Moment-Implied Densities: Properties and Applications

Eric GHYSELS

Department of Economics, University of North Carolina at Chapel Hill and Department of Finance, Kenan-Flagler Business School at UNC, Chapel Hill, NC 27599 (*eghysels@unc.edu*)

Fangfang WANG

Department of Information and Decision Sciences, University of Illinois at Chicago, Chicago, IL 60607 (*ffwang@uic.edu*)

Suppose one uses a parametric density function based on the first four (conditional) moments to model risk. There are quite a few densities to choose from and depending on which is selected, one implicitly assumes very different tail behavior and very different feasible skewness/kurtosis combinations. Surprisingly, there is no systematic analysis of the tradeoff one faces. It is the purpose of the article to address this. We focus on the tail behavior and the range of skewness and kurtosis as these are key for common applications such as risk management.

KEY WORDS: Affine jump-diffusion model; Feasible domain; Generalized skewed t distribution; Normal inverse Gaussian; Risk neutral measure; Tail behavior; Variance Gamma.

1. INTRODUCTION

Suppose a researcher cares about the (conditional) moments of returns, in particular variance, skewness, and kurtosis. In addition, assume that he or she wants to use a parametric density function with the first four (conditional) moments given. The idea of keeping the number of moments small and characterizing densities by those moments has been suggested in various articles. This practice is popular among practitioners who use risk-neutral densities: see Madan, Carr, and Chang (1998); Theodossiou (1998); Aas, Haff, and Dimakos (2005); Eriksson, Ghysels, and Wang (2009); among others.[1] Depending on which density is selected, one implicitly assumes very different tail behavior and very different feasible skewness/kurtosis combinations.

The bulk of the academic literature has focused on estimating an asset return model and then uses its associated option pricing model. However, the bulk of practitioner's implementations do not involve estimating parameters via statistical methods, but rather via calibration. Black-Scholes implied volatilities are the most celebrated example of this practice. The common practice is of implied parameters, especially volatilities, being "plugged into" formulas. Our article tries to provide a deeper understanding of the common practice of calibration.

Many appealing distributions commonly used in financial modeling belong to a larger class of densities called the generalized hyperbolic (GH) class of distributions which, as a rich family, has wide applications in risk management and financial modeling.[2] The GH class is characterized by five parameters which, when further narrowed down to subclasses of

four-, three-, or two-parameter distributions, yields widely used distributions such as the normal inverse Gaussian distribution, the hyperbolic distribution, the variance gamma distribution, the generalized skewed t distribution, the student t distribution, the gamma distribution, the Cauchy distribution, the normal distribution, etc.

In this article we focus on the normal inverse Gaussian (NIG) distribution, the variance gamma (VG) distribution, and the generalized skewed t (GST) distribution, because they are fully characterized by their first four moments, and are extensively used in the risk management literature.[3] We use the distributions to model the risk neutral density of asset returns, with moments extracted from derivative contracts. In particular, the risk neutral moments are formulated by a portfolio of the out-of-money European Call/Put options indexed by their strikes (Bakshi, Kapadia, and Madan 2003). The risk neutral density is important to price derivative contracts. Focusing on risk neutral densities

---

[3]The application of the NIG distribution in the context of risk management appears in Eberlein and Keller (1995a), Barndorff-Nielsen 1997a, 1997b, Rydberg (1997), Eberlein (2001b), Venter and de Jongh (2002), Aas et al. (2006), Kalemanova (2007), Eriksson, Ghysels, and Wang (2009), among others. The VG distribution was introduced in Madan and Seneta (1990). It was further studied by Madan, Carr, and Chang (1998), Carr et al. (2002), Konikov and Madan (2002), Ribeiro and Webber (2004), Hirsa and Madan (2004), Seneta (2004), Avramidis and L'Ecuyer (2006), Moosbrucker (2006), among others. Finally, the GST distribution also has had many applications, including Frecka and Hopwood (1983), Theodossiou (1998), Prause (1999), Wang (2000), Bams, Lehnert, and Wolff (2005), Bauwens and Laurent (2005), Aas, Haff, and Dimakos (2005), Kuester, Mittnik, and Paolella (2006), Rosenberg and Schuermann (2006), and Bali and Theodossiou (2007).

---

[1]Figlewski (2007) provided a comprehensive literature review of various approaches to derive risk neutral densities—we focus on moment-based methods.

[2]See, for example, Eberlein, Keller, and Prause (1998), Rydberg (1999), Eberlein (2001a), Eberlein and Prause (2002), Eberlein and Hammerstein (2004), Bibby and Sørensen (2003), Chen, Hardle, and Jeong (2008), among many others.

allows us to use relevant parameter settings—given the widely availability of options data and its applications. Note that Aït-Sahalia and Lo (2000a) argued in more general terms for the use of risk neutral distributions for the purpose of risk management. Yet, many of the issues we address pertain to distributions in general—not just risk neutral ones. In particular, we focus on the tail behavior and so-called feasible domain, that is, the skewness/kurtosis combinations that are feasible for each of the densities. We derive closed-form expressions for the moments (which can be used for the purpose of estimation) for the three aforementioned classes of distributions.

Note that there are alternatives to the GH class of distributions, such as Edgeworth and Gram-Charlier expansions, SNP distributions or mixtures of normals. While some of these alternatives have a wider range of feasible skewness-kurtosis values, they go beyond characterizing the first four moments and typically involve considerable more parameters (in some cases an unbounded number).[4] Because the first four moments have straightforward interpretations they are commonly used by practitioners. Our article studies the consequences of focusing on those moments and fitting a density either for option pricing or characterizing value-at-risk. The former involves risk-neutral densities, whereas the latter involves physical densities. We study the use of four-moment-based densities both in a risk neutral option pricing setting and physical value-at-risk setting. To appraise how well the various density approximations perform, we consider affine jump-diffusion and GARCH models, which yield closed form expressions for the risk neutral density. This allows us to appraise how well the various classes of distributions approximate the density implied by realistic jump diffusions and GARCH models and their resulting derivative contracts. We study the moment-implied density approach in the pricing of options. In addition, we also discuss the transformation from the risk neutral to the physical measure within the class of GH densities and study their use in value-at-risk calculations.

The rest of this article is outlined as follows: we start with a review on the GH family of distributions in Section 2, and then study their tail behavior and moment-based parameter estimation. In Section 3 we characterize the feasible domains of various distributions using S&P 500 index options data and report findings of a simulation study based on jump diffusion processes and GARCH option pricing model. An option pricing exercise with real data is discussed in Section 4. Section 5 discusses the transformation from the risk neutral to the physical measure within the class of GH densities and study their use in value-at-risk calculations. Concluding remarks appear in Section 6. The technical details are in an Appendix.

2. MOMENT CONDITIONS OF THE GENERALIZED HYPERBOLIC DISTRIBUTION

The GH distribution was introduced by Barndorff-Nielsen (1977) to study aeolian sand deposits, and it was first applied in

a financial context by Eberlein and Keller (1995b). In this section we will give a brief review of the GH family of distributions and then discuss their tail behavior and moments.

