Which Volatility Model for Option Valuation?

Peter Christoffersen and Kris Jacobs*
McGill University and CIRANO

April 2, 2002

Abstract

Characterizing asset return dynamics using volatility models is an important part of empirical finance. The existing literature favors some rather complex volatility specifications whose relative performance is usually assessed through their likelihood based on a time-series of asset returns. This paper compares a range of volatility models along a different dimension, using option prices and returns under the risk-neutral as well as the physical probability measure. We judge the relative performance of various models by evaluating an objective function based on option prices. In contrast with returns-based inference, we find that our option-based objective function favors a relatively parsimonious model. Specifically, when evaluated out-of-sample, our analysis favors a model that besides volatility clustering only allows for a standard leverage effect. This empirical analysis is part of a growing literature suggesting that discrete-time option pricing with time-varying volatility is practical and insightful.

JEL Classification: G12

Keywords: option pricing, GARCH, risk-neutral pricing, parsimony, forecasting, out-of-sample.

*We would like to thank Jerome Detemple, Jin Duan, Rene Garcia, Steve Heston, Nour Meddahi, and Eric Renault for helpful comments. John Kozlowski, Ines Chaieb and Vadim di Pietro provided expert research assistance. We are grateful to FCAR, IFM^2 and SSHRC for financial support. Any remaining inadequacies are ours alone. Correspondence to: Kris Jacobs, Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montreal, Canada H3A 1G5; Tel: (514) 398-4025; Fax: (514) 398-3876; E-mail: jacobs@management.mcgill.ca.

1 Introduction

The poor empirical performance of the classic Black-Scholes (1973) option-valuation model is well documented. One of the main reasons for this poor empirical performance is the assumption of constant volatility. The problematic nature of this assumption is evidenced by the well-known smile and smirk patterns of implied volatilities, which indicate that market prices can only be reconciled with volatilities that vary across maturities and exercise prices. While it is entirely possible that these patterns could disappear by modifying some other building blocks of the model, it is not surprising that researchers have attempted to modify the constant volatility assumption. For instance, Dupire (1994) and Derman and Kani (1994) have proposed deterministic volatility models whose performance in turn has been questioned by Dumas, Fleming and Whaley (1998). Dumas et al. found that option prices from a simple OLS regression of implied volatility on a polynomial in strike price and maturity (the so-called Practitioner's Black-Scholes) outperformed models whose diffusion term is a deterministic function of the strike price and maturity.

A more structural approach to improving the empirical performance is to model volatility as stochastically time-varying and to explicitly derive option prices under that maintained assumption. Interestingly, this literature has developed along two related but somewhat different avenues. The early literature on option valuation with time-varying volatility is largely specified in continuous time (Hull and White (1987), Johnson and Shanno (1987), Melino and Turnbull (1990), Scott (1987), and Wiggins (1987)). The use of continuous time models has certain advantages. Importantly, for certain continuous-time option valuation models with time-varying volatility, it is possible to find analytical solutions for option prices (see Heston (1993)). The disadvantage of continuous-time models is that estimating the model parameters is not necessarily straightforward as volatility is modelled as an unobserved factor, and as real-world data is of course recorded at discrete intervals. Nevertheless, econometricians have made great strides in this domain, and the literature on the estimation of continuous-time option valuation models has exploded in recent years.[1]

An alternative and intuitive approach is to directly specify models of time-varying volatility at a discrete frequency matching that of the observed data. Following the work of Engle (1982) and Bollerslev (1986), a voluminous econometric literature has developed on volatility estimation and forecasting using discrete time GARCH processes (see the overview in Bollerslev, Chou and Kroner (1992)).[2] Duan (1995) characterizes the transition between the

[1] For overviews on econometric techniques to estimate continuous time models, see Campbell, Lo and MacKinlay (1997), Ghysels, Harvey and Renault (1996), Melino (1994), Renault (1997) and Tauchen (1997). For empirical applications of these to options prices see Andersen, Benzoni, and Lund (1999), Bakshi, Cao and Chen (1997), Benzoni (1998), Chernov and Ghysels (2000) , Bates (1996a), Eraker (2000), Jiang (1998), Jones (2000) and Pan (2000).

[2] Some important papers that estimate GARCH models for asset returns dynamics are Bollerslev, Engle and Wooldridge (1988), French, Schwert and Stambaugh (1987), Glosten, Jagannathan and Runkle (1993),

physical and risk-neutral probability distributions if the dynamic of the primitive security is given by a GARCH process, and thus establishes the foundation for the valuation of European options. Garcia and Renault (1998) discuss hedging in GARCH models. Ritchken and Trevor (1999) in turn construct trinomial trees to price American options under GARCH. Several papers have investigated aspects of the empirical performance of these models. Amin and Ng (1993) estimate GARCH parameters under the physical distribution and investigate the option valuation implications for individual stocks. Engle and Mustafa (1992) back out implied volatility paths from option prices and estimate GARCH model parameters to match these paths, and Duan (1996) estimates GARCH parameters under the risk neutral distribution using a single cross-section of option prices and investigates the potential of this model to explain the smile. In a recent paper, Heston and Nandi (2000) formulate a particular GARCH specification that yields an analytical solution and provide an empirical analysis of the model. They convincingly demonstrate that the inclusion of a leverage effect as well as volatility clustering are of great importance in improving valuation performance.[3]

While these papers make important contributions, it is clear that many issues remain unexplored. Most glaringly, the literature that estimates GARCH processes using time series on asset returns contains a wealth of evidence on specifications that generalize the fairly simple GARCH processes that are used in the option valuation literature. It seems natural that the large amount of work researchers have put into modeling the volatility of foreign exchange, bond, equity and index returns, should also be relevant for fitting the prices of options written on these underlying assets. Curiously however, no paper so far has assessed the relative performance of different GARCH models in valuing options across strike prices and times-to-maturity. This paper intends to fill this gap. We apply the GARCH specifications of Ding, Granger and Engle (1993) and Hentschel (1995). These specifications summarize the differences between different GARCH models in terms of differences in the news impact function (Pagan and Schwert (1990) and Engle and Ng (1993)). Through judicious choice of the specification of the news impact function, these models manage to nest a number of existing GARCH specifications. Adopting these specifications therefore allows us to investigate the pricing performance of a number of existing models while keeping the empirical analysis manageable.

We evaluate the valuation performance of the GARCH specifications both in- and out-of-sample, and we estimate model parameters under the physical as well as the risk-neutral probability distributions. In doing this, we attempt to demonstrate that option valuation in this framework is feasible. Unfortunately, except for in a limited class of processes (Heston and Nandi (2000)), analytical solutions are not available and we have to resort to numerical solutions. However, advances in computational power have rendered numerical techniques

Pagan and Schwert (1990) and Schwert (1989).

[3]Several studies have also studied the relationship between volatility implied in option prices and the conditional volatility in GARCH models estimated from returns. See Day and Lewis (1992), Jorion (1995) and Lamoureux and Lastrapes (1993).

much more feasible, and furthermore, a defining benefit of these discrete time models is that one-period ahead conditional volatility is deterministic. This constitutes a great advantage over continuous time models where volatility has to be extracted after the parameters of a given model have been estimated.

Our results are surprising. While most of the GARCH literature that investigates returns data using maximum likelihood favors relatively complex models, we find that for the purpose of option valuation, one should not look beyond a simple GARCH model that allows for volatility clustering and a leverage effect. While the less parsimonious models we investigate achieve a better fit when estimating using returns data and when evaluating option prices in-sample, they perform worse than the parsimonious model out-of-sample. Moreover, the improvement in in-sample fit is small in some cases, even though the extra parameters are usually significantly estimated.

Given that the specifications we investigate nest a number of GARCH models, our results have potentially serious implications. Because much of the literature on complex GARCH models does not consider economically motivated objective functions, its value could be put into question if other objective functions yield similar results. Moreover, there are equally serious consequences for the continuous-time stochastic volatility literature. While the empirical research on discrete and continuous-time models has developed independently, we know that most continuous-time models can be formulated as the limit cases of discrete-time models (Nelson and Foster (1994), Duan (1997), Heston and Nandi (2000)). The implication of our results for the continuous time stochastic volatility literature is therefore that there may be little to gain from moving beyond the simple volatility specification in Heston (1993), which allows for volatility clustering and a leverage effect, and which offers the distinct advantage of an analytical solution.[4]

The remainder of the paper is organized as follows. In Section 2 we discuss the estimation methodology under the physical as well as the risk-neutral measure. Section 3 introduces the data set, Section 4 discusses the empirical results, and Section 5 concludes and points out directions for future work.

2 Methodology

In this section we discuss the different methodologies used to estimate parameters for the option valuation models. We first outline the methodology used to estimate the parameters of the volatility process under the physical probability measure. This methodology exclusively

[4]Even if there may be little to gain from changing the volatility dynamic in continuous-time stochastic volatility models, it may prove worthwhile to change other building blocks of these models, such as the specification of the jumps in prices and volatility (see Andersen, Benzoni and Lund (1999), Bakshi, Cao and Chen (1997), Bates (1996a, 2000), Duffie, Pan and Singleton (2000), Eraker (2000), Jiang (1998) and Pan (2000)).

uses the returns on the underlying asset. Subsequently we discuss the estimation of the parameters of the volatility process under the risk neutral probability measure. We estimate the risk neutral measure by combining data on returns with data on option prices. To clarify the relationship between the physical and risk neutral probability measures, we indicate the relationship between the parameters obtained under the physical and risk-neutral volatility dynamics.

2.1 Estimating the Volatility Process Under The Physical Probability Measure Using Asset Returns

We investigate the standard class of GARCH models pioneered by Engle (1982) and Bollerslev (1986). GARCH models are discrete-time models that have been used to estimate a variety of financial time series such as stock returns, interest rates and foreign exchange rates. See Bollerslev, Engle and Nelson (1994) for an overview of GARCH models and Bollerslev, Chou and Kroner (1992) and Campbell, Lo and MacKinlay (1997) for an overview of the use of GARCH models for financial time series. We assume that the logarithm of stock returns under the physical probability measure P follows the dynamic

$$\ln(S_t/S_{t-1}) \equiv R_t = r + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \sqrt{h_t}z_t, \quad z_t \sim N(0,1), \tag{1}$$

where r is the risk-free rate, and λ is the price of risk. Notice that it follows that the conditional expectation of gross returns is

$$E^P[\exp(R_t)|\Omega_{t-1}] = \exp(r + \lambda\sqrt{h_t}). \tag{2}$$

The specification of returns in (1) is common to all GARCH models we investigate. The differences between the models concern the specification of the volatility dynamic h_t. Ding, Granger and Engle (1993) and Hentschel (1995) provide very general specifications of the volatility dynamic that nest most existing work. Motivated by these studies, we first write the volatility dynamic h_t as follows[5]

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 h_{t-1} f(z_{t-1}) \tag{3}$$

where $z_t \sim N(0,1)$. Different GARCH models are mainly characterized by differences in the innovation functions f. We consider the following specifications of f

$$\begin{aligned} \textit{Simple} &: \quad f(z_{t-1}) = z_{t-1}^2 \\ \textit{Leverage} &: \quad f(z_{t-1}) = (z_{t-1} - \theta)^2 \end{aligned}$$

[5]The vast majority of papers in the literature find little or no support for higher-order GARCH models, thus we restrict attention to first order models here.

$$\begin{aligned} News &: \quad f(z_{t-1}) = \{|z_{t-1} - \theta| - \kappa(z_{t-1} - \theta)\}^2 \\ Power &: \quad f(z_{t-1}) = (z_{t-1} - \theta)^{2\gamma} \\ News\&Power &: \quad f(z_{t-1}) = \{|z_{t-1} - \theta| - \kappa(z_{t-1} - \theta)\}^{2\gamma} \end{aligned}$$

We then generalize to allow for nonlinear volatility dynamics as follows

$$Box - Cox : h_t^{\psi} = \beta_0 + \beta_1 h_{t-1}^{\psi} + \beta_2 h_{t-1}^{\psi} f(z_{t-1}), \text{ with } f(z_{t-1}) = (z_{t-1} - \theta)^{2\psi} \tag{4}$$

Notice that the $News\&Power$ model nests the preceding models, while the $News$, $Power$ and $Box - Cox$ models nest the *Leverage* and *Simple* models. To appreciate the importance of the innovation function, Figure I plots four of these innovation functions for certain (arbitrarily chosen) parameter combinations. All panels plot the innovation function $f(z_{t-1})$ as a function of the $i.i.d.$ "shock" z_{t-1}. In every panel the broken line depicts the symmetric innovation function associated with the *Simple* model, and the solid line depicts the innovation function for the alternative model, which has extra parameters. It can be seen that the *Leverage* parameter θ "shifts" the innovation function, the $News$ parameter κ "tilts" the innovation function, and the $Power$ parameters, γ and ψ flatten or steepen the innovation function. Similar effects are described in Pagan and Schwert (1990), Engle and Ng (1993), Ding, Granger and Engle (1993) and Hentschel (1995).

