An Empirical Comparison of GARCH Option Pricing Models

K. C. Hsieh* Peter Ritchken†

September 14, 2000

*Charles Schwab, SF120KNY-18-282,101 Montgomery St., San Francisco, CA 94104, Tel: (415)636-5708, Fax:(415)636-3637,E-mail: kchsieh@pacbell.net

†Case Western Reserve University, WSOM, 10900 Euclid Ave., Cleveland, OH 44106-7235, Phone: (216) 368–3849, Fax: (216) 368–4776, E-mail: phr@pyrite.som.cwru.edu

Abstract

Heston and Nandi (2000) provide considerable empirical support for their GARCH option pricing model. Their model has the advantage that analytical solutions are available for pricing European options. This article takes a closer look at this model and compares its performance with the NGARCH option model of Duan (1995). We confirm Heston and Nandi's findings, namely that their model can explain a significant portion of the volatility smile. However, we show that the NGARCH model is superior in removing biases from pricing residuals for all moneyness and maturity categories especially for out-the-money contracts. The out-of-sample performance of both GARCH models is closely examined, and the NGARCH model is shown to have very attractive properties. The NGARCH model continues to perform well, even when the parameters of the model are not re-estimated for long periods of time. Given the existence of relatively efficient algorithms for pricing American claims and exotics under NGARCH processes, we recommend that traders and risk managers consider the NGARCH model.

There is overwhelming empirical evidence that return innovations in stocks influence future volatilities. For example, large absolute returns are more likely to be followed by large absolute returns, with volatility being persistent. In addition, if the news is bad, volatility expands more than if the news is good. This implies that there is a negative correlation between asset return innovations and volatility innovations.[1] These time series properties of volatility are important features that should be captured in models that specify the dynamics of prices over time. This feedback effect between returns and volatility is also important to capture in option pricing. Duan (1995) shows how this can be accomplished by pricing options under GARCH processes that have this property. Indeed, computing option prices under GARCH processes is now very well understood. Unfortunately, analytical solutions for prices of options are not generally available and hence numerical procedures have to be invoked. Efficient martingale simulation methods have been developed by Duan and Simonato (1998) for pricing European claims. Duan and Simonato (2000) and Ritchken and Trevor (1999) also develop numerical schemes for pricing American claims. Finally, Heston and Nandi (2000) (hereafter, HN) have developed "closed form" solutions for European options under very specific GARCH like volatility updating schemes and Duan, Gauthier and Simonato (1999) have established analytical approximations for the NGARCH process.

The importance of GARCH option pricing has recently expanded due to their linkage with stochastic volatility models. Indeed, even if one finds GARCH models a bit mechanical, the methodology is useful since their diffusion limits contain many well known stochastic volatility models.[2] From an estimation perspective, GARCH models may have distinct advantages over stochastic volatility models. Continuous time stochastic volatility models are difficult to implement, because, with discrete observations on the underlying asset price process, the volatility is not readily identifiable. If the volatility level cannot be established, option prices cannot be computed. To overcome this problem implied volatilities are often established from concurrent option prices. Indeed, a common technique for estimating stochastic volatility models, as adopted by Bakshi, Cao, and Chen (1997) for example, is to use a cross section of option data to estimate *all* the parameters, including volatility, on a daily basis. If the parameters of the process are required to be constant through time, then a time series of daily option records are used in the analysis and a *daily sequence* of implied volatilities has to be estimated. Since the number of unknown volatilities increases linearly with the number of days, the computational effort involved in the optimization problem soon becomes severe. This approach has been used by Bates (1996) and Nandi (1998), for example. In contrast, GARCH models have the advantage that the volatility is observable from the history of asset prices. Consequently, it is possible to

[1] For discussions on these features see Black (1976), Bollerslev, Chou and Kroner (1992), and Engle, and Ng (1993), for example.

[2] Duan (1996, 1997), Corradi (2000) and Nelson (1990), provide details on the relationships between univariate GARCH models and bivariate stochastic volatility models.

price options, solely on the basis of observable history of the underlying asset process, without requiring information on derivative prices. As a result, option prices can be generated in illiquid markets where concurrent information on derivative prices may not exist. In addition, as emphasized by HN, with GARCH models, only a finite number of parameters need to be estimated irrespective of the length of the time series, thereby considerably simplifying the estimation procedure.

HN perform extensive empirical tests that provide convincing support for their model. In particular, they show that their model provides a substantial improvement over the ad-hoc Black Scholes model of Dumas, Fleming and Whaley (1998) that uses a separate implied volatility for each option to fit the volatility "smile". In contrast, previous empirical tests, conducted by Dumas et. al., showed that the implied binomial tree-deterministic volatility models were outperformed by the ad hoc Black Scholes model. HN conclude that the improvements provided by their model is due to the ability of their model to capture the correlation of volatility with returns and the path dependence in volatility. Their results bring GARCH models to the forefront of viable pricing models.

The primary purpose of this paper is to further examine the empirical performance of the HN model and to compare this model to Duan's (1995) NGARCH option model. We are keen to establish if the HN model dominates the NGARCH model. If the HN model does dominate Duan's NGARCH model, then it might be worthwhile to design specialized algorithms for pricing American options and exotics under the HN specification. However, if the model has some shortfalls, then the fact that it offers closed form solutions may not be that significant, especially given the fact that there are analytical approximations for European contracts for the NGARCH model and relatively efficient pricing mechanisms exist for pricing American claims for most GARCH processes. The paper proceeds as follows. In section 1 we discuss the two GARCH option models. In section 2 we describe the experimental design for evaluating the performance of the two models. We carefully describe the data and discuss the pricing methodology that we use in the study. We especially focus on the computational issues used to generate the option prices. In section 3 we present the empirical results, and section 4 concludes.

I The Option Pricing Models

(a) Duan's NGARCH Model

Let S_t be the asset price at date t, and let h_t be the conditional volatility of the logarithmic return over the period $[t, t+1]$, which is a day. The dynamics are assumed to follow the process:

$$\ln \frac{S_{t+1}}{S_t} = r_f + \lambda h_t - \frac{1}{2} h_t^2 + h_t \epsilon_{t+1} \tag{1}$$

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2(\epsilon_{t+1} - \gamma)^2 \tag{2}$$

where r_f is the risk free rate, λ is the unit risk premium for the asset, ϵ_{t+1} is a standard normal random variable, and γ is a nonnegative parameter that captures the negative correlation between return and volatility innovations. To ensure that the conditional volatility stays positive β_0, β_1 and β_2 should be nonnegative.

Under suitable preference restrictions, Duan (1995) has derived the following risk neutral probability measure under which discounted claims are martingales:

$$\ln \frac{S_{t+1}}{S_t} = r_f - \frac{1}{2}h_t^2 + h_t \nu_{t+1} \tag{3}$$

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2[\nu_{t+1} - \omega]^2 \tag{4}$$

where $\omega = \gamma + \lambda$ and ν_{t+1} is a standard normal random variable.

The above option model has four parameters, β_0, β_1 and β_2 and ω that need to be estimated, together with the initial volatility, h_0. If time series information is to be incorporated as well, then the additional parameter λ can be identified.

(b) Heston and Nandi Model

Heston and Nandi postulate the following dynamics:

$$\ln \frac{S_{t+1}}{S_t} = r_f + \lambda h_t^2 + h_t \epsilon_{t+1} \tag{5}$$

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2(\epsilon_{t+1} - \gamma h_t)^2 \tag{6}$$

The process is stationary with finite mean and variance if $\beta_1 + \beta_2\gamma^2 < 1$. The variance updating equation is similar to the NGARCH model, with the biggest difference being the fact that the β_2 term is not multiplied by the local variance. That is, the last term is determined to a large degree by the *normalized* residual, rather than the residual.

For pricing purposes, the risk neutralized measure is given by

$$\ln \frac{S_{t+1}}{S_t} = r_f - \frac{1}{2}h_t^2 + h_t \nu_{t+1} \tag{7}$$

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2[\nu_{t+1} - \omega h_t]^2 \tag{8}$$

where $\omega = \gamma + \lambda + \frac{1}{2}$.

Heston and Nandi show that, for this particular structure, the moment generating function of the logarithmic price at date T takes on a log linear form. As a result risk neutral probabilities can be computed and European call option prices can be computed. Like the previous model, for purposes of option pricing, there are 4 unknown parameters in this model, together with the initial local volatility.

II Experimental Design

In this section we describe the option data, discuss the estimation methodology and lay out the types of analysis that are performed. Since option data are used to extract out parameter values, non linear optimization methods are invoked that require large sets of option contracts to be frequently repriced. As a result, a high demand is placed on the pricing routines and efficient schemes are crucial. We therefore discuss the numerical pricing mechanism that we adopted in some detail.

II.1 Description of Data

The $S\&P500$ index options are European options that exist with maturities in the next two calendar months, and also for the time periods corresponding to the expiration dates of the futures. Our price data on the options covered the five year period from January 1991 to December 1995. We collected data on Wednesdays and excluded contracts with maturities fewer than six days. We only used options with bid/ask price quotes during the last half hour of trading. For these contracts we also captured the reported concurrent stock index level associated with each option trade.

In order to price the call options we need to adjust the index level according to the dividends paid out over the time to expiration. We follow Harvey and Whaley (1992), and Bakshi, Cao and Chen (1997), and use the actual cash dividend payments made during the life of the option to proxy for the expected dividend payments. The present value of all the dividends is then subtracted from the reported index levels to obtain the contemporaneous adjusted index levels. This procedure assumes that the reported index level is not stale and reflects the actual price of the basket of stocks representing the index. Since intra day data and not the end of the day option prices are used, the problem with the index level being stale is not severe.[3] Since we used the actual contemporaneous index level associated with each option trade that was reported in the data base, the actual adjusted index level would vary slightly among the individual contracts depending on their time of trade. We normalize all option and strike prices so that the adjusted index price is exactly \$1 for every contract. This transformation is helpful, since all contracts can now be priced relative to the same constant underlying price. Finally, we used the T-Bill term structure to extract the appropriate discount rates.

[3] There are other methods for establishing the adjusted index level. The first is to compute the mid points of call and put options with the same strikes and then to use put-call parity to imply out the value of the underlying index. Of course, this method has its own problems, since with non negligible bid ask spreads, put call parity only holds as an inequality. An alternative approach is to use the stock index futures price to back out the implied dividend adjusted index level. This leads to one stock index adjusted value that is used for all option contracts. For a discussion of these approaches see Jackwerth and Rubinstein (1996).

