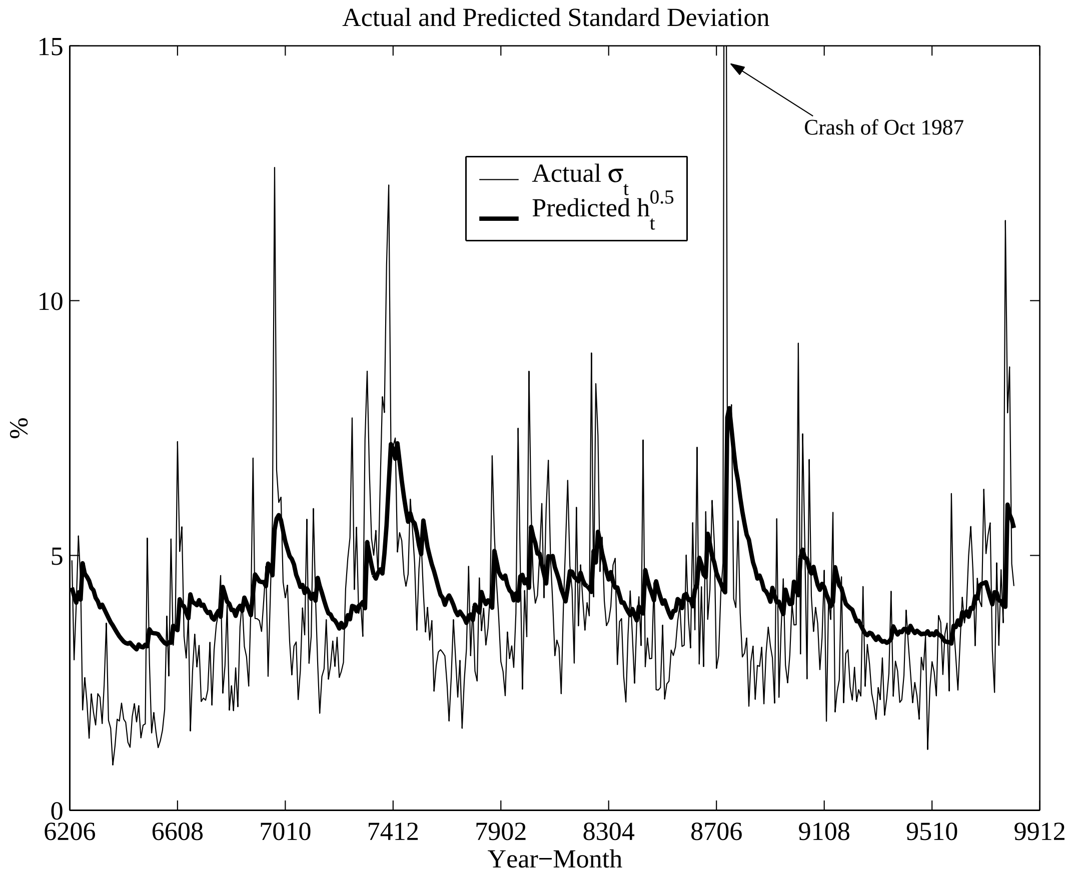


Figure 3: Confidence Intervals of Volatility Estimates

This figure plots the confidence intervals for the actual and predicted standard deviation from 1962.07 to 1998.12. For details on the construction of the confidence intervals, please refer to the main text.