2.1   The Generalized Hyperbolic Distribution

The GH distribution is a normal variance-mean mixture where the mixture is a Generalized Inverse Gaussian (GIG) distribution. Suppose that Y is GIG distributed with density

$$f(y; \psi, \chi, \lambda) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\psi\chi})} y^{\lambda-1} \exp\left[-\frac{1}{2}(\psi y + \chi y^{-1})\right],$$
$$y > 0,$$

where $K_\lambda(z) = 1/2 \int_0^\infty y^{\lambda-1} \exp[-1/2 z(y + y^{-1})]dy$ (for $z > 0$) is a modified Bessel function of the third kind with index λ. We then write $Y \overset{\mathcal{L}}{=} \text{GIG}(\psi, \chi, \lambda)$. The parameter space of $\text{GIG}(\psi, \chi, \lambda)$ is $\{\psi > 0, \chi > 0, \lambda = 0\} \cup \{\psi > 0, \chi \geq 0, \lambda > 0\} \cup \{\psi \geq 0, \chi > 0, \lambda < 0\}$.

A GH random variable is constructed by allowing for the mean and variance of a Normal random variable to be GIG distributed. Namely, a random variable X is said to be GH distributed, or $X \overset{\mathcal{L}}{=} \text{GH}(\alpha, \beta, \mu, b, p)$, if $X \overset{\mathcal{L}}{=} \mu + \beta Y + \sqrt{Y} Z$ where $Y \overset{\mathcal{L}}{=} \text{GIG}(\alpha^2 - \beta^2, b^2, p)$, $Z \overset{\mathcal{L}}{=} N(0, 1)$, and Y is independent of Z. The density function of X is therefore

$$f_{\text{GH}}(x; \alpha, \beta, \mu, b, p) = \frac{\alpha^{1/2-p}(\alpha^2 - \beta^2)^{p/2} e^{(x-\mu)\beta}}{\sqrt{2\pi b} K_p(b\sqrt{\alpha^2 - \beta^2})} K_{p-1/2}$$

$$\times \left(\alpha b \sqrt{1 + \frac{(x-\mu)^2}{b^2}}\right)$$

$$\times \left(1 + \frac{(x-\mu)^2}{b^2}\right)^{p/2-1/4} \quad (1)$$

with parameters satisfying $\alpha > |\beta|, b > 0, p \in \mathbb{R}, \mu \in \mathbb{R}$.

The GH distribution is closed under linear transformation, which is a desirable property notably in portfolio management. Note that for $X \overset{\mathcal{L}}{=} \text{GH}(\alpha, \beta, \mu, b, p)$, $tX + l$ is $\text{GH}(\alpha/|t|, \beta/t, t\mu + l, |t|b, p)$ distributed for $t \neq 0$ due to the scaling property of the GIG distribution, that is, if $Y \overset{\mathcal{L}}{=} \text{GIG}(\psi, \chi, \lambda)$, then $tY \overset{\mathcal{L}}{=} \text{GIG}(\psi/t, t\chi, \lambda)$ for $t > 0$. It follows that the set $\{\mu + \beta Y + \sigma\sqrt{Y}Z : Y \overset{\mathcal{L}}{=} \text{GIG}(\alpha^2 - \beta^2, b^2, p), Z \overset{\mathcal{L}}{=} N(0, 1), Y \perp Z, \alpha > |\beta|, b > 0, p \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0\}$ is equivalent to the set $\{\mu + \beta Y + \sqrt{Y}Z : Y \overset{\mathcal{L}}{=} \text{GIG}(\alpha^2 - \beta^2, b^2, p), Z \overset{\mathcal{L}}{=} N(0, 1), Y \perp Z, \alpha > |\beta|, b > 0, p \in \mathbb{R}, \mu \in \mathbb{R}\}$ under an affine transform. Therefore, it is sufficient to characterize the GH distribution with five parameters. It also follows that the GH distribution is infinitely divisible, a property that yields GH Lévy processes by subordinating Brownian motions. However, the GH distribution is not closed under convolution in general except when $p = -1/2$.

Various subclasses of the GH distribution can be derived by confining the parameters to a subset of the parameter space. The widely used distributions, which form subclasses of the GH distribution, are the symmetric GH distribution $\text{GH}(\alpha, 0, \mu, b, p)$, the hyperbolic distribution $\text{GH}(\alpha, \beta, \mu, b, 1)$, and the normal inverse Gaussian (NIG) distribution $\text{GH}(\alpha, \beta, \mu, b, -1/2)$. Note that the parameter space of the GH distribution excludes $\{\alpha >$

---

[4]The so called Hamburger theorem proves the existence of a distribution for any given feasible skewness kurtosis values, however it does not show how to construct the density (see e.g., Widder 1946; Chihara 1989). In addition, with SNP distributions the space spanned by the skewness and kurtosis values is bounded for any finite expansion; see Figure 1 and Section 2.3 of León, Mencía, and Sentana (2009).

$|\beta|, b = 0, p > 0\}$ and $\{\alpha = |\beta|, b > 0, p < 0\}$ which are permitted by the GIG distribution. If we allow parameters to take values on the boundary of parameter space, we can obtain various limiting distributions. These include the (1) variance gamma distribution, (2) generalized skewed T distribution, (3) skewed T distribution, (4) noncentral student T distribution, (5) Cauchy distribution, (6) normal distribution, among others (see, e.g., Bibby and Sørensen 2003; Eberlein and Hammerstein 2004; Haas and Pigorsch 2007).

2.2 Tail Behavior

The GH family covers a wide range of distributions and therefore exhibits various tail patterns. We discuss its tail behavior in general.[5] We write $A(x) \sim B(x)$ as $x \to \infty$ for functions A and B if $\lim_{x \to \infty} A(x)/B(x) = c$ for some constant c.

Note that $f_{\text{GH}}(x; \alpha, \beta, \mu, b, p) \sim |x - \mu|^{p-1} \exp(-\alpha|x - \mu| + \beta(x - \mu))$ (see Haas and Pigorsch 2007). An application of L'Hôpital's rule yields the following:

Proposition 2.1. Suppose that X is GH$(\alpha, \beta, \mu, b, p)$ distributed with $\alpha > |\beta|$, $b > 0$, $p \in \mathbb{R}$, $\mu \in \mathbb{R}$. Its right tail and left tail satisfy $P(X - \mu > x) \sim x^{p-1}e^{-(\alpha-\beta)x}$ and $P(X - \mu < -x) \sim x^{p-1}e^{-(\alpha+\beta)x}$, respectively, where $x > 0$ is sufficiently large.

Therefore, α, β, and p control tail behavior. A small α and a large p yield heavy tails. β pertains to skewness. The right tail is heavier when $\beta > 0$, whereas the left tail is heavier when $\beta < 0$. $\beta = 0$ yields a symmetric distribution. The tails of subclasses of the GH distribution can be derived from Proposition 2.1 directly.

Next we look at the tails of various "limiting" distributions. One obtains the Variance Gamma (VG) distribution from the GH distribution by letting b go to 0 and keeping p positive. Hence, the VG has the same tail behavior as the GH distribution. Set $\alpha = |\beta|$ in (1), and we will have the Generalized Skewed T (GST) distribution.

Corollary 2.1. Suppose that X is GST(β, μ, b, p) distributed with $|\beta| > 0$, $\mu \in \mathbb{R}$, $b > 0$, and $p < 0$. Then for $x > 0$ sufficiently large, $P(X - \mu > x) \sim x^p$ and $P(X - \mu < -x) \sim x^{p-1}e^{-2\beta x}$ for $\beta > 0$, while $P(X - \mu > x) \sim x^{p-1}e^{2\beta x}$ and $P(X - \mu < -x) \sim x^p$ for $\beta < 0$.

The skewed T distribution is derived from the GST distribution by allowing $p = -b^2/2$. Letting β go to 0 in the GST distribution yields the noncentral student T distribution with $-2p$ degrees of freedom. Its density behaves like $f_{\text{GH}}(x; 0, 0, \mu, b, p) \sim |x - \mu|^{2p-1}$ for sufficiently large $|x - \mu|$, and hence $P(|X - \mu| > x) \sim x^{2p}$ when $x > 0$ is sufficiently large. The tail property of the Cauchy distribution, as a special case of the noncentral student T distribution, is obtained by $p = -\frac{1}{2}$.