II.2 Estimation

It is possible to use the time series of the underlying $S\&P$ 500 index to establish the maximum likelihood estimates for all the parameters for both models. However, such an analysis ignores the information content of the option prices that complement the time series of underlying prices. In our analysis we wanted to incorporate the time series properties of prices, together with the cross sectional information provided by option prices.

Our objective function and methodology is similar to Bakshi, Cao and Chen (1997), Dumas, Fleming and Whaley (1998), and others, who minimize the sum of squared errors between theoretical and actual prices using a non-linear least squares procedure. These studies have been conducted in the context of continuous time stochastic volatility option pricing. Since our underlying process is a GARCH process, our exact methodology is similar to Heston and Nandi, and is briefly reviewed.

In a GARCH setting there are two state variables, namely the underlying asset price and the local variance. The fact that the local variance is determined by the history of asset innovations makes this problem considerably easier to solve relative to the estimation problems of continuous time stochastic volatility models.[4]

Let $e_{i,t}$ represent the difference between the model price and the actual price of contract i at date t. Heston and Nandi use the following criterion function:

$$\text{Minimize SSE}(\theta) = \sum_{t=1}^{T} \sum_{i=1}^{N_t} e_{i,t}^2$$

Here, T denotes the number of weeks (Wednesdays) in the sample, N_t is the number of options traded on the Wednesday of week t, and θ represents the parameter set, $\theta = \{\beta_0, \beta_1, \beta_2, \lambda, \gamma, h_0\}$.

Notice that in order to price options under a specific parameter set, for each week we need to have values for the two state variables. The asset price is known, but the local volatility has to be determined from its value at the end of the previous week. Its new value will depend on the daily innovations of prices over the past week. Given the sequence of daily moves, the volatility updates can be performed, and the time series for the second state variable is determined, conditional on its beginning week value. The initial value for h_0 enters the analysis as a parameter to be optimally determined from the data.

We split up each year of our 5 years of data into two 6 month intervals, this giving us 5 non overlapping data sets. For each of these data sets we use the time series of daily asset prices, together with weekly option prices, to estimate the parameters using the minimum sum

[4] For discussions on some stochastic volatility models and issues of estimating parameters see Andersen and Lund (1997), Bakshi, Cao and Chen (1997), Heston (1993), and Hull and White (1987). Heston and Nandi (2000) provide convincing arguments in favor of GARCH models over continuous time stochastic volatility models.

of squares principle.

Given the parameter estimates, together with the initial volatility, we use the daily time series of actual index prices to generate a daily time series of local volatilities over the entire year. Given, the index and local volatility at each week, theoretical option prices can be generated and compared with actual option prices. Over the first 6 months of each year, the residuals we generate are referred to as *in-sample residuals*. However, over the last 6 months of each year, the theoretical prices are based on parameter estimates that have not used information on concurrent option prices. Since these theoretical prices do not use any option information over the last six months, these residuals are referred to as *out-of-sample residuals*. We therefore obtain 5 sets of parameter estimates, 5 sets of in sample residuals and 5 sets of out-of-sample residuals. The residuals from our two GARCH models can be examined to identify whether a strike price bias or a maturity bias exists. The residuals can also be compared with each other.

Our benchmark model is the Black Scholes model. The Black Scholes residuals are generated each week by identifying the volatility as the number that minimizes the sum of squared errors of all option prices that are available at that date. The residuals generated by the Black Scholes model in all weeks are in-sample residuals that use the concurrent option data to establish an optimal implied volatility. We emphasize that over the last 6 months of each year our GARCH models do not use any option data to estimate parameters, whereas the Black Scholes model uses all the option data each week to estimate the best volatility. By comparing the "in sample" residuals from Black and Scholes with the "out of sample" residuals of the GARCH models, we are requiring a higher hurdle for assessing the performance of our GARCH models.[5]

II.3 Computational Schemes

The optimization problem encountered in each of the in sample problems is highly nonlinear in the parameter values, and is a non trivial problem to solve. Since there are no analytical solutions for the gradient, numerical optimization techniques have to be used that require hundreds, if not thousands of function evaluations. Since each function call requires large sets of option prices to be computed, we need an efficient scheme for pricing.

At a given date, t, we have a collection of call option contracts. Let C_i, be the price of contract i, with strike X_i, and maturity, T_i, where $i = 1, 2..., N_t$. Let G_i be the dividend adjusted index for the i^{th} contract. For contracts that have the same maturity the G_i values

[5] It is well known that the Black Scholes model produces large errors in pricing. Indeed, HN compare their GARCH model to an ad-hoc Black Scholes model that uses a *separate* implied volatility for each option, as in Dumas, Fleming and Whaley (1998). HN show that their GARCH model produces substantially smaller out-of-sample residuals than the ad-hoc model. Our primary reason for presenting the simple Black Scholes residuals will be to give some sense of the magnitude of improvement that GARCH models have over Black Scholes.

will typically be close, if not identical. As discussed earlier, since option contracts do not all trade at the same time, the underlying index prices might not be identical. As a first step, we recognize that options are homogeneous of degree 1 in the underlying price and strike. Hence, for computational purposes, we normalize all the prices so that the underlying price is exactly 1.

We then use simulation to price the contracts. Since the initial price of the underlying is the same for all the contracts, we can generate one path over time, and at the appropriate expiration dates, compute the exercise value of all the terminating contracts. Thus, each path gives rise to N_i option prices. After K paths are generated we have all our option prices. Simulation is particularly attractive when the number of contracts is large. In our case, on each Wednesday we typically have over 30 contracts. To reduce the standard errors, we used a control variate method. In particular, we used Duan and Simonato's (1997) efficient martingale simulation method. This method generates all K paths at once, and adjusts the sample paths so that the sample process is a martingale. Among other things, this ensures that the computed call and put values satisfy put call parity. We used 5000 replications in our pricing module. Before selecting this number we performed extensive computational tests, and for a wide array of parameter values we concluded that 2000 replications produced tight enough confidence intervals for the true prices.

While there is a question that a small bias in results could result from using a small number of sample paths, we wanted to ensure that the bias in prices were the same for both models. To accomplish this the sequence of random numbers used in both the Duan and HN models were identical. In this regard, the experimental results are performed under identical conditions.

An alternative approach would have been to use a computational scheme like Ritchken and Trevor (1999) or Duan and Simonato (2000) to price European claims. Their algorithms are extremely useful for pricing American claims, when the parameter values are reasonable. However, in the search for an optimal set of parameters, our experience has been that certain configurations of parameters can cause the algorithms to slow down considerably. We found the simulation procedure to be much more efficient and robust than using other computational schemes. In addition, using different numerical procedures for pricing HN and NGARCH models, results in additional errors in the analysis. To the extent that we have used common streams of random numbers for pricing HN and NGARCH models, the biases in prices will be common.

III Empirical Results

Both the HN model and the NGARCH model have 6 unobservable values, including the 5 parameters, $\beta_0, \beta_1, \beta_2\ \omega$, and λ, as well as the initial local volatility, h_0. In our first optimizations, we found that the surface was fairly flat around the optimal solution, and estimates of the parameters could fluctuate widely, without significant improvement in the SSE. By assuming the market price of risk, λ, to be zero, lead to very little change in the objective function. As a result, in what follows, we report the results when our optimizations were conducted over the 5 remaining values.

Table 1a reports the parameter estimates for the HN Model for each of the 10 six month periods, from January 1991 to the last 6 month period in 1995. Table 1b shows the same results for the NGARCH model. For both models we report the estimates of the 4 parameters, namely β_0, β_1, β_2 and ω, as well as the implied estimate of the initial local volatility, h_0. Our results for the HN model are generally in the same range as those established by HN. For example, their parameter estimates were $b_0 = 5.02E - 06$, $b_1 = 0.58$, $b_2 = 1E - 06$ $\omega = 421$, and the market price of risk, λ was not significant.

Insert Tables 1a and 1b Here

The tables also report the stationary volatility, and the final local volatility that exists at the end of the period. The long run stationary volatility estimates produced by both the NGARCH model and the HN model appear to be very stable, not deviating too far from about 17% per year. Figure 1 shows the time series of these "local" volatilities for each year. The parameters for the process are estimated using the first six months of data, and then these values are used with the actual daily time series of the index to generate the local volatilities over the entire year. The time series produced by these two models over each of the 5 years are shown in the first two panels of Figure 1. In most years the time series of local volatility in the NGARCH model appears to be less volatile than in the HN model. The third panel in Figure 1 shows the time series of implied volatilities obtained by minimizing the sum of squared errors in option prices using the Black Scholes model for each of the 52 weeks in the year. This time series is the least volatile.

Insert Figure 1 Here

Since our primary goal is to investigate the performance of our two GARCH models in pricing options, we now turn attention to examining the residuals associated with our models.

In what follows we define moneyness as $M = (S_t - X)/S_t$, where X is the strike price. Deep-out-the money options are defined as $M < -0.04$; out-the-money contracts are defined as $-0.01 < M$ 0.04; at-the-money contracts have $-0.01 < M$ 0.01; in the money contracts have $0.01 < M$ 0.04, and deep in the money contracts have $M > 0.04$. Expiration dates are bucketed into 3 groups: near term contracts have maturities between 10 and 45 days; mid term contracts have maturities between 46 and 90 days; and long term contracts have maturities between 91 and 200 days. Residuals are computed as theoretical prices less actual prices. Unless stated otherwise, all contract prices have been normalized so that the underlying asset price is \$1.0.

All the residuals over the in-sample periods are first analyzed to assess whether the models are misspecified. Figure 2 shows box and whisker plots for residuals generated by the three models categorized by moneyness and maturity. If there were no systematic biases in the models, the residuals should be centered around zero, for each moneyness-maturity category.

Insert Figure 2 Here

The plots reveal large volatility skew and smile patterns associated with the Black Scholes model. The biases for this model are particularly large on average. In particular, on average, deep in-the-money contracts are priced too low and deep out-the-money contracts are priced too high with the bias increasing, in raw dollar terms, with maturity. That is, on average, at-the-money and out-the-money Black Scholes option prices are higher than actual prices while deep in-the-money contracts are priced too low.