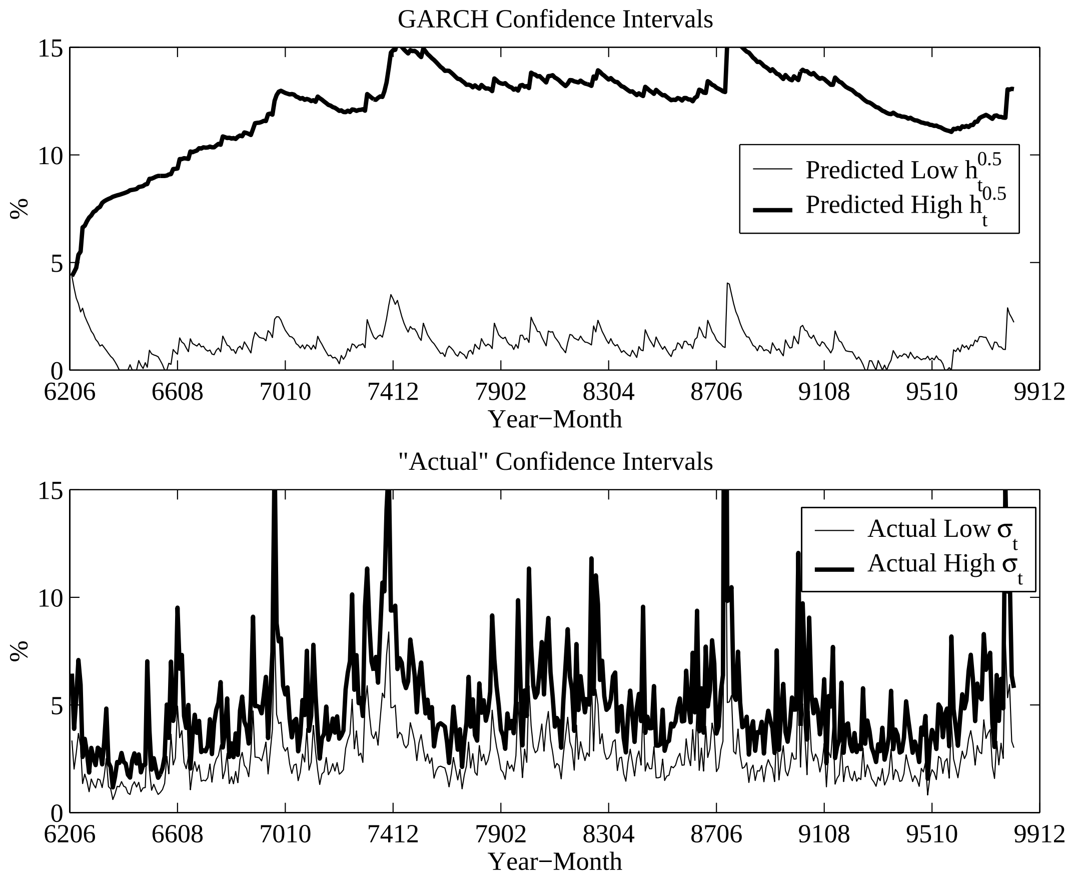


Table 3: Estimation of GARCH Models with Different Volatility Specifications

This table presents results for the estimation of various GARCH models without using the instruments. The return series (r_t) is the CRSP value weighted monthly returns (including dividends). Sample period is 1962.07 to 1998.12 (T=438). Four Models are estimated – GARCH, EGARCH, AGARCH and GJRGARCH. First row in each estimation is the coefficient, second row (in parenthesis) is the t-statistic computed from the outer product of the scores and the third row [in brackets] is the Bollerslev and Wooldridge (1992) corrected t-statistic. Panel B presents various summary statistics from GARCH estimation. LB-1 is the Ljung-Box(12) statistic for the standardized residuals ($\epsilon_t/\sqrt{h_t}$) while LB-2 is the Ljung-Box(12) statistic for the square of standardized residuals (ϵ_t^2/h_t). logL is the log likelihood value.

$$
\begin{aligned}
r_t &= \mu + \delta\, h_t + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \qquad \text{(GARCH)} \\
\log\, h_t &= \gamma_1 + \gamma_2 \log\, h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right] \qquad \text{(EGARCH)} \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 (\epsilon_{t-1} + \gamma_4)^2 \qquad \text{(AGARCH)} \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 + \gamma_4 S_{t-1}^- \epsilon_{t-1}^2 \qquad \text{(GJRGARCH)}
\end{aligned}
$$