The persistence of the models under consideration can be written as

$$Persistence = \beta_1 + \beta_2 E\{f(z_{t-1})\} \tag{5}$$

where $E\{f(z_{t-1})\}$ can be calculated analytically for the *Simple* model to be 1, for the *Leverage* model to be $1 + \theta^2$, and for the $News$ model to be

$$\left(1+\theta^2\right)\left(1+\kappa^2\right) + 2\kappa[2\phi(\theta) + \left(1+\theta^2\right)(2\Phi(\theta) - 1)]$$

where $\phi(\theta)$ and $\Phi(\theta)$ denote the probability density function and the cumulative distribution function for the standard normal. For the three other models, $Power$, $News\&Power$ and $Box - Cox$, the persistence measure can be computed easily using simulation. Notice that in the $Box - Cox$ model, persistence is with respect to the power of volatility, h_t^{ψ} rather than of volatility itself. It is the case for all models that restricting the persistence to be below unity guarantees covariance stationarity of the model.

One way to price options under GARCH dynamics for returns is to proceed in two steps. In a first step, one estimates the parameters of (1) and (3) under the physical probability measure using asset returns. In a second step, one maps these parameters into the parameters for the risk-neutral distribution. For GARCH models, this approach is followed for instance by Amin and Ng (1993), Bollerslev and Mikkelsen (1996), and Hardle and Hafner (2000). This second step is described in more detail in the next subsection. To implement the first step, a convenient approach is to estimate the model parameters using maximum likelihood.

Denoting the vector of volatility parameters by δ, we maximize the following log likelihood function (conditioning on the first observation)

$$\ln L = -(T-1)\ln(2\pi)/2 - \sum_{t=2}^{T} \ln\left(h_t(\delta)\right)/2 - \sum_{t=2}^{T}\left(R_t - r - \lambda\sqrt{h_t(\delta)} + \frac{1}{2}h_t(\delta)\right)^2 / \left(2h_t(\delta)\right) \quad (6)$$

2.2 Transforming the Physical Return Dynamic to a Risk Neutral Dynamic

To use the parameters obtained under the physical probability measure for option valuation, we have to be explicit about the relationship between the physical and risk neutral dynamic. This relationship amounts to a choice of pricing kernel or equivalently a choice of the utility function of the representative agent. In a GARCH context, Duan (1995), building on the work of Rubinstein (1976) and Brennan (1979), provides a locally risk-neutral valuation relationship ($LRNVR$) which is satisfied by a measure Q if

$$E^Q[\exp(R_t)|\Omega_{t-1}] = \exp(r), \quad (7)$$

and

$$Var^Q[R_t|\Omega_{t-1}] = Var^P[R_t|\Omega_{t-1}] = h_t. \quad (8)$$

The $LRNVR$ implies that under the risk neutral measure Q, the return process evolves according to

$$R_t = r - \frac{1}{2}h_t + \sqrt{h_t}z_t^*, \quad z_t^* \sim N(0,1). \quad (9)$$

The volatility process (3) becomes under the risk-neutral measure[6]

$$h_t(\delta,\lambda) = \beta_0 + \beta_1 h_{t-1} + \beta_2 h_{t-1} f\left(z_{t-1}^* - \lambda, \delta\right) \quad (10)$$

To provide some more insight into this result, notice that solving for z_{t-1}^* from the risk neutral dynamic (9) yields

$$z_{t-1}^* = \left(R_{t-1} - r + \frac{1}{2}h_{t-1}\right)/\sqrt{h_{t-1}}, \quad (11)$$

whereas solving for z_{t-1} from the physical dynamic (1) we get

$$z_{t-1} = \left(R_{t-1} - r - \lambda\sqrt{h_{t-1}} + \frac{1}{2}h_{t-1}\right)/\sqrt{h_{t-1}} = \left[\left(R_{t-1} - r + \frac{1}{2}h_{t-1}\right)/\sqrt{h_{t-1}}\right] - \lambda. \quad (12)$$

[6]This specification only covers the *Simple*, *Leverage*, *News*, *Power* and *News&Power* models. The transformation for the $Box-Cox$ model in (4) is similar.

It is therefore clear from comparing (11) and (12) that for a general innovation function f, we have

$$f\left(z^*_{t-1} - \lambda\right) = f\left(z_{t-1}\right), \forall f. \tag{13}$$

Given the risk neutral dynamic in (9) and (10), the price of a European call option can be computed as

$$C(h_t\left(\delta, \lambda\right)) = \exp(-r(T-t))E^Q[\max(S_T - K, 0)|\Omega_{t-1}], \tag{14}$$

However, we do not have an analytical expression for this option price. Because the multi-period distribution of the GARCH process is unknown, we need to compute the conditional expectation using Monte Carlo simulation, the details of which will be discussed below.

As mentioned above, we can implement the option valuation models in two steps. In a first step, we obtain estimates $\hat{\delta}$ and $\hat{\lambda}$ for the parameters δ and λ under the physical probability measure using a time series of returns as described in section 2.1. In a second step, we use these parameters in the risk neutral dynamic (10) and compute the option price using (14). The fit of a volatility model can then be assessed using standard loss functions. Following much of the existing literature, we use mean-squared dollar errors.[7]

$$\$MSE = \frac{1}{N}\sum_i \left(C_i - C_i(h_t(\hat{\delta}, \hat{\lambda})\right)^2 \tag{15}$$

where C_i is the market price of option i, $C_i\left(h_t(\hat{\delta}, \hat{\lambda})\right)$ is the corresponding model price, and N is the total number of contracts in the sample. The sample can consist of a number of contracts on a given day, or a number of contracts on different days.

2.3 Estimating the Volatility Process Under the Risk-Neutral Probability Measure Using Option Prices and Asset Returns

In this section we describe an alternative approach to estimating the parameters of the volatility process and testing the option valuation models. The approach described above that estimates the parameters under the physical probability measure is conceptually straightforward and computationally easy. However, for the purpose of option valuation, it may be preferable to estimate the parameters directly using (a different set of) option prices. One possible alternative is to use just one cross-section of option prices at some point in time to estimate. For instance, if we use a mean-squared dollar objective function, we can obtain model parameters by minimizing

$$\$MSE = \frac{1}{N_t}\sum_i \left(C_{i,t} - C_{i,t}(h_t(\delta^*)\right)^2 \tag{16}$$

[7]The choice of loss function is an important and often ignored topic. See Renault (1997) for a careful discussion.

where δ^* is the vector of risk-neutral parameters to be estimated and N_t is the number of option prices present in the sample at time t. Alternatively, if we choose to use more than one cross-section of option prices to estimate the parameters, then we can link volatility on different dates using the time-series of stock returns. The objective is again to minimize the sum of squared option valuation errors in

$$\$MSE = \frac{1}{N_T} \sum_{t,i} (C_{i,t} - C_{i,t}(h_t(\delta^*))^2 \tag{17}$$

where $N_T = \sum_{t=1}^{T} N_t$ and T is the total number of days included in the sample. The updating from h_t to h_{t+1} is done using the observed daily returns, R_t on the underlying asset, by substituting (12) in (3), which yields an updating function that exclusively involves observables

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 h_{t-1} f\left(\left[\left(R_{t-1} - r + \frac{1}{2} h_{t-1}\right) / \sqrt{h_{t-1}}\right] - \lambda\right) \tag{18}$$

There is a technical issue regarding identification when estimating under the risk-neutral probability measure. This issue is best illustrated using a specific model rather than using the general notation involving the innovation function $f(z_{t-1})$. For example, in the case of the *Leverage* model, the option valuation formula involves the returns formula (9) and the risk-neutral volatility process is

$$h_t(\delta, \lambda) = \beta_0 + \beta_1 h_{t-1} + \beta_2 h_{t-1} \left(z^*_{t-1} - \theta - \lambda\right)^2 \tag{19}$$

Notice that we cannot separately identify λ and θ from (9) and (19), as we can only estimate $\theta^* \equiv \theta + \lambda$. It is often thought that by including several cross-sections of options in the sample, one can separately identify λ and θ, because of the use of the volatility updating formula, which is in obtained under the physical probability measure. However, notice that for the leverage model the volatility updating formula (18) becomes

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 h_{t-1} \left(\left[\left(R_{t-1} - r + \frac{1}{2} h_{t-1}\right) / \sqrt{h_{t-1}}\right] - (\lambda + \theta)\right)^2 \tag{20}$$

So clearly we cannot separately identify λ and θ when estimating under the risk-neutral probability measure, as we only estimate $\theta^* \equiv \theta + \lambda$.[8] This is not a problem for option valuation, because all we need is the combination $\theta + \lambda$, but it is an important difference compared to the technique described above, where we separately identify λ and θ using the physical probability measure and then use the sum of the two parameters for option

[8] We choose to estimate $\theta^* \equiv \theta + \lambda$. If one only estimates under the risk neutral measure one can equivalently assume that $\lambda = 0$, which trivially identifies θ (see Hsieh and Ritchken (2000)).

valuation. This remark applies for all of the models we investigate, as they all involve the same functional form in θ.[9]

We have described two approaches to estimating parameters for use in option valuation models. The first approach consists of using the time series of asset returns to estimate the parameters under the physical probability measure. Subsequently these parameters are mapped into the risk-neutral parameters, and plugged into the option valuation formulas. The second approach is to estimate the risk-neutral parameters directly from option prices. If we use just one cross-section of option prices, this approach is relatively straightforward. If we use multiple cross-sections, we have to use a volatility updating rule.[10]

A priori we would expect the second approach to work better, for several reasons. First, option prices contain forward looking information over and beyond historical returns, and thus using option prices to find parameters can have an important advantage simply from the perspective of the data used. Second, when using maximum likelihood to estimate parameters under the physical measure, it is clear that the loss function is quite different from an out of sample loss function which could be something like the mean-squared dollar errors in (15).

3 Data

We conduct our empirical analysis using four years of data on S&P 500 call options. First, the three-year period between June 1, 1988 and May 31, 1991, which we denote Sample A, is used exclusively for in-sample estimation. This sample closely corresponds to the data used by Bakshi, Cao and Chen (1997).[11] We subsequently use a fourth year of data covering the period from June 1, 1991 to May 31, 1992, which we refer to as Sample B. We apply several filters to Sample B that are identical to the ones used in Bakshi, Cao and Chen (1997), and we refer the reader to that study for the details.

In both samples, we only use options data for Wednesdays.[12] If Wednesday is a holiday,

[9]It is important to remember that this issue is simply a consequence of the way in which the parameter λ enters the innovation function, and therefore a consequence of the assumptions on the discount factor (or the price of risk). Also, it must be emphasized that in principle one could identify λ and θ separately, for instance by specifying a joint likelihood function for returns and options prices. By using the updating rule (18), returns data are only implicitly taken into account in the objective function, and this does not identify λ.

[10]When we use multiple cross-sections of options data, we therefore use joint information on returns and options prices, but the information on returns only enters the objective function indirectly. This approach is also followed by Heston and Nandi (2000) and Hsieh and Ritchken (2000) in a discrete-time environment. In a continuous-time environment, Eraker (2000) and Jones (2000) use a likelihood-based approach that combines options and returns data, and Chernov and Ghysels (2000) and Pan (2000) use a GMM approach to combine both types of data.

[11]The three years of data was graciously provided to us by Gurdip Bakshi.

[12]This choice is to some extent motivated by constraints. Because we cannot compute option prices

we use the next trading day. Using only Wednesday data allows us to study a fairly long time-series which is useful considering the highly persistent volatility processes. An additional motivation for only using Wednesday data is that following the work of Dumas, Fleming and Whaley (1998), several studies have used this setup (see for instance Heston and Nandi (2000)). For our empirical work, this data selection criterion leaves us with 156 cross-sections of options data in the June 1, 1988 to May 31, 1991 period and 52 cross–sections of options data in the June 1, 1991 to May 31, 1992 period.