The box and whisker plots clearly reveal that the two GARCH models remove a significant fraction of the strike price bias, for each maturity bucket. There are still patterns, on average in the Heston-Nandi model. Specifically, away from the money contracts are priced too low, while at-the money contracts are priced a bit too high. The NGARCH model appears to remove more of the strike price bias. Moreover the interquartile range of residuals appears to be tighter as does the 95% confidence intervals.

By normalizing the index price to be \$1.0 over the entire five years, the option price residuals have a very natural interpretation. An error of 0.01 for example, can be viewed as a error of one cent, or 1% of the underlying. However, perhaps a better criterion to assess the fit of the models is to examine the *percentage* error in pricing. Since out-the-money options have small actual prices, this criterion magnifies the ability of different models in explaining the prices of out-the-money contracts.

Figure 3 shows the plots of these percentage errors. Since the percentage errors of deep-in-the money contracts are so small relative to the other 4 moneyness buckets, the figure only shows the pattern over the remaining 4 categories.

Insert Figure 3 Here

The results are very revealing. First, note the huge biases in Black Scholes prices. For deep in, in, and at-the-money contracts, the interquartile range of theoretical prices is within 3 percentage points of actual prices. However, the bias in percentage errors increases as the contract moves out of the money. For example, on average, deep out-the-money options are mispriced by almost 50%. More than one in four contracts in this category were mispriced by at least 100%, and the 95% confidence interval extended to 200%. As Figure 3 illustrates, the bias holds true for all maturity buckets.

The two GARCH models perform much better, with the NGARCH model doing fairly well, even for the deep-out-the money contracts. For example, this model shows almost no moneyness bias for each of the three maturity buckets. Of course the interquartile range expands as we move out of the money, but this is to be expected, since the denominator is getting smaller. The HN model produces intermediate results. Our results here confirm other studies, such as those by Duan (1995) and Heston and Nandi (2000) that have shown that GARCH models are capable of explaining a significant portion of the volatility strike price bias.

We now investigate the out-of-sample performance on the models using models that were estimated over the first six months, and residuals generated from the last 6 months of each year. The pattern of these residuals in the out of sample period are similar to the in sample period for the first six months. For example, Figure 4 compares the out-of-sample box and whisker plots of the two GARCH models.

Insert Figure 4 Here

In comparing the residuals of the NGARCH and HN models, their magnitudes appear to be of the same order, although for out-the-money options there does appear to be more bias in the HN model. To establish which of the two GARCH models is better, we compare the out-of-sample predictions, contract by contract. We compute the absolute error for each contract produced by each model, and compute the difference between the two values. Figure 5 provides histograms of these differences by maturity and moneyness. Negative values indicate that the NGARCH model produced more precise values.

Insert Figure 5 Here

Table 2 summarizes the proportion of contracts for which the HN model outperformed the NGARCH model in the out of sample periods. For almost all categories the NGARCH model produces smaller residuals than the HN model. The most dramatic differences occur in the deep out the money contracts. For these contracts the HN model is unable to explain the volatility strike bias and systematically underperforms.

Insert Table 2 Here

We next investigate how the out of sample errors behave as the time since calibration increases. Specifically, it seems plausible that the conditional forecasts of option prices one week after the parameters are estimated might be relatively small compared to conditional option prices generated several months after the parameters are estimated. To address this issue we first grouped all the residuals produced by a model into moneyness and maturity buckets and then looked at the distribution of the percentage pricing errors as the time since estimation increased, from one week through 25 weeks. Figure 6 compares the distribution of residuals for 6 different out-of-sample periods. In particular, a box and whisker plot is provided for each month, from the first out of sample month to the last.

Figure 6*a* shows the results for the short maturity contracts for the HN model while Figure 6*b* shows the results for the NGARCH model. The average bias, as indicated by the average deviation from 0, seems to remain fairly steady as the time since estimation increases. In addition, the quartiles do not expand over time.

Insert Figure 6a and 6b Here

The figures indicate that the magnitude of the errors are not strongly related to the time since the estimation was conducted . That the bias of residuals does not appear to expand over the six month periods after the model was estimated appears to be surprising, especially since the parameter estimates for the GARCH models in successive years were not that similar. To look at this more closely, for each moneyness-maturity bucket, and for each year, we computed several statistics of the residuals as the time horizon expanded from one to twenty weeks. Table 3*a* summarizes the findings for contracts with less than 10 weeks to maturity, and for 3 moneyness categories for the HN model. Table 3*b* presents similar results for the NGARCH model.

Insert Table 3a and 3b Here

In particular, the table reports the average error, the average absolute error and the standard deviation of errors for each category. The first statistic gives a measure of bias; the second gives a measure of accuracy, while the third gives a measure of precision. For ease of presentation we only have presented these statistics for selected time periods, namely for $1, 2, 5, 10$ and 20 weeks after the parameters were estimated. Below each of the tables, there are three typical time series plots of the average absolute errors of prices over the last 6 months in the middle year, 1993, for the different moneyness categories.

The results confirm that the bias, accuracy, and precision generated from models calibrated using data from 6 months earlier are not that dissimilar from measures obtained by models that are calibrated with more recent data. This indicates that the GARCH models may be capturing important elements of the time series properties of asset and option prices.

As a final analysis we considered the NGARCH model estimated using our data set over the first 6 months of 1991. We computed the out of sample residuals for this model over the following four and a half years, without reestimating the parameters. Figure 7 shows the box and whisker plots of the percentage errors over each quarter since the model was estimated.

Insert Figure 7a and 7b Here

The figure shows the bias in the NGARCH model over time. There appears to be very little deterioration of the model over time. The figure only reports results for the near term contracts with less than 45 days to expiration; the pattern of the plots is very similar for the mid term and longer dated contracts. While there are some quarters where the bias increases, overall there is very little trend in the biases. Figure 7*b* contrasts the time series of percentage errors produced by the NGARCH model with the *in-sample* percentage errors produced by the Black Scholes model. The scales of the exhibits in the two figures are the same so as to facilitate easy comparisons. The enormous biases produced by the Black Scholes model are especially pronounced in the out-the-money contracts. For these contracts, the out-of-sample NGARCH model consistently produces better results, even more than four years after the NGARCH model is calibrated. For example, over 19 consecutive out of sample quarters, the mean absolute percentage error of out the money contracts never exceeded 50%, whereas the mean absolute percentage error of in-sample Black Scholes errors exceeded 50% on 18 of the 19 quarters. Indeed, in comparing figures 7*a* and 7*b* the out-of-sample performance of the NGARCH model appears to hold its own or dominate the in sample performance of the Black Scholes model, even after many years have passed since the parameters were calibrated. Moreover, unlike the Black Scholes model, where there is a persistent bias in the direction of the residuals, for the NGARCH model, the bias is generally smaller and tends to shift around 0.

IV Conclusion

This article has investigated the performance of two GARCH models, namely Duan's NGARCH model and the HN model. The NGARCH model is important in its own right and also serves as an approximation for particular stochastic volatility models generated by two orthogonal Wiener processes. The HN model is also interesting since its volatility updating structure permits analytical solutions to be generated for European options. This article has compared these two models relative to each other and relative to the Black Scholes model. The results indicate that both GARCH models are capable of explaining a significant amount of the maturity and strike price bias associated with the Black Scholes model. The NGARCH model appears to outperform the HN model, especially in its ability to price deep out-the-money contracts.

The out-of-sample performance of the GARCH models, and especially the NGARCH model is encouraging. The fact that models estimated using old option data are still capable of explaining option prices for significant time periods indicates that the underlying models are capturing important elements of the option pricing process. In extreme cases, where the NGARCH parameters are not reestimated for quarters, or even years, for in and at the money options the model continues to perform at levels usually no worse than in-sample Black Scholes, and, for out-the-money options, the NGARCH model continues to do significantly better. The cost of reestimating the parameters of a GARCH process is not that high. We therefore do not recommend using the model to price options based on parameter estimates that have been estimated over a distant time horizon. Our point here, is that if frequent updates of the model are not made, then the performance of the GARCH models is still adequate, especially relative to the performance of a Black Scholes model. In addition, the prolonged good performance of an NGARCH model indicates that it must be capturing essential elements that determine option prices.

Since American options and exotics can be efficiently priced using numerical procedures developed by Duan and Simonato (2000) and Ritchken and Trevor (1999), this article suggests that GARCH models, perhaps as a proxy for true stochastic volatility models, significantly improves upon the performance of the Black Scholes model, and, in light of the relative ease in pricing American claims under these processes, these models should be given closer scrutiny by the trading community.

References

Andersen T. and J. Lund, 1997, "Estimating Continuous time Stochastic Volatility Models of the Short Term Interest Rate," *Journal of Econometrics,* 77, 343-377

Bakshi, G., C. Cao and Z. Chen 1997, Empirical Performance of Alternative Option Pricing Models, *The Journal of Finance*, 53, 499-547.

Bates, D., 1996, "Jumps and Stochastic Volatility: Exchange Rate processes Implicit in Deutchemark Options," *Review of Financial Studies,* 9, 69-107.

Black, F. 1976, Studies in Stock Price Volatility Changes, *Proceedings of the 1976 Meeting of the Business and Economics Statistics Section*, American Statistical Association, 177–181.

Bollerslev, T, R. Chou and K. Kroner, 1992, "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52,1-59.

Corradi, V.,2000,"Reconsidering the Continuous Time Diffusion Limit of the GARCH(1,1) Process," *Journal of Econometrics*, 96, 145-153.

Duan, J. 1995, "The GARCH option pricing model," *Mathematical Finance* 5, 13-32.

Duan, J. 1996, A Unified Theory of Option Pricing under Stochastic Volatility - from GARCH to Diffusion, unpublished manuscript, Hong Kong University of Science and Technology.

Duan, J., 1997, "Augmented GARCH(p,q) Process and its Diffusion Limit," *Journal of Econometrics*, 79, 97-127.

Duan, J., G. Gauthier, and J. Simonato, 1999, "An Analytical Approximation for the GARCH Option Pricing Model," *Journal of Computational Finance*, 2, 75-116.

Duan J. and J.G. Simonato, 1998, "Empirical Martingale Simulation for Asset Prices", *Management Science*, 44, 1218-1233.

Duan J. and J.G. Simonato, 2000, "American Option Pricing under GARCH by a Markov Chain Approximation", *Journal of Economic Dynamics and Control* , forthcoming.

