Conditionally Fat-Tailed Distributions and the Volatility Smile in Options

Jin-Chuan Duan*

This Draft: February 22, 1999

Abstract

Volatility smile refers to the convex pattern in the Black-Scholes implied volatility when plotted against the option's exercise price. This paper provides a theoretical analysis of volatility smile using a generalized GARCH option pricing model in which the asset return innovation is conditionally leptokurtic and skewed. This generalization accommodates the empirical evidence on conditional leptokurtosis and also allows for other features of financial data such as long memory in volatility. Our analysis using this generalized GARCH option pricing model suggests that conditional leptokurtosis, leverage effect and asset risk premium together determine the shape of the volatility smile.

Key words: GARCH, Volatility Smile, Leptokurtosis, Leverage Effect, Long Memory, Risk Premium, Local Risk Neutralization.

JEL classification: C32, G1

*Correspondence to: J.C. Duan, Department of Finance, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong; Tel: 852-2358 7671; Fax: 852-2358 1749; E-mail: jcduan@ust.hk; Web: http://www.bm.ust.hk/~jcduan. Financial support provided by the Research Grants Council of Hong Kong and the SSHRC of Canada is gratefully acknowledged. The author also acknowledges the support received as Senior Wei Lun Fellow at the Hong Kong University of Science and Technology.

1 Introduction

The seminal work of Black and Scholes (1973) and Merton (1973) on option pricing theory, commonly known as the Black and Scholes model, has not only spawned a huge literature on derivative contracts but also transformed the business practice in the financial industry. This influential option pricing model is not without problems, however. Many empirical studies, including the empirical work in Black and Scholes (1972), have shown that the Black and Scholes model exhibits systematic pricing biases. More recent empirical studies typically focus on the pricing biases in terms of implied volatilities, and the bias phenomenon is referred to as "volatility smile." Volatility smile pertains to the phenomenon that the Black-Scholes implied volatilities for stock call options often exhibit a downward sloping, convex pattern when plotted against their exercise prices.[1] This persistent feature of option price data contradicts the prediction of the Black-Scholes model, because the theory suggests a horizontal line for the implied volatility. The volatility smile phenomenon motivated many extensions of the Black and Scholes model.[2]

Duan (1995) proposed a different approach to extending the Black-Scholes model, in which the underlying asset return is assumed to follow the GARCH dynamic. The ARCH family of models was first introduced by Engle (1982), and the most notable generalization was by Bollerslev (1986). The generalized ARCH (GARCH) model has in the recent years gained prominence for modeling financial time series. Financial data typically exhibit time-varying volatility, which has particularly important implications for pricing derivative contracts. Since derivative claims are sensitive to volatility, it is of paramount importance to use a good model for the volatility dynamics of financial assets. The derivation of the GARCH option pricing model in Duan (1995) utilizes the standard competitive equilibrium asset pricing framework. Recently, Kallsen and Taqqu (1998) have come up with a continuous-time version of the model so that an arbitrage-free argument can be used to establish the same pricing result. The numerical methods for this class of option pricing models have been emerging; for example, Hanke (1997), Duan, *et al* (1998), Ritchken and Trevor (1998), Duan and Simonato (1998a&b), and Heston and Nandi (1998). The GARCH option pricing model has so far experienced some empirical successes: see, for example, Amin and Ng (1994), Heynen, *et al* (1994), Duan (1996a), and Heston and Nandi (1998).

One of the key assumptions for the GARCH option pricing model is conditional normality. This assumption is needed to ensure a distributional invariance when the asset return innovation undergoes risk neutralization. This assumption is, however, at odds with the direct empirical evidence on the asset return dynamic as summarized, for example, in Bollerslev, *et al* (1992). A typical finding concerning the return characteristics is that one-period asset return, conditional on the most up-to-date information, continues to exhibit a fat-tail behavior. This fat-tail phenomenon is sometimes referred to as conditional leptokurtosis. This evidence raises a question concerning the appropriateness of the conditional normality assumption. Since the current GARCH option pricing theory does not permit conditional leptokurtosis, it is unclear as to what effects conditional lep-

[1] The volatility smile phenomenon is also present in currency option markets. Although the curve is still convex, it sometimes slopes upwards or has no skew at all.

[2] See, for example, the earlier papers by Cox (1975), Merton (1976), Geske (1979), Rubinstein (1983), and more recent ones by Hull and White (1987), Johnson and Shanno (1987), Scott (1987), Wiggins (1987), Melino and Turnbull (1990), Stein and Stein (1991), Madan and Milne (1991) and Heston (1993), and many others.

tokurtosis could have on volatility smile. This paper generalizes the GARCH option pricing model by specifically allowing the asset return to have a conditionally fat-tailed and skewed distribution. The generalized option pricing theory is then used to study the effect of conditional leptokurtosis on volatility smile.

Empirical evidence in the literature has recently been emerging to indicate a long memory property in the volatility dynamics. This long memory phenomenon was reported, for example, in Ding, et al (1993). Recently, long memory in volatility has been modeled as a fractionally differenced system, which allows volatility innovation to die out at a hyperbolic, instead of exponential rate.[3] Crouhy and Rockinger (1997) recently recognized that the hysteresis effect is also present in the conditional volatility of asset return. They proposed the use of a GARCH model with hysteresis effect to describe asset returns. Our generalization of the GARCH option pricing theory incorporates these two features of financial data. More importantly, our generalization encompasses any imaginable parametric form of the GARCH dynamic.

Our strategy for developing the generalized version of the GARCH option pricing model begins with a search for a transformation that is capable of converting the fat-tailed and/or skewed random variables into normally distributed ones. Applying the local risk-neutral valuation principle to these transformed variables, we have effectively generalized Duan's (1995) model. We also adopt the general error function as the conditional distribution function to demonstrate the effect of conditional leptokurtosis. The numerical analysis is carried out by varying the tail-fatness parameter of the generalized error distribution. The results suggest that the shape of volatility smile is jointly determined by the extent of conditional tail-fatness, the degree of leverage effect and the magnitude of asset risk premium. Our findings also offer several interesting theoretical predictions useful for future empirical studies.

2 Local risk neutralization under conditionally fat-tailed and skewed distributions

This section provides a development of the option pricing theory for the GARCH model with non-normally distributed innovations. Consider an asset in a discrete-time economy with price at time t denoted by X_t and dividend yield from time $(t-1)$ to t by d_t. The one-period one plus the rate of return for this asset, i.e., $\frac{X_{t+1}}{X_t} + d_{t+1}$, is assumed to obey, under the data generating probability measure P, the following dynamic:

$$B(\frac{X_{t+1}}{X_t} + d_{t+1}; \delta) = \mu_{t+1} + \sqrt{h_{t+1}}\varepsilon_{t+1} \quad (1)$$

$$h_{t+1} = f(h_\tau, \varepsilon_\tau; -\infty < \tau \leq t, \theta) \quad (2)$$

$$\varepsilon_{t+1}|\mathcal{F}_t \sim D(0,1) \quad (3)$$

$$B(z;\delta) = \frac{z^\delta - 1}{\delta} \text{ for } z > 0 \text{ and } \delta \geq 0 \quad (4)$$

[3] See, for example, Baillie, *et al* (1996), Bollerslev and Mikkelsen (1996) and McCurdy and Michaud (1996).

where $\mathcal{F}_t$ is the information set containing all information up to and including time t. In equations, $D(0,1)$ denotes some distribution function that is continuous over its support with mean 0 and variance 1, where the support is assumed to be a connected set. ($D(a)$ will sometimes be used to denote the cumulative distribution function evaluated at a). The set of parameters, θ, governs the volatility dynamic. The conditional mean, μ_{t+1}, is any predictable process which means measurable with respect to the information set $\mathcal{F}_t$. The conditional variance, h_{t+1}, is by definition measurable with respect to the information set $\mathcal{F}_t$. Finally, δ is restricted to take on values that ensure a finite conditional second moment for $\frac{X_{t+1}}{X_t} + d_{t+1}$ under measure P.

$B(\cdot;\delta)$ is the well-known Box-Cox transformation so that $B(\frac{X_{t+1}}{X_t}+d_{t+1};0) = \ln(\frac{X_{t+1}}{X_t}+d_{t+1})$ and $B(\frac{X_{t+1}}{X_t} + d_{t+1};1) = \frac{X_{t+1}-X_t}{X_t} + d_{t+1}$. The model in (1)-(4) describes the continuously compounded return when $\delta = 0$, and describes the simple rate of return when $\delta = 1$. The value of δ cannot be arbitrary, however. If, for example, ε_{t+1} has a conditionally fat-tailed distribution such as the t-distribution, then δ cannot equal zero because this combination makes the expected simple rate of return unbounded (clearly not a sensible situation). Note that δ should not be viewed as an unknown parameter; rather it is a modeling choice made by the user.

Our general specification allows the asset return innovation to have a conditionally fat-tailed and skewed distribution. This feature is important because mounting empirical evidence has indicated that high frequency data continues to exhibit conditional tail-fatness after allowing for the GARCH effect (see Bollerslev, *et al*, 1992). Conditional leptokurtosis was directly modeled as a t-distribution in Bollerslev (1987). Alternatively, the generalized error distribution can be used to model conditional leptokurtosis as in Nelson (1991).