Table I presents the number of contracts used in the empirical work for Sample A and B by moneyness and maturity. It can be seen that the relative importance of the different cells is fairly similar across periods. Tables II and III present average prices and implied volatilities for both sample periods. Given the differences between option prices in different cells in Table II, it is clear that different options will receive different weights when using the mean-squared dollar objective function (15). Table III indicates that in both sample periods we find implied volatility patterns across maturity and moneyness that are comparable to those found in other studies. In-the-money calls (and therefore also out-of-the-money puts) are expensive relative to the Black-Scholes model.

To give an idea of the differences in returns over the sample, Figure II plots the S&P 500 daily log returns and daily absolute log returns from CRSP for the option sample period. The vertical line separates the Sample A and B periods. It can be seen that the patterns in returns are fairly similar over the two periods, and that there are occasional outliers in returns. Figure III presents a similar plot for the average implied volatilities extracted from the options on a week-by-week basis. It can be seen that just as is the case with returns, implied volatility varies considerably. Interestingly, the level of implied volatility in Sample B is lower than in Sample A. Also, it is evident from Figure III that there is substantial clustering in implied volatilities.

4 Empirical Implementation and Results

In this section we discuss our empirical results for Sample A and B. We also discuss some of the details related to implementation that were not discussed above. First, we present the results from maximum likelihood estimation of the parameters under the physical probability measure. Second, we discuss the performance of the option valuation models using those physical parameters. Third, we discuss the estimation of parameters using options from Sample A (June 1, 1988 to May 31, 1991) and returns. We also discuss the performance of the different option valuation models for this in-sample exercise. Fourth, we use the parameters estimated using Sample A options and asset returns to price options in Sample B (June 1, 1991 to May 31, 1992). Fifth, we evaluate the performance of the option valuation

analytically, the optimization problems are fairly time-intensive, and limiting the number of options reduces the computational burden.

models in Sample B in a different way, by re-estimating the models every week and valuing the options one week out-of-sample.

4.1 Maximum Likelihood Estimation of Physical Processes Using Returns

Table IV shows parameter estimates obtained by maximization of the log-likelihood function (6). Robust standard errors are computed according to White (1982). For the riskless interest rate, we have assumed a constant 5% yearly rate leading to a daily rate of .05/365=0.000137. It is well-known that it is difficult to estimate GARCH parameters precisely from returns data unless long time-series are used. For comparable models, Hentschel (1995) and Ding, Granger and Engle (1993) use more than 50 years of daily data. Here we present results obtained using a twelve year period beginning at the start of our options data set and ending on December 31, 1999.[13]

The results are comparable with standard findings in the literature on GARCH processes. First, consider the *Simple* and *Leverage* models. Parameter estimates for β_1 and β_2 are very precise and roughly of the same order of magnitude as in the existing literature. In the *Leverage* model the estimate of the parameter θ is positive, indicating negative skewness.[14] The persistence of the process implied by these parameter estimates is 0.9823, indicating a very persistent process, again in accordance with the literature. Finally, the standard deviation of returns implied by the model parameters is 0.1786, which is reasonable.

While implied persistence and standard deviation for the three more richly parameterized models *News*, *Power* and *Power&News* are very similar to the *Leverage* model, other estimation results differ between the models. While in the *News* model the additional parameter κ is not estimated significantly, the power parameter γ is estimated significantly smaller than one in the *Power* and *Power&News* models. Interestingly, the parameter κ is estimated significantly in the *Power&News* model, but the estimate of the θ parameter in this model is very different from the other models and no longer significant. The combination of κ and γ is apparently able to capture part of the leverage effect previously estimated by θ. Note also that the $Box-Cox$ model has a power parameter ψ which is about two standard deviations below 1.

Finally, and probably most importantly, the likelihood ratio tests indicates that any model is preferred to the *Simple* model. When testing against the *Leverage* model, the

[13]Estimation with shorter sample periods yielded roughly comparable point estimates but lower t-statistics. We therefore start the sample at the beginning of the options data sample and include all subsequent returns available through CRSP.

[14]This estimate also indicates a negative relationship between shocks to returns and volatility, labeled the "leverage effect" by Black (1976). This effect has been documented by a large number of studies that estimate stock returns. Using option prices the presence of this effect has been confirmed among others by Benzoni (1998), Chernov and Ghysels (2000), Eraker (2000), Heston and Nandi (2000) and Nandi (1998).

$Power$ and $Power\&News$ models are found to be significantly better, but the $News$ and $Box-Cox$ models are not.[15]

4.2 Option Valuation with ML Estimates of Physical Processes

Table V presents results obtained by using the parameters obtained in Section 4.1 to price options in Sample A (June 1, 1988 to May 31, 1991) as well as in Sample B (June 1, 1991 to May 31, 1992). To assess the models' fit, we present the mean squared dollar error loss as well as its square root, which expresses the valuation error in dollar terms

$$\$RMSE = \sqrt{\frac{1}{N_T}\sum_{t,i}\left(C_{i,t} - C_{i,t}(h_t(\hat{\delta},\hat{\lambda})\right)^2} \qquad (21)$$

where $C_{i,t}$ is the market price of contract i on day t, $C_{i,t}(h_t(\hat{\delta},\hat{\lambda}))$ is the corresponding model price, and N_T denotes the total number of contracts available. The evidence can be summarized very briefly. First, in comparison with the average option prices reported in Table II, the $\$RMSE$ errors are very large. Second, the more complex models perform even worse than the *Simple* and *Leverage* model, which are nested in the more complex models. Especially in light of its poor performance using the Maximum Likelihood criterion above, it is indeed shocking that the *Simple* model performs the best in terms of option valuation. Thus our first conclusions emerge: Parameters from MLE on returns only should not be used for option valuation, and furthermore, inference procedures from MLE criteria are not reliable to rank the models' performance in option valuation.[16]

4.3 NLS Estimation of Risk Neutral Processes Using Options and Returns

In this section we present estimation results obtained using 156 Wednesdays of options data from Sample A. For all models, the objective function used is the (square root) dollar mean squared error loss function (21). Model option prices are obtained using (9), (10) and (14),

[15] It must be noted that the *Simple* model in this paper sets the price of risk λ equal to zero. Allowing for a nonzero λ in the *Simple* model will induce a leverage effect under the risk neutral measure even when none exists under the physical measure. To avoid any type of leverage affect under any measure for the *Simple* model, we implement the model with $\lambda = \theta = 0$. Our estimates therefore imply a symmetric news impact function under both measures.

[16] These results are somewhat different from existing ones in the literature. Amin and Ng (1993) perform a similar analysis using data on single stock options and report much smaller option valuation errors. Hardle and Hafner (2000) estimate model parameters under the objective distribution but use relative valuation errors. Bollerslev and Mikkelsen (1996) analyze leaps. Their results are therefore difficult to compare to the ones in this paper.

and volatility is updated using (18). In order to compute the option prices numerically, we use 1000 simulated paths,[17] and we use the empirical martingale method of Duan and Simonato (1999) to increase numerical efficiency. We also use stratified random numbers, antithetic variates, and a control variate technique using the Black-Scholes price as the control.

There are some important issues regarding the implementation of this empirical exercise that deserve elaboration. First, there are several ways to treat the initial conditional volatility. For instance, Bakshi, Cao and Chen (1997) treat the initial conditional volatility as an extra parameter in the implementation of a continuous-time option valuation model. This approach could also be followed for the estimation of discrete-time processes, but instead we follow Heston and Nandi (2000) and set the initial volatility equal to the unconditional volatility 250 days before the first Wednesday that is included in the sample. Subsequently, we use the volatility updating rule (18) for the next 250 days to obtain the conditional volatility for the first Wednesday used in the sample.

A second issue regarding implementation is the size, nature and length of the option valuation sample. Our motivation in selecting the sample is essentially to use the longest time span possible, while limiting the number of contracts to limit computation time. There are several good reasons for preferring a long sample. First, as emphasized by Heston and Nandi (2000), if the option valuation model is a reliable one, it should price options well over any time horizon, because it has implications for the returns process as well as for option prices. Second, from a practical perspective, a long sample of option prices is preferable for the same reason a long sample of returns is preferable, because the time variation in prices and volatility will help identify the model's parameters. We experimented with several types of samples for option prices. First, following the work of Bakshi, Cao and Chen (1997) and Duan (1996), we experimented with estimating the parameters using a single cross-section of option prices. This is a convenient approach but unfortunately for the models we investigated it is not satisfactory. The objective function is not sufficiently well behaved and numerical problems are so serious that the results become impossible to interpret. A second approach is to use multiple cross-sections of option prices. We experimented with six-month samples of options, following the example of Heston and Nandi (2000). While this approach was more satisfactory, it still occasionally gave rise to numerical problems. We therefore decided to use a long, three-year sample of options data to estimate the model's parameters. Extensive sensitivity analysis demonstrated that the results for this approach are robust. As mentioned above, we only include Wednesdays in the sample following the setup in Dumas, Fleming and Whaley (1998) and Heston and Nandi (2000). It would be prohibitively expensive in terms of computer time to use all trading days in the three-year sample. The design of our empirical experiment can therefore be summarized as follows:

[17] To verify that 1000 draws are adequate, we repeated our analysis for a limited number of cases using 5000 draws and obtained identical results.

we believe that for estimation purposes, it is preferable to use 156 consecutive Wednesdays rather than 156 consecutive trading days.

Table VI presents the results for the nonlinear least squares estimation of the risk-neutral parameters. As in Table IV, the riskless rate of return is set at 5% yearly leading to a daily rate of .05/365=0.000137. A number of results are noteworthy. First, the persistence for all models is extremely high and roughly the same for the four models. Persistence is higher than the persistence for the physical process in Table IV, even when transforming the physical parameters to risk neutrality. The implied standard deviation of the four models is roughly comparable and reasonable.

To understand these results, consider the simple *Leverage* model. Leverage under the physical measure is simply θ, but under the risk neutral measure it is $\theta + \lambda$ which is greater if the price of risk is positive. Volatility persistence under the physical measure is $\beta_1 + \beta_2(1 + \theta^2)$ and under the risk-neutral measure it is $\beta_1 + \beta_2(1 + (\theta + \lambda)^2)$ which is higher than the physical persistence for a positive θ (implying negative skewness) and a positive price of risk λ.

When comparing the estimated parameter values across models in Table VI, it is of interest to remember that in many cases it is combinations of parameters that determine the model's most important characteristics, and not necessarily the individual parameters. For instance, the model's persistence is an important model characteristic and persistence is determined by a combination of all the model's parameters. Nevertheless, it is tempting to compare individual parameter estimates across models. For instance, the parameter β_0, which determines the model's unconditional volatility, is estimated in a fairly narrow range across the four models. Another important parameter is the $\theta^* \equiv \theta + \lambda$ combination, which indicates the size of the leverage effect. This effect is forced to zero in the *Simple* model. It is seen that θ^* is estimated at 2.9939 for the *Leverage* model, and somewhat lower for the more richly specified models. We therefore conclude that some of the effects captured by the parameters κ and γ are captured by the leverage effect if κ and γ are omitted. It is also noteworthy that when κ and γ are both estimated in the *Power&News* model, their estimates are not very different from the ones obtained in the *News* and *Power* models respectively. Finally we see that the power parameter in the $Box - Cox$ model is much higher compared to the estimate in Table IV, but still roughly two standard deviations below 1.