The normal distribution can be viewed as a limiting case of the GH law (the hyperbolic distribution in particular, i.e., $p = 1$) as well with mean $\mu + \beta\sigma^2$ and variance σ^2, where $\sigma^2 = \lim_{\alpha \to \infty}(b/\alpha)$.

It is known from the above discussion that the tails of the GH and VG distributions are exponentially decaying, and is slower than the normal distribution but faster than the GST and the skewed T distribution. The GH and VG distributions are therefore also referred to as semiheavy tailed (see Barndorff-Nielsen and Shephard 2001), and they possess moments of arbitrary order. The GST distribution does not have moments of arbitrary order. The rth moment exists if and only if $r < -p$. However, tails of the GST distribution are a mixture of polynomial and exponential decays—one heavy tail and one semiheavy tail, which distinguishes the GST law from the others (see Aas and Haff 2006).

2.3 Skewness and Kurtosis

We are interested in the first four moments, in particular, the space spanned by the squared skewness and excess kurtosis. In this section, we will first present general results regarding the GH distribution to characterize the mapping between moments and parameters.

For a centered GH distribution (i.e., $\mu = 0$), the moments of arbitrary order can be expanded as an infinite series of Bessel functions of the third kind with gamma weights (see, e.g., Barndorff-Nielsen and Stelzer 2005). Using this representation, we have the following proposition:

Proposition 2.2. Suppose that X is GH$(\alpha, \beta, 0, b, p)$ distributed. The first four moments, $m_n = EX^n$ for $n = 1, 2, 3, 4$, can be explicitly expressed as

$$m_1 = \frac{b\beta K_{p+1}(b\gamma)}{\gamma K_p(b\gamma)},$$

$$m_2 = \frac{b K_{p+1}(b\gamma)}{\gamma K_p(b\gamma)} + \frac{\beta^2 b^2 K_{p+2}(b\gamma)}{\gamma^2 K_p(b\gamma)},$$

$$m_3 = \frac{3\beta b^2 K_{p+2}(b\gamma)}{\gamma^2 K_p(b\gamma)} + \frac{\beta^3 b^3 K_{p+3}(b\gamma)}{\gamma^3 K_p(b\gamma)},$$

$$m_4 = \frac{\beta^4 b^4 K_{p+4}(b\gamma)}{\gamma^4 K_p(b\gamma)} + \frac{6\beta^2 b^3 K_{p+3}(b\gamma)}{\gamma^3 K_p(b\gamma)} + \frac{3b^2 K_{p+2}(b\gamma)}{\gamma^2 K_p(b\gamma)},$$

(2)

where $\gamma = \sqrt{\alpha^2 - \beta^2}$.

Proof. See Appendix A. □

Therefore, the mean M, variance V, skewness S, and excess kurtosis K of a GH$(\alpha, \beta, \mu, b, p)$ distribution can be expressed explicitly as functions of the five parameters, which yield a mapping (call it $\mathcal{T}$) from the parameter space to the space spanned by (M, V, S, K).[6] Note, however, that $\mathcal{T}$ may not be a bijection and therefore its inverse may not exist. Since this article is aimed at modeling financial returns which are skewed and leptokurtic and aimed at building densities based on skewness and (excess) kurtosis, we restrict our attention to subclasses of the GH family which have a four-parameter characterization and

---

[5]Bibby and Sørensen (2003) discussed the tail behavior of the GH family using a different parameterization.

[6]The moment-generating function and characteristic function of an arbitrary GH$(\alpha, \beta, \mu, b, p)$ can be found in for instance Prause (1999). Particularly, Prause (1999) also gave explicit expressions of the mean and variance.

yield a bijection between moments and parameters. It is impossible to express in general parameters explicitly via the first four moments due to the presence of Bessel functions. We will focus in the remainder of the article on the cases where we have an explicit mapping between moments and parameters, namely we will focus on the NIG, VG, and GST distributions.

2.3.1 The Normal Inverse Gaussian Distribution. One obtains the NIG distribution from the GH distribution by letting $p = -1/2$. Therefore, as an application of Proposition 2.2, we have the following:

Proposition 2.3. Denote by M, V, S, K the mean, variance, skewness, and excess kurtosis of a NIG(α, β, μ, b) random variable with $\alpha > |\beta|$, $\mu \in \mathbb{R}$, and $b > 0$. Then we have the following: (1)

$$M = \mu + \frac{\beta b}{\gamma}, \quad V = \frac{b\alpha^2}{\gamma^3}, \quad S = \frac{3\beta}{\alpha\sqrt{b\gamma}},$$

$$K = \frac{3(4\beta^2 + \alpha^2)}{b\alpha^2\gamma}, \tag{3}$$

where $\gamma = \sqrt{\alpha^2 - \beta^2}$. (2) If $D \equiv 3K - 5S^2 > 0$,

$$\alpha = 3\frac{\sqrt{D + S^2}}{D}V^{-1/2}, \quad \beta = \frac{3S}{D}V^{-1/2},$$

$$\mu = M - \frac{3S}{D + S^2}V^{1/2}, \quad b = \frac{3\sqrt{D}}{D + S^2}V^{1/2}.$$

The proof of (1) uses the fact that $K_{n+1/2}(z) = K_{1/2}(z)$ $(1 + \sum_{i=1}^{n}[(n+i)!2^{-i}z^{-i}]/[i!(n-i)!])$. The formal derivation of (2) appears in Eriksson, Forsberg, and Ghysels (2004).

2.3.2 The Variance Gamma Distribution. Recall that the VG distribution is obtained by keeping $\alpha > |\beta|$, $\mu \in \mathbb{R}$, $p > 0$ fixed and letting b go to 0. Proposition 2.2 is stated for the GH distribution. For the limiting cases, similar results are derived by applying the dominant convergence theorem. Particularly, when b approaches 0, we have the following result:

Proposition 2.4. Denote by M, V, S, K the mean, variance, skewness, and excess kurtosis of a VG(α, β, μ, p) random variable with $\alpha > |\beta|$, $\mu \in \mathbb{R}$, and $p > 0$. Then, (1)

$$M = \mu + \frac{\beta p}{\eta}, \quad V = \frac{p}{\eta^2}(\eta + \beta^2), \quad S = \frac{\beta(3\eta + 2\beta^2)}{(\eta + \beta^2)^{3/2}p^{1/2}},$$

$$K = \frac{3(\eta^2 + 4\eta\beta^2 + 2\beta^4)}{p(\eta + \beta^2)^2}, \tag{4}$$

where $\eta = \frac{\alpha^2 - \beta^2}{2} > 0$. (2) If $2K > 3S^2$, letting $C = 3S^2/2K$, the equation $(C - 1)R^3 + (7C - 6)R^2 + (7C - 9)R + C = 0$ has a unique solution in $(0, 1)$, denoted by R, and

$$\alpha = \frac{2\sqrt{R}(3 + R)}{\sqrt{V}|S|(1 - R^2)}, \quad \beta = \frac{2R(3 + R)}{\sqrt{V}S(1 - R^2)},$$

$$p = \frac{2R(3 + R)^2}{S^2(1 + R)^3}, \quad \mu = M - \frac{2\sqrt{V}R(3 + R)}{S(1 + R)^2}.$$