It is quite common in the literature to adopt a special functional form for $f(\cdot)$. In the case of the non-linear asymmetric GARCH(1,1) model of Engle and Ng (1993), $h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\varepsilon_t - c)^2$. If we let $c = 0$, the model becomes the linear GARCH(1,1) of Bollerslev (1986) and Taylor (1986). To obtain the exponential GARCH(1,1) model of Nelson (1991), one specifies $h_{t+1} = \exp\{\beta_0 + \beta_1 ln(h_t) + \beta_2(|\varepsilon_t| + \gamma\varepsilon_t)\}$. For the version of GARCH model proposed by Glosten, *et al* (1993), one simply specifies $h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t \varepsilon_t^2 + \beta_3 h_t \max(0, -\varepsilon_t)^2$. More general GARCH models such as Higgins and Bera (1992), Hentschel (1995) and Duan (1997) are also special cases of the general dynamic in (1)-(4).

The volatility dynamic in (2) also includes the long memory model of Baillie, *et al* (1996). To obtain their FIGARCH(1,d,1) model, we specify $h_{t+1} = \beta_0 + \beta_1 h_t + [1 - \beta_1 L - (1-\theta L)(1-L)^d] h_{t+1}\varepsilon_{t+1}^2$ with L denoting the lag operator and where $0 \le d \le 1$. The long memory models of Bollerslev and Mikkelsen (1996) and McCurdy and Michaud (1996) are also special cases of the general volatility dynamic. The GARCH model with hysteresis effect proposed by Crouhy and Rockinger (1997) can be captured by the dynamic in (2). A simple version of the NGARCH with hysteresis can, for example, be formulated as $h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\varepsilon_t - c_1)^2 + (\beta_3 \chi_{\{\varepsilon_\tau > c_2, t-k+1 \le \tau \le t\}} + \beta_4 \chi_{\{\varepsilon_\tau < c_2, t-k+1 \le \tau \le t\}})(\sum_{\tau=t-k+1}^{t} \varepsilon_\tau - kc_2)^2$ where $\chi_{\{\cdot\}}$ denotes the indicator function.

The GARCH option pricing model as derived in Duan (1995) critically relies on the assumption that the asset returns exhibit conditionally lognormal distributions (or conditional normality for the continuously compounded return).[4] With conditionally leptokurtic asset return innovations,

[4]Kallsen and Taqqu (1998) connected the discrete-time GARCH process by many geometric Brownian motion

one seems to need an entirely different approach to option pricing. The key to our derivation in this paper is to recognize that the approach of Duan (1995) remains essentially valid when conditional normality is restored by transformation.

Let $\phi(\cdot)$ denote the standard normal distribution function. A one-to-one transformation can be designed in such a way that it converts ε_t to a standard normal random variable.[5] The desired transformation is given in the following proposition.

Proposition 1. (Conditional normality by transformation)

$$\Psi(\varepsilon_t) \equiv \phi^{-1}[D(\varepsilon_t)] \tag{5}$$

is, conditional on $\mathcal{F}_{t-1}$, a standard normal random variable with respect to measure P.

Since the proof for Proposition 1 is fairly intuitive, we skip the formal proof and instead provide a sketch of the idea. The idea resembles the process of generating normal random numbers. Standard normal random numbers can be generated by plugging uniformly distributed random numbers into the inverse of the standard normal distribution function. Since ε_t has a continuous distribution over its connected support, $D(\varepsilon_t)$ must be a P-uniform random variable with support $[0,1]$. Inverting the uniformly distributed random variable, $D(\varepsilon_t)$, by the standard normal distribution function thus creates a P-standard normal random variable. The transformation in (5) is a one-to-one relationship because of the monotonicity of $\phi(\cdot)$ and the distributional assumptions for ε_t.

In the continuous-time complete market option pricing literature, the pricing measure is obtained by the use of the Girsanov theorem and by the assumption of continuous trading (see, for example, Harrison and Pliska, 1981). In the discrete-time GARCH framework, Duan (1995) characterized the conditions under which the equilibrium price measure can be regarded as a locally risk-neutralized probability. Following the approach of Duan (1995), we modify the local risk-neutral valuation principle to suit the current setting of conditional non-normality. We consider two arbitrary finite integer time points t_l and t_u such that $t_l < t_u$. These two finite time points are used to ensure the existence of the locally risk-neutralized pricing measure. Since they are arbitrary, any option pricing horizon of interest can be covered by a suitable choice of t_l and t_u.

Let $U(C_t)$ and C_t denote the strictly increasing utility function of the representative agent and the aggregate consumption at time t, respectively, in a time separable and additive exchange economy. The impatience factor is denoted by parameter ρ. The standard expected utility maximization argument leads to the following Euler equation:

$$W_{t-1} = E^P[e^{-\rho}\frac{U'(C_t)}{U'(C_{t-1})}(W_t + D_t)|\mathcal{F}_{t-1}], \tag{6}$$

for any traded financial asset whose price and dividend at time t are denoted by W_t and D_t. To ensure that existence of the expectation the price and dividend corresponding to the asset and the

processes. More specifically, the price dynamic between any two adjacent discrete time points is a constant volatility diffusion. The one-period asset return in their model is thus conditionally lognormally distributed. Amin and Ng (1994) directly assumed a jointly lognormal conditional distribution for the asset return and state price. Their approach clearly precludes conditional leptokurtosis.

[5] Thanks to Ronald Gallant for pointing out such a transformation.

marginal rate of substitution are required to have finite second moments, conditional on $\mathcal{F}_{t_l}$ and under measure P. We are now ready to define the equilibrium pricing measure.

Proposition 2. (Equilibrium (or locally risk-neutralized) pricing measure)

Define a measure Q over the time interval $[t_l, t_u]$ as

$$dQ = \exp[-\rho(t_u - t_l) + \sum_{s=t_l+1}^{t_u} (r_{s-1,s} + \ln \frac{U'(C_s)}{U'(C_{s-1})})]dP, \tag{7}$$

where $r_{t,t+1}$ denotes the one-period risk-free interest rate (continuously compounded) at time t. Then, measure Q satisfies two conditions:

(i) Q is a probability measure;

(ii) for $t_l + 1 \leq t \leq t_u$, consider any traded asset whose time-t price is denoted by W_t and dividend at time t by D_t. If $E^P\{(W_t + D_t)^2|\mathcal{F}_{t_l}\} < \infty$ and $E^P\{(\frac{U'(C_t)}{U'(C_{t-1})})^2|\mathcal{F}_{t_l}\} < \infty$, then its time-$(t-1)$ price equals

$$W_{t-1} = e^{-r_{t-1,t}} E^Q \left\{(W_t + D_t)|\mathcal{F}_{t-1}\right\}. \tag{8}$$

The equilibrium pricing measure is equivalent to the notion of the Arrow-Debreu state price density, although the market analyzed in this paper is inherently incomplete. In a complete-market setup, the specific form of the utility function is not important because the traded asset prices are sufficient for identifying the state price density. In our setup, however, one must rely on further assumption on preferences and aggregate consumption to characterize the pricing system. Here, we generalize the local risk-neutral valuation principle established in Duan (1995).

Definition. (Generalized local risk-neutral valuation relationship, GLRNVR)

The equilibrium pricing measure Q, defined over $[t_l, t_u]$, is said to satisfy the generalized local risk-neutral valuation relationship (GLRNVR) if, for $t_l \leq t \leq t_u - 1$,

(i) measure Q is mutually absolutely continuous with respect to measure P;

(ii) there exists a predictable process λ_t such that $\Psi(\varepsilon_{t+1}) + \lambda_{t+1}$, conditional on $\mathcal{F}_t$, is a standard normal random variable with respect to measure Q;

(iii) $E^Q\{\frac{X_{t+1}}{X_t} + d_{t+1}|\mathcal{F}_t\} = \exp(r_{t,t+1})$.

The first condition is due to a technical requirement that two probability measures agree on the events of measure zero. The second condition states that risk-neutralization has an invariance property which means that the nature of the distribution for the transformed innovation remains unchanged. Risk neutralization merely causes the transformed innovation to undergo a shift in mean. The magnitude of mean shift, i.e., λ_t, is determined by the third condition. The precise expression of λ_t depends on the transformation $\Psi(\cdot)$, which in turn is determined by the specific distribution adopted for $D(0,1)$. Mean shift is the standard result in option pricing when the conditional distribution is normal. For the familiar diffusion model, the innovation is driven by a Wiener process with locally normal distributions. The risk neutralization for that class of models is typically carried out by absorbing the risk premium into the innovation term. The standard martingale pricing result then suggests that the newly formed innovation term is again a Wiener process (relative to the martingale measure) by the Girsanov theorem. Our condition is hence similar to the standard result except that the absorption is carried out only after transforming the innovation term into a standard normal random variable.

One may argue that the GLRNVR is a sensible generalization of the traditional risk-neutral valuation concept because the definition has an intuitive appeal on its own. We choose instead to provide a classical equilibrium justification to the GLRNVR in the following proposition.

Proposition 3. (Sufficient conditions for the GLRNVR)

If the representative agent is an expected utility maximizer and the utility function is time separable and additive, then the GLRNVR holds under any of the following three conditions:

(i) the utility function is of constant relative risk aversion and changes in the logarithmic aggregate consumption, i.e., $(\ln C_t - \ln C_{t-1})$, and the transformed asset return innovation, i.e., $\Psi(\varepsilon_t)$, follow a P-bivariate normal distribution, conditional on $\mathcal{F}_{t-1}$;

(ii) the utility function is of constant absolute risk aversion and changes in the aggregate consumption, i.e., $(C_t - C_{t-1})$, and the transformed asset return innovation, i.e., $\Psi(\varepsilon_t)$, follow a P-bivariate normal distribution, conditional on $\mathcal{F}_{t-1}$;

(iii) the utility function is linear.