Panel A: Estimation

	γ_1	γ_2	γ_3	γ_4	μ	δ
GARCH-M	1.061	0.872	0.076		-0.085	0.066
	(1.750)	(19.258)	(2.479)		(-0.122)	(1.692)
	[1.786]	[16.952]	[3.106]		[-0.138]	[1.877]
GARCH	0.950	0.873	0.083		1.052	
	(1.826)	(19.730)	(2.498)		(5.280)	
	[1.608]	[17.164]	[3.288]		[5.018]	
EGARCH-M	0.259	0.912	-0.085	0.178	0.157	0.050
	(2.047)	(20.998)	(-2.278)	(2.094)	(0.298)	(1.666)
	[2.962]	[29.805]	[-1.936]	[3.171]	[0.323]	[1.780]
EGARCH	0.230	0.923	-0.097	0.188	0.986	
	(1.916)	(22.285)	(-2.467)	(2.156)	(4.503)	
	[2.374]	[28.739]	[-2.076]	[3.162]	[5.465]	
AGARCH-M	1.360	0.821	0.076	-2.920	-0.044	0.058
	(1.511)	(12.797)	(1.790)	(-1.739)	(-0.072)	(1.682)
	[1.424]	[14.413]	[2.906]	[-1.750]	[-0.076]	[1.841]
AGARCH	0.994	0.824	0.081	-3.450	0.942	
	(0.927)	(11.700)	(1.747)	(-1.799)	(4.313)	
	[0.776]	[12.294]	[2.613]	[-1.951]	[4.785]	
GJRGARCH-M	1.924	0.831	0.015	0.098	-0.006	0.059
	(2.017)	(11.596)	(0.371)	(1.958)	(-0.009)	(1.558)
	[2.464]	[15.394]	[0.387]	[1.769]	[-0.009]	[1.697]
GJRGARCH	1.925	0.830	0.000	0.140	1.012	
	(1.898)	(10.094)	(0.000)	(2.231)	(4.636)	
	[2.194]	[14.189]	[0.000]	[2.500]	[5.177]	

Table continued on next page ...

Panel B: Diagnostics from GARCH Estimation

	Mean	SDev	Min	Max	Skew	Kurt	Mean	SDev	LB-1	LB-2	logL
	Standardized Residuals ($\epsilon_t/\sqrt{h_t}$)						Volatility (h_t)				
GARCH-M	-0.03	1.00	-5.52	2.68	-0.65	5.60	19.18	7.65	8.81	6.00	1255.14
GARCH	-0.01	1.00	-5.41	2.61	-0.65	5.48	19.67	8.44	8.73	5.55	1258.19
EGARCH-M	-0.01	1.00	-5.87	2.45	-0.76	6.00	19.02	9.34	8.88	5.90	1247.32
EGARCH	0.00	1.00	-5.86	2.46	-0.76	5.99	19.29	9.57	9.19	5.57	1249.87
AGARCH-M	0.00	1.00	-6.00	2.52	-0.80	6.40	19.12	8.71	9.38	5.43	1251.52
AGARCH	0.02	1.00	-6.05	2.52	-0.80	6.48	19.34	8.83	9.76	5.01	1254.32
GJRGARCH-M	-0.01	1.00	-5.95	2.51	-0.77	6.32	19.13	8.72	9.53	5.58	1253.42
GJRGARCH	-0.00	1.00	-6.04	2.55	-0.78	6.48	19.43	9.47	9.90	5.12	1255.80

Table 4: Performance of GARCH Models with Different Volatility Specification

h_t is the predicted volatility as predicted by various models in Table 3. σ_t^2 is the actual estimate of volatility calculated using the daily returns. The following two regressions are estimated.

$$\begin{aligned} \sigma_t^2 &= a + b\, h_t + u_{1,t} \\ r_t &= c + d\, \hat{\sigma}_t^2 + e\,(\sigma_t^2 - \hat{\sigma}_t^2) + u_{2,t} \end{aligned}$$

The first equation is estimated using OLS. Second equation uses the predicted values ($\hat{\sigma}_t^2$) from the first equation and is estimated using WLS where the weights are σ_t. White's (1980) heteroskedasticity consistent t-statistics are in brackets below the coefficient estimates. Adjusted R^2 is in percent.