Proof. See Appendix A. $\square$

2.3.3 The Generalized Skewed T Distribution. The GST distribution is obtained from the GH distribution by $\alpha \to |\beta|$, with parameters satisfying $\beta \in \mathbb{R}$, $\mu \in \mathbb{R}$, $b > 0$, and $p < -4$ so that the 4th moment exists. An application of the dominant convergence theorem to Proposition 2.2 yields the following:

Proposition 2.5. Denote by M, V, S, K the mean, variance, skewness and excess kurtosis of a GST(β, μ, b, p) random variable with $\beta \in \mathbb{R}$, $\mu \in \mathbb{R}$, $b > 0$, and $p < -4$. Then, (1)

$$M = \mu + \frac{b^2\beta}{v - 2}$$

$$V = \frac{b^2}{v - 2} + \frac{2b^4\beta^2}{(v - 2)^2(v - 4)}$$

$$S = \left[6(v - 2) + \frac{16b^2\beta^2}{v - 6}\right]\frac{b\beta(v - 4)^{1/2}}{[(v - 2)(v - 4) + 2b^2\beta^2]^{3/2}}$$

$$K = \left[\frac{8b^4\beta^4(5v - 22)}{(v - 6)(v - 8)} + \frac{16b^2\beta^2(v - 2)(v - 4)}{v - 6}\right.$$

$$\left. + (v - 2)^2(v - 4)\right]\frac{6}{[(v - 2)(v - 4) + 2b^2\beta^2]^2}, \tag{5}$$

where $v = -2p > 8$. (2) If (K, S^2) satisfy

$$\frac{32(y_U + 4)}{(y_U + 2)^2} > S^2 \quad \text{where}$$

$$y_U \equiv \frac{\sqrt{K^2 + 156K + 900} - K + 30}{K} > 0, \tag{6}$$

the following system of equations:

$$0 = 2\rho[3(v - 6) + 4(v - 4)\rho]^2 - S^2(v - 4)(v - 6)^2(1 + \rho)^3$$

$$0 = 12(v - 4)(5v - 22)\rho^2 + 48(v - 4)(v - 8)\rho$$

$$+ 6(v - 6)(v - 8) - K(v - 4)(v - 6)(v - 8)(1 + \rho)^2 \tag{7}$$

has a unique solution satisfying $\rho > 0$ and $v > 8$, denoted by (ρ, v). We then have

$$\beta = sig(S)\sqrt{\frac{\rho(1 + \rho)(v - 4)}{2V}}, \quad \mu = M - sig(S)\sqrt{\frac{\rho(v - 4)V}{2(1 + \rho)}},$$

$$b = \sqrt{\frac{V(v - 2)}{1 + \rho}}, \quad p = -v/2.$$

Proof. See Appendix A. $\square$

2.3.4 Feasible Domain. It follows from Proposition 2.3 that the range of excess kurtosis and skewness admitted by the NIG distribution is $D_{\text{nig}} \equiv \{(K, S^2) : 3K > 5S^2\}$, which will be referred to as the *feasible domain* of the NIG distribution. The *feasible domain* of the VG distribution read from Proposition 2.4 is $D_{\text{vg}} \equiv \{(K, S^2) : 2K > 3S^2\}$ and it includes D_{nig}. The GST distribution has a *feasible domain* of $D_{\text{gst}} \equiv \{(K, S^2) : S^2 < \frac{32(y_U + 4)}{(y_U + 2)^2}\}$ where $y_U = \sqrt{1 + 156/K + 900/K^2} - 1 + 30/K$. Note that D_{gst} is included in the set $\{(K, S^2) : S^2 < \min(32, 8K/15)\}$, which is a subset of D_{nig} and D_{vg}. The GST distribution has bounded skewness. The VG distribution has the largest feasible domain among the three. It should also be noted that $\{(K, S^2) : 2K = 3S^2\}$ is the skewness-kurtosis combination

of the Pearson Type III distribution. Further details regarding feasible domains appear in Section 3.2.

3. RISK NEUTRAL MOMENTS AND IMPLIED DENSITIES

We are interested in modeling asset returns with NIG, VG, and GST distributions for the purpose of risk management. It is therefore natural to think in terms of the risk neutral distribution since it plays an important role in derivative pricing. In this section, we present risk neutral moment-based estimation methods using the European put and call contracts. Affine jump-diffusion models and GARCH option pricing models will be used to evaluate the performance of the NIG, VG, and GST approximations.

3.1 Estimating Moments of Risk Neutral Distributions

Given an asset price process $\{S_t\}$, Bakshi, Kapadia, and Madan (2003) showed that the risk neutral conditional moments of τ−period log return $R_t(\tau) = \ln(S_{t+\tau}) - \ln(S_t)$, given time t information, can be written as an integral of out-of-the-money (OTM) call and put option prices. In particular, the arbitrage-free prices of the volatility contract $V(t, \tau) = E_t^Q(e^{-r\tau} R_t(\tau)^2)$, cubic contract $W(t, \tau) = E_t^Q(e^{-r\tau} R_t(\tau)^3)$ and quartic contract $X(t, \tau) = E_t^Q(e^{-r\tau} R_t(\tau)^4)$ at time t can be expressed as

$$V(t, \tau) = \int_{S_t}^{\infty} \frac{2(1 - \ln(K/S_t))}{K^2} C(t, \tau; K) dK$$
$$+ \int_0^{S_t} \frac{2(1 - \ln(K/S_t))}{K^2} P(t, \tau; K) dK \quad (8)$$

$$W(t, \tau) = \int_{S_t}^{\infty} \frac{6\ln(K/S_t) - 3(\ln(K/S_t))^2}{K^2} C(t, \tau; K) dK$$
$$+ \int_0^{S_t} \frac{6\ln(K/S_t) - 3(\ln(K/S_t))^2}{K^2} P(t, \tau; K) dK \quad (9)$$

$$X(t, \tau) = \int_{S_t}^{\infty} \frac{12(\ln(K/S_t))^2 - 4(\ln(K/S_t))^3}{K^2} C(t, \tau; K) dK$$
$$+ \int_0^{S_t} \frac{12(\ln(K/S_t))^2 - 4(\ln(K/S_t))^3}{K^2} P(t, \tau; K) dK \quad (10)$$

where Q represents the risk neutral measure, r is risk-free rate, while $C(t, \tau; K)$ and $P(t, \tau; K)$ are the prices of European calls and puts written on the underlying asset with strike price K and expiration τ periods from time t. Therefore, the time t conditional risk neutral moments (mean, variance, skewness, and excess kurtosis) of $\ln(S_{t+\tau})$ are:

$$\text{Mean}(t, \tau) = \mu(t, \tau) + \ln(S_t) \quad (11)$$

$$\text{var}(t, \tau) = e^{r\tau} V(t, \tau) - \mu(t, \tau)^2 \quad (12)$$

$$\text{Skew}(t, \tau) = \frac{e^{r\tau} W(t, \tau) - 3\mu(t, \tau)e^{r\tau} V(t, \tau) + 2\mu(t, \tau)^3}{[e^{r\tau} V(t, \tau) - \mu(t, \tau)^2]^{3/2}}$$
$$\quad (13)$$

$$\text{EKurt}(t, \tau) = (e^{r\tau} X(t, \tau) - 4\mu(t, \tau)e^{r\tau} W(t, \tau)$$
$$+ 6e^{r\tau} \mu(t, \tau)^2 V(t, \tau) - 3\mu(t, \tau)^4)$$
$$/ ([e^{r\tau} V(t, \tau) - \mu(t, \tau)^2]^2) - 3 \quad (14)$$

where $\mu(t, \tau) = e^{r\tau} - 1 - e^{r\tau} V(t, \tau)/2 - e^{r\tau} W(t, \tau)/6 - e^{r\tau} X(t, \tau)/24$.