The most important conclusion from Table VI is obtained by comparing the estimation results with those of Table IV, obtained under the physical probability measure. For each model, the resulting two sets of parameter estimates are related by virtue of the precise form of the risk neutralization discussed in Section 2.2. It can be seen that implied parameter estimates are quite different. The most important differences are the following. First, for all models, the estimates of β_0 are not just different, but of a different order of magnitude. Second, the estimate of $\theta^* = \theta + \lambda$ implied by Table IV is positive for all models except for the *Power&News* model. However, it is always much smaller than the corresponding estimate in Table VI, implying a much smaller leverage effect even after adding λ for proper comparison. A likely culprit is that λ is poorly estimated under the physical measure where it only enters the conditional mean equation. Third, the estimates of κ in Table IV have a different sign than the ones in Table VI, even though the parameter is not significantly estimated for the *News* model. On the other hand, the estimates of γ in Table IV are not too different from those in Table VI. Nevertheless, it is clear that if the parameter estimates in Table VI are fairly accurate, it is perhaps not surprising that the parameter estimates from Table IV price options so poorly, as documented in Table V and Section 4.2.[18]

4.4 Option Valuation with NLS Estimates of Risk Neutral Processes

When comparing the six models in-sample, the most important criterion is the value of the minimized objective function. These values are presented in the last row of Table VI and repeated with more detail in the first panel of Table VII. First, we note that all GARCH models outperform the Black-Scholes model as well as the Practitioner's Black-Scholes (PBS) model from Dumas, Fleming and Whaley (1998) in Sample A.[19] Table VII.A presents two versions of the PBS model, where all parameters are kept constant throughout the in-sample period. The PBS model is estimated using either OLS on implied volatility or NLS minimizing \$MSE.[20]

It is clear that the extra parameters in the *News*, *Power* and $Box-Cox$ models do not improve the fit of the model, as the \$RMSE is only slightly lowered from 1.0445 to 1.0410, 1.0400 and 1.0440 respectively. However, the combination of κ and γ in the *Power&News* model lowers the \$RMSE to 1.0106. Interestingly, this pattern is similar to the one obtained in Table IV, when estimating under the physical probability measure. In Table IV, the change in the log likelihood for the *News* and $Box-Cox$ models is very small, and the *Power&News* model yields a bigger change in the log likelihood.

Figure IV provides more insight into the valuation differences between the models by comparing the option prices from each model with the Black-Scholes price for the same

[18] It is not clear whether this lack of consistency between the objective and risk-neutral parameters is a consequence of a badly specified volatility model or a badly specified price of risk. Chernov and Ghysels (2000) make the same point and formally test the mapping from the objective to the risk-neutral measure.

[19] The PBS model simply assumes a second order polynomial in strike price and maturity for the implied Black-Scholes volatility. It must be noted that the implementation of the PBS model reported on in Table VI is different from the one in Dumas, Fleming and Whaley (1998). The coefficients in the polynomial used for Table VI are constant across the sample to ensure comparability of the results with the implementation of the GARCH models in this table. In Dumas, Fleming and Whaley (1998), the coefficients of the polynomial are constant for a given day only. We present an empirical analysis that is identical to the one in Dumas, Fleming and Whaley (1998) below.

[20] The PBS model is traditionally implemented using an OLS setup (see Dumas, Fleming and Whaley (1998)). Christoffersen and Jacobs (2001) demonstrate that the pricing errors can be lowered by using a NLS setup.

option. For each model, the risk-neutral parameter estimates in Table VI are used. The call price from each model is plotted against moneyness for three maturities. Similar figures are shown in Heston (1993) for his stochastic volatility model. The initial conditional volatility in each GARCH model is set to its unconditional value as implied by the assumed parameters, and the volatility used for Black-Scholes valuation is set to this conditional volatility. It must be noted that this setup ensures that the Black-Scholes price has the best possible chance of matching the GARCH prices, because the initial volatilities are the same across models. We see that the *Simple* model yields very small deviations from the Black-Scholes model across maturities. The other models, particularly the $Box-Cox$ model, display systematic differences from Black-Scholes, enabling the models to fit observed patterns in the data such as implied volatility smirks. The differences between the GARCH models and Black-Scholes are smaller for short maturities, partly as a result of the lower values of these options.

Table VIII.A further elaborates on these findings by presenting \$RMSEs by moneyness and maturity for the four models (for Sample A). A first interesting finding from these tables is that in certain cells there are notable differences between the *Leverage*, $News$ and *Power* models. The differences between these three models are perhaps surprising, because on the basis of the overall fit in Table VII, one might have concluded that the three models yield nearly identical option prices. Table VIII.A indicates that this is not the case.

It can also be seen from Table VIII.A that even though the overall fit of the $Power\&News$ model is better than that of the other three models, this does not mean that this model does a better job of valuing all options. It is not a surprise that the most important improvements over the other three models are made for short and medium maturity options. The reason for this is that differences between GARCH models are more likely to be significant for short and medium-horizon forecasts, while they even out over long horizons where the conditional volatility approaches the unconditional. Furthermore, for the short and medium maturities improvements are made primarily for options that are in the money. The reason for this is that in-the-money options are more expensive and therefore carry more weight in the objective function.

All of the evidence discussed above pertains to in-sample evaluation of the option valuation models. However, from a practitioner's perspective out-of-sample valuation performance is much more important. The second panel of Table VII.A investigates out-of-sample valuation performance by using the model parameters estimated in Table VI (using Sample A) to evaluate the models' performance in Sample B. The most important conclusion is that the ranking of the models is reversed compared to the in-sample exercise. While the most richly parameterized model ($Power\&News$) has the best in-sample performance, it has the worst out-of-sample performance except for the *Simple* model. Conversely, the most parsimonious model (*Leverage*) has the worst in-sample performance, but performs better than the other models out-of-sample. Table VIII.B breaks down model performance across moneyness and maturity. By simply comparing the two models with the best and worst overall out-of-sample fit, it becomes clear that comparing the models' valuation performance is less than straightforward. While the overall fit of the *Power&News* model is significantly worse than that of the *Leverage* model ($1.1571 versus $0.9777 in $RMSE terms), it still outperforms the *Leverage* model for some deep in-the-money options. However, for long-maturity in-the-money options, the performance of the *Power&News* model is relatively poor.

Table VII.B compares the accuracy of the models using the Diebold and Mariano (1995) test on the weekly $RMSEs from Sample B, keeping the parameters fixed at their Sample A values. We test each model against the *Leverage* model. Not surprisingly, the *Simple* model is rejected as are the Black-Scholes and the PBS(OLS) model. None of the alternative GARCH models are significantly different from the *Leverage* model.

The following broad conclusions emerge from the results in Tables VII and VIII: First, even when using relatively large samples (for example, sample A contains more than 8,000 option contracts) and for the parsimonious models (3-7 parameters) studied here, results from in-sample estimation do not carry over to out-of-sample experiments. Second, when using options to estimate the parameters, the *Simple* model without leverage effect and price of volatility risk performs poorly. This is in stark contrast to the returns-based analysis in Table IV where the *Simple* model performed the best for option valuation. Third, all GARCH models (except for the *Simple* model) outperform the Black-Scholes model as well as the Practitioner Black-Scholes model from Dumas, Fleming and Whaley (1998) with constant parameters throughout the period.

4.5 Option valuation with Weekly Updating

The conclusion from the out-of-sample exercise is that one should use the parsimonious *Leverage* model rather than the more richly parameterized models. However, it may be argued that the out-of-sample exercise we conduct is very different from the way these models are typically used by practitioners. A typical critique on our analysis of the model's performance would be that it is unrealistic to assume that the model's parameters are constant over a four-year period (the three years of the in-sample exercise and the one year of the out-of-sample exercise).[21] We now investigate the robustness of our conclusions in a situation where we allow the model parameters to change over time. We work exclusively with the data in Sample B (June 1, 1991 to May 31, 1992) and conduct the following exercise: for each model we estimate different parameter values every Wednesday and use these parameter values to price the options the next Wednesday. In addition to using one cross-section of option prices, we also use the volatility updating rule (18) for the 250 days predating the Wednesday used in the estimation exercise. It turns out that by (indirectly) incorporating

[21] Our estimation and testing methodology is similar to that of Heston and Nandi (2000), who estimate parameters over a six-month period and subsequently evaluate the model out-of-sample over the next six months. Obviously our estimation period is longer, as explained above, because we did not always obtain satisfactory results using the six month periods.

this extra information on returns in the objective function, the optimization problem is much better behaved.

Table IX contains the results of the weekly estimation analysis. The most important part of the Table are panels B and C, which report the results of the one-week ahead exercise. For each of the 52 Wednesdays, we evaluate the fit of each cross section using parameters estimated from the previous Wednesday and returns data for the 250 days preceding it.[22] A first important conclusion is that the $RMSEs in Table IX are much lower than the corresponding numbers in Table VII. For example, the Sample B average dollar error for the *Leverage* model is 0.7081 with weekly updating, compared to 0.9777 when parameter estimates are not updated. When testing the average weekly $RMSEs in panel C of Table IX, the conclusions from the out-of-sample analysis in Table VII are reinforced. The parsimonious *Leverage* model is not outperformed by any of the alternative models and it outperforms several other GARCH models as well as Black-Scholes and the PBS(OLS) model, which is now updated every week. Remarkably, the *Leverage* model is not significantly outperformed by the PBS(NLS) model, which is estimated minimizing $MSE.

For completeness, the panel A of Table IX presents the in-sample fits using parameters that are updated weekly. In addition, Figure V presents the $RMSEs for each of the 52 Wednesdays in Sample B, both for the in-sample and out-of-sample analysis. It can be seen that out-of-sample, the relative performance of the four models can differ substantially from week to week.

In summary, the empirical out-of-sample results with weekly updating confirm the earlier out-of-sample results. When considering a family of GARCH models that includes a large number of existing models, there seems to be no good reason to look beyond a model that allows for volatility clustering and a leverage effect. While in-sample analysis on options data and estimation on returns data may suggest more richly parameterized models, the out-of-sample performance of those models is rather disappointing. Conversely, the comparisons between the GARCH option valuation models and the PBS models can be interpreted as being rather positive for GARCH models. All GARCH models, bar the *Simple* one, outperform the PBS(OLS) model in and out of sample in our paper. The equally simple but tougher competitor PBS(NLS) is only marginally better than the best GARCH model.

5 Summary and Directions for Future Work

This paper compares the performance of a number of GARCH models using specifications related to those of Ding, Granger and Engle (1993) and Hentschel (1995). In the existing literature, different GARCH are typically judged by comparing the log-likelihood based on a time-series of asset returns. In this paper, we investigate the performance of these models

[22] Obviously daily returns data between the two Wednesdays are also used to obtain the initial conditional volatility.

for the purpose of valuing options. We find that a comparison of the in-sample fit (based on dollar squared errors) favors the most richly parameterized model, as does a likelihood-based comparison that uses the time-series of asset returns. However, when assessing the performance of the models out-of-sample (again using dollar squared errors), the data favor the more parsimonious model that only contains a leverage effect.

The potential consequences of these findings for the GARCH literature are far-reaching. The advantage of the specifications of Ding, Granger and Engle (1993) and Hentschel (1995) is that they nest a large number of existing GARCH models. While these specifications do not nest all existing models, and while it is of course possible that we can find different GARCH specifications that perform more satisfactorily for the purpose of option valuation, it is safe to conclude that these findings question somewhat the value of the extensive literature that formulates complex alternatives to GARCH models that contain volatility clustering and a leverage effect. The reason that this finding is not articulated in the existing literature is that existing comparisons are often in-sample, and most often based on the model's likelihood value, which may not have an obvious relationship with more relevant objective functions based on hedging or speculation.

Our findings also have consequences for other literatures. In particular, because continuous-time stochastic volatility models can be thought of as the limits of GARCH processes, our results suggest that little may be gained in that literature from volatility dynamics different from models such as Heston (1993), which contains a leverage effect and allows for volatility clustering. Instead, more may be gained from changing the specification of other building blocks of the model, such as jumps.

Our results suggest that most likely, more is to be gained by more radical departures from the modeling assumptions used in a traditional simple GARCH model. For instance, modeling deviations from normality in the innovations process may prove instructive. Also, all models investigated in this paper share the same specification for the price of risk. It may be worthwhile to model the price of risk differently, especially because a comparison of the parameters estimated under the physical and risk-neutral probability distributions indicates such significant differences.

References

[1] Amin, K. and V. Ng (1993), "ARCH Processes and Option Valuation", Manuscript, University of Michigan.

[2] Andersen, T., Benzoni, L. and J. Lund (1999), "Estimating Jump-Diffusions for Equity Returns," Manuscript, Northwestern University.

[3] Bakshi, C., Cao, C. and Z. Chen (1997), "Empirical Performance of Alternative Option Pricing Models," Journal of Finance, 52, 2003-2049.

[4] Bates, D. (1996a), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," Review of Financial Studies, 9, 69-107.

[5] Bates, D. (1996b), "Testing Option Pricing Models," in G.S. Maddala and C.R. Rao, eds.: Handbook of Statistics, Vol 15: Statistical Methods in Finance (North-Holland, Amsterdam), 567-611.

[6] Bates, D. (2000), "Post-'87 Crash Fears in S&P 500 Futures Options," Journal of Econometrics, 94, 181-238.