Proof: See Appendix.

We have, by adding Proposition 2 to the development, made explicit the equilibrium valuation process for the contingent claims that was implicit in the GARCH option pricing theory developed in Duan (1995). The GLRNVR is a generalization of the local risk-neutral valuation principle of Duan (1995), which is in turn a generalization of the risk-neutral valuation relationship developed by Rubinstein (1976) and Brennan (1979). Our style of proof differs considerably from that of Rubinstein (1976) and Brennan (1979), which is due to the complexity related to the time-varying volatility. The message is, however, similar to that of Rubinstein (1976) and Brennan (1979). That is, even if the asset market is incomplete, one is still able to characterize the pricing system by avoiding the difficult task of identifying the preference parameter, and therefore retain to some extent the simplicity of Arrow-Debreu pricing. Although some conditions on the utility function and the aggregate consumption are imposed, there is no need to be particularly specific about their parameter values. For example, the sufficient condition in (i) of Proposition 3 does not need the knowledge of either the relative risk aversion coefficient or the distributional parameters of the aggregate consumption.

Assuming that the GLRNVR holds, the asset return process can be characterized by a simple risk-neutralized dynamic. The following proposition results from a substitution using the system in (1)-(4) and the definition of the GLRNVR. Since the proof is somewhat obvious, it is skipped here.

Proposition 4. (Option pricing system)

Assume that the GLRNVR holds. For $t_l \leq t \leq t_u - 1$,

$$B(\frac{X_{t+1}}{X_t} + d_{t+1}; \delta) = \mu_{t+1} + \sqrt{h_{t+1}}\Psi^{-1}(Z_{t+1} - \lambda_{t+1}) \quad (9)$$

$$h_{t+1} = f(h_\tau, \varepsilon_\tau; -\infty < \tau \leq t, \theta) \quad (10)$$

$$\varepsilon_\tau = \Psi^{-1}(Z_\tau - \lambda_\tau) \text{ if } \tau \geq t_l + 1 \quad (11)$$

where Z_{t+1}, conditional on $\mathcal{F}_t$, is a Q-standard normal random variable. Moreover, λ_{t+1} is the solution to

$$E^Q\{B^{-1}(\mu_{t+1} + \sqrt{h_{t+1}}\Psi^{-1}(Z_{t+1} - \lambda_{t+1}); \delta)|\mathcal{F}_t\} = \exp(r_{t,t+1}). \quad (12)$$

Propositions 2 and 4 together provide an operational basis for pricing any derivative claim written on X_T. Specifically, for any contingent claim $g(X_T)$, one can be certain that its time-t value equals $e^{-r(T-t)}E^Q\{g(X_T)|\mathcal{F}_t\}$ by part (ii) of Proposition 2 and the law of iterated expectations, if the interest rate is assumed to be constant. Since the dynamic of X_T with respect to measure Q is completely characterized in Proposition 4, the valuation problem reduces to the task of computing expectation using the system in Proposition 4. In actual implementation, for example, one can use Monte Carlo simulation to generate many sample paths in accordance with the system in Proposition 4 and then take the discounted average of the contract payoff to yield the price for the derivative claim in question. If the interest rates are stochastic, the valuation result, due to the law of iterated expectations, becomes $E^Q\{\exp(-\sum_{s=t+1}^{T} r_{s-1,s})g(X_T)|\mathcal{F}_t\}$. A dynamic for interest rates also needs to be specified, and this interest rate dynamic must undergo risk-neutralization before option valuation can be implemented. A specific model pertaining to the case of stochastic interest rates under GARCH is available in Duan (1996b).

The option pricing system in Proposition 4 can be conveniently specialized to different model specifications. For example, we can adopt the non-linear asymmetric GARCH(1,1) (NGARCH) model of Engle and Ng (1993) to describe the asset price dynamic. Note that the NGARCH(1,1) model is the standard linear GARCH(1,1) model with the allowance for leverage effect. Parameter c in the following corollary captures the leverage effect; that is, a positive c gives rise to a negative correlation between the innovations in the asset return and its conditional volatility. In order to see that the option pricing system in Proposition 4 is indeed a generalization of the GARCH option pricing model of Duan (1995), we consider the continuously compounded return, i.e., $\delta = 0$, and impose the conditional normality assumption on ε_{t+1}, i.e., $\Psi(\cdot)$ is an identity mapping.

Corollary 1. (NGARCH(1,1) with conditional normality)

If ε_{t+1}, conditional on $\mathcal{F}_t$, is a P-standard normal random variable, $\delta = 0$, and

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\varepsilon_t - c)^2, \tag{13}$$

$\beta_0 > 0, \beta_1 \geq 0, \beta_2 \geq 0$ and $\beta_1 + \beta_2(1 + c^2) < 1$, then, for $t_l \leq t \leq t_u - 1$,

$$\ln(\frac{X_{t+1}}{X_t} + d_{t+1}) = r_{t,t+1} - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}Z_{t+1} \tag{14}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(Z_t - c - \lambda_t)^2 \tag{15}$$

$$\lambda_t = \frac{\mu_t - r_{t-1,t} + \frac{1}{2}h_t}{\sqrt{h_t}} \tag{16}$$

where Z_{t+1}, conditional on $\mathcal{F}_t$, is a Q-standard normal random variable.

Since Corollary 1 is essentially that of Duan (1995) and is implied by Proposition 4, our option pricing system can be regarded as a generalized version of the GARCH option pricing model. Equations (14)-(16) form a self-contained system and can, for example, be used to price European options by simulation in a straightforward manner. The system also serves as the theoretical basis for the methods developed by Ritchken and Trevor (1998) and Duan and Simonato (1998b) for pricing American options in the GARCH framework.

Several recent studies indicate that the conditional volatilities of many stock index returns exhibit long-memory property: see for example, Baillie, *et al* (1996), Bollerslev and Mikkelsen (1996) and McCurdy and Michaud (1996). Long-memory in volatility pertains to the fact that shocks to conditional volatility die away at a hyperbolic, instead of exponential rate. The option pricing system in Proposition 4 can be applied to the long-memory GARCH models. We use, as an example, the FIGARCH(1,d,1) model of Baillie, *et al* (1996) in the following corollary. Similar results can be obtained for other long-memory GARCH models.

Corollary 2. (FIGARCH(1,d,1) with conditional normality)

If ε_{t+1}, conditional on $\mathcal{F}_t$, is a P-standard normal random variable, $\delta = 0$, and

$$h_{t+1} = \beta_0 + \beta_1 h_t + [1 - \beta_1 L - (1 - \theta L)(1 - L)^d]h_{t+1}\varepsilon_{t+1}^2, \tag{17}$$

$0 \leq d \leq 1, \beta_0 > 0, \beta_1 \geq 0$, and $0 \leq \theta \leq 1$, then, for $t_l \leq t \leq t_u - 1$,

$$\ln(\frac{X_{t+1}}{X_t} + d_{t+1}) = r_{t,t+1} - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}Z_{t+1} \tag{18}$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + [1 - \beta_1 L - (1 - \theta L)(1 - L)^d]h_{t+1}\varepsilon_{t+1}^2 \tag{19}$$

$$\varepsilon_\tau = Z_\tau - \lambda_\tau \text{ if } \tau \geq t_l + 1 \tag{20}$$

$$\lambda_\tau = \frac{\mu_\tau - r_{\tau-1,\tau} + \frac{1}{2}h_\tau}{\sqrt{h_\tau}} \tag{21}$$

where Z_{t+1}, conditional on $\mathcal{F}_t$, is a Q-standard normal random variable.

Implementing the FIGARCH option pricing model is more complicated because it requires conditioning on an infinite sequence of lagged error terms.[6] For practical reasons, however, one must truncate the infinite sequence at some large lag, say 500. Although Monte Carlo simulations can still be used to price options, the methods developed by Ritchken and Trevor (1998) and Duan and Simonato (1998b) are not applicable to the long-memory GARCH model because the underlying system cannot be expressed as a low-dimensional Markov process.

3 Conditional leptokurtosis and volatility smile

There are several ways of analyzing the impact of conditional leptokurtosis. Although, the use of conditional t-distributions has been popular among many researchers, it presents a conceptual problem if it is used to model the continuously compounded rate of return. This conceptual problem can be understood if we set $\delta = 0$ in Equation (1). Suppose that the innovation term in (1) is a t-distributed random variable with κ degrees of freedom. This combination of assumptions implies that the continuous compounded rate of return is modeled as a conditionally t-distributed random variable. Under this assumption, computing the conditional expected simple rate of return is equivalent to evaluating the moment generating function of a t-distributed random variable. Since the moment generating function of a t-distribution with any finite degrees of freedom does not exist, using t-distributions for asset innovations amounts to assuming an unbounded expected simple asset return; that is, $E^Q\{\exp(\mu_{t+1} + \sqrt{h_{t+1}}\varepsilon_{t+1})|\mathcal{F}_t\} = \infty$. Intuitively, this implication is clearly not sensible. Technically, it violates the finite second moment condition for the asset payoff that was required as in part (ii) of Proposition 2. In short, using t-distributions to model continuously compounded asset returns is inappropriate.

One way of dealing with these conceptual and technical problems is to directly use t-distributions in describing the simple rate of return in order to avoid exponentiation of the t-distributed random variable. In terms of our notation, this approach amounts to setting $\delta = 1$ in Equation (1). This modeling strategy, unfortunately, departs from the tradition of using continuously compounded returns in the derivatives pricing literature. The use of t-distributions to directly model the simple rate of return also permits the occurrence of negative asset prices, which is not a desirable feature either.