	a	$b^\star$	c	d	e	R^2
GARCH-M	-1.046	1.031				4.483
	[-0.327]	[0.208]				
			0.526	0.030	-0.042	6.335
			[1.411]	[1.361]	[-2.757]	
GARCH	0.951	0.904				4.180
	[0.334]	[-0.751]				
			0.527	0.030	-0.042	6.277
			[1.396]	[1.344]	[-2.711]	
EGARCH-M	-1.929	1.086				7.576
	[-0.531]	[0.530]				
			0.647	0.023	-0.051	6.772
			[2.021]	[1.176]	[-3.176]	
EGARCH	-1.276	1.037				7.232
	[-0.354]	[0.235]				
			0.626	0.024	-0.051	6.841
			[1.947]	[1.227]	[-3.172]	
AGARCH-M	-1.872	1.078				6.449
	[-0.450]	[0.424]				
			0.441	0.034	-0.050	6.984
			[1.252]	[1.612]	[-3.221]	
AGARCH	-1.661	1.055				6.347
	[-0.398]	[0.304]				
			0.413	0.036	-0.051	7.051
			[1.143]	[1.650]	[-3.227]	
GJRGARCH-M	-0.403	1.000				5.535
	[-0.097]	[0.000]				
			0.244	0.045	-0.048	6.954
			[0.586]	[1.834]	[-3.127]	
GJRGARCH	0.891	0.918				5.504
	[0.219]	[-0.459]				
			0.233	0.045	-0.048	6.877
			[0.525]	[1.759]	[-3.078]	

$\star$ t-statistic of b is for null of $b = 1$

Table 5: Summary Statistics of Instruments

There are three information variables. "Dvy" is the dividend yield on the CRSP value weighted index. "Term" is the term premium calculated as difference between 10 year Treasury Bond yield and 3-month Treasury Bill yield. "Def" is the default premium calculated as the difference between yields on BAA and AAA rated corporate bonds. The sample period covers monthly observations from 1962.07 to 1998.12 (438 observations). Last three columns refer to stochastically detrended variables where stochastic detrending is carried out by subtracting the lagged moving average of the past 12 months (Lamont (1998)). Mean and Std are in percent. $\rho(n)$ is the autocorrelation coefficient at n'th lag. LB(12) is the Ljung-Box statistic of order 12. ADF is the Augmented Dickey-Fuller statistic with 12 lags. Critical value of LB at 95% level is 21.03 while that of ADF is -2.869. Panel B shows the cross-correlations between these variables

Panel A: Individual Statistics

	Dvy	Term	Def	Dvy	Term	Def
		Level			Detrended	
Mean	0.292	0.115	0.085	-0.0018	-0.0011	0.0001
Std	0.172	0.106	0.038	0.1586	0.0701	0.0192
$\rho(3)$	0.944	0.813	0.902	0.9346	0.5267	0.5700
$\rho(6)$	0.925	0.664	0.820	0.9133	0.1807	0.2411
$\rho(9)$	0.901	0.582	0.743	0.8875	0.0689	0.0042
$\rho(12)$	0.906	0.444	0.662	0.8995	-0.0892	-0.1450
LB(12)	1699.4	2521.6	3557.2	2187.1	793.7	910.3
ADF(12)	-0.938	-2.808	-2.503	-5.159	-4.182	-4.881

Panel B: Cross-Correlations $\times$ 100

Level

	Term	Def
Dvy	3.15	27.95
Term		24.05

Detrended

	Term	Def
Dvy	0.56	-0.07
Term		25.62

Table 6: OLS Regression of Stock Returns on Instrumental Variables

The return series (r_t) is the CRSP value weighted monthly returns (including dividends). The stochastically detrended information variables are described in Table 5. Sample period is 1962.07 to 1998.12 (T=438). Panel A reports three rows for an OLS regression. First row is the estimated coefficient, second row (in parenthesis) is normal OLS t-statistic and the third row [in brackets] is Newey-West (1987) heteroskedasticity and autocorrelation corrected t-statistic. Second panel reports some summary statistics on the returns and residuals. Third panel reports autocorrelation coefficients and the Ljung-Box(12) statistic.