[7] Benzoni, L. (1998), "Pricing Options Under Stochastic Volatility: An Econometric Analysis," Manuscript, University of Minnesota.

[8] Black, F. (1976), "Studies of Stock Price Volatility Changes," in: Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association, 177-181.

[9] Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81, 637-659.

[10] Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," Journal of Econometrics, 31, 307-327.

[11] Bollerslev, T., Chou, R. and K. Kroner (1992), "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," Journal of Econometrics, 52, 5-59.

[12] Bollerslev, T., Engle, R. and D. Nelson (1994), "ARCH Models," in: R. Engle and D. McFadden (eds.), Handbook of Econometrics, Vol. IV, Elsevier, Amsterdam.

[13] Bollerslev, T., Engle, R. and J. Wooldridge (1988), "A Capital Asset Pricing Model with Time-Varying Covariances," Journal of Political Economy, 96, 116-131.

[14] Bollerslev, T. and H.O. Mikkelsen (1996), "Long-Term Equity AnticiPation Securities and Stock Market Volatility Dynamics," Journal of Econometrics, 92, 75-99.

[15] Brennan, M. (1979), "The Pricing of Contingent Claims in Discrete-Time Models," Journal of Finance, 34, 53-68.

[16] Campbell, J., and L. Hentschel (1992), "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," Journal of Financial Economics, 31, 281-318.

[17] Campbell, J., Lo, A. and C. MacKinlay (1997), "The Econometrics of Financial Markets," Princeton University Press.

[18] Chernov, M. and E. Ghysels, (2000) "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Option Valuation," Journal of Financial Economics, 56, 407-458.

[19] Christoffersen,P. and K. Jacobs (2001), "The Importance of the Loss Function in Option Pricing", working paper, McGill University and CIRANO.

[20] Day, T. and C. Lewis (1992), "Stock Market Volatility and the Information Content of Stock Index Options, Journal of Econometrics, 52, 267-287.

[21] Derman, E., and I. Kani (1994), "Riding on the Smile," Risk, 7, 32-39.

[22] Diebold, F.X. and R. Mariano (1995), "Comparing Predictive Accuracy," Journal of Business and Economic Statistics, 13, 253-263.

[23] Ding, Z., C.W.J. Granger, and R.F. Engle (1993), "A Long Memory Property of Stock Market Returns and a New Model," Journal of Empirical Finance, 83–106.

[24] Duan, J.-C. (1995), "The GARCH Option Pricing Model," Mathematical Finance, 5, 13-32.

[25] Duan, J.-C. (1996), "Cracking the Smile", Risk, 9, 55-59.

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

[27] Duan, J.-C. and J.-G. Simonato (1998), "Empirical Martingale Simulation for Asset Prices", Management Science, 44, 1218-1233.

[28] Duffie, D., Pan, J. and K. Singleton (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," Econometrica, 68, 1343-1376.

[29] Dumas, B., Fleming, F. and R. Whaley (1998), "Implied Volatility Functions: Empirical Tests," Journal of Finance, 53, 2059-2106.

[30] Dupire, B. (1994), "Pricing with a Smile," Risk, 7, 18-20.

[31] Engle, R. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation," Econometrica, 50, 987-1008.

[32] Engle, R. and C. Mustafa (1992), "Implied ARCH Models from Options Prices," Journal of Econometrics, 52, 289-311.

[33] Engle, R. and V. Ng (1993), "Measuring and Testing the Impact of News on Volatility," Journal of Finance, 48, 1749-1778.

[34] Eraker, B. (2000), "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," Working Paper, University of Chicago.

[35] French, K., G. W. Schwert, and R. Stambaugh (1987), "Expected Stock Returns and Volatility," Journal of Financial Economics, 19, 3-30.

[36] Garcia, R. and E. Renault (1998), "A Note on Heging in ARCH and Stochastic Volatility Option Pricing Models", Mathematical Finance, 8, 153-161.

[37] Glosten, L., R. Jagannathan, and D. Runkle (1993), "On the Relation between the Expected Value and the Volatility fo the Nominal Excess Return on Stocks," Journal of Finance, 48, 1779-1801.

[38] Hardle, W. and C. Hafner (2000), "Discrete Time Option Pricing with Flexible Volatility Estimation", Finance and Stochastics, 4, 189-207.

[39] Hentschel, L. (1995), "All in the Family: Nesting Symmetric and Asymmetric GARCH Models," Journal of Financial Economics, 39, 71-104.

[40] Heston, S. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," Review of Financial Studies, 6, 327-343.

[41] Heston, S. and S. Nandi (2000), "A Closed-Form GARCH Option Pricing Model," Review of Financial Studies, 13, 585-626.

[42] Hsieh, K. and P. Ritchken (2000), "An Empirical Comparison of GARCH Option Pricing Models", manuscript, Case Western Reserve University

[43] Hull, J. and A. White (1987), "The Pricing of Options with Stochastic Volatilities," Journal of Finance, 42, 281-300.

[44] Jiang, G. (1998), "Testing Option Pricing Models with Stochastic Volatility, Random Jump and Stochastic Interest Rate-Open the "Black Box", Working Paper, University of Groningen.

[45] Johnson, H.E., and D. Shanno (1987) ,"Option Pricing when the Variance is Changing," Journal of Financial and Quantitative Analysis, 22, 143-151.

[46] Jones, C. (2000), "The Dynamics of Stochastic Volatility," Manuscript, University of Rochester.

[47] Jones, C. (2001), "A Nonlinear Analysis of S&P 500 Index Option Returns," Manuscript, Simon School of Business, University of Rochester.

[48] Jorion, P. (1995), "Predicting Volatility in the Foreign Exchange Market," Journal of Finance, 50, 507-528.

[49] Lamoureux, C. and W. Lastrapes (1993), "Forecasting Stock Return Variance: Towards an Understanding of Stochastic Implied Volatilities," Review of Financial Studies, 5, 293-326.

[50] Melino, A. and S. Turnbull (1990), "Pricing Foreign Currency Options with Stochastic Volatility," Journal of Econometrics, 45, 239-265.

[51] 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 Options Market," Journal of Banking and Finance 22, 589-610.

[52] Nelson, D. and D. Foster (1994), "Asymptotic Filtering Theory for Univariate ARCH Models," Econometrica, 62, 1-41.

[53] Pagan, A. and G.W. Schwert (1990), "Alternative Models for Conditional Stock Volatility," Journal of Econometrics, 45, 267-290.

[54] Pan, J. (2000), "Integrated Time-Series Analysis of Spot and Option Pricing," Manuscript, MIT Sloan School of Management.

[55] Renault, E. (1997), "Econometric Models of Option Pricing Errors," Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress. Volume 3. (Kreps, D. and K. Wallis, eds.), Econometric Society Monographs, no. 28. Cambridge; New York and Melbourne: Cambridge University Press, 223-78.

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

[57] Rubinstein, M. (1976), "The Valuation of Uncertain Income Streams and The Pricing of Options," Bell Journal of Economics, 7, 407-425.

[58] Schwert, G.W. (1989), "Why Does Stock Market Volatility Change over Time?" Journal of Finance, 44, 1115-1153.

[59] Scott, L. (1987), "Option Pricing when the Variance Changes Randomly: Theory, Estimators and Applications," Journal of Financial and Quantitative Analysis, 22, 419-438.

[60] Tauchen, G. (1997), "New Minimum Chi-Square Methods in Empirical Finance," in: Kreps, D. and K. Wallis (eds.): Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Vol. 3, Cambridge University Press.

[61] White, H. (1982), "Maximum Likelihood Specification of Misspecified Models," Econometrica, 50, 1-25.

[62] Wiggins, J.B. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Evidence," Journal of Financial Economics, 19, 351-372.

Table I: Number of Contracts Across Moneyness and Maturity

	Sample A					Sample B			
	DTM < 60	60<DTM<180	180 < DTM	Total		DTM<60	60<DTM<180	180<DTM	Total
S/X < .94	164	641	515	1,320	S/X < .94	49	222	173	444
.94 < S/X < .97	436	413	207	1,056	.94 < S/X < .97	170	201	104	475
.97 < S/X < 1.00	558	400	202	1,160	.97 < S/X < 1.00	233	193	82	508
1.00 < S/X < 1.03	523	357	172	1,052	1.00 < S/X < 1.03	214	187	97	498
1.03 < S/X < 1.06	474	312	137	923	1.03 < S/X < 1.06	199	160	63	422
1.06 < S/X	1015	985	570	2,570	1.06 < S/X	346	332	179	857
Total	3,170	3,108	1,803	8,081	Total	1,211	1,295	698	3,204

Table II: Average Quoted Price Across Moneyness and Maturity

	Sample A					Sample B			
	DTM < 60	60<DTM<180	180 < DTM	All		DTM<60	60<DTM<180	180<DTM	All
S/X < .94	1.52	5.05	10.20	6.62	S/X < .94	0.82	3.36	9.40	5.43
.94 < S/X < .97	2.63	9.69	18.93	8.59	.94 < S/X < .97	1.81	7.80	17.80	7.85
.97 < S/X < 1.00	5.35	14.97	24.93	12.08	.97 < S/X < 1.00	4.86	13.50	24.00	11.23
1.00 < S/X < 1.03	11.14	21.38	31.84	18.00	1.00 < S/X < 1.03	11.80	20.90	31.90	19.13
1.03 < S/X < 1.06	18.58	28.35	37.01	24.62	1.03 < S/X < 1.06	20.60	29.00	40.20	26.71
1.06 < S/X	40.35	50.39	62.65	49.14	1.06 < S/X	40.00	48.30	60.30	47.46
All	18.92	25.53	33.53	24.72	All	18.12	22.78	31.33	22.88

Table III: Average Implied Volatility Across Moneyness and Maturity

	Sample A					Sample B			
	DTM < 60	60<DTM<180	180 < DTM	All		DTM<60	60<DTM<180	180<DTM	All
S/X < .94	0.1786	0.1706	0.1668	0.1701	S/X < .94	0.1339	0.1377	0.1506	0.1423
.94 < S/X < .97	0.1652	0.1719	0.1765	0.1700	.94 < S/X < .97	0.1346	0.1461	0.1607	0.1452
.97 < S/X < 1.00	0.1712	0.1821	0.1858	0.1775	.97 < S/X < 1.00	0.1445	0.1566	0.1722	0.1536
1.00 < S/X < 1.03	0.1901	0.1941	0.1976	0.1927	1.00 < S/X < 1.03	0.1626	0.1704	0.1781	0.1685
1.03 < S/X < 1.06	0.2172	0.2047	0.1942	0.2096	1.03 < S/X < 1.06	0.1847	0.1824	0.1896	0.1846
1.06 < S/X	0.3131	0.2363	0.2179	0.2626	1.06 < S/X	0.2574	0.2104	0.2037	0.2280
All	0.2262	0.1992	0.1912	0.2080	All	0.1847	0.1707	0.1756	0.1771

Notes: We report various descriptive statistics on our two subsamples of options. Sample A denotes June 1, 1988 - May 31, 1991, and Sample B denotes June 1, 1991 - May 31, 1992. Each statistic is reported for three maturity bins and six moneyness bins.

Table IV: Maximum Likelihood Estimates of Physical Processes

	Simple	Leverage	News	Power	Power&News	Box-Cox
r	1.37E-04	1.37E-04	1.37E-04	1.37E-04	1.37E-04	1.37E-04
λ		0.0452	0.0455	0.0465	0.0456	0.0464
		0.0185	0.0185	0.0186	0.0185	0.0186
β_0	1.84E-06	2.24E-06	2.30E-06	2.05E-06	1.98E-06	8.42E-06
	2.27E-07	1.82E-07	1.85E-07	1.92E-07	1.85E-07	5.44E-06
β_1	0.8873	0.8524	0.8601	0.8397	0.8531	0.8615
	0.0054	0.0052	0.0101	0.0107	0.0126	0.0050
β_2	0.0984	0.0867	0.0771	0.1217	0.1881	0.0990
	0.0026	0.0053	0.0102	0.0186	0.0260	0.0084
θ		0.7061	0.5605	0.6328	-0.1056	0.6381
		0.0845	0.1508	0.0757	0.0728	0.0778
κ			0.1267		0.7025	
			0.1093		0.0840	
γ				0.8505	0.5335	
				0.0626	0.0481	
ψ						0.8541
						0.0701
Persistence*	0.9857	0.9823	0.9821	0.9877	0.9931	0.9810
Annual Std. Dev.	0.1799	0.1786	0.1796	0.2049	0.2696	0.1681
LogLikelihood	10590.3	10639.0	10639.6	10640.8	10651.3	10639.7
LR P-Value (Simple)		0.0000	0.0000	0.0000	0.0000	0.0000
LR P-Value (Leverage)			0.2690	0.0561	0.0000	0.2319

Notes: We estimate the six GARCH models using Maximum Likelihood on daily S&P500 returns from June 1, 1987 through December 31, 1999 for a total of 3,182 observations. Standard errors from White (1982) appear below each estimate. The persistence and the annualized standard deviation implied by each model is reported at the bottom of the table. *) For the Box-Cox model persistence refers to the power of volatility rather than volatility itself. The LR P-values refer to Likelihood ratio tests of the null hypothesis that a particular model fits the return data no better than the *Simple* GARCH model or the *Leverage* GARCH model respectively.