In this section, we follow the tradition of modeling the continuously compounded rate of return. To study the effect of conditional tail-fatness, we use the generalized error distribution (GED) as in Nelson (1991). The density function of a GED random variable, after normalizing to yield a zero

[6] There is an alternative way of expressing the conditional variance dynamic in (19) which can be useful in practice; that is,

$$h_{t+1} = \beta_0 + \beta_1 h_t(1 - \varepsilon_t^2) + \sum_{k=0}^{\infty}(\theta\pi_k - \pi_{k+1})L^k h_t \varepsilon_t^2$$

where $\pi_0(d) = 1$ and $\pi_k(d) = (-1)^k \Pi_{i=1}^k \frac{d-i+1}{i}$ for $k \geq 1$. This expression for h_{t+1} is a result of $(1-L)^d = \sum_{k=0}^{\infty} \pi_k(d) L^k$.

mean and unit variance, is

$$g(z;v) = \frac{v\exp(-\frac{1}{2}|\frac{z}{\eta}|^v)}{2^{(1+1/v)}\eta\Gamma(1/v)} \text{ for } 0 < v \leq \infty \tag{22}$$

where

$$\eta = \left(\frac{2^{-\frac{2}{v}}\Gamma(1/v)}{\Gamma(3/v)}\right)^{\frac{1}{2}}$$

and $\Gamma(\cdot)$ is the gamma function. Parameter v determines the tail-fatness of the density function. The standard normal density function is a special case that $v = 2$. For $v > 2$, the density function has a tail thinner than the normal distribution. If $v < 2$, the fat-tail phenomenon occurs. The case that $v = 1$ yields the double exponential distribution. This distribution is of particular interest to us because $E[\exp(qZ)] < \infty$ if $|q| < \sqrt{2}$. This implies that if ε_{t+1} is of the GED with $v = 1$, $E^Q\{\exp(\mu_{t+1} + \sqrt{h_{t+1}}\varepsilon_{t+1})|\mathcal{F}_t\}$ is still finite for $h_{t+1} < 2$. In general, the expected simple return exists if $v > 1$. Thus, the double exponential distribution is used as a bound on the set of permissible fat-tailed distributions. The use of GED allows us to analyze the effect of tail-fatness without departing from the tradition of modeling the continuously compounded return.[7]

For the volatility dynamic, we choose to employ the NGARCH(1,1) model so that the effect of leverage in conjunction with other factors can be analyzed. The option pricing system in Proposition 4 can be specialized to the case of GED as follows.

Corollary 3. (NGARCH(1,1) with conditional leptokurtosis)

If ε_{t+1}, conditional on $\mathcal{F}_t$ and with respect to measure P, is a GED random variable with $v > 1$, $\delta = 0$, and

$$h_{t+1} = \beta_0 + \beta_1 h_t + \beta_2 h_t(\varepsilon_t - c)^2, \tag{23}$$

$\beta_0 > 0, \beta_1 \geq 0, \beta_2 \geq 0$ and $\beta_1 + \beta_2(1 + c^2) < 1$, then, for $t_l \leq t \leq t_u - 1$,

$$\begin{aligned} \ln(\frac{X_{t+1}}{X_t} + d_{t+1}) &= r_{t,t+1} - \ln\left[E^Q\left(\exp\{\sqrt{h_{t+1}}G^{-1}[\phi(Z_{t+1} - \lambda_{t+1}); v]\}|\mathcal{F}_t\right)\right] + \\ &\quad \sqrt{h_{t+1}}G^{-1}[\phi(Z_{t+1} - \lambda_{t+1}); v] \quad (24) \\ h_{t+1} &= \beta_0 + \beta_1 h_t + \beta_2 h_t\left(G^{-1}[\phi(Z_t - \lambda_t); v] - c\right)^2 \quad (25) \end{aligned}$$

where Z_{t+1}, conditional on $\mathcal{F}_t$, is a Q-standard normal random variable, and $G^{-1}[\cdot; v]$ stands for the inverse GED cumulative distribution function with parameter v.

[7] If an asymmetric conditional distribution is preferred, the approach proposed by Fernandez and Steel (1998) can be used to construct such a density function. Specifically, let

$$g^*(z; v, \gamma) \equiv \frac{2}{\gamma + \frac{1}{\gamma}}\left\{g(\frac{z}{\gamma}; v)\chi_{\{z\in[0,\infty)\}} + g(\gamma z; v)\chi_{\{z\in(-\infty,0)\}}\right\}$$

where $\chi_{\{\cdot\}}$ is an indicator function. If $\gamma \neq 1$, $g^*(z; v, \gamma)$ is a skewed and leptokurtic probability density function. The random variable following this distribution can then be normalized to have a zero mean and unit variance.

The option pricing system in Corollary 3 looks considerably more complicated when compared to the case of conditional normality in Corollary 1. The complexity, of course, arises from conditional leptokurtosis. Specifically, the functions $G(\cdot)$ (corresponding to the GED) and $\phi(\cdot)$ (corresponding to the standard normal distribution) are different so that $G^{-1}(\cdot)$ and $\phi(\cdot)$ do not cancel each other. The presence of risk premium parameter λ_t further complicates the matter. Although the option pricing system with conditional leptokurtosis is analytically complicated, it is still completely characterized as a self-contained pricing system and can be numerically implemented by Monte Carlo simulation.

For simplicity, we assume in this section that the asset risk premium λ_t and $r_{t-1,t}$ are constant. If interest rates are stochastic, the valuation problem becomes considerably more complicated. As discussed earlier, a dynamic for interest rates will need to be specified, and this interest rate dynamic must also undergo risk-neutralization before option valuation can be implemented. A European call option with maturity τ and exercise price K has a time-t theoretical price equal to $e^{-r\tau}E^Q[\max(X_{t+\tau}-K,0)|\mathcal{F}_t]$. This price can be numerically assessed by repeating the following empirical martingale Monte Carlo simulation steps:[8]

1. Identify an observable two-dimensional sufficient statistic at time t ; that is (X_t, h_{t+1}).

2. Generate N standard normal random numbers, $Z_{t+1}^{(i)}, i = 1, \cdots, N$, to advance the dynamics in (24) and (25) to time $t+1$, and then make an empirical martingale adjustment. (Specifically, we first compute the discounted sample average of simulated asset price for time $t+1$, and then multiply each of the N simulated asset prices by the ratio of the initial asset price over the discounted average. This adjustment ensures that the simulated sample has an empirical martingale property.)

3. Repeat steps 1–2 until arriving at N simulated asset prices, $X_{t+\tau}^{(i)}, i = 1, \cdots, N$.

4. Compute each of N option payoffs. Average N option payoffs, and discount the average, using the risk-free rate, back to the time of option valuation.

For the numerical analyses reported in this section, we use 10,000 Monte Carlo sample paths. The numerical procedure for evaluating $\phi(\cdot)$ is a standard one. For $G^{-1}[\cdot; v]$ and $E^Q(\exp\{\sqrt{h_{t+1}}\, G^{-1}[\phi(Z_{t+1}-\lambda); v]\}|\mathcal{F}_t)$, the procedures are described in Appendix.

We conduct a numerical analysis to examine the Black-Scholes implied volatility smile pattern exhibited by the GARCH option prices under the influence of conditional leptokurtosis. This analysis is performed by computing the generalized GARCH option prices for a group of option contracts and then converting their prices into the corresponding Black-Scholes implied volatilities. This practice is similar to the typical empirical approach of taking the market price of an option and converting it into the Black-Scholes implied volatility. The difference between our use in this paper

[8]The empirical martingale simulation method was developed by Duan and Simonato (1998a). This simulation method has been shown to be much more efficient because it takes advantage of the fact that the discounted asset price is a martingale under the risk-neutral probability measure. The empirical martingale simulation method can also ensure that option pricing bounds are respected. This property is particularly useful for this paper because the Black-Scholes implied volatility cannot be computed if option pricing bounds are violated by a price estimate.

and the empirical studies in the option literature is that the prices are generated by a theoretical model instead of the actual market prices of options. The use of the Black-Scholes implied volatility to study option markets is a common practice in the literature; for example, Rubinstein (1985), Sheikh (1991), Canina and Figlewski (1993) and Duan (1996a).

To estimate the NGARCH(1,1) model with the conditional GED, this study uses the S&P 500 index daily returns (continuously compounded) from January 3, 1994 to March 19, 1997, totaling 838 observations. The estimation is carried out by restricting the conditional mean to a constant, a. This specification for the conditional mean makes the estimation easier, but leaves the parameter λ unidentified. To be completely consistent with Corollary 3, the conditional mean μ_{t+1} should equal $r - \ln\left[E^Q\left(\exp\{\sqrt{h_{t+1}}G^{-1}[\phi(Z_{t+1} - \lambda); v]\}|\mathcal{F}_t\right)\right]$. Such a quantity can be numerically evaluated when parameter values are known. It, however, becomes extremely demanding in computing time if one needs to repeat the procedure in search of the maximum likelihood parameter estimates. Our empirical estimation based on the restricted model produces the maximum likelihood parameter estimates and the standard errors in the following table.

Table 1. The estimated GARCH and conditional tail-fatness parameters using the S&P 500 index daily returns from January 3, 1994 to March 19, 1997.