Dumas, B., J. Fleming and B. Whaley, 1998, Implied Volatility Functions: Empirical Tests, *Journal of Finance*, 53.

Engle, R. and V. Ng, 1993, Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48, 1749-1778.

Harvey, C. and R. Whaley, 1992, Dividends and *S&P* 100 Index Option Valuation, *Journal of Futures Markets*, 12, 123-137.

Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility," *Review of Financial Studies* 6, 327-344.

Heston, S., and S. Nandi, 2000, "A Closed Form GARCH Option Pricing Model, " *The Review of Financial Studies,* 13, 585-625.

Hull, J. and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatility, *Journal of Finance*, 42, 281-300.

Jackwerth and M. Rubinstein, 1996, Recovering probability Distributions from Option Prices, *Journal of Finance*, 51, 1611-1631.

Nandi, S., 1998, "How Important is the Correlation Between Returns and Volatility in a Stochastic Volatility Model? Empirical Evidence from Pricing and Hedging in the S&P 500 Index Option Market," *Journal of Banking and Finance,* 22, 589-610.

Nelson, D., 1990, "ARCH Models as Diffusion Approximations," *Journal of Econometrics,* 45, 7-38.

Ritchken, P. and R. Trevor, 1999, "Pricing Options under Generalized GARCH and Stochastic Volatility Processes," *Journal of Finance,* 54, 377- 402

Table 1a: Parameter Estimates for the HN Model

Time Period	No. of Weeks	No. of Contracts	SSE	b_0	b_1	b_2	w	Initial Volatility	Final Volatility	Stationary Volatility
1991 1st Half	25	953	1264.51	1.19E-06	8.73E-03	3.04E-06	555.57	0.4215	0.1741	0.16982
1991 2nd Half	25	961	390.56	2.24E-06	1.15E-01	3.13E-06	509.48	0.2676	0.1162	0.16463
1992 1st Half	25	1175	415.93	1.04E-06	3.10E-01	8.34E-07	888.10	0.2013	0.1435	0.14711
1992 2nd Half	25	1035	335.28	4.60E-07	5.63E-01	4.37E-07	984.03	0.1216	0.1046	0.15645
1993 1st Half	25	1492	941.14	3.01E-11	3.60E-01	3.12E-06	430.06	0.0741	0.1553	0.13386
1993 2nd Half	24	1466	702.71	9.66E-10	2.02E-01	2.55E-06	541.71	0.1167	0.1182	0.13564
1994 1st Half	25	1419	1727.78	8.25E-07	3.18E-02	1.42E-06	806.96	0.1301	0.1870	0.13709
1994 2nd Half	25	1598	1089.49	7.91E-07	1.90E-09	7.15E-07	1166.70	0.1306	0.1400	0.14486
1995 1st Half	25	1774	996.58	3.67E-07	1.42E-01	1.27E-06	805.13	0.1541	0.1250	0.13493
1995 2nd Half	25	1702	1135.83	9.00E-07	1.39E-01	1.27E-06	803.30	0.1391	0.1157	0.13618

Table 1a shows the parameter estimates for the HN model for each of ten periods. Over each time period the sum of squared error was computed as discussed in the text and is reported in the fourth column. All volatility values are reported in an annualized basis.

Table 1b: Parameter Estimates for the NGARCH Model

Time Period	No. of Weeks	No. of Contracts	SSE	b_0	b_1	b_2	w	Initial Volatility	Final Volatility	Stationary Volatility
1991 1st Half	25	953	582.2	6.77E-07	0.8576	0.0172	2.6128	0.2766	0.1615	0.1747
1991 2nd Half	25	961	378.02	1.45E-06	0.6807	0.0240	3.4134	0.1897	0.0749	0.1858
1992 1st Half	25	1175	404.78	2.23E-06	0.2660	0.0192	5.9645	0.2181	0.1349	0.1580
1992 2nd Half	25	1035	323.73	7.42E-07	0.6589	0.0108	5.4495	0.1189	0.1060	0.1761
1993 1st Half	25	1492	622.32	1.88E-06	0.6319	0.1651	1.0424	0.0792	0.1450	0.1707
1993 2nd Half	24	1466	580.23	3.71E-06	0.2627	0.2359	1.3973	0.1072	0.1246	0.1822
1994 1st Half	25	1419	1189.73	2.28E-06	0.2618	0.0636	3.1907	0.0177	0.1940	0.1743
1994 2nd Half	25	1598	860.07	1.29E-06	0.2599	0.0206	5.8437	0.0905	0.1269	0.1753
1995 1st Half	25	1774	662.78	1.35E-06	0.6384	0.0905	1.6692	0.1750	0.1086	0.1615
1995 2nd Half	25	1702	852.35	1.69E-06	0.6017	0.0455	2.6789	0.0073	0.0965	0.1520

Table 1b shows the parameter estimates for the NGARCH model for each of ten periods. Over each time period the sum of squared error was computed as discussed in the text, and is reported in the fourth column. All volatility values are reported in an annualized basis.

Table 2: Comparison of HN and NGARCH Prediction Errors in the Out-of-Sample Periods*

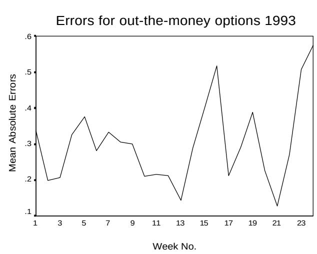
Maturity	Moneyness: deep in		in		at		out		deep out	
Short	746	245	373	154	254	90	285	77	52	7
		0.3284		0.4129		0.3543		0.2702		0.1346
Middle	1060	443	422	178	289	122	383	150	287	68
		0.4179		0.4218		0.4221		0.3916		0.2369
Long	1338	535	375	183	229	110	306	136	363	137
		0.3999		0.488*		0.4803*		0.4444*		0.3774

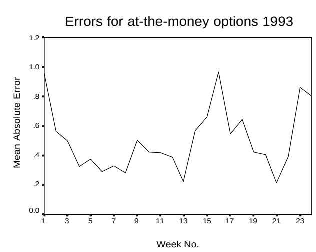
*The table compares the absolute residual of the NGARCH model with the absolute error of the HN model. In each cell there are three entries. The first entry is the number of contracts in that category. The second entry is the number of times the absolute value of the residual for the HN model is lower than the absolute value of the NGARCH model. The final value is the proportion of times the HN model beat the NGARCH model. For example, consider the short term, deep in the money contracts. In 746 out-of-sample predictions, the HN model gave a smaller absolute error, than the NGARCH model. That is, the HN model won 32.84% of the time.

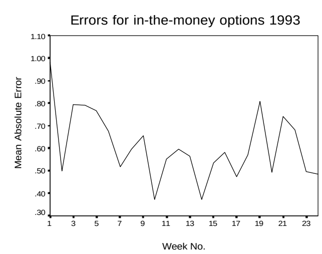
At the 5% level of significance, the NGARCH model outperforms the HN model for all short term, and middle term contracts, and for all deep out the money contracts. For long term contracts, near-the-money, and indicated with a * above the fraction of wins, the NGARCH model wins more times, but the differences are not statistically significant, at the 5% level

Table 3a
Out of Sample Performance Over Time for HN Model

Year	Week No.	Out of the Money Option: # Contracts	Out of the Money Option: Average Error($)	Out of the Money Option: Average Absolute Error ($)	Out of the Money Option: Standard Deviation ($)	At the Money Option: # Contracts	At the Money Option: Average Error	At the Money Option: Average Absolute Error ($)	At the Money Option: Standard Deviation ($)	In the Money Option: # Contracts	In the Money Option: Average Error($)	In the Money Option: Average Absolute Error ($)	In the Money Option: Standard Deviation ($)
	1	39	-1.063	1.063	0.405	70	-0.751	0.874	0.651	56	-0.146	0.402	0.480
	2	36	-1.364	1.364	0.273	57	-1.220	1.220	0.343	58	-0.822	0.822	0.386
1991	5	21	-1.096	1.096	0.338	43	-0.950	0.950	0.341	55	-0.280	0.318	0.296
	10	23	-0.317	0.393	0.298	42	0.163	0.533	0.620	40	0.445	0.506	0.422
	20	38	1.151	1.157	0.690	55	1.659	1.663	0.924	37	1.854	1.854	0.856
	1	19	-0.182	0.267	0.251	47	0.223	0.398	0.438	41	0.021	0.354	0.425
	2	35	-0.230	0.308	0.251	62	-0.075	0.299	0.350	76	-0.158	0.462	0.565
1992	5	26	-0.741	0.741	0.202	52	-0.880	0.880	0.373	52	-0.640	0.640	0.335
	10	6	0.048	0.255	0.286	21	0.628	0.657	0.334	13	0.057	0.242	0.306
	20	24	-0.446	0.446	0.179	54	-0.153	0.292	0.305	64	-0.360	0.433	0.393
	1	30	0.302	0.338	0.312	63	0.944	0.958	0.586	68	0.976	0.982	0.480
	2	28	0.124	0.199	0.222	49	0.534	0.566	0.448	77	0.105	0.499	0.596
1993	5	31	-0.357	0.376	0.217	72	-0.210	0.377	0.367	105	-0.574	0.768	0.876
	10	24	-0.064	0.210	0.250	47	0.304	0.423	0.442	68	0.146	0.373	0.460
	20	24	-0.005	0.226	0.268	58	0.303	0.405	0.399	82	-0.302	0.493	0.646
	1	29	0.282	0.346	0.390	62	0.829	0.859	0.657	83	0.463	0.931	0.938
	2	28	0.248	0.331	0.343	52	0.747	0.767	0.561	66	1.084	1.092	0.437
1994	5	8	-0.199	0.291	0.244	36	0.238	0.470	0.515	33	0.457	0.576	0.536
	10	29	0.437	0.493	0.443	61	1.032	1.055	0.637	60	1.177	1.189	0.463
	20	39	-0.531	0.551	0.237	65	-0.308	0.469	0.422	66	-0.060	0.348	0.420
	1	27	-1.403	1.403	0.682	77	-0.680	1.111	1.080	144	-0.416	0.654	0.771
	2	38	-1.674	1.674	0.650	76	-1.288	1.339	0.859	151	-0.108	0.562	0.637
1995	5	25	-1.112	1.112	0.392	63	-0.313	0.768	0.865	60	0.624	0.625	0.459
	10	34	-1.411	1.411	0.415	79	-1.142	1.142	0.490	117	-0.321	0.381	0.295
	15	5	-0.269	0.557	0.572	52	0.086	0.616	0.722	189	-0.361	0.563	0.537



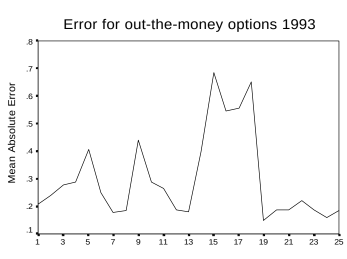


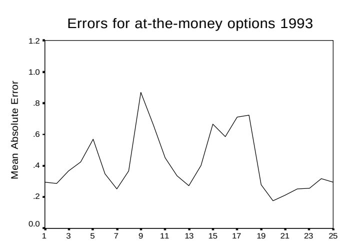


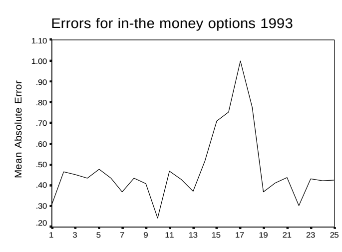
This table reports statistics on the errors in pricing as the number of weeks from the estimation period increases. The errors are actual dollar errors in prices of contracts. The results are presented for contracts with less than 70 days to maturity, and for three moneyness buckets. The number of contracts in each bucket are reported as well as the average error, average absolute error, and std. deviation of errors. The fact that the HN model does not deteriorate significantly over time indicates that the model may be capturing important aspects of the true dynamics.