Table V: Fit of Option Prices from ML Estimates of Physical Processes

Sample A	Simple	Leverage	News	Power	Power&News	Box-Cox
$MSE	3.7623	4.9827	4.7751	8.7663	19.7077	6.6071
$RMSE	1.9397	2.2322	2.1852	2.9608	4.4393	2.5704

Sample B	Simple	Leverage	News	Power	Power&News	Box-Cox
$MSE	5.6770	9.0601	8.6060	16.0743	32.9882	12.3743
$RMSE	2.3827	3.0100	2.9336	4.0093	5.7435	3.5177

Notes: We transform the physical MLE estimates from Table IV into risk neutral parameters and calculate GARCH option prices for each contract in Sample A (June 1, 1988 - May 31, 1991; 8,081 contracts) and in Sample B (June 1, 1991 - May 31, 1992; 3,204 contracts). Using the actual observed market price for each option, we then calculate $MSE and $RMSE for every model on both samples.

Table VI: NLS Estimates of Risk-Neutral Processes. Sample A

	Simple	Leverage	News	Power	Power&News	Box-Cox
r	1.37E-04	1.37E-04	1.37E-04	1.37E-04	1.37E-04	1.37E-04
β_0	4.89E-07	5.92E-07	5.79E-07	5.92E-07	5.66E-07	9.64E-07
	1.75E-08	9.74E-09	9.63E-09	9.75E-09	9.82E-09	2.65E-07
β_1	0.9699	0.8629	0.8666	0.8649	0.8253	0.8735
	0.0007	0.0037	0.0041	0.0030	0.0038	0.0058
β_2	0.0250	0.0133	0.0209	0.0291	0.1832	0.0163
	0.0007	0.0003	0.0245	0.0023	0.0166	0.0017
$\theta+\lambda$		2.9939	2.9190	2.3321	2.4186	2.7391
		0.0737	0.0790	0.0784	0.0525	0.1381
κ			-0.1933		-0.6217	
			0.4760		0.0285	
γ				0.8306	0.7193	
				0.0184	0.0157	
ψ						0.9466
						0.0294
Persistence*	0.9950	0.9959	0.9962	0.9961	0.9962	0.9953
Annual Std. Dev.	0.1562	0.1912	0.1949	0.1950	0.1950	0.1688
$MSE	2.6028	1.0910	1.0836	1.0817	1.0214	1.0900
$RMSE	1.6133	1.0445	1.0410	1.0400	1.0106	1.0440

Notes: We estimate the risk-neutral dynamics for each GARCH model directly by fitting the observed option prices using a nonlinear least squares routine to minimize $MSE. Only options in Sample A (June 1, 1988 - May 31, 1991; 8,081 contracts) are used in estimation. Standard errors are reported below each parameter estimate. The bottom of the table reports the risk-neutral volatility persistence and the risk-neutral annualized standard deviation implied by the GARCH parameters. *) For the Box-Cox model persistence refers to the power of volatility rather than volatility itself. We also report the $MSE and $RMSE at the parameter optima.

Table VII.A: Fit of Option Prices from NLS Risk-Neutral Estimates.

Sample A Fit	Simple	Leverage	News	Power	Power&News	Box-Cox	Black-Scholes	PBS(OLS)	PBS(NLS)
Sample $MSE	2.6028	1.0910	1.0836	1.0817	1.0214	1.0900	3.7351	3.6157	2.9307
Sample $RMSE	1.6133	1.0445	1.0410	1.0400	1.0106	1.0440	1.9326	1.9015	1.7119
Sample B Fit									
Sample $MSE	2.4250	0.9777	0.9684	1.0197	1.1571	0.9825	4.0112	3.5982	1.4511
Sample $RMSE	1.5573	0.9888	0.9841	1.0098	1.0757	0.9912	2.0028	1.8969	1.2046

Table VII.B: Comparing Weekly Predictive Accuracy with the Leverage Model

Sample B	Simple	Leverage	News	Power	Power&News	Box-Cox	Black-Scholes	PBS(OLS)	PBS(NLS)
Mean Weekly $RMSE	1.5169	0.9032	0.9002	0.9278	1.0020	0.9062	1.9535	1.7662	1.1298
DM-Test Value	6.9936		-1.4427	1.5817	1.3143	1.2179	15.2805	6.3207	1.7858
P-Value (2-sided)	0.0000		0.1491	0.1137	0.1887	0.2233	0.0000	0.0000	0.0741

Notes: We compute the $MSE and $RMSE using the parameter estimates in Table VI on both Sample A and Sample B. The Sample A numbers for the six GARCH models are identical to those reported in Table VI. We also report Sample A fits for three other models: First, the standard Black-Scholes model with volatility estimated using nonlinear least squares minimizing $MSE. Second, the Practitioner Black-Scholes, PBS(OLS), from Dumas, Fleming and Whaley (1998), who regress implied volatility on a second order polynomial in strike price and time to maturity. Third, PBS(NLS) which estimates the PBS polynomial using nonlinear least squares to minimize $MSE. We keep all parameters constant across the 156 weeks in Sample A. We also report the out-of-sample fits for the nine models in Sample B, using the parameter estimates from Sample A. In Table VII.B we calculate the average of the weekly $RMSE across the 52 weeks in Sample B. Due to the concavity of the square-root function, the average weekly $RMSEs are lower than the Sample $RMSE calculated over all the 52 weeks. We use the weekly $RMSE sequences to test the significance of the difference in fits across models by applying the Diebold-Mariano (1995) test. The DM test is implemented allowing for autocorrelation of up to four weeks in the $RMSE difference sequence. We take the *Leverage* GARCH model to be the benchmark in the pairwise DM tests. The DM test has a standard normal distribution and we report the 2-sided P-values.

Table VIII.A: $RMSE on Sample A from NLS Estimates on Sample A. By Moneyness and Maturity

Simple Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.9616	1.5246	2.2977
.94 < S/X < .97	0.9841	1.5720	2.0373
.97 < S/X < 1.00	0.9155	1.5963	1.9129
1.00 < S/X < 1.03	0.9991	1.8677	2.1563
1.03 < S/X < 1.06	1.2191	2.1494	1.7587
1.06 < S/X	1.0448	1.8351	2.1288

Leverage Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6470	0.9089	1.3186
.94 < S/X < .97	0.7692	0.9915	1.3204
.97 < S/X < 1.00	0.8277	1.0415	1.3636
1.00 < S/X < 1.03	0.8304	1.1628	1.3381
1.03 < S/X < 1.06	0.8701	1.2720	1.2241
1.06 < S/X	0.8033	1.1245	1.3125

News Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6531	0.9123	1.3120
.94 < S/X < .97	0.7705	0.9917	1.3126
.97 < S/X < 1.00	0.8236	1.0383	1.3575
1.00 < S/X < 1.03	0.8221	1.1579	1.3354
1.03 < S/X < 1.06	0.8635	1.2674	1.2168
1.06 < S/X	0.8004	1.1218	1.3093

Power Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6281	0.8915	1.2989
.94 < S/X < .97	0.7477	0.9878	1.3223
.97 < S/X < 1.00	0.8174	1.0478	1.3682
1.00 < S/X < 1.03	0.8329	1.1715	1.3419
1.03 < S/X < 1.06	0.8784	1.2642	1.2481
1.06 < S/X	0.8076	1.1059	1.3169

News&Power Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6322	0.9033	1.2228
.94 < S/X < .97	0.7377	1.0099	1.2986
.97 < S/X < 1.00	0.7993	1.0601	1.3427
1.00 < S/X < 1.03	0.7839	1.1573	1.3421
1.03 < S/X < 1.06	0.8193	1.1995	1.2692
1.06 < S/X	0.7640	1.0129	1.3344

Box-Cox Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6446	0.9095	1.3184
.94 < S/X < .97	0.7659	0.9951	1.3193
.97 < S/X < 1.00	0.8254	1.0441	1.3629
1.00 < S/X < 1.03	0.8303	1.1648	1.3343
1.03 < S/X < 1.06	0.8723	1.2716	1.2233
1.06 < S/X	0.8052	1.1233	1.3086

Notes: We report $RMSE from various GARCH options valuation models on Sample A, which denotes June 1, 1988 - May 31, 1991. The $RMSE is reported for three maturity bins and six moneyness bins. The parameters in the GARCH option valuation models are estimated minimizing $MSE on Sample A itself.

Table VIII.B: $RMSE on Sample B from NLS Estimates on Sample A. By Moneyness and Maturity

Simple Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.9970	1.9700	2.7700
.94 < S/X < .97	1.1400	1.8800	1.8100
.97 < S/X < 1.00	0.9930	1.3800	1.1200
1.00 < S/X < 1.03	0.5710	1.1200	1.3600
1.03 < S/X < 1.06	0.7550	1.4200	1.9200
1.06 < S/X	0.9460	1.8200	2.3700

Leverage Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.5350	0.9330	1.4000
.94 < S/X < .97	0.7840	1.1300	1.3700
.97 < S/X < 1.00	0.8880	1.1900	1.2000
1.00 < S/X < 1.03	0.7390	0.9870	1.0200
1.03 < S/X < 1.06	0.6400	0.9710	0.9520
1.06 < S/X	0.8010	1.1000	0.8880

News Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.5458	0.9376	1.3806
.94 < S/X < .97	0.7941	1.1359	1.3354
.97 < S/X < 1.00	0.8978	1.1883	1.1797
1.00 < S/X < 1.03	0.7395	0.9830	0.9948
1.03 < S/X < 1.06	0.6353	0.9657	0.9389
1.06 < S/X	0.7979	1.0971	0.8786

Power Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.5498	0.9624	1.4468
.94 < S/X < .97	0.8083	1.1729	1.4406
.97 < S/X < 1.00	0.9241	1.2323	1.2659
1.00 < S/X < 1.03	0.7595	1.0090	1.1040
1.03 < S/X < 1.06	0.6327	0.9609	1.0110
1.06 < S/X	0.7849	1.0523	0.9251

News&Power Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.6792	1.0682	1.3759
.94 < S/X < .97	0.9694	1.3163	1.4072
.97 < S/X < 1.00	1.1400	1.3925	1.3475
1.00 < S/X < 1.03	0.9278	1.1518	1.1585
1.03 < S/X < 1.06	0.6577	1.0184	1.1197
1.06 < S/X	0.7260	0.9629	1.0038

Box-Cox Model

	DTM<60	60<DTM<180	180<DTM
S/X < .94	0.5411	0.9414	1.4098
.94 < S/X < .97	0.7894	1.1407	1.3742
.97 < S/X < 1.00	0.8941	1.1947	1.2105
1.00 < S/X < 1.03	0.7396	0.9868	1.0272
1.03 < S/X < 1.06	0.6373	0.9654	0.9583
1.06 < S/X	0.7984	1.0921	0.8886

Notes: We report $RMSE from various GARCH options valuation models on Sample B, which denotes June 1, 1991 - May 31, 1992. The $RMSE is reported for three maturity bins and six moneyness bins. The parameters in the GARCH option valuation models are estimated minimizing $MSE on Sample A, which denotes June 1, 1988 - May 31, 1991.