Typically we cannot implement directly Equations (8), (9), and (10) since we do not have a continuum of strike prices available. Hence, the integrals are replaced by approximations involving weighted sums of OTM puts and calls across (a subset of) available strike prices. While the approximation entails a discretization bias, Dennis and Mayhew (2002) reported that such biases are typically small even with a small set of puts and calls.[7] In particular, the integrals in Equations (8), (9), and (10) are evaluated by a trapezoid approximation method described in Conrad, Dittmar, and Ghysels (2013). Therefore, the above formulas—computed using discrete weighted sums—yield estimates of the mean, variance, skewness, and excess kurtosis of the risk neutral density. In the empirical work, we follow the practical implementation discussed by Conrad, Dittmar, and Ghysels (2013).

Note that the approach pursued here is different from statistical analysis based on return-based estimation via sample counterparts of population moments. The use of derivative contracts for the purpose of pricing and risk management is widespread in the financial industry. We follow exactly this strategy, by computing moments of risk neutral densities obtained from extracting information from existing derivatives contracts. Then we will use parametric densities based on the extracted moments to compute various objects of interest, ranging from pricing other derivative contracts to value-at-risk computations, etc.

3.2 Risk Neutral Moments and Feasible Domains

The very first question we address is whether the range of moments that are extracted from market data fall within the feasible domain of the respective densities. Figure 1 plots daily skewness and kurtosis extracted from S&P 500 index options with 5–14 days to maturity for a sample covering January 4, 1996, to October 30, 2009.[8] There are 1258 dots in Figure 1(a). Each dot represents a daily combination of squared skewness and excess kurtosis calculated via expressions (13) and (14). The pairs of OTM call/put options involved in the calculation of risk neutral moments per day ranges from 3 to 85, with an average of 21.26 (see Table 1).

Superimposed on the dots are the feasible domains of the NIG, VG, and GST distributions, that is, the area under the curves with labels "VG," "NIG," and "GST," respectively. The line labeled "Upper bound" represents the largest possible

---

[7]As noted by Dennis and Mayhew (2002) and Conrad, Dittmar, and Ghysels (2013), it is critical to select a set of puts and calls that are symmetric in strike prices. In contrast, discretely weighted sums of asymmetrically positioned puts/calls result in biases.

[8]The data are obtained from Optionmetrics through Wharton Research Data Services. A number of data filters are applied to screen for recording errors - these are standard in the literature and described elsewhere, see, for example, in Battalio and Schultz (2006) and Conrad, Dittmar, and Ghysels (2013), among others. Notably we use filters to try to ensure that our results are not driven by stale or misleading prices. In addition to eliminating option prices below 50 cents and performing robustness checks with additional constraints on option liquidity, we also remove options with less than one week to maturity, and eliminate days in which closing quotes on put-call pairs violate no-arbitrage restrictions. We encountered some days with two pairs of OTM call/put options. The literature often puts a lower bound of three pairs to avoid noisy moment estimates.
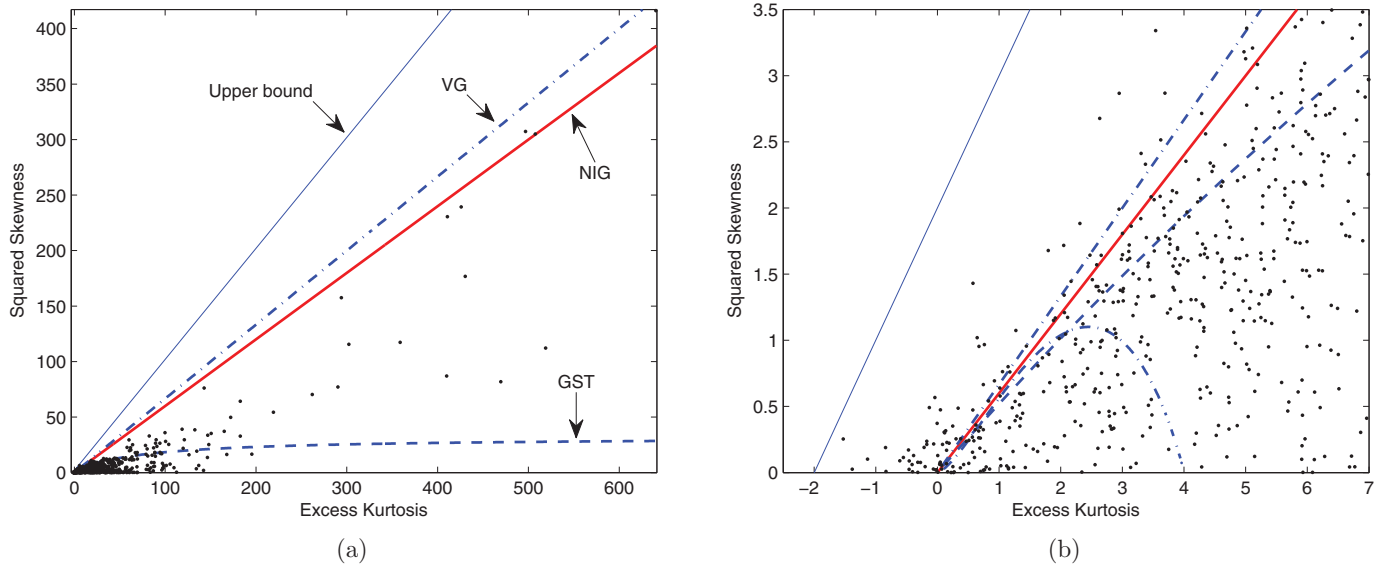
Figure 1. Daily squared skewness and kurtosis of SPX from January 1996 to October 2009. Panel (a) plots the daily squared skewness and excess kurtosis extracted from the S&P 500 index options with 5–14 days to maturity for a sample covering January 4, 1996, to October 30, 2009. Superimposed are the feasible domains of the NIG, VG, and GST distributions, that is, the area under the curves with labels "VG," "NIG," and "GST," respectively. The line labeled "Upper bound" represents the largest possible skewness-kurtosis combination of an arbitrary random variable. When zooming in on Panel (a) we obtain Panel (b). The area under the curve with a range of excess kurtosis from 0 and 4 is the feasible domain of A-type Gram-Charlier series expansion (GCSE).

skewness-kurtosis combination of an arbitrary random variable, and it is obtained by the formula $S^2 = K + 2$. The region above the bound is the so-called impossible region. As reported in Table 1, the VG-feasible region covers 94.54% of the data points, and the NIG feasible region covers 92.59%. Finally, the coverage of the GST feasible region is 81.44% of the data points.

Table 1 summarizes coverage of VG, NIG, and GST distributions, and number of contracts used to compute the risk neutral moments. Though the VG, NIG, and GST distributions cannot accommodate any arbitrary skewness-kurtosis combinations, most of the S&P 500 options with short maturities feature skewness-kurtosis combinations within the feasible region of the VG and NIG distributions. The number of contracts listed in Table 1 deserves some clarification. The smallest number of contracts is 3. This does not mean that we fit four moments with three prices. The header in the table says "put/call pairs." Three prices, means 6 contracts, 3 puts, and 3 calls. This prompts the question as to how many contracts are needed to obtain reliable moment estimates. This is not so straightforward to answer, but is discussed notably in Dennis and Mayhew (2002). The accu-

racy depends not only on the number of contracts, but also how well they cover the range over which to compute the discrete approximations to the integral formulas discussed earlier.