Parameter	Estimate	Standard error
a	6.6080×10^{-4}	1.2488×10^{-4}
β_0	2.6627×10^{-6}	1.2873×10^{-6}
β_1	0.8232	0.0675
β_2	0.0582	0.0199
c	0.9849	0.4184
v	1.2627	0.0869

The parameter v determines the extent of conditional tail-fatness. A smaller value for v implies fatter tails. We consider three values for v: 1, 2 and 1.2627, with the last value being the parameter estimate obtained in the empirical analysis. Recall that $v = 1$ gives rise to the double exponential distribution. For $v = 2$, the asset return innovation becomes normally distributed. The actual parameter estimate for v at 1.2627 implies a fat-tailed distribution.

The option maturity for our numerical analyses is set to three months (63 trading days) and later to six months (126 trading days). The interest rate is fixed at 0% so that there is no ambiguity as to whether an option is at-the-money. To conserve space, we only consider the case that the initial conditional volatility is equal to the stationary standard deviation under measure P. Similar to Theorem 3.1 of Duan (1995), the stationary standard deviation of the model can be shown to equal $\sqrt{\beta_0[1 - \beta_1 - \beta_2(1 + c^2)]^{-1}}$. The parameter values in Table 1 yield a stationary standard deviation (annualized) equal to 10.4%. The comparison is repeated for different moneyness positions which are expressed as a ratio of the exercise price to the underlying asset price.

The results are summarized in Figures 1-4. Figure 1 reports the implied volatility curves corresponding to three different values of v with the risk premium parameter λ set to 0. All other parameters are at their estimated values. The typical volatility smile pattern emerges. The effect of conditional leptokurtosis is interesting. If conditional leptokurtosis increases, the implied volatility becomes higher for the out-of-the-money options, but has little effect on other options. The downward sloping feature of these curves is primarily a result of a positive leverage parameter, i.e., $c = 0.9849$. Although not reported here, a negative c will cause the volatility smile curve to slope upwards. If the option maturity is increased to six months, conditional leptokurtosis has little impact on the level of implied volatility (see Figure 2). An increase in maturity has the effect of flattening the volatility smile curve, however. This maturity effect is similar to that established earlier by Duan (1995), and is consistent with the empirical evidence documented in the literature.

Corollary 3 can be reduced to Corollary 1 when conditional normality is assumed, i.e., $v = 2$. In such a case, the unit risk premium of the asset return, λ, can be directly added to the leverage parameter, c, to determine the overall leverage effect under the locally risk-neutralized probability measure Q. This result suggests that for option pricing, either a positive unit risk premium or the presence of leverage effect can produce a downward skewed volatility smile pattern. In other words, the curve in either Figure 1 or 2, corresponding to conditional normality, can also be produced when $c = 0$ but $\lambda = 0.9849$.

Conditional leptokurtosis, however, destroys this additivity. Figures 3 and 4 are presented to study its effects. In Figure 3, we set both the risk premium parameter, λ, and leverage parameter, c, to 0. Since c affects the magnitude of the stationary variance under measure P, we must make a compensating adjustment to parameter β_0 so that the stationary variance can remain unchanged after c is altered. The volatility smile curve corresponding to the GED with the estimated tail-fatness parameter value is fairly flat for most part. The curve corresponding to conditional normality is lower for both in-the-money and out-of-the-money options. In other words, the effect of conditional leptokurtosis is to increase the value for in-the-money and out-of-the-money options.

In Figure 4, the leverage parameter, c, continues to be zero whereas the risk premium parameter, λ, is increased to 0.5. Notice that the smile curves are all skewed downwards under this scenario, which means that the skewed smile need not come from the leverage parameter. A positive risk premium can also cause the volatility smile to skew toward the out-of-the-money options. The effect of conditional leptokurtosis does show up prominently when compared to the patterns in Figure 1. The effect of the risk premium parameter differs from that of the leverage parameter if the asset return innovation has conditional fat-tails. When the underlying asset contains a positive risk premium, fatter-tails simply make option more valuable across the board. Its effects are particularly pronounced for in-the-money options.

Comparison of Figures 3 and 4 also leads to another interesting conclusion. Recall that the stationary standard deviation under the data generating probability P, based on the estimated parameter values, equals 10.4%. The implied volatilities in Figure 3 are all close to this value, whereas those in Figure 4 are much higher due to a positive λ. This result is actually predicted by our option pricing theory. This is because local risk-neutralization transforms the dynamic of conditional volatility in such a way that a positive λ increases the level of stationary volatility under measure Q. Empirical studies on option prices typically conclude that the implied volatility is higher than the corresponding historical volatility or "realized" volatility. This kind of empirical findings thus lends support to the option pricing theory developed in this paper.

Our results are particularly interesting when compared to Figure 2 of Rubinstein (1994). Rubinstein's plot is a typical volatility smile of the S&P500 index options for the post-1987 stock market crash period, in which the smile is close to being a straight line. Rubinstein (1994) also refers to a paper by Shimko (1991) which reported a very high negative correlation between the implied volatility innovation of the S&P100 index option and the return innovation of that index for the post-crash period between 1987 and 1989. The negative correlation of this sort can be viewed as a reflection of the joint leverage effect under the locally risk-neutralized measure Q. In our theoretical framework, the joint leverage effect is captured by a positive value for c and/or λ. Rubinstein's empirical finding is thus not surprising from the standpoint of the GARCH option pricing theory.

4 Conclusion

A generalized GARCH option pricing model is developed in this article. This generalized theory allows for conditional leptokurtosis and many other well-documented features of the asset return dynamics. The GARCH option pricing model of Duan (1995) becomes a special case of our model. Importantly, our model offers additional flexibility in a parsimonious manner in describing the volatility smile phenomenon.

The ability of the generalized GARCH option pricing model to describe the observed volatility smile is important. It constitutes a more comprehensive theoretical model and leads to many new theoretical insights. This option pricing model can also be used to generate reasonable implied risk-neutralized probabilities. Similar to Shimko (1993) and Rubinstein (1994), these implied risk-neutralized probabilities can be used to price exotic options and for achieving better hedging positions. In a limited-scale study by Duan (1996a), the GARCH pricing model with conditional normality has been found to be successful in describing the volatility smile of the FT-SE 100 index options. An exploration of practical benefits of this generalized GARCH option pricing model is left for future research.

5 Appendix

Proof of Proposition 2.

(i) It is clear that Q is a positive measure because marginal utility is by definition positive. We need to show that $\int 1dQ = 1$ to conclude that Q is a probability measure.

$$\begin{aligned}
&\int 1dQ \\
=\ & E^P\left\{\exp\left[-\rho(t_u - t_l) + \sum_{s=t_l+1}^{t_u}(r_{s-1,s} + \ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_{t_l}\right\} \\
=\ & E^P\left\{\begin{array}{c}\exp\left[-\rho(t_u - t_l - 1) + \sum_{s=t_l+1}^{t_u-1}(r_{s-1,s} + \ln\frac{U'(C_s)}{U'(C_{s-1})})\right] \\ \exp(r_{t_u-1,t_u})E^P\left[\exp(-\rho)\frac{U'(C_{t_u})}{U'(C_{t_u-1})}|\mathcal{F}_{t_u-1}\right]|\mathcal{F}_{t_l}\end{array}\right\}
\end{aligned}$$

$$= E^P\left\{\exp\left[-\rho(t_u-t_l-1)+\sum_{s=t_l+1}^{t_u-1}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_{t_l}\right\}.$$

The last equality is due to

$$E^P\left[\exp(-\rho)\frac{U'(C_{t_u})}{U'(C_{t_u-1})}|\mathcal{F}_{t_u-1}\right]=\exp(-r_{t_u-1,t_u}).$$

That is, one dollar payable at time t_u has its time-(t_u-1) value equal to $\exp(-r_{t_u-1,t_u})$. Continuing the process, one obtains $\int 1dQ=1$.

(ii) Using the asset pricing equation in (6) we only need to show

$$e^{-r_{t-1,t}}E^Q((W_t+D_t)|\mathcal{F}_{t-1})=E^P\left\{e^{-\rho}\frac{U'(C_t)}{U'(C_{t-1})}(W_t+D_t)|\mathcal{F}_{t-1}\right\}.$$

Let

$$\varphi_t\equiv E^P\left\{\exp\left[-\rho(t_u-t)+\sum_{s=t+1}^{t_u}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_t\right\}.$$

It follows that $\varphi_t=1$ for any t by the argument identical to part (i). Consider any bounded $\mathcal{F}_{t-1}$-measurable function p_{t-1}. By the definition of conditional expectation and measure Q, we have

$$\begin{aligned}
& E^Q\left\{p_{t-1}E^Q\left[(W_t+D_t)|\mathcal{F}_{t-1}\right]|\mathcal{F}_{t_l}\right\}\\
= & E^Q\left\{p_{t-1}(W_t+D_t)|\mathcal{F}_{t_l}\right\}\\
= & E^P\left\{p_{t-1}(W_t+D_t)\exp\left[-\rho(t_u-t_l)+\sum_{s=t_l+1}^{t_u}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_{t_l}\right\}\\
= & E^P\left\{\begin{array}{c} p_{t-1}\exp\left[-\rho(t-t_l-1)+\sum_{s=t_l+1}^{t-1}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]\\ E^P\left((W_t+D_t)\exp\left[-\rho(t_u-t+1)+\sum_{s=t}^{t_u}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t_l}\end{array}\right\}
\end{aligned}$$