Panel A: OLS Estimation

Cnst	Dvy	Term	Def
1.084	0.676	0.741	0.368
(5.250)	(0.518)	(0.243)	(3.304)
[5.166]	[0.578]	[0.218]	[3.386]

Adjusted R^2=2.11%

Panel B: Residuals

	Mean	Std	Min	Max	Skew	Kurt
r_t	1.085	4.367	-22.488	16.561	-0.470	5.564
ϵ_t	0.000	4.306	-23.110	14.015	-0.626	5.509
	$\rho(3)$	$\rho(6)$	$\rho(9)$	$\rho(12)$	LB	
r_t	-0.013	-0.059	0.005	0.020	10.249	
ϵ_t	0.004	-0.034	0.021	0.024	8.280	

Table 7: Estimation of GARCH Models with Instrumental Variables

This table presents results for the estimation of various GARCH models. The return series (r_t) is the CRSP value weighted monthly returns (including dividends). Sample period is 1962.07 to 1998.12 (T=438). Three instruments (x_t)are chosen, viz. dividend yield, term premium and default premium. Term premium is the difference between 10 year Treasury notes and 3 month treasury bill. Default premium is the difference between BAA and AAA rated corporate bonds. All three instruments are stochastically detrended as in Lamont (1998). Four Models are estimated – GARCH, EGARCH, AGARCH and GJRGARCH. First row in each estimation is the coefficient, second row (in parenthesis) is the standard error computed from the Hessian and the third row (in brackets) is the Bollerslev and Wooldridge (1992) corrected standard error. Panel B presents various summary statistics from GARCH estimation. LB-1 is the Ljung-Box(12) statistic for the standardized residuals ($\epsilon_t/\sqrt{h_t}$) while LB-2 is the Ljung-Box(12) statistic for the square of standardized residuals (ϵ_t^2/h_t). logL is the log likelihood value.

$$
\begin{aligned}
r_t &= x_t'\beta + \delta\, h_t + \epsilon_t \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \qquad \text{(GARCH)} \\
\log\, h_t &= \gamma_1 + \gamma_2 \log\, h_{t-1} + \gamma_3 \frac{\epsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma_4 \left[\frac{|\epsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right] \qquad \text{(EGARCH)} \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 (\epsilon_{t-1} + \gamma_4)^2 \qquad \text{(AGARCH)} \\
h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 + \gamma_4 S_{t-1}^- \epsilon_{t-1}^2 \qquad \text{(GJRGARCH)}
\end{aligned}
$$

Panel A: Estimation

	γ_1	γ_2	γ_3	γ_4	Cnst	Dvy	Term	Def	δ
GARCH-M	0.956	0.882	0.070		0.281	0.341	-0.804	0.300	0.047
	(1.610)	(18.765)	(2.309)		(0.384)	(0.243)	(-0.287)	(2.425)	(1.138)
	[1.615]	[17.739]	[3.017]		[0.427]	[0.296]	[-0.222]	[2.213]	[1.252]
GARCH	0.991	0.876	0.076		1.097	0.298	-0.106	0.341	
	(1.705)	(17.845)	(2.278)		(5.344)	(0.214)	(-0.039)	(2.853)	
	[1.546]	[16.496]	[3.088]		[5.165]	[0.258]	[-0.030]	[2.618]	
EGARCH-M	0.253	0.913	-0.085	0.172	0.548	0.259	-1.294	0.276	0.031
	(2.040)	(21.254)	(-2.517)	(2.026)	(0.892)	(0.185)	(-0.423)	(1.992)	(0.899)
	[2.616]	[27.330]	[-1.499]	[2.355]	[1.062]	[0.246]	[-0.442]	[2.582]	[1.072]
EGARCH	0.254	0.913	-0.095	0.175	1.078	0.237	-0.939	0.319	
	(2.008)	(20.714)	(-2.720)	(2.082)	(4.933)	(0.173)	(-0.313)	(2.509)	
	[2.389]	[25.883]	[-1.637]	[2.287]	[5.863]	[0.223]	[-0.324]	[3.105]	
AGARCH-M	6.023	0.051	0.173	-7.318	1.636	-0.004	3.101	0.352	-0.033
	(2.082)	(0.659)	(3.036)	(-3.812)	(3.993)	(-0.004)	(1.139)	(3.690)	(-1.296)
	[3.059]	[0.554]	[3.183]	[-5.512]	[5.004]	[-0.004]	[1.161]	[3.397]	[-1.692]
AGARCH	5.981	0.048	0.161	-7.579	1.159	0.133	2.818	0.335	
	(2.036)	(0.659)	(3.012)	(-3.623)	(6.078)	(0.117)	(1.108)	(3.882)	
	[2.912]	[0.491]	[2.843]	[-4.916]	[6.372]	[0.131]	[1.114]	[3.281]	
GJRGARCH-M	2.452	0.807	0.000	0.111	0.457	0.723	0.022	0.263	0.035
	(1.874)	(8.496)	(0.000)	(2.001)	(0.574)	(0.506)	(0.007)	(2.019)	(0.795)
	[2.110]	[10.655]	[0.000]	[1.773]	[0.595]	[0.643]	[0.007]	[2.293]	[0.865]
GJRGARCH	2.424	0.805	0.000	0.121	1.083	0.644	0.010	0.303	
	(1.832)	(8.091)	(0.000)	(2.062)	(4.969)	(0.454)	(0.004)	(2.542)	
	[2.046]	[10.267]	[0.000]	[1.865]	[5.398]	[0.579]	[0.003]	[2.787]	