When we zoom in Figure 1(a) we obtain the next plot (b). The area under the curve with a range of excess kurtosis from 0 and 4 is the feasible domain of A-type Gram-Charlier series expansion (GCSE). The A-type GCSE and the Edgeworth expansion have been studied by Madan and Milne (1994), Rubinstein (1998), Eriksson, Ghysels, and Wang (2009), among others as a way to approximate the unknown risk neutral density. Since the Edgeworth expansion admits a smaller feasible region than the A-type GCSE (see Barton and Dennis 1952 for more detail), we only draw the feasible domain of the A-type GCSE in Figure 1. Clearly, most of the dots are outside the feasible region of the A-type GCSE and, hence, outside the feasible region of the Edgeworth expansion.

Figure 2 plots the squared skewness and excess kurtosis using options with longer time to expiration: 17–31 days (around 1 month), 81–94 days (around 3 months), and 171–199 days (around 6 months). Table 1 contains again the numerical values

Table 1. Coverage of VG, NIG, and GST distributions

	Total	Pairs of call/put			VG	NIG	GST	Impossible region
		min	max	average				
5–14	1258	3	85	21.26	94.54%	92.59%	81.44%	0
17–31	1855	3	90	22.22	74.45%	69.06%	50.08%	0
81–94	1220	3	61	13.96	23.89%	15.66%	3.63%	34/1220
171–199	1294	3	39	11.54	8.06%	5.43%	3.47%	25/1294

The table reports the number of contracts used to compute the daily risk neutral moments, and the coverage of the VG, NIG, and GST distributions for the S&P 500 index options with 5–14, 17–31, 81–94, and 171–199 days to maturity, from January 4, 1996 to October 30, 2009.

(a) TTM: 17∼31 days

(b) TTM: 17∼31 days

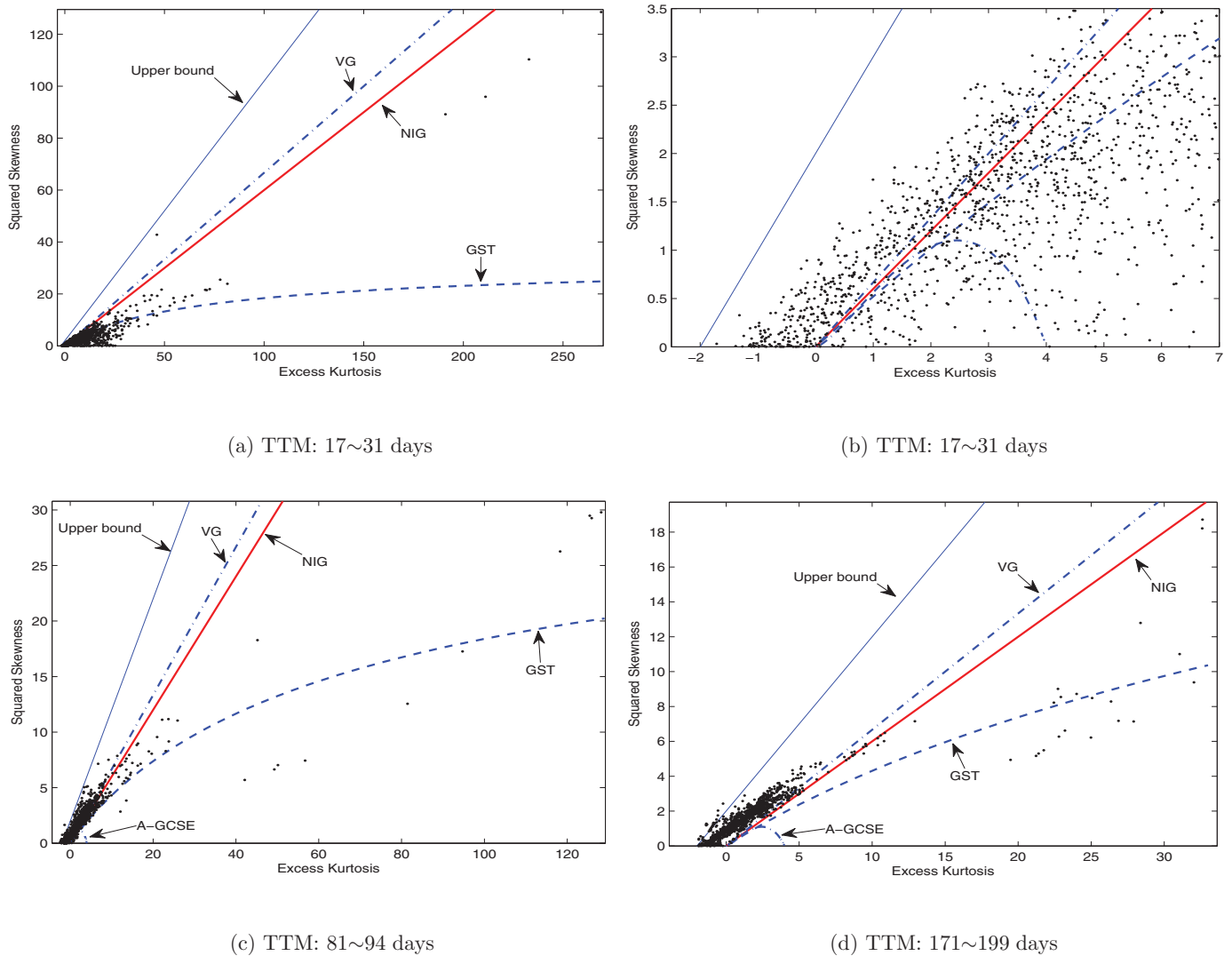(c) TTM: 81∼94 days

(d) TTM: 171∼199 days

Figure 2. Daily squared skewness and kurtosis of SPX from January 1996 to October 2009, continued.

of the feasible domain coverage for the various distributions. The VG feasible region covers 74.45% of the data points (out of 1855 contracts) for 17–31 days to maturity, and 23.89% (out of 1220) for 81–94 days to maturity. Moreover, when the time-to-maturity increases to 171–199 days, the coverage drops to 8.06% (out of 1294). These observations are consistent with the fact that the returns are more leptokurtic when sampled more frequently. It should also be noted that a few data points in Figure 2(c)–(d) are in the impossible region – 2.78% (81–94) and 1.93% (171–199), respectively, according to the figures in the last column of Table 1. This could be due to estimation error in the moments. Moreover, we plot in Figure 3 the skewness and excess kurtosis for the various time-to-maturities. This is consistent with the study of Bakshi, Kapadia, and Madan (2003) that "the risk neutral distribution of the index is generally left skewed."

The overall picture that emerges from our analysis so far is that the classes of distributions we examine have appealing properties in terms of the skewness-kurtosis coverage—in particular, when compared with approximating densities (e.g., A-type GCSE and the Edgeworth expansion) proposed in the

prior literature. The feasible domain does become more restrictive for longer term maturities beyond 3 months. We, therefore, focus on options with maturities up to 1 month.