Table IX: Fit of Option Prices from Weekly NLS Estimates on Sample B

A. Current Week

	Simple	Leverage	News	Power	Power&News	Box-Cox	Black-Scholes	PBS(OLS)	PBS(NLS)
Sample $RMSE	1.4639	0.4223	0.3963	0.3882	0.3802	0.4197	1.4859	0.9257	0.3557

B. One Week Ahead

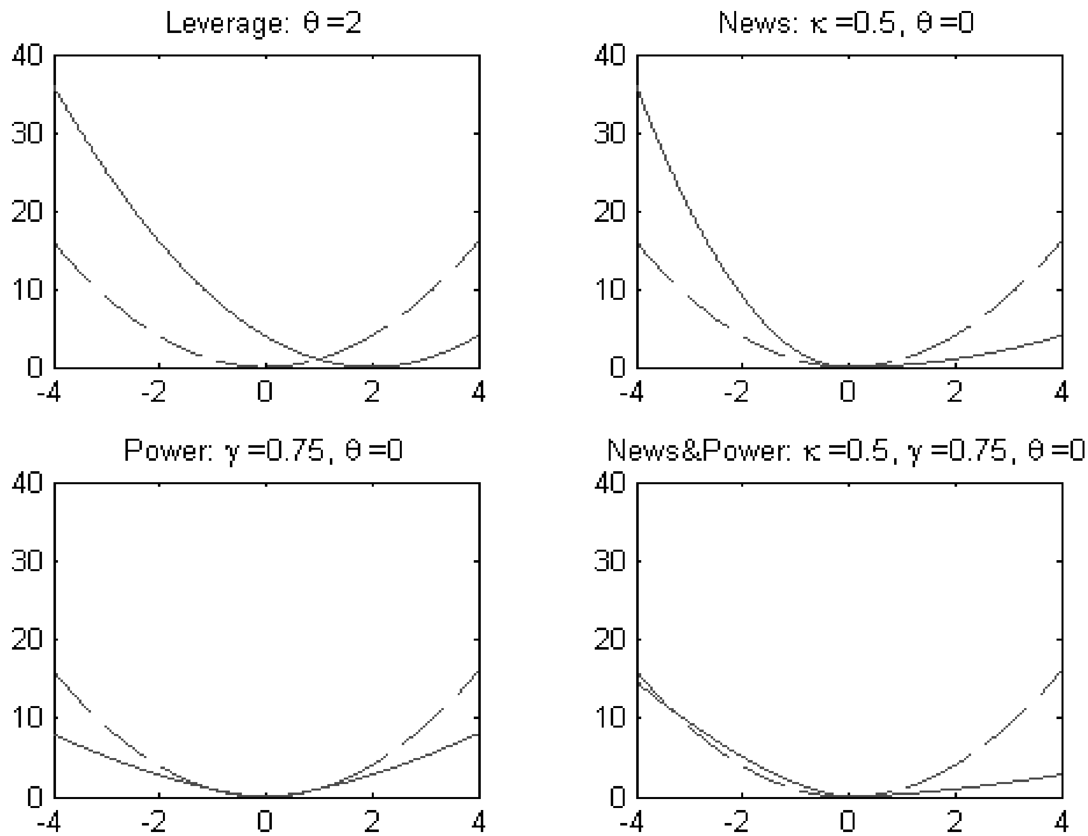
	Simple	Leverage	News	Power	Power&News	Box-Cox	Black-Scholes	PBS(OLS)	PBS(NLS)
Sample $RMSE	1.5540	0.7081	0.7348	0.7135	0.7454	0.6978	1.5656	1.0952	0.6351

C. Comparing Weekly Predictive Accuracy with the Leverage Model

	Simple	Leverage	News	Power	Power&News	Box-Cox	Black-Scholes	PBS(OLS)	PBS(NLS)
Mean Weekly $RMSE	1.5295	0.5838	0.5980	0.6104	0.6558	0.5875	1.5429	0.8764	0.5395
DM Test Value	17.2292		1.5564	6.6388	1.9602	0.1746	20.5065	3.8026	-1.6054
P-Value (2-sided)	0.0000		0.1196	0.0000	0.0500	0.8614	0.0000	0.0001	0.1084

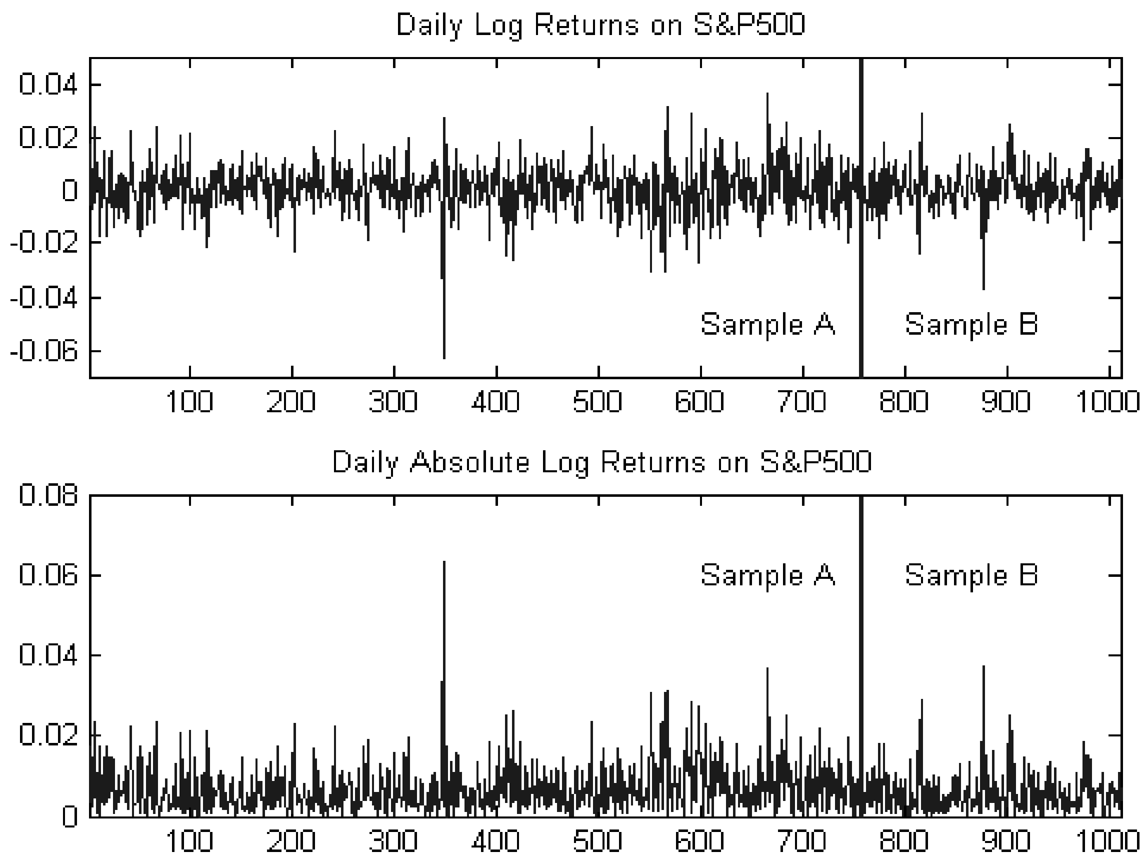
Notes: On every Wednesday in Sample B, we estimate a new set of GARCH parameters for each model based only on options data from that particular day and returns from the past 250 days. We then report the Sample $RMSE for all the current Wednesdays (Current Week) as well as the Sample $RMSE for the same Wednesdays using the previous Wednesdays' GARCH estimates (One Week Ahead). We also report Sample $RMSE for three other models: First, the standard Black-Scholes model with volatility estimated using nonlinear least squares minimizing $MSE. Second, the Practitioner Black-Scholes, PBS(OLS), from Dumas, Fleming and Whaley (1998), who regress implied volatility on a second order polynomial in strike price and time to maturity. Third, PBS(NLS) which estimates the PBS polynomial using nonlinear least squares to minimize $MSE. In these three models, the parameters are reestimated weekly as well. In Table IX.C we calculate the mean of the weekly $RMSE across the 52 weeks in Sample B. Due to the concavity of the square-root function, the average weekly $RMSEs are lower than the Sample $RMSEs calculated over all the 52 weeks. We use the weekly $RMSE sequences to test the significance of the difference in fits across models by applying the Diebold-Mariano (1995) test. The DM test is implemented allowing for autocorrelation of up to four weeks in the $RMSE difference sequence. We take the *Leverage* GARCH model to be the benchmark in the pairwise DM tests. The DM test has a standard normal distribution and we report the 2-sided P-values.

Figure I: Stylized Innovation Functions, f(z). Various Models



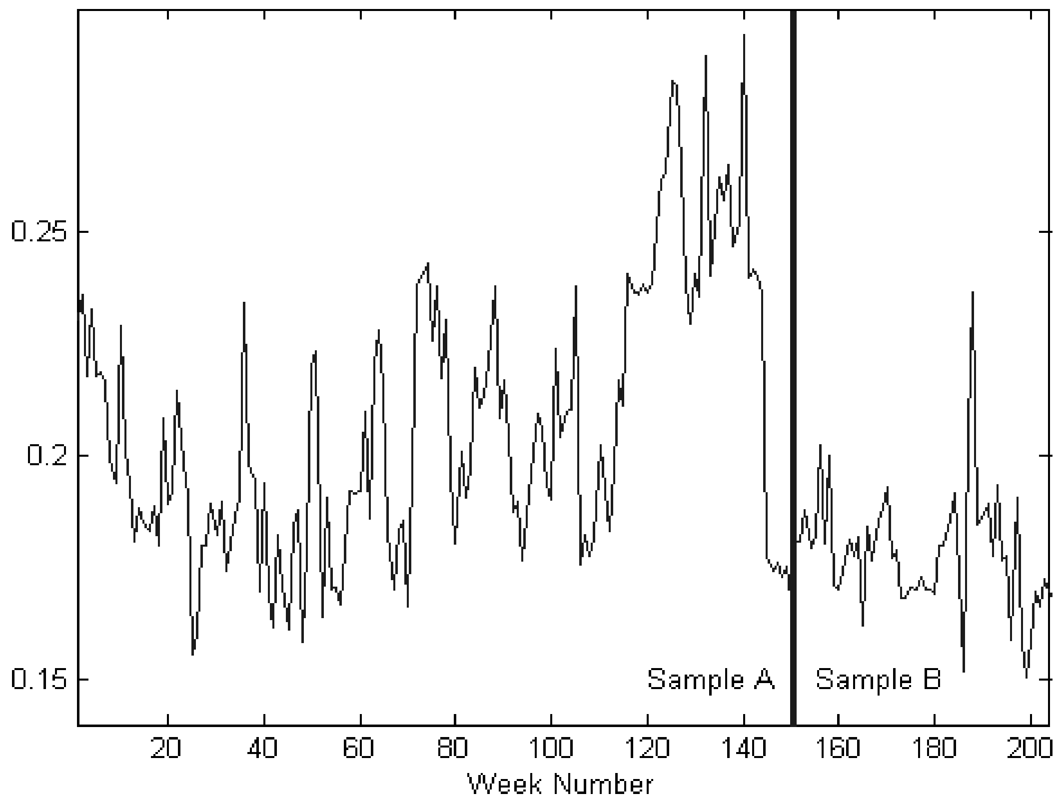
Notes: The innovation function from each model (solid line) is superimposed on the symmetric squared innovation function from the *Simple* model (dashed line). The innovation function conveys the impact on volatility from a particular standard normal innovation, which is given on the horizontal axis.