Table 3b
Out of Sample Performance Over Time for NGARCH Model

Year	Week No.	Out of the Money Option # Contracts	Average Error ($)	Average Absolute Error ($)	Standard Deviation ($)	At the Money Option # Contracts	Average Error ($)	Average Absolute Error ($)	Standard Deviation ($)	In the Money Option # Contracts	Average Error ($)	Average Absolute Error ($)	Standard Deviation ($)
	1	39	-0.349	0.359	0.234	70	-0.408	0.449	0.379	56	-0.707	0.710	0.382
	2	36	-0.336	0.340	0.257	57	-0.473	0.478	0.383	58	-1.253	1.253	0.445
1991	5	21	-0.112	0.156	0.153	43	-0.152	0.208	0.226	55	-0.582	0.585	0.242
	10	23	0.148	0.181	0.150	42	0.170	0.214	0.210	40	-0.468	0.475	0.289
	20	38	0.051	0.261	0.295	55	-0.104	0.320	0.387	37	-0.694	0.779	0.573
	1	19	-0.135	0.208	0.200	47	0.114	0.295	0.348	41	0.005	0.310	0.374
	2	35	-0.090	0.240	0.266	62	-0.058	0.288	0.330	76	-0.138	0.465	0.573
1992	5	26	-0.406	0.406	0.182	52	-0.542	0.568	0.453	52	-0.456	0.477	0.424
	10	6	0.271	0.287	0.248	21	0.658	0.663	0.291	13	0.100	0.242	0.316
	20	24	-0.171	0.188	0.123	54	0.004	0.175	0.209	64	-0.265	0.412	0.417
	1	30	0.236	0.267	0.351	63	0.509	0.526	0.535	68	0.786	0.786	0.361
1993	2	28	0.753	0.754	0.549	49	1.163	1.163	0.771	77	0.554	0.849	0.835
	5	31	-0.157	0.194	0.167	72	-0.208	0.288	0.267	105	-0.584	0.741	0.807
	10	24	0.082	0.183	0.206	47	0.215	0.332	0.383	68	0.053	0.327	0.398
	20	24	-0.321	0.323	0.268	58	-0.512	0.513	0.381	82	-0.765	0.773	0.532
	1	29	0.112	0.197	0.283	62	0.501	0.546	0.568	83	0.553	0.882	0.826
1994	2	28	0.117	0.194	0.218	52	0.406	0.432	0.396	66	1.025	1.026	0.419
	5	8	0.100	0.118	0.123	36	0.335	0.364	0.345	33	0.571	0.626	0.463
	10	29	0.496	0.497	0.308	61	0.889	0.890	0.451	60	1.242	1.261	0.397
	20	39	-0.630	0.630	0.201	65	-0.646	0.693	0.360	66	-0.081	0.428	0.544
	1	27	-1.012	1.012	0.454	77	-0.594	0.834	0.730	144	-0.350	0.598	0.711
1995	2	38	-1.342	1.342	0.403	76	-1.162	1.162	0.469	151	-0.136	0.435	0.484
	3	38	-0.448	0.687	0.715	72	0.301	1.047	1.184	110	0.865	0.895	0.734
	5	25	-0.957	0.957	0.331	63	-0.583	0.706	0.570	60	0.389	0.419	0.379
	10	34	-0.819	0.819	0.207	79	-0.736	0.736	0.262	117	-0.253	0.338	0.289
	15	5	-1.308	1.308	0.757	52	-1.466	1.468	0.768	189	-0.788	0.813	0.487







This table reports statistics on the errors in pricing as the number of weeks from the estimation period increases. The errors are actual dollar errors in prices of contracts. The results are presented for contracts with less than 70 days to maturity, and for three moneyness buckets. The number of contracts in each bucket are reported as well as the average error, average absolute error, and std. deviation of errors. The fact that the NGARCH model does not deteriorate significantly over time indicates that the model may be capturing important aspects of the true dynamics. The graphs show the weekly evolution of the average absolute error over the out of sample weeks.

Figure 1
Time Series of Local Volatility

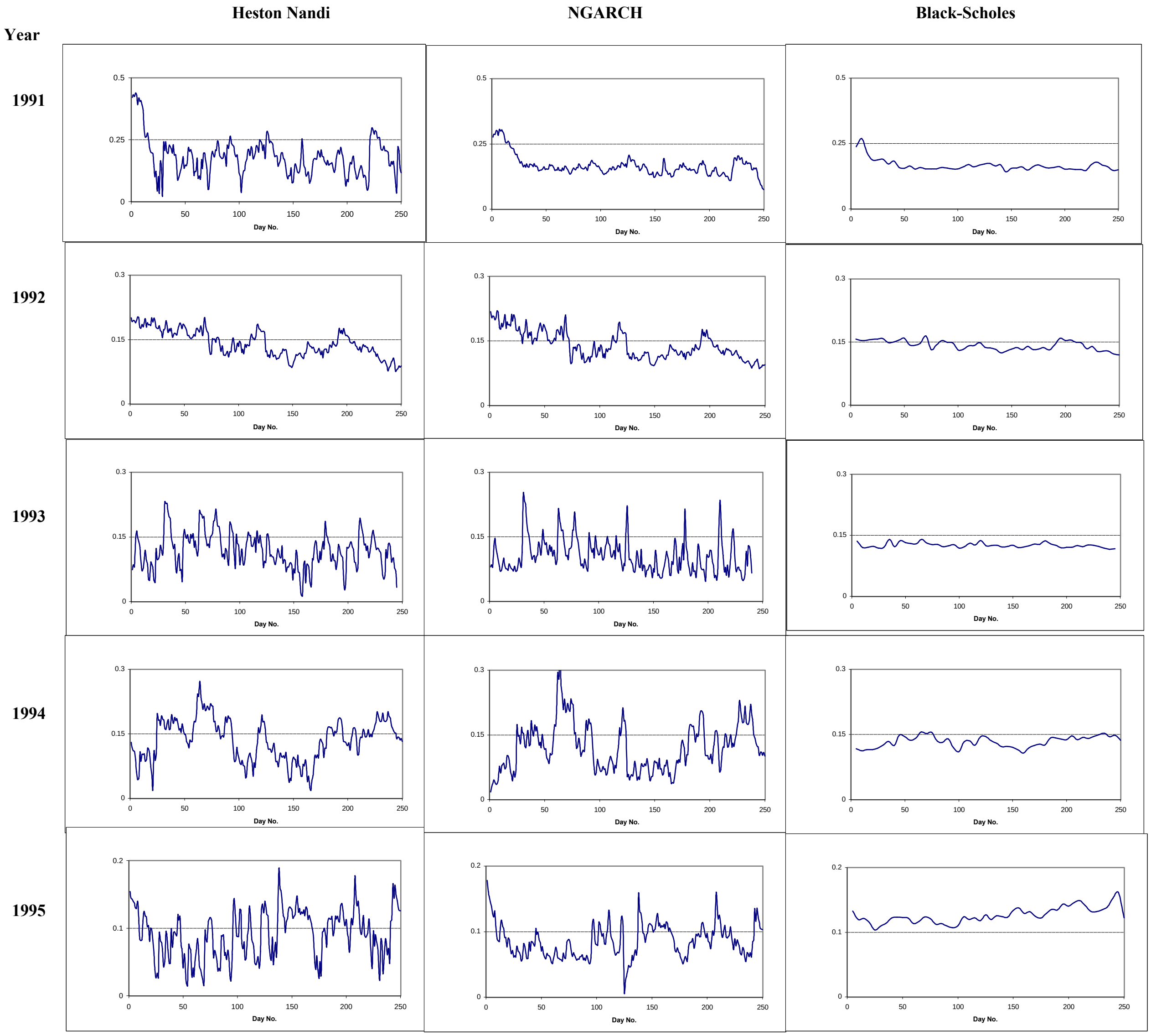


Figure 1 shows the time series of local volatilities for each year. For the first two models the parameters are estimated using the first six months of data. The time series of local volatilities is then updated daily based on the time series of the underlying index. For the Black Scholes model the implied volatility is extracted weekly using the option data. The implied volatility is estimated using all option contracts, with the criterion being minimizing the sum of squared errors.