3.3 Simulation Evidence

We want to assess the accuracy of the various distributions via a simulation experiment. To do so we characterize risk neutral densities with a commonly used framework in financial asset pricing and risk management, namely continuous time jump diffusion processes. We also consider the GARCH option pricing models for the numerical evaluation. In particular, we simulate the log prices using either affine jump-diffusions or a GARCH(1,1), which yield explicit expressions for the risk neutral density and option prices. We select parameters that are empirically relevant, allowing us to assess numerically how accurate the various approximating distributions are in realistic settings. The discussion will focus again on the VG, NIG, and GST distributions.
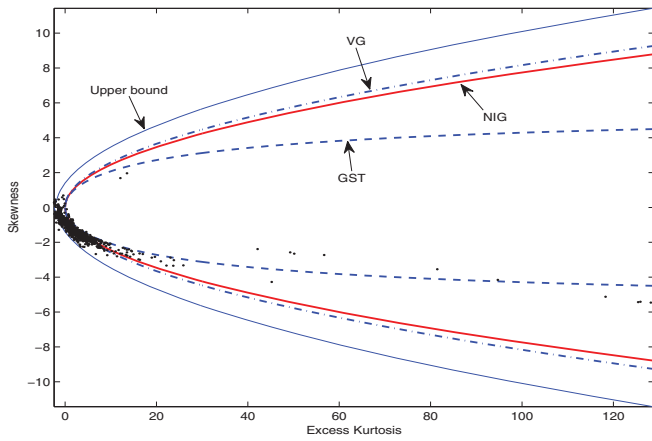
3.3.1 Affine Jump-Diffusion Models. Suppose that the log price process $Y_t = \ln(S_t)$ is generated from the following affine
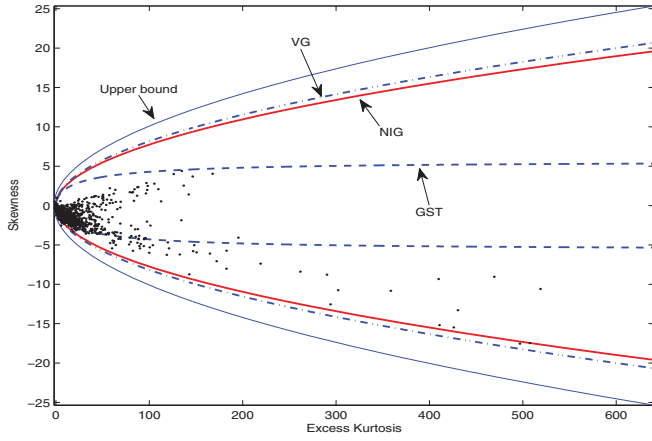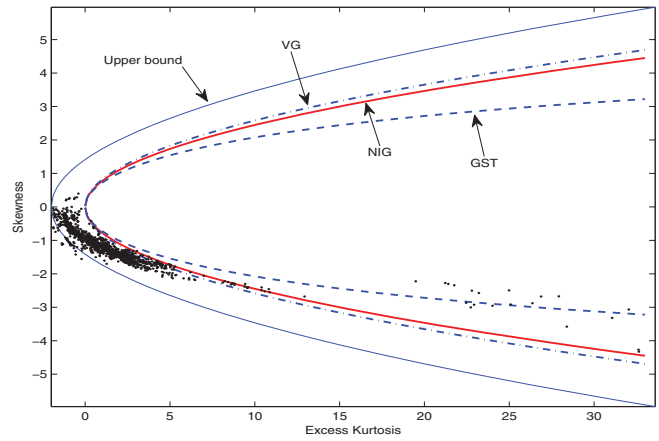
(a) 5~14 days
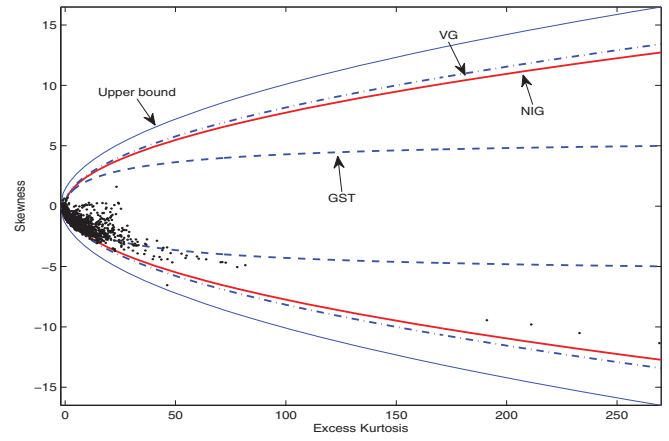
(b) 17~31 days

(c) 81~94 days

(d) 171~199 days

Figure 3. Daily skewness and kurtosis of SPX from January 1996 to October 2009.

jump-diffusion model under the risk neutral measure:

$$dY_t = \left(r - \lambda_J \bar{\mu} - \frac{1}{2} V_t\right)dt + \sqrt{V_t}\,dW_t^1 + dN_t$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\rho\sqrt{V_t}\,dW_t^1 + \sigma\sqrt{1-\rho^2}\sqrt{V_t}\,dW_t^2,$$
(15)

where N_t is a compound Poisson process with Lévy measure $v(dy) = \lambda_J f(dy)$, and $\bar{\mu} = \int_\mathbb{R} e^y f(dy) - 1$ is the mean jump size. W^1, W^2 are two independent Brownian motions, and are independent of N_t as well.

For any $u \in \mathbb{C}$, the conditional characteristic function of Y_T at time t is

$$\Psi_t(u; T, x_t) \doteq E(e^{uY_T}|\mathcal{F}_t) = \exp(\psi_1(u, T-t)$$
$$+ \psi_2(u, T-t)V_t + uY_t),$$
(16)

where $x_t = (Y_t, V_t)$, $\psi_1(u, \tau) = ru\tau - \kappa\theta(\frac{\gamma+b}{\sigma^2}\tau + \frac{2}{\sigma^2}\ln[1 - \frac{\gamma+b}{2\gamma}(1 - e^{-\gamma\tau})]) - \lambda_J\tau(1 + \bar{\mu}u) + \lambda_J\tau\int_\mathbb{R} e^{uy}f(dy)$, $\psi_2(u, \tau) = -\frac{a(1-e^{-\gamma\tau})}{2\gamma-(\gamma+b)(1-e^{-\gamma\tau})}$, and $b = \sigma\rho u - \kappa$, $a = u(1-u)$, and

$\gamma = \sqrt{b^2 + a\sigma^2}$ (see Duffie, Pan, and Singleton 2000). The density function of Y_T conditional on information up to time t is therefore $f_t(y; T, x_t) = \frac{1}{\pi}\int_0^\infty e^{-iuy}\Psi_t(iu; T, x_t)du$, which follows from inverse Fourier transform. The price of European call option written on Y with maturity T and strike price K is defined as $C_t(K; T, x_t) = E((e^{Y_T} - K)^+|\mathcal{F}_t)$, and it has an explicit expression: $C_t(K; T, x_t) = P_1(K, t, T, x_t) - KP_2(K, t, T, x_t)$ where $P_1(K, t, T, x_t) = \frac{1}{2}s_t - \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \text{Im}[\frac{e^{iv\ln(K)}\Psi_t(1-iv; T, x_t)}{v}]dv$, $P_2(K, t, T, x_t) = \frac{1}{2}e^{-r(T-t)} - \frac{e^{-r(T-t)}}{\pi}\int_0^\infty \text{Im}[\frac{e^{iv\ln(K)}\Psi_t(-iv; T, x_t)}{v}]dv$ and Im represents the imaginary part of a complex number (see Duffie, Pan, and Singleton 2000). The put price follows from the put/call parity. In the numeric calibration, we apply the fast Fourier transform of Carr and Madan (1999) to both $f_t(y; T, x_t)$ and $C_t(K; T, x_t)$ (see also Lee 2004). The details are in Appendix B.