Figure II: S&P 500 Returns Across Option Sample Period



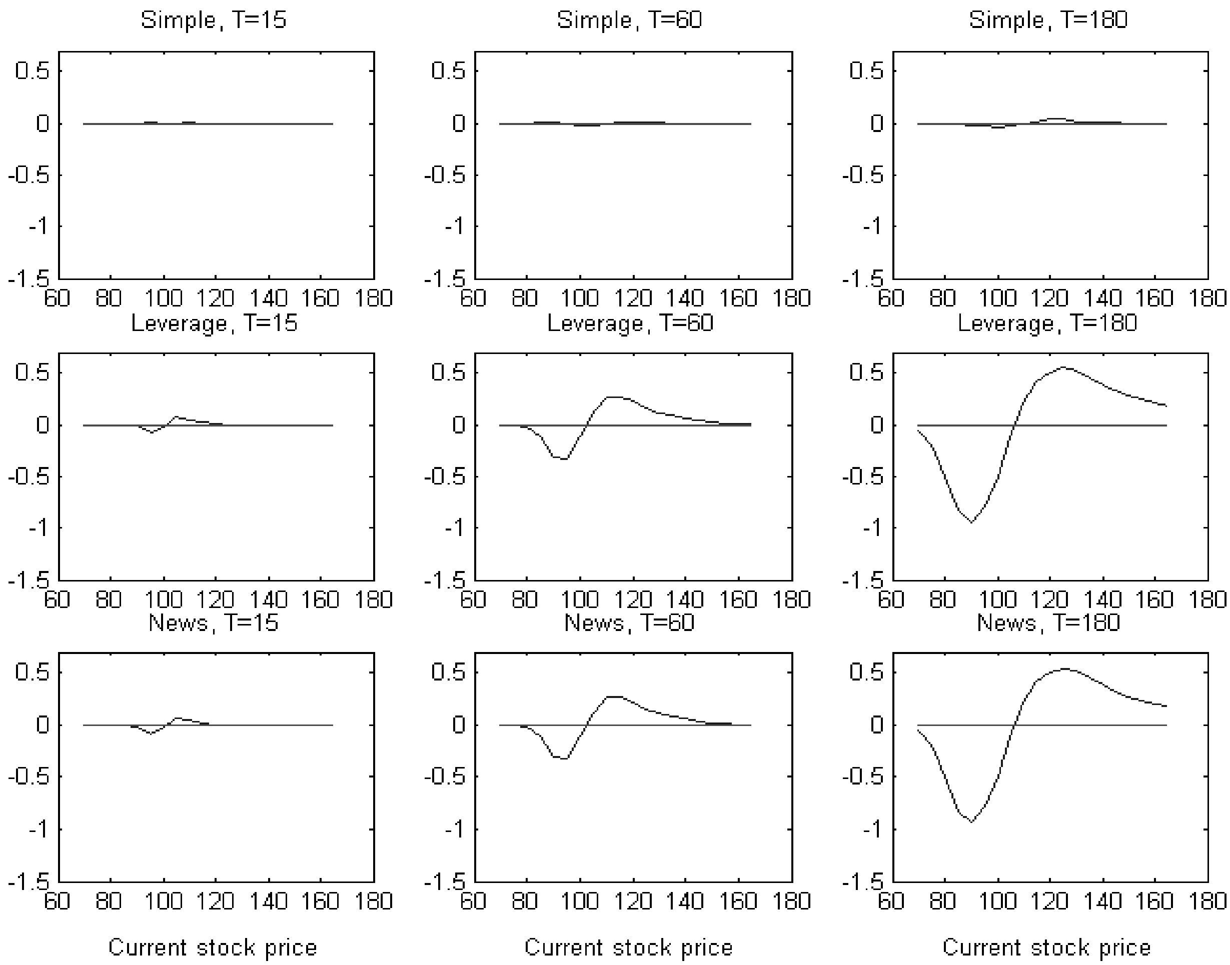
Notes: The top panel shows the daily log returns on the S&P500 index. The data is shown from June 1, 1988 through May 31, 1992. The vertical line delimits Sample A, which corresponds to June 1, 1988 through May 31, 1991, from Sample B, which corresponds to June 1, 1991 through May 31, 1992. The bottom panel shows the daily absolute returns on the S&P500 index for the same periods.

Figure III: Average Implied Volatility of S&P 500 Index Options



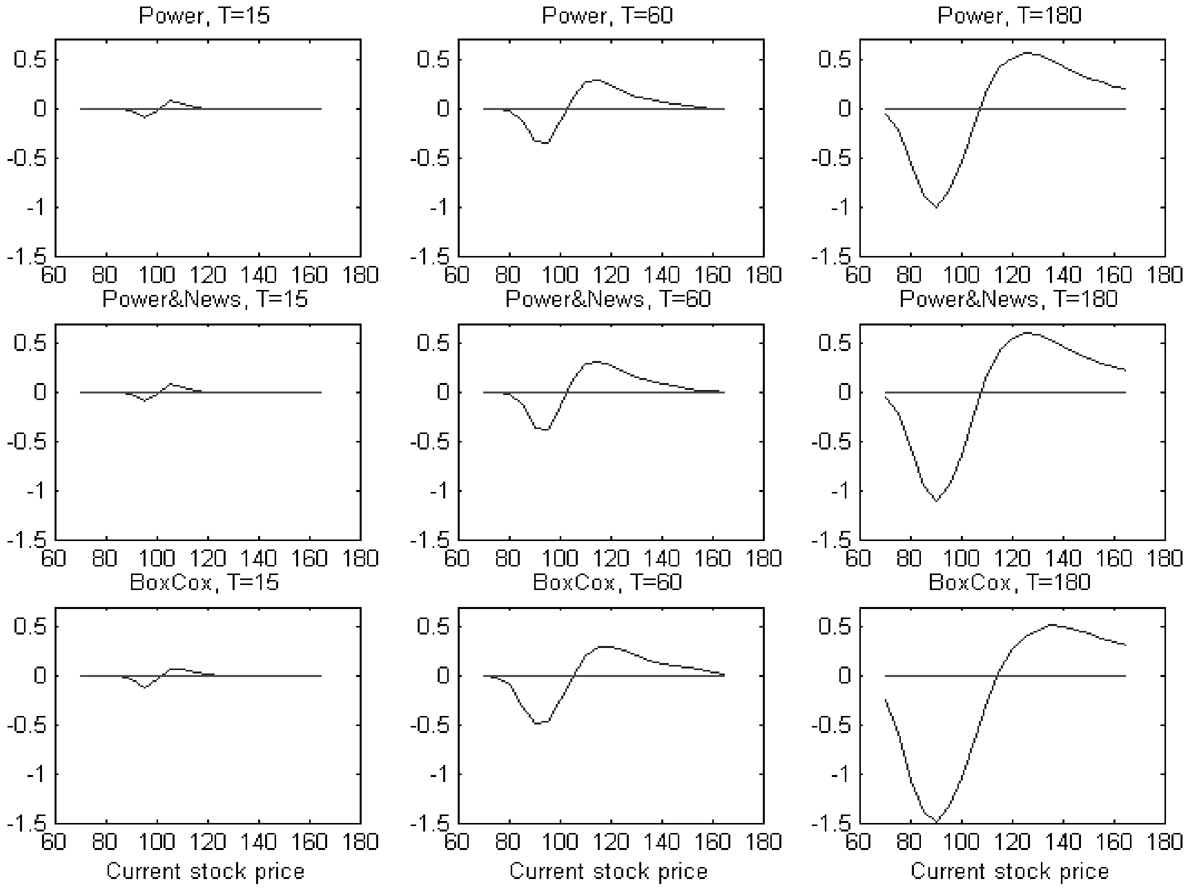
Notes: On each Wednesday, we plot the simple average implied Black-Scholes volatility across the S&P500 index option contracts observed at the close of trading. The average implied volatilities are shown for each Wednesday from June 1, 1988 through May 31, 1992. The vertical line delimits Sample A, which corresponds to June 1, 1988 through May 31, 1991, from Sample B, which corresponds to June 1, 1991 through May 31, 1992.

Figure IV.A: Model Price Less Black-Scholes Price. Various Models



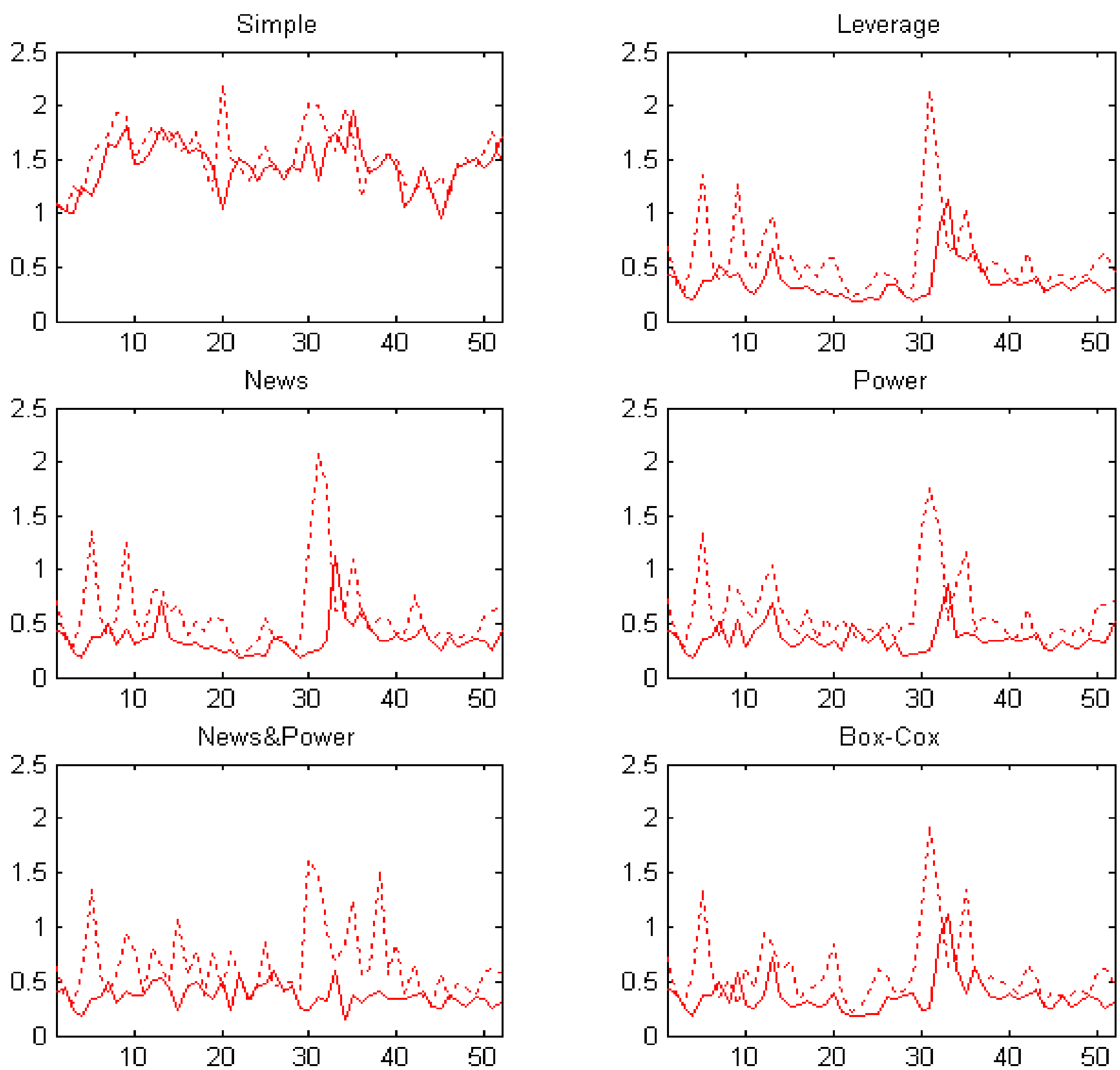
Notes: Each line shows the option price from a particular GARCH model less the Black-Scholes price of the same option. Horizontal lines are plotted at zero. The horizontal axis denotes the current price of the underlying asset, and the strike price is fixed at 100. The parameters of each GARCH model are set equal to the NLS estimates from Table VI. The initial conditional variance in the GARCH model and the volatility in the Black-Scholes model are both set equal to the unconditional value implied by the parameters in the GARCH model. The columns of plots correspond to 15, 60 and 180 days to maturity respectively. Each row of plots corresponds to a particular GARCH model.

Figure IV.B: Model Price Less Black-Scholes Price. Various Models (cont.)



Notes: Each line shows the option price from a particular GARCH model less the Black-Scholes price of the same option. Horizontal lines are plotted at zero. The horizontal axis denotes the current price of the underlying asset, and the strike price is fixed at 100. The parameters of each GARCH model are set equal to the NLS estimates from Table VI. The initial conditional variance in the GARCH model and the volatility in the Black-Scholes model are both set equal to the unconditional value implied by the parameters in the GARCH model. The columns of plots correspond to 15, 60 and 180 days to maturity respectively. Each row of plots corresponds to a particular GARCH model.

Figure V: In and Out of Sample $RMSE. Weekly Estimates



Notes: On every Wednesday in Sample B, we estimate a new set of GARCH parameters for each model based only on options data from that particular day and returns from the past 250 days. We then plot the $RMSE for the current Wednesday (In-Sample) as well as the $RMSE for the same Wednesday using last Wednesday's GARCH estimates (Out-of-Sample). The solid lines denote In-Sample $RMSE for each model and the dashed lines denote Out-of-Sample $RMSE.