Table continued on next page ...

Panel B: Diagnostics from GARCH Estimation

	Mean	SDev	Min	Max	Skew	Kurt	Mean	SDev	LB-1	LB-2	logL
	Standardized Residuals ($\epsilon_t/\sqrt{h_t}$)						Volatility (h_t)				
GARCH-M	-0.02	1.00	-5.37	2.59	-0.71	5.50	18.85	7.18	8.38	5.86	1251.86
GARCH	-0.01	1.00	-5.32	2.60	-0.72	5.44	19.05	7.43	7.74	5.39	1253.61
EGARCH-M	-0.02	1.00	-5.85	2.47	-0.81	6.07	18.74	8.99	8.26	5.51	1244.34
EGARCH	-0.01	1.00	-5.89	2.47	-0.81	6.11	18.81	9.01	7.99	5.31	1245.67
AGARCH-M	0.01	1.00	-4.57	2.61	-0.43	4.06	19.32	14.12	6.97	8.96	1233.53
AGARCH	-0.01	1.00	-4.74	2.65	-0.46	4.21	19.37	14.02	6.35	8.74	1235.06
GJRGARCH-M	-0.01	1.00	-5.99	2.58	-0.82	6.44	18.69	7.72	8.90	5.29	1250.57
GJRGARCH	-0.01	1.00	-6.01	2.59	-0.82	6.45	18.77	7.91	8.58	5.12	1251.79

Table 8: Performance of Various GARCH Models with Instrumental Variables

h_t is the predicted volatility as predicted by various models in Table 7. σ_t^2 is the actual estimate of volatility calculated using the daily returns. The following two regressions are estimated.

$$\sigma_t^2 = a + b\, h_t + u_{1,t}$$
$$r_t = c + d\, \hat{\sigma}_t^2 + e\, (\sigma_t^2 - \hat{\sigma}_t^2) + u_{2,t}$$

The first equation is estimated using OLS. Second equation uses the predicted values ($\hat{\sigma}_t^2$) from the first equation and is estimated using WLS where the weights are σ_t. White's (1980) heteroskedasticity consistent t-statistics are in brackets below the coefficient estimates. Adjusted R^2 is in percent.