Since

$$\begin{aligned}
& E^P\left((W_t+D_t)\exp\left[-\rho(t_u-t+1)+\sum_{s=t}^{t_u}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right]|\mathcal{F}_{t-1}\right)\\
= & E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}\varphi_t|\mathcal{F}_{t-1}\right) \qquad \text{(by conditioning on } \mathcal{F}_t)\\
= & E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right) \qquad (\varphi_t=1)
\end{aligned}$$

we have

$$
\begin{aligned}
& E^Q\left\{p_{t-1}E^Q\left[(W_t+D_t)|\mathcal{F}_{t-1}\right]|\mathcal{F}_{t_l}\right\} \\
= \;& E^P\left\{\begin{array}{c} p_{t-1}\exp\left[-\rho(t-t_l-1)+\sum_{s=t_l+1}^{t-1}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right] \\ E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t_l}\end{array}\right\} \\
= \;& E^P\left\{\begin{array}{c} p_{t-1}\varphi_{t-1}\exp\left[-\rho(t-t_l-1)+\sum_{s=t_l+1}^{t-1}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right] \\ E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t_l}\end{array}\right\} \\
= \;& E^P\left\{\begin{array}{c} p_{t-1}\exp\left[-\rho(t_u-t_l)+\sum_{s=t_l+1}^{t_u}(r_{s-1,s}+\ln\frac{U'(C_s)}{U'(C_{s-1})})\right] \\ E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t_l}\end{array}\right\} \\
= \;& E^Q\left\{p_{t-1}E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right)|\mathcal{F}_{t_l}\right\}.
\end{aligned}
$$

Since p_{t-1} is any bounded $\mathcal{F}_{t-1}$-measurable function, it follows that

$$E^Q\left[(W_t+D_t)|\mathcal{F}_{t-1}\right]=E^P\left(\exp(-\rho+r_{t-1,t})(W_t+D_t)\frac{U'(C_t)}{U'(C_{t-1})}|\mathcal{F}_{t-1}\right).$$

This in turn implies

$$e^{-r_{t-1,t}}E^Q\left[(W_t+D_t)|\mathcal{F}_{t-1}\right]=E^P\left\{e^{-\rho}\frac{U'(C_t)}{U'(C_{t-1})}(W_t+D_t)|\mathcal{F}_{t-1}\right\}.$$

Note that $E^P\left\{e^{-\rho}\frac{U'(C_t)}{U'(C_{t-1})}(W_t+D_t)|\mathcal{F}_{t-1}\right\}$ is finite because the Cauchy-Schwarz inequality and the finite second moments condition given in the statement of the proposition. The proof is thus complete.□

Proof of Proposition 3.

It is clear that measure Q is mutually absolutely continuous with respect to measure P. To prove the remaining two requirements are met, we use the following intermediate lemma.

Lemma. Let $Y_t=\ln\frac{U'(C_t)}{U'(C_{t-1})}$. If Y_t and $\Psi(\varepsilon_t)$ have a bivariate normal distribution under measure P and conditional on $\mathcal{F}_{t-1}$, then

(a) $E^Q(\frac{X_t}{X_{t-1}}+d_t|\mathcal{F}_{t-1})=\exp(r_{t-1,t})$, and

(b) there exists a predictable process λ_t such that $\Psi(\varepsilon_t)+\lambda_t$, conditional on $\mathcal{F}_{t-1}$, is a Q-standard normal random variable.

Proof: Part (a) is easily verified:

$$\begin{aligned}
& E^Q\left(\frac{X_t}{X_{t-1}}+d_t|\mathcal{F}_{t-1}\right) \\
= \;& E^P\left[\left(\frac{X_t}{X_{t-1}}+d_t\right)\exp(r_{t-1,t}-\rho+Y_t)|\mathcal{F}_{t-1}\right] \\
= \;& \frac{1}{X_{t-1}}E^P[(X_t+d_tX_{t-1})\exp(-\rho+Y_t)|\mathcal{F}_{t-1}]\exp(r_{t-1,t}) \\
= \;& \exp(r_{t-1,t}).
\end{aligned}$$

To prove (b), we consider the conditional moment generating function of $\Psi(\varepsilon_t)$ under Q:

$$E^Q[\exp(q\Psi(\varepsilon_t))|\mathcal{F}_{t-1}] = E^P[\exp(q\Psi(\varepsilon_t)+r_{t-1,t}-\rho+Y_t)|\mathcal{F}_{t-1}].$$

Since, under measure P, $\Psi(\varepsilon_t)$ and Y_t are conditionally bivariate normal random variables, it follows from a linear projection that $Y_t=\alpha_t-\lambda_t\Psi(\varepsilon_t)+U_t$, where U_t is P-independent of $\Psi(\varepsilon_t)$. Note that α_t and λ_t are $\mathcal{F}_{t-1}$-measurable. Thus,

$$\begin{aligned}
& E^Q[\exp(q\Psi(\varepsilon_t))|\mathcal{F}_{t-1}] \\
= \;& \exp(\alpha_t+r_{t-1,t}-\rho)E^P\{\exp[(q-\lambda_t)\Psi(\varepsilon_t)+U_t]|\mathcal{F}_{t-1}\} \\
= \;& \exp[\alpha_t+r_{t-1,t}-\rho+\frac{1}{2}E^P(U_t^2|\mathcal{F}_{t-1})]E^P\{\exp[(q-\lambda_t)\Psi(\varepsilon_t)]|\mathcal{F}_{t-1}\} \\
= \;& \exp[\alpha_t+r_{t-1,t}-\rho+\frac{1}{2}E^P(U_t^2|\mathcal{F}_{t-1})+\frac{\lambda_t^2}{2}+\frac{q^2}{2}-\lambda_t q].
\end{aligned}$$

Let $q=0$, and use the fact that $E^Q(1|\mathcal{F}_{t-1})=1$ to yield

$$E^Q[\exp(q\Psi(\varepsilon_t))|\mathcal{F}_{t-1}] = \exp(\frac{q^2}{2}-\lambda_t q).$$

This in turn implies

$$\Psi(\varepsilon_t)|\mathcal{F}_{t-1}\sim N(-\lambda_t,1)$$

under measure Q.□

With the intermediate lemma in place, the three assertions in the proposition can be proved easily.

(i) $\ln\frac{U'(C_t)}{U'(C_{t-1})}=(q_1-1)\ln\frac{C_t}{C_{t-1}}$ where q_1 is the constant relative risk-aversion coefficient. Since $\ln\frac{C_t}{C_{t-1}}$ and $\Psi(\varepsilon_t)$ have a P-bivariate normal distribution conditional on $\mathcal{F}_{t-1}$, the result is immediately established.

(ii) $\ln\frac{U'(C_t)}{U'(c_{t-1})}=-q_2(C_t-C_{t-1})$ where q_2 is the constant absolute-risk-aversion coefficient. The assertion is true because (C_t-C_{t-1}) and $\Psi(\varepsilon_t)$ have a P-bivariate normal distribution conditional on $\mathcal{F}_{t-1}$.

(iii) This result holds trivially because the ratio of marginal utilities equals one.□

Numerical Algorithm.

A. Finding $y = G^{-1}(x; v)$

Consider a sequence of values: $k\Delta y$ where $k = 1, 2, \cdots$. Let $G(0; v) = 0.5$ and

$$G[k\Delta y; v] = G[(k-1)\Delta y; v] + \frac{\Delta y}{2}\left(g[k\Delta y; v] + g[(k-1)\Delta y; v]\right)$$

Let n be the smallest k such that $G[k\Delta y; v] \geq 1 - \epsilon$ where ϵ is some small positive real number.

For $x \in (0.5, 1-\epsilon]$, there exists a unique positive integer k such that $G[(k-1)\Delta y; v] < x \leq G[k\Delta y; v]$. This allows us to define an interim function for $x \in (0.5, 1]$ as follows:

$$H(x; v) = \begin{cases} \left(k - 1 + \frac{x - G[(k-1)\Delta y; v]}{G[k\Delta y; v] - G[(k-1)\Delta y; v]}\right)\Delta y & \text{for } x \in (0.5, 1-\epsilon] \\ n\Delta y & \text{for } x \in (1-\epsilon, 1] \end{cases}$$

Finally,

$$G^{-1}(x; v) = \begin{cases} H(x; v) & \text{if } x \in (0.5, 1] \\ 0 & \text{if } x = 0.5 \\ -H(1-x; v) & \text{if } x \in [0, 0.5) \end{cases}$$

For the results reported in this article, we set $\epsilon = 0.00001$ and $\Delta y = 0.005$.

B. Computing $E^Q\left(\exp\{\sqrt{h_{t+1}}G^{-1}[\phi(Z_{t+1} - \lambda); v]\}|\mathcal{F}_t\right)$

This quantity can be computed using a "vector" Monte Carlo simulations. Suppose that we are at time t and want to simulate for future time points. Instead of simulating the whole sample and then repeating N times, we first simulate a vector of N independent standard normal random variates for time $t+1$, i.e., $\{Z_{t+1}^{(i)}, i = 1, 2, \cdots, N\}$. Compute $G^{-1}[\phi(Z_{t+1}^{(i)} - \lambda); v]$ according to the procedure described in **A**. Note that h_{t+1} is measurable with respect to $\mathcal{F}_t$. For the k-th element of the vector containing simulated h_{t+1}, denoted by $h_{t+1}^{(k)}$, we can compute the following quantity:

$$\frac{1}{N}\sum_{i=1}^{N}\exp\{h_{t+1}^{(k)}G^{-1}[\phi(Z_{t+1}^{(i)} - \lambda); v]\}$$

to approximate $E^Q\left(\exp\{\sqrt{h_{t+1}^{(k)}}G^{-1}[\phi(Z_{t+1} - \lambda); v]\}|\mathcal{F}_t\right)$. Repeating for $k = 1, 2, \cdots, N$ completes the "vector" simulation for time $t+1$.□

References

[1] Amin, K. and V. Ng (1994): "A Comparison of Predictable Volatility Models Using Option Data," unpublished manuscript, University of Michigan.