Figure 2

Box and Whisker Plots of Residuals vs. Moneyness

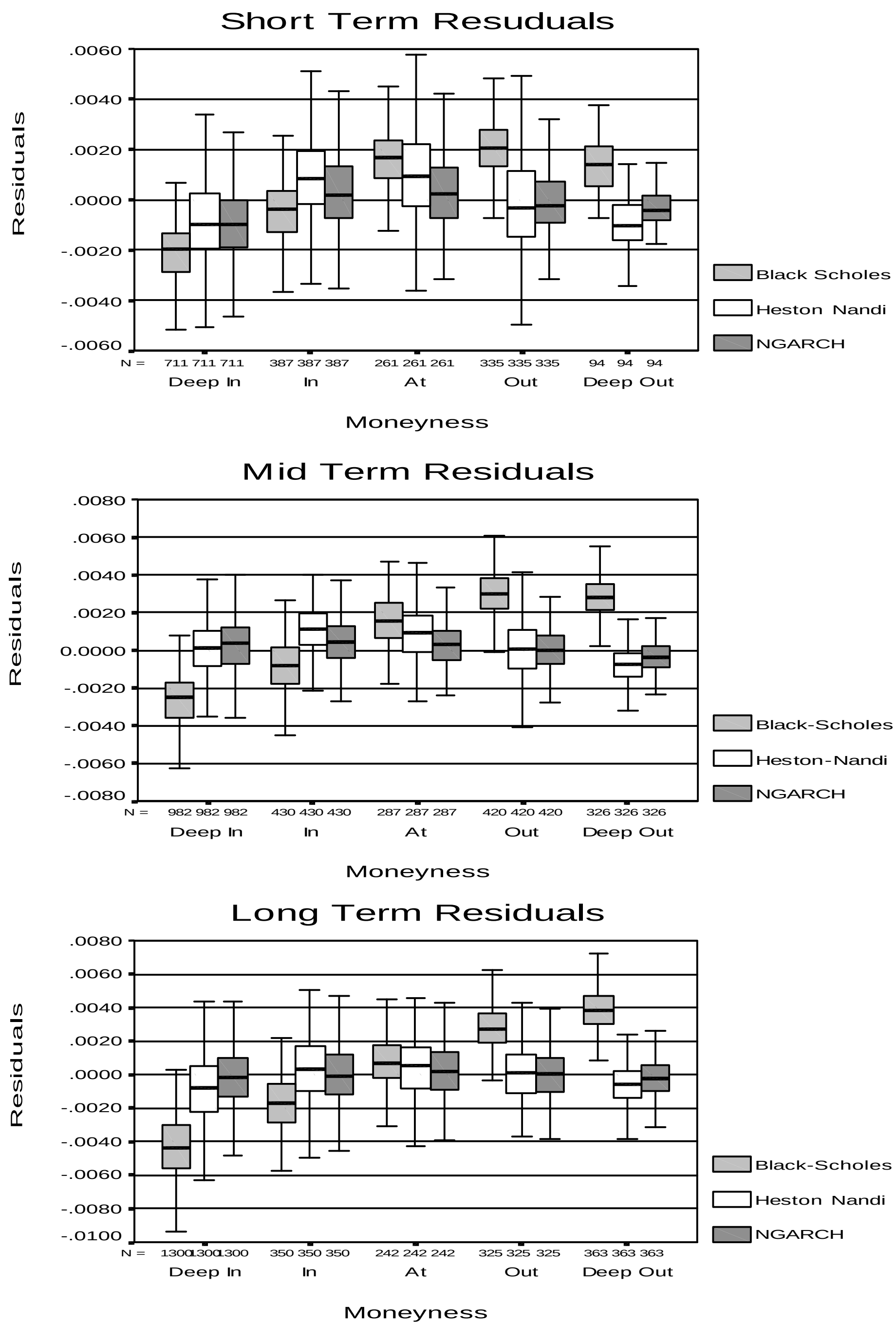


Figure 2 shows the distribution of residuals, in the form of box and whisker plots, for each moneyness-maturity bucket, for the three models. The underlying equity price is normalized to $1.0.

Figure 3
Box and Whisker Plots of Percentage Errors vs. Moneyness

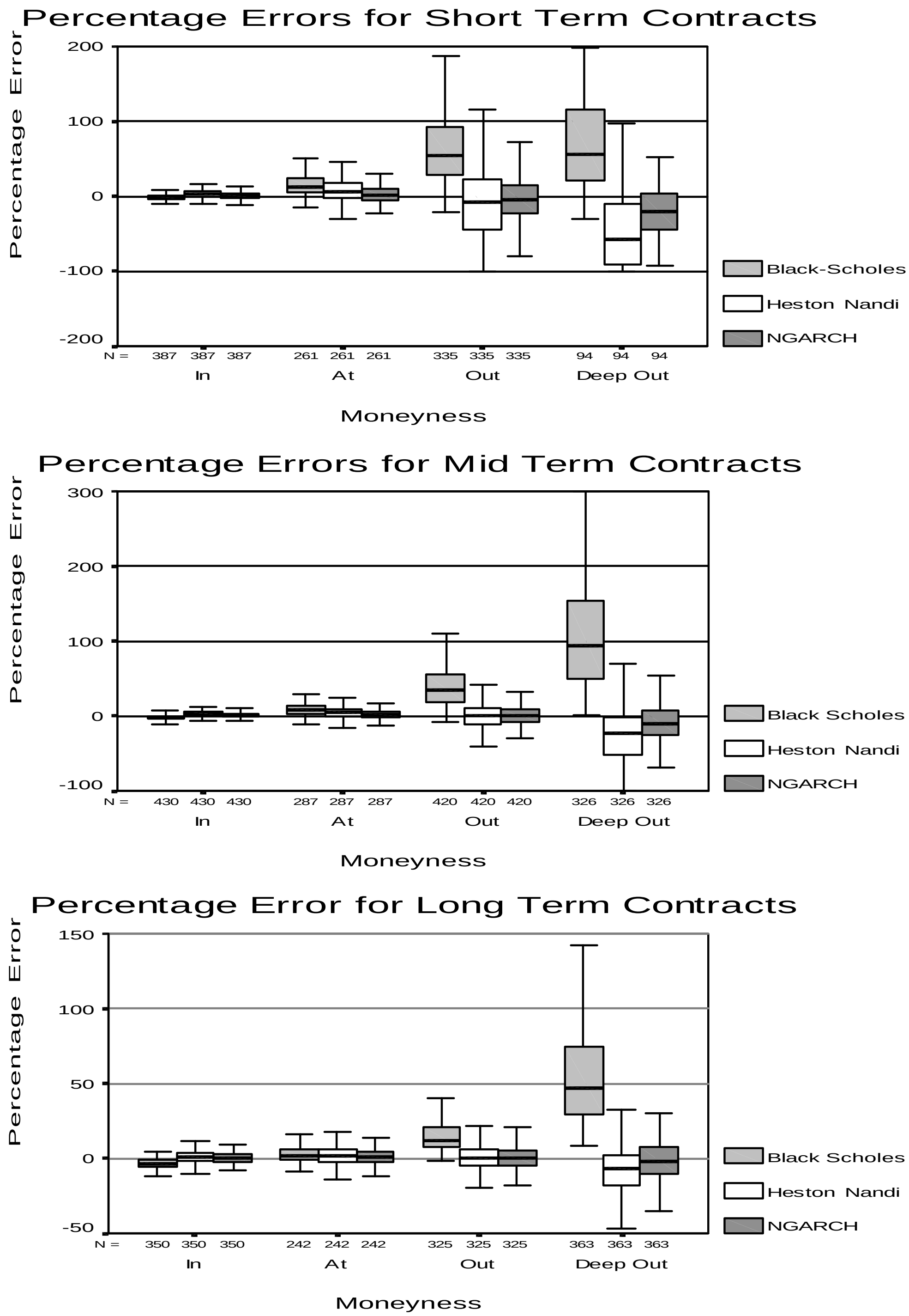
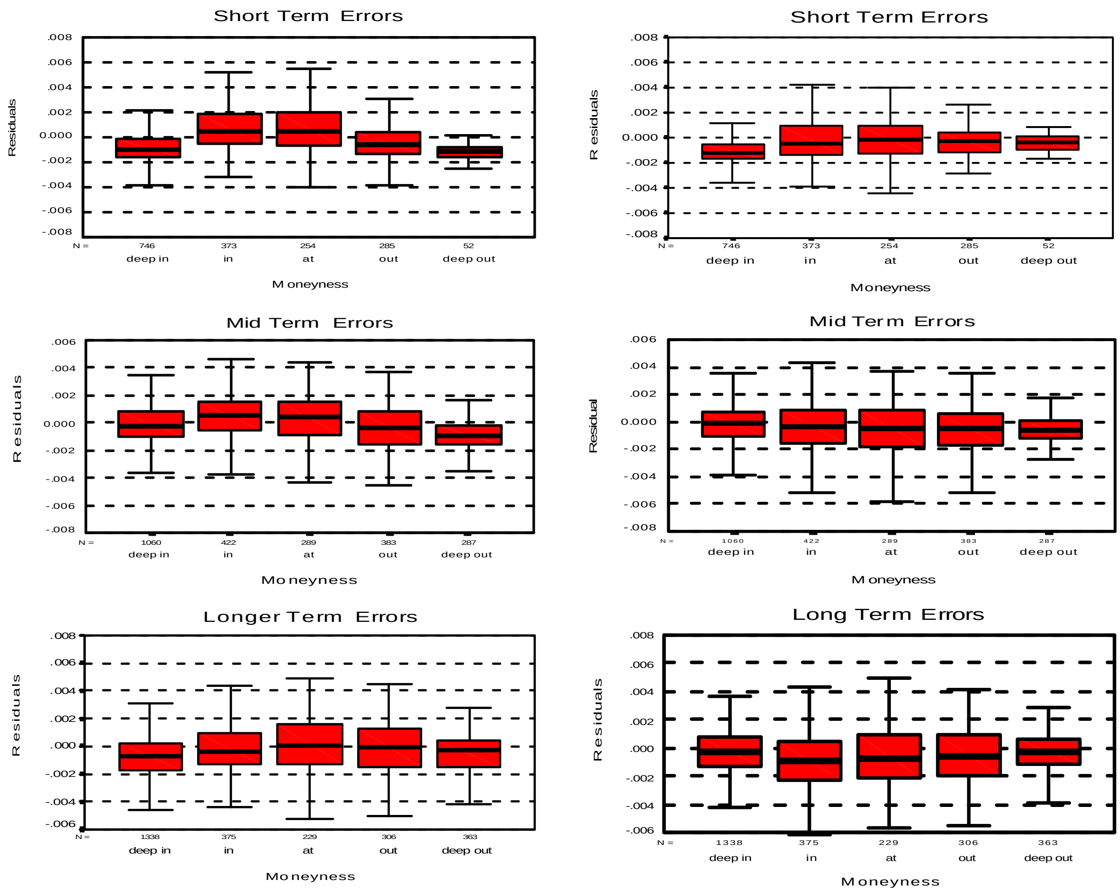


Figure 3 shows the distribution of percentage errors, in the form of box-whisker plots for each moneyness-maturity bucket, for the three models.