	a	$b^\star$	c	d	e	R^2
GARCH-M	-1.966	1.098				4.477
	[-0.632]	[0.625]				
			0.593	0.026	-0.041	6.195
			[1.610]	[1.208]	[-2.683]	
GARCH	-1.455	1.060				4.478
	[-0.487]	[0.401]				
			0.591	0.026	-0.041	6.220
			[1.610]	[1.215]	[-2.682]	
EGARCH-M	-2.278	1.121				7.478
	[-0.608]	[0.714]				
			0.542	0.029	-0.054	7.322
			[1.707]	[1.470]	[-3.425]	
EGARCH	-2.199	1.113				7.383
	[-0.580]	[0.664]				
			0.500	0.031	-0.055	7.508
			[1.564]	[1.568]	[-3.477]	
AGARCH-M	6.939	0.610				5.396
	[2.835]	[-2.661]				
			0.815	0.014	-0.030	5.464
			[1.690]	[0.511]	[-2.103]	
AGARCH	6.557	0.629				5.660
	[2.631]	[-2.579]				
			0.776	0.016	-0.031	5.565
			[1.640]	[0.592]	[-2.165]	
GJRGARCH-M	-2.955	1.160				5.851
	[-0.624]	[0.729]				
			0.075	0.054	-0.052	7.446
			[0.172]	[2.102]	[-3.341]	
GJRGARCH	-2.565	1.135				5.872
	[-0.548]	[0.624]				
			0.076	0.054	-0.052	7.472
			[0.171]	[2.070]	[-3.339]	

$\star$ t-statistic of b is for null of $b = 1$

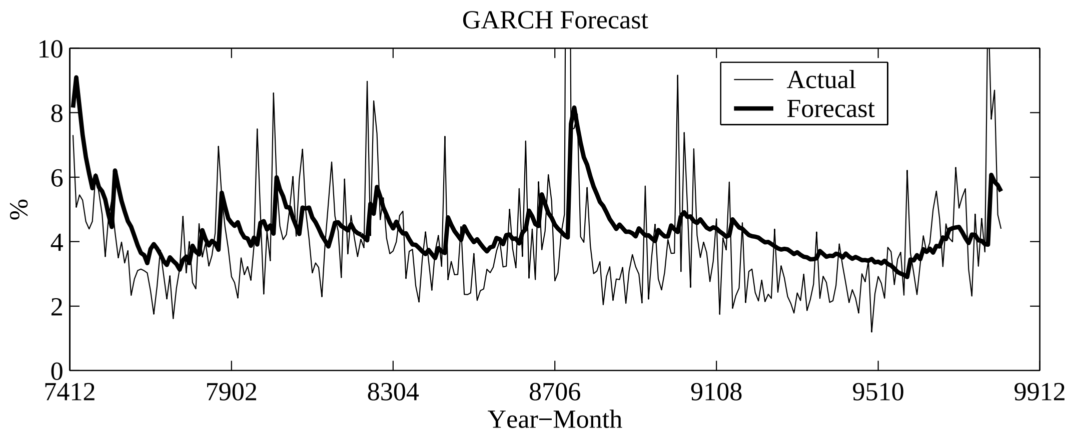
Figure 4: Out-of-Sample Forecasts of Standard Deviation

This figure plots out-of-sample forecasts of monthly standard deviation. The out of sample forecasts start in 1975.01 after an initial phase in period.
GARCH forecasts are one period ahead forecasts obtained by recursively estimating the equation

$$\begin{aligned} r_t &= \mu + \delta\, h_t + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned}$$

ARMA forecasts are one period ahead forecasts obtained by recursively estimating the equation