[2] Baillie, R., T. Bollerslev and H. Mikkelsen (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 74, 3-30.

[3] Black, F. and M. Scholes (1972): "The Valuation of Option Contracts and a Test of Market Efficiency," *Journal of Finance* 27, 399-417.

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

[5] Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, 307–327.

[6] Bollerslev, T. (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics* 69, 542–547.

[7] Bollerslev, T. and H. Mikkelsen (1996): "Modelling and Pricing Long-Memory in Stock Market Volatility," *Journal of Econometrics* 73, 151-184.

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

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

[10] Canina, L. and S. Figlewski (1993): "The Information Content of Implied Volatility," *Review of Financial Studies* 6, 659-681.

[11] Cox, J. (1975): "Notes on Option Pricing I: Constant Elasticity of Variance Diffusions," unpublished manuscript, Stanford University.

[12] Crouhy, M. and M. Rockinger (1997): "Volatility Clustering, Asymmetry and Hysteresis in Stock Returns: International Evidence," *Financial Engineering and Japanese Markets* 4.

[13] Ding, Z., C. Granger and R. Engle (1993) "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance* 1, 83–106.

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

[15] Duan, J.-C. (1996a): "Cracking the Smile," *RISK* 9 (December), 55-59.

[16] Duan, J.-C. (1996b): "Term Structure and Bond Option Pricing under GARCH," working paper, Hong Kong University of Science & Technology.

[17] Duan, J.-C. (1997): "Augmented GARCH(p,q) Process and it's Diffusion Limit," *Journal of Econometrics* 79, 97-127.

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

[19] Duan, J.-C. and J.-G. Simonato (1998b): "American Option Pricing under GARCH by a Markov Chain Approximation," unpublished manuscript, Hong Kong University of Science and Technology.

[20] Duan, J.-C., G. Gauthier and J.-G. Simonato (1998): "An Analytical Approximation for the GARCH Option Pricing Model," unpublished manuscript, Hong Kong University of Science and Technology.

[21] Engle, R. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica* 50, 987–1008.

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

[23] Fernandez, C. and M. Steel (1998): "On Bayesian Modeling of Fat Tails and Skewness," *Journal of the American Statistical Association* 93, 359-371.

[24] Geske, R. (1979): "The Valuation of Compound Options," *Journal of Financial Economics* 7, 63-81.

[25] Glosten, L., R. Jagannathan and D. Runkle (1993): "Relationship between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance* 48, 1779–1801.

[26] Hanke, M. (1997): "Neural Network Approximation of Option Pricing Formulas for Analytically Intractable Option Pricing Models," *Journal of Computational Intelligence in Finance* 5, 20-27.

[27] Harrison, M. and S. Pliska (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Applications* 11, 215-260.

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

[29] 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.

[30] Heston, S. and S. Nandi (1998): "A Closed-Form GARCH Option Pricing Model," working paper, Goldman Sachs & Company.

[31] Heynen, R., A. Kemna and T. Vorst (1994): "Analysis of the Term Structure of Implied Volatilities," *Journal of Financial and Quantitative Analysis* 29, 31–56.

[32] Higgins, M. and A. Bera (1992): "A Class of Nonlinear ARCH Models," *International Economic Review* 33, 137–158.

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

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

[35] Kallsen, J. and M. Taqqu (1998): "Option Pricing in ARCH-Type Models," *Mathematical Finance* 8, 13-26.

[36] Madan, D. and F. Milne (1991): "Option Pricing with VG Martingale Components," *Mathematical Finance* 1, 39-56.

[37] McCurdy, T. and P. Michaud (1996): "Capturing Long Memory in the Volatility of Equity Returns: a Fractionally Integrated Asymmetric Power ARCH Model," unpublished manuscript, Queen's University.

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

[39] Merton, R. (1973): "The Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science* 4, 141-183.

[40] Merton, R. (1976): "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics* 3, 125-144.

[41] Nelson, D. (1991): "Conditional Heteroskedasticity in Asset Returns: a New Approach," *Econometrica* 59, 347–370.

[42] Ritchken, P. and R. Trevor (1998): "Pricing Options under Generalized GARCH and Stochastic Volatility Processes," to appear in *Journal of Finance*.

[43] Rubinstein, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics and Management Science* 7, 407-425.

[44] Rubinstein, M. (1983): "Displaced Diffusion Option Pricing," *Journal of Finance* 38, 213-217.

[45] Rubinstein, M. (1985): "Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978," *Journal of Finance* 40, 455–480.

[46] Rubinstein, M. (1994): "Implied Binomial Trees," *Journal of Finance* 49,771–818.

[47] Scott, L. (1987): "Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application," *Journal of Financial and Quantitative Analysis* 22, 419-438.

[48] Sheikh, A. (1991): "Transaction Data Tests of S & P 100 Call Option Pricing," *Journal of Financial and Quantitative Analysis* 26, 459–475.

[49] Shimko, D. (1993): "Bounds of Probability," *RISK* 6, 33–37.

[50] Shimko, D. (1991): "Beyond Implied Volatility: Probability Distribution and Hedge Ratios Implied by Option Prices," unpublished manuscript, University of Southern California.

[51] Stein, E. and J. Stein (1991): "Stock Price Distributions with Stochastic Volatility: An Analytic Approach," *Review of Financial Studies* 4, 727-752.

[52] Taylor, S. (1986): Modelling Financial Time Series (Wiley, New York, NY).

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

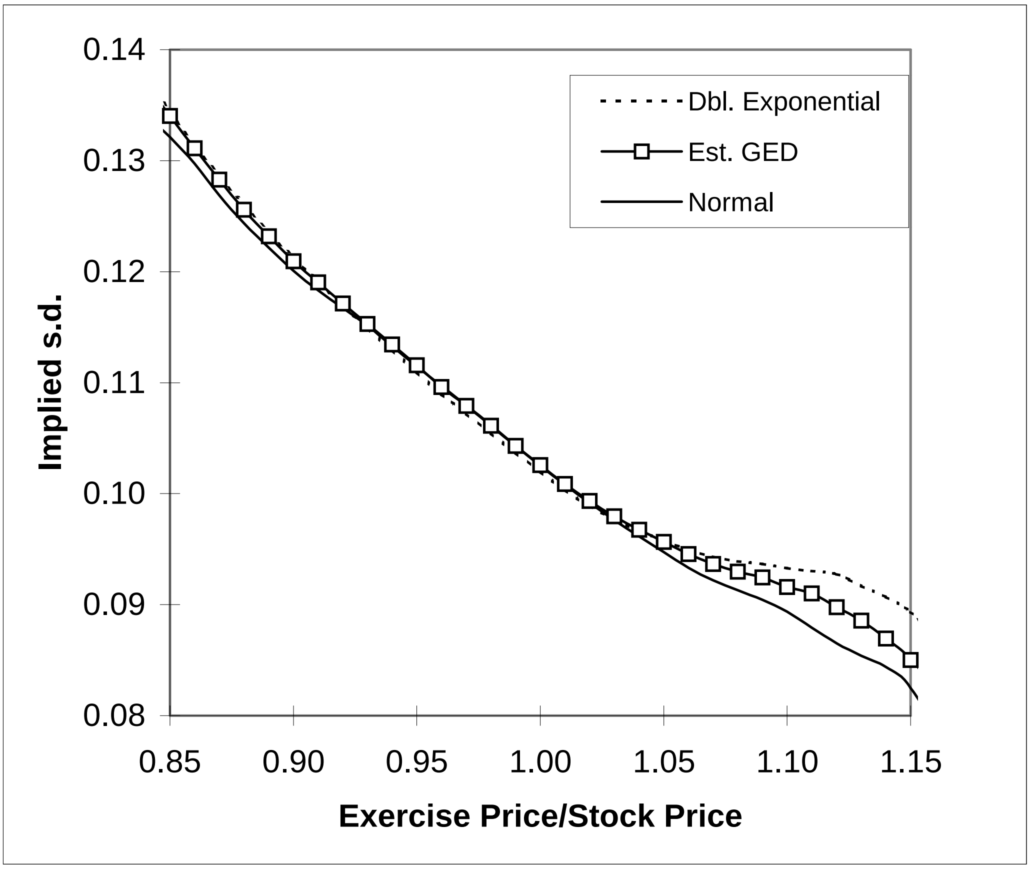

Figure 1. Impact of conditional leptokurtosis using the Black-Scholes implied standard deviations (annualized) of 3-month European call options with the risk premium parameter $\lambda = 0$. The estimated tail-fatness parameter of the GED equals 1.2627, which is estimated using the S&P 500 daily index returns from January 3, 1994 to March 19, 1997. The other two conditional distributions are assumed with other model parameters fixed.