Figure 4
Out of Sample Box and Whisker Plots*

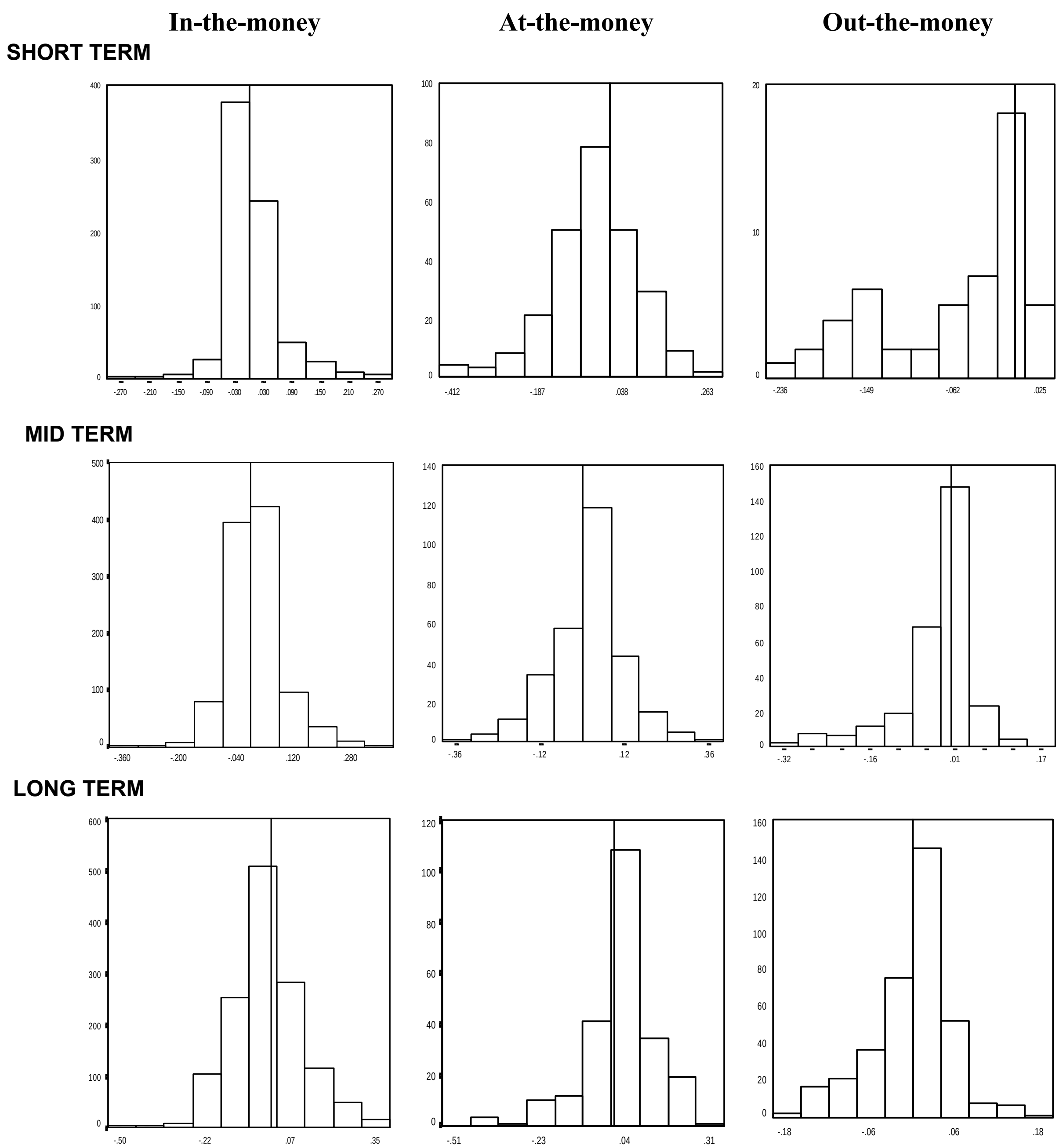
HN MODEL **NGARCH MODEL**



* Figure 4 compares the out of sample residuals for the two GARCH models by moneyness and by maturity.

Figure 5

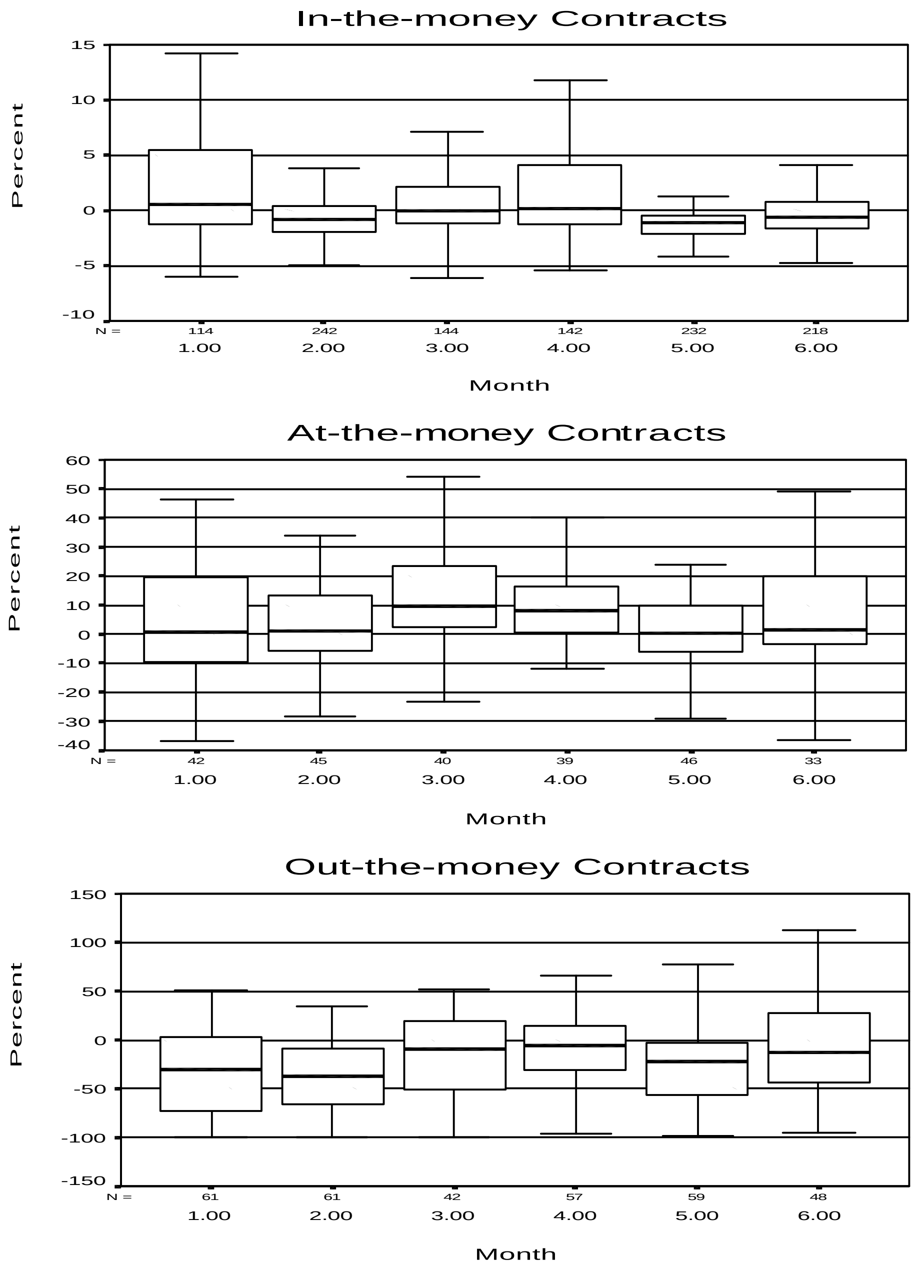
Difference in Absolute Value of NGARCH Residuals and HN Residuals



Each histogram has a vertical line placed at zero. Negative values indicate that the NGARCH model produces more precise values; positive values indicate that the HN model produces more precise values. The in-the-money contracts include deep in the money; the out-the-money contracts include deep out the money.

Figure 6a

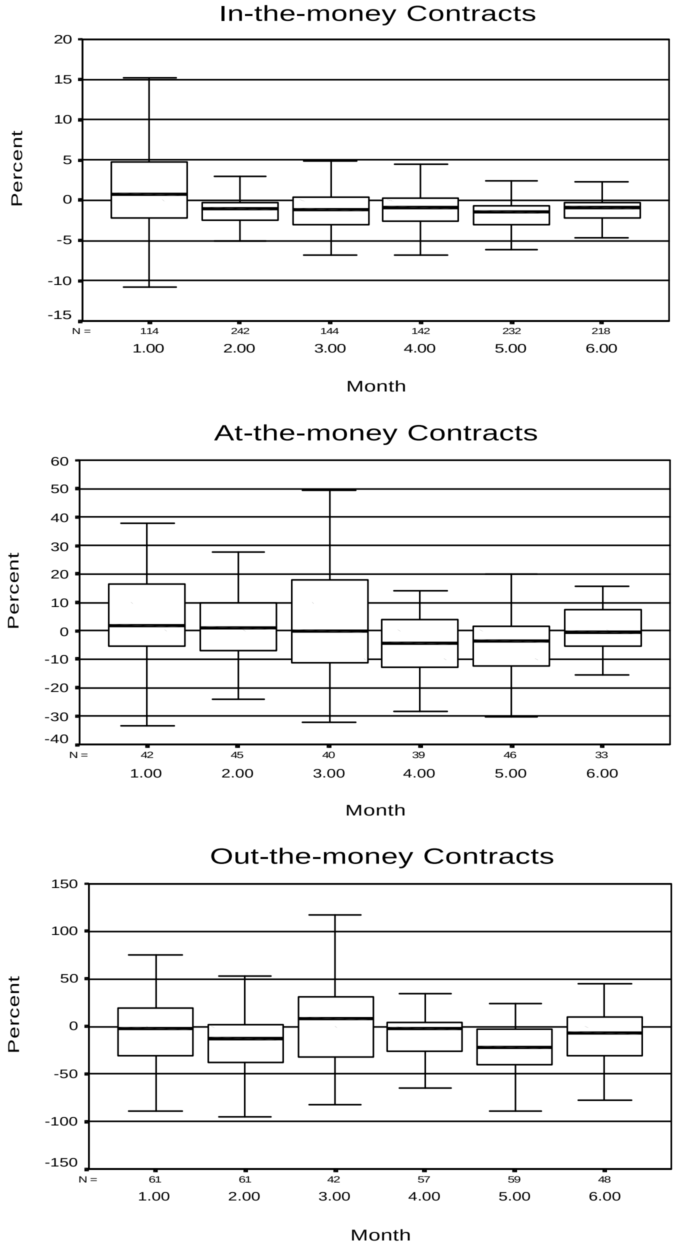
Out of Sample Percentage Errors by Moneyness for the HN Model



The figure shows the out of sample performance of the HN model for the short term contracts. The parameters are estimated using the first 6 months of data in each year. The residuals are established weekly for the last 6 months of each year. The figure shows the distributions, in the form of box and whisker plots of the percentage errors for adjacent months. For example, for short term at the money contracts, the second plot summarizes the 45 contracts that were in the 2 month category, which actually extends from 4 weeks to under 8 weeks after the parameters were estimated. The in the money contracts include the deep in the money contracts, and the out the money contracts include the deep out the money contracts.