$$(1 - \theta_1 L)\sigma_t^2 = \phi_0 + (1 + \phi_1 L + \phi_2 L^2) u_t$$



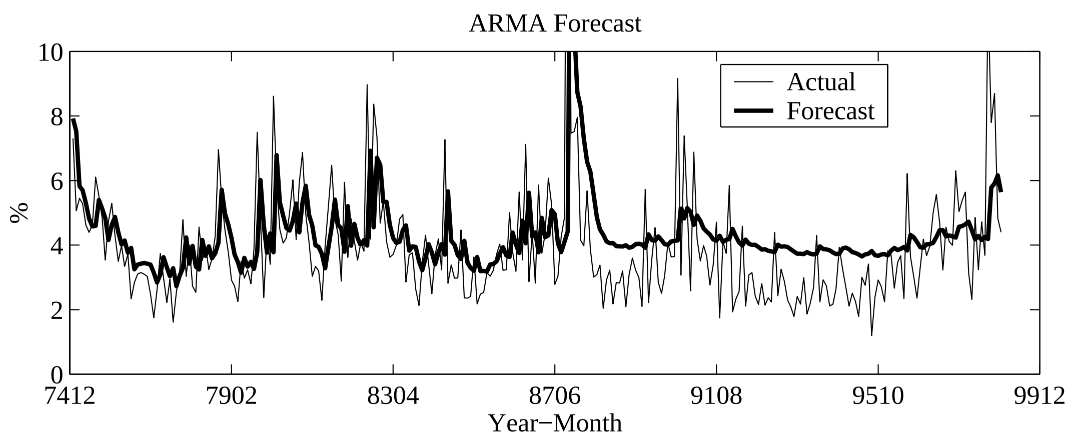


Table 9: Out-of-Sample Performance of GARCH and ARMA Models

This table gives summary statistics of out-of-sample forecast errors. The out of sample forecasts start in 1975.01 after an initial phase in period. Actual volatility is the monthly variance calculated from daily data using the equation

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=2}^{N_t} r_{it} r_{i-1t}$$

GARCH forecasts ($\hat{h}_t$)are one period ahead forecasts obtained by recursively estimating the equation

$$\begin{aligned} r_t &= \mu + \delta\, h_t + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned}$$

ARMA forecasts ($\hat{\sigma}_t^2$)are one period ahead forecasts obtained by recursively estimating the equation

$$(1 - \theta_1 L)\sigma_t^2 = \phi_0 + (1 + \phi_1 L + \phi_2 L^2) u_t$$

	GARCH $\sigma_t^2 - \hat{h}_t$	ARMA $\sigma_t^2 - \hat{\sigma}_t^2$
Mean	-2.62	-1.87
Median	-5.72	-4.92
Min	-57.08	-51.55
Max	118.74	116.44
SDev	15.41	15.13
RMSE	15.60	15.22

Table 10: GARCH Estimation of Exchange Rate Data

This table presents results of estimation of GARCH model of exchange rate data. Base data is half hourly exchange rates between US$ and DM for year 1996. Using this base data, daily returns are the continuously compounded returns for the day using the 48 half hourly observations for that day. Daily variance is calculated using the equation

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it}r_{i-1t} + 2\sum_{i=3}^{N_t} r_{it}r_{i-2t}$$

where N_t is usually 48. $\rho(n)$ is the autocorrelation a nth lag.

Panel A: Summary Statistics

	Half Hourly Returns N=12,574	Daily Returns N=262	Daily Variance N=262
Mean	5.10×10^{-5}	2.41×10^{-3}	1.65×10^{-5}
Max	8.26×10^{-3}	1.17×10^{-2}	1.77×10^{-4}
Min	-7.86×10^{-3}	-1.98×10^{-2}	0
SDev	7.26×10^{-3}	4.15×10^{-2}	2.18×10^{-5}
Skew	-0.249	-0.380	4.176
Kurt	16.122	4.892	25.38
$\rho(1)$	-0.158	-0.080	0.066
$\rho(2)$	0.012	0.017	-0.013
$\rho(3)$	0.018	0.017	0.170
$\rho(4)$	-0.014	0.104	0.024
$\rho(5)$	0.005	-0.046	-0.001

Panel B: GARCH Estimation using Daily Data

$$r_t = \underset{[0.828]}{2.05 \times 10^{-3}} + \epsilon_t$$

$$h_t = \underset{[0.480]}{1.51 \times 10^{-6}} + \underset{[5.240]}{0.878}\ h_{t-1} + \underset{[1.358]}{0.045}\ \epsilon_{t-1}^2$$

Figure 5: Actual and Predicted Standard Deviation from GARCH Estimation of Exchange Rate Data

This figure plots the actual and predicted standard deviation from 01/01/1996 to 12/31/1996. Actual standard deviation is calculated using intraday data and the formula

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2\sum_{i=2}^{N_t} r_{it}r_{i-1t} + 2\sum_{i=3}^{N_t} r_{it}r_{i-2t}$$

The predicted standard deviation is the volatility predicted using GARCH Model of Panel B of Table 10. In other words it is the square root of h_t estimated from

$$\begin{aligned} r_t &= \mu + \epsilon_t \\ h_t &= \gamma_1 + \gamma_2 h_{t-1} + \gamma_3 \epsilon_{t-1}^2 \end{aligned}$$