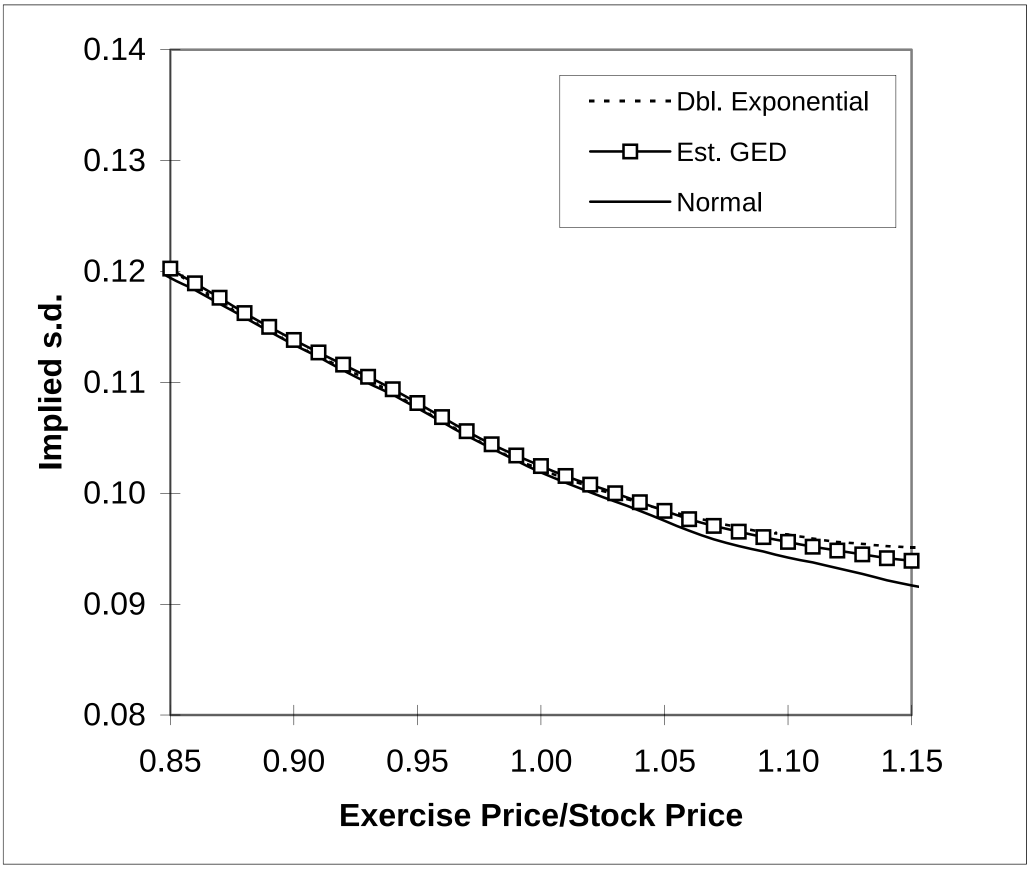

Figure 2. Impact of conditional leptokurtosis using the Black-Scholes implied standard deviations (annualized) of 6-month European call options with the risk premium parameter $\lambda = 0$. The estimated tail-fatness parameter of the GED equals 1.2627, which is estimated using the S&P 500 daily index returns from January 3, 1994 to March 19, 1997. The other two conditional distributions are assumed with other model parameters fixed.

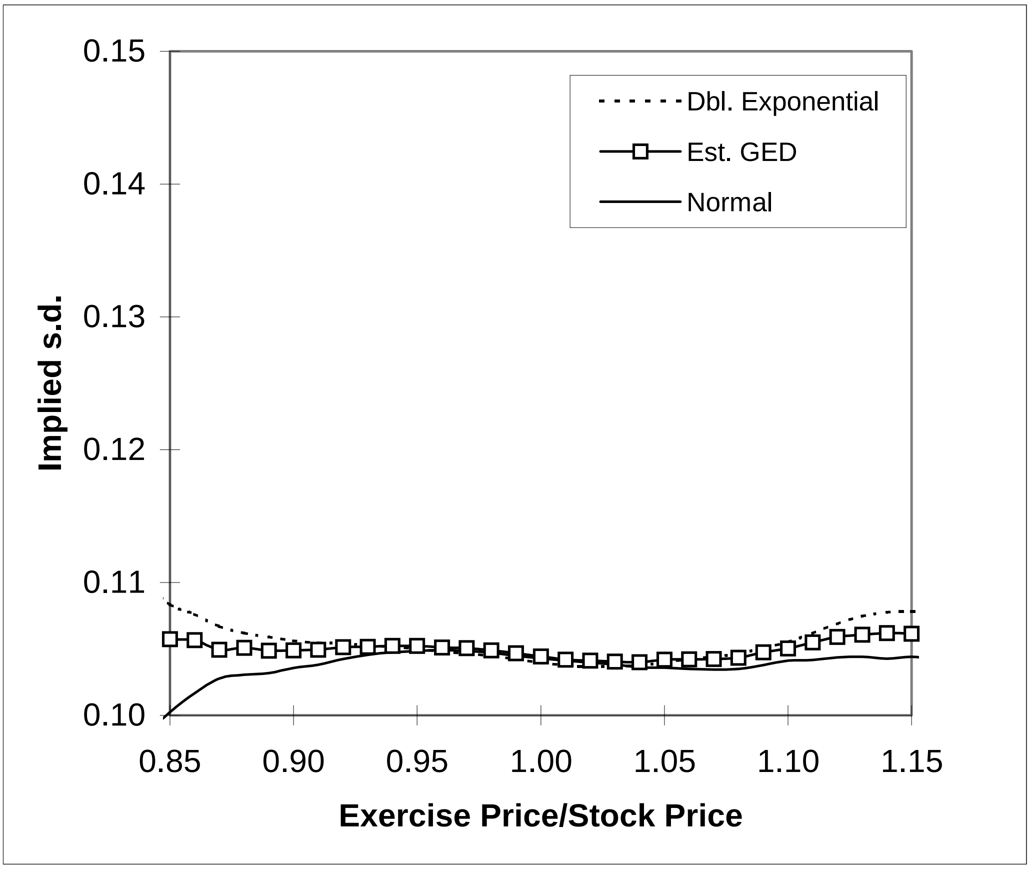

Figure 3. Impact of conditional leptokurtosis using the Black-Scholes implied standard deviations (annualized) of 3-month European call options with the leverage parameter $c = 0$ and the risk premium parameter $\lambda = 0$. The estimated tail-fatness parameter of the GED equals 1.2627, which is estimated using the S&P 500 daily index returns from January 3, 1994 to March 19, 1997. The other two conditional distributions are assumed with other model parameters fixed.

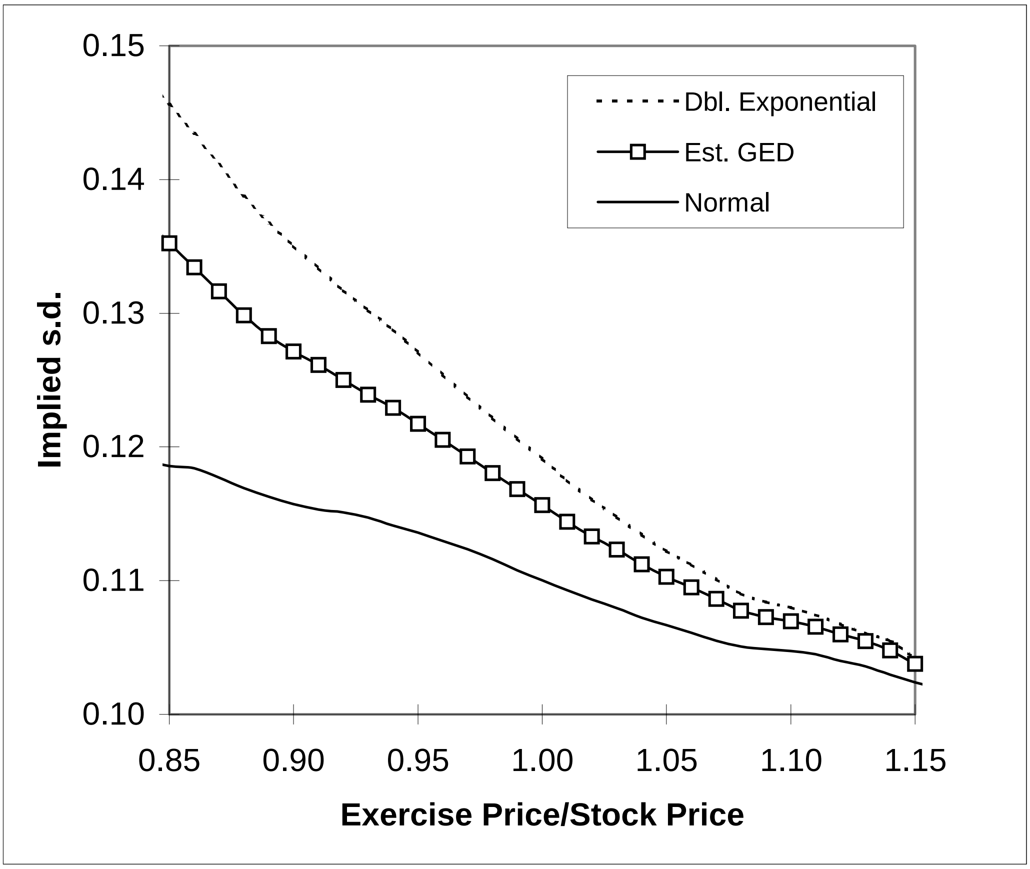

Figure 4. Impact of conditional leptokurtosis using the Black-Scholes implied standard deviations (annualized) of 3-month European call options with the leverage parameter $c = 0$ and the risk premium parameter $\lambda = 0.5$. The estimated tail-fatness parameter of the GED equals 1.2627, which is estimated using the S&P 500 daily index returns from January 3, 1994 to March 19, 1997. The other two conditional distributions are assumed with other model parameters fixed.