Figure 6b

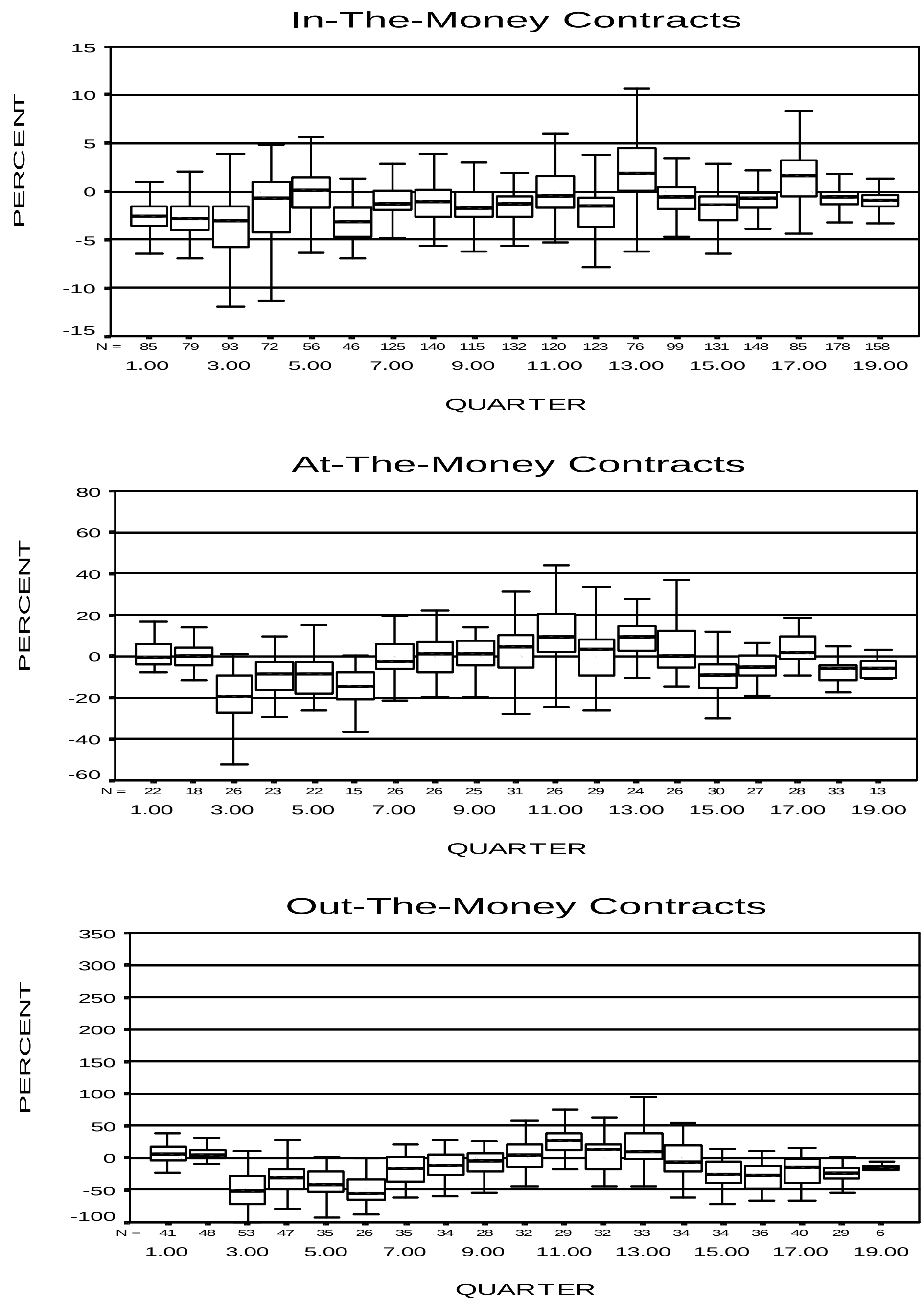
Out of Sample Percentage Errors by Moneyness for the NGARCH Model



The figure shows the out of sample performance of the NGARCH model for the short term contracts. The parameters are estimated using the first 6 months of data in each year. The residuals are established weekly for the last 6 months of each year. The figure shows the distributions, in the form of box and whisker plots of the percentage errors for adjacent months. For example, for short term at the money contracts, the second plot summarizes the 45 contracts that were in the 2 month category, which actually extends from 4 weeks to under 8 weeks after the parameters were estimated. The in the money contracts include the deep in the money contracts, and the out the money contracts include the deep out the money contracts.

Figure 7a

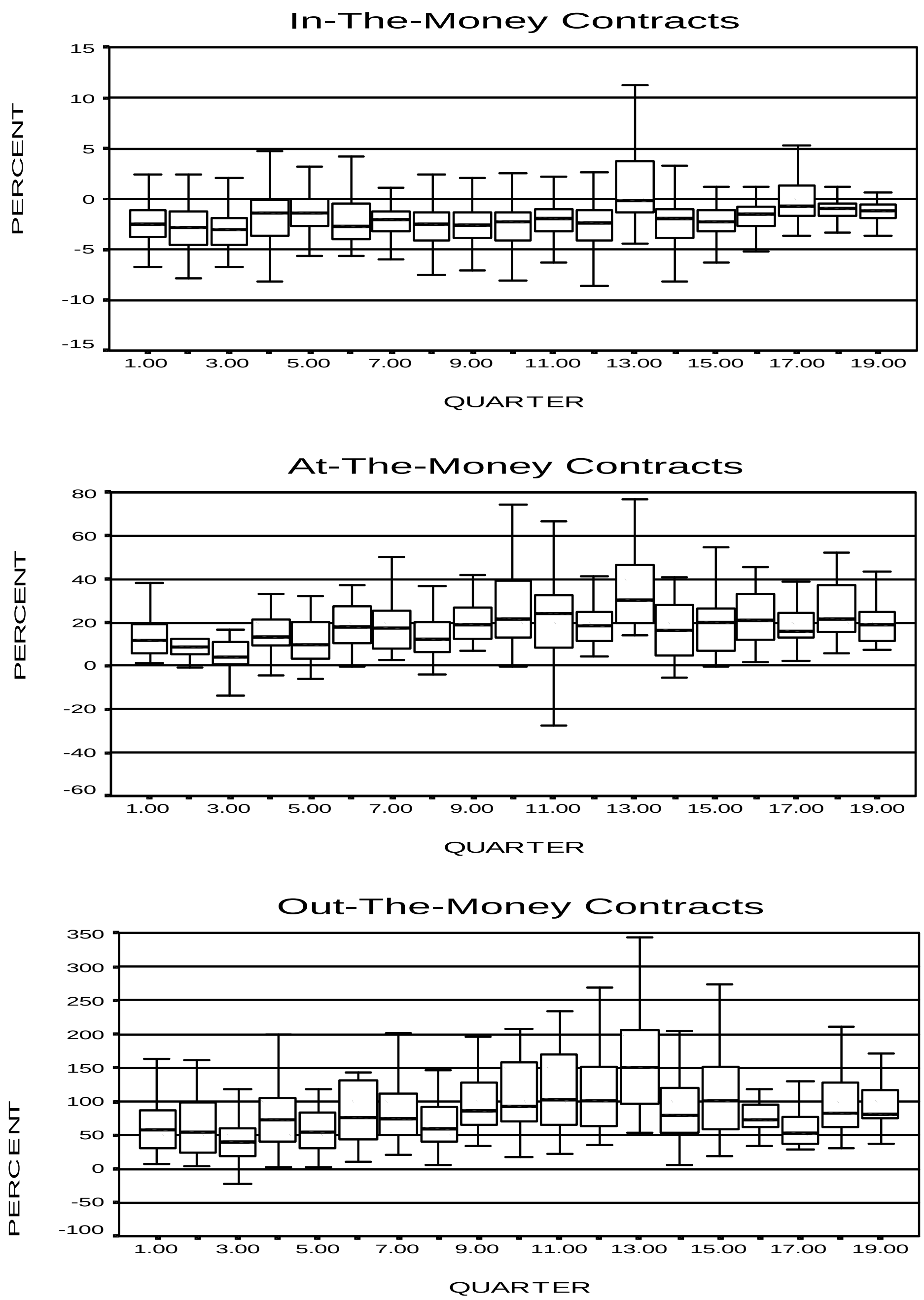
Time Series of Out-of-Sample Percentage Errors by Moneyness for NGARCH Short Term Contracts



Each exhibit shows a time series of box and whisker plots of quarterly NGARCH percentage errors for a particular moneyness category. In (out) the money contracts include deep in (out) the money contracts. All three exhibits are for short term maturities. The parameters for the model were estimated over the first six months of 1991. The residualsare from June 1991 through December 1995.

Figure 7b

Time Series of In-Sample Percentage Errors by Moneyness for Black Scholes Short Term Contracts



Each exhibit shows a time series of box and whisker plots of quarterly Black Scholes percentage errors for a particular moneyness category. In (out) the money contracts include deep in (out) the money contracts. All three exhibits are for short term maturities.