We focus exclusively on two cases involving jumps: Gaussian jumps, that is, $\int_\mathbb{R} e^{uy}f(dy) = \exp(\mu_J u + \sigma_J^2 u^2/2)$, and exponential jumps, that is, $\int_\mathbb{R} e^{uy}f(dy) = (1 - u\mu_J)^{-1}$, as they represent the most realistic processes. The model parameters are taken from Duffie, Pan, and Singleton (2000): $r = 3.19\%$,
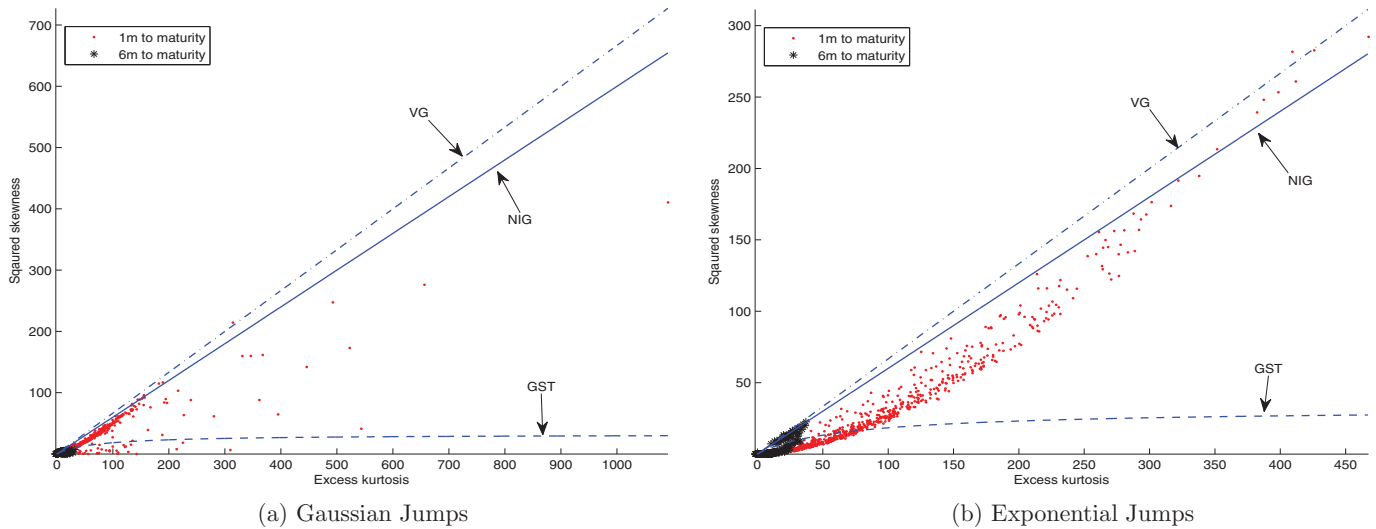
(a) Gaussian Jumps

(b) Exponential Jumps

Figure 4. Skewness and kurtosis of noisy option prices generated by the AJD model. The figure displays the squared skewness and excess kurtosis calculated from the noisy option prices which are constructed by multiplying a Gamma-distributed noise to the option prices calculated from affine jump-diffusion (15) with Gaussian Jumps (see Panel (a)) and Exponential Jumps (see Panel (b)). Superimposed, we draw the feasible domains of the VG, NIG, and GST distributions.

$\rho = -0.79$, $\theta = 0.014$, $\kappa = 3.99$, $\sigma = 0.27$, $\lambda_J = 0.11$. We consider $\mu_J = -0.14$, and $\sigma_J = 0.15$ for Gaussian jumps, while $\mu_J = 0.14$ for Exponential jumps. To determine the value of $x_0 = (y_0, v_0)$, we simulate the log price process starting from value 0 and draw from the invariant distribution of the volatility process which is a Gamma distribution with characteristic function $\phi(u) = (1 - iu\sigma^2/(2\kappa))^{-2\theta\kappa/\sigma^2}$ (see, e.g., Keller-Ressel 2011). We simulate 1000 sample paths. For each, we drop the first 1000 observations and use the 1001th observation from simulation as the value of x_0.

Note that in reality, the observed option prices might deviate from the true prices due to mistakes in the recording of data, bid-ask spread, nonsynchronicity, liquidity premia, or other market frictions. We add observational errors to the simulated call/put prices. The "observed" call/put prices are constructed by multiplying a noise ϵ to the theoretical call/put prices, that is, $C_0^*(K; T, x_0) = \epsilon C_0(K; T, x_0)$ and $P_0^*(K; T, x_0) = \epsilon P_0(K; T, x_0)$. The noise ϵ is assumed to be Gamma distributed with mean 1 and variance s where s is the bid-ask spread. Note that the spread is larger for the in-the-money options than for the OTM options. We set $s = \min(M(C_0(K; T, x_0)), M(C_0(K; T, x_0) + K - y_0))$, which is the maximum spread allowed by the Chicago Board Options Exchange. $M(\cdot)$ is a piecewise linear function on the interval [0, 50], with knots $M(0) = 1/8$, $M(2) = 1/4$, $M(5) = 3/8$, $M(10) = 1/2$, $M(20) = 3/4$, and $M(q) = 1$ for $q \geq 50$.[9] We eliminate option prices which violate the arbitrage-free bounds.

We use the approximating densities to price European call options and compare them with the observed prices $C_0^*(K; T, x_0)$. Two measures of pricing errors are considered. The first is based on absolute price difference, denoted by L_a. The second is defined in terms of relative price difference, denoted by L_r. The

two measures are defined as follows:

$$L_a = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(C_i^{\text{model}} - C_i^{\text{obs}} \right)^2} \quad \text{and}$$

$$L_r = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\frac{C_i^{\text{model}} - C_i^{\text{obs}}}{C_i^{\text{obs}}} \right)^2}, \tag{17}$$

where C_i^{obs} and C_i^{model} represent the "observed" price and the price estimated from the approximating distribution. The sum is taken over a range of strikes (or moneyness, which is the ratio of asset price S_t and strike price K_i) for contracts written on the same security. These measures are, respectively, absolute and relative pricing errors.

Figure 4 displays the skewness and kurtosis calculated from the noisy option prices which are generated from AJD models with Gaussian jumps and Exponential jumps, respectively, for $T = 1/12$ (in years) and $1/2$ (in years).[10] Table 2 reports the mean pricing errors $\bar{L}_a$ and $\bar{L}_r$, an average of L_a and L_r over 1000 iterations, for three different ranges of moneyness. Figure 4 shows that the most of the skewness-kurtosis combinations are within the feasible domain of the NIG and VG distributions, but the GST is inadequate in terms of feasible domain coverage—for both the Gaussian and Exponential jump cases. This is why we cover in Table 2 only the NIG and VG cases. The pricing errors of the two approximating densities are quite similar—particularly in absolute terms. In relative measures (using the L_r loss function) it seems that the NIG has a slight edge for short maturities (1 month) while the reverse is true for the longer maturity (6 months). Yet, in terms of feasible domain and pricing

---

[9]See Bondarenko (2003) for more details.

[10]It is worth noting that we applied Kolmogorov-Smirnov tests to see whether there were significant differences between the model-based densities and the approximating ones. In almost all cases we rejected the null of identical distributions. The Kolmogorov-Smirnov test statistics reflect overall fit, however, while we will focus on the parts of the distributions that matter for risk management.

Table 2. Mean pricing errors for the noisy call prices generated by the AJD model

| | Gaussian jump | | | | | | Exponential jump | | | | | |
| | (0.7, 0.8) | | (0.8, 0.9) | | (0.97, 1.03) | | (0.7, 0.8) | | (0.8, 0.9) | | (0.97, 1.03) | |
	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$
						1 month						
NIG	0.0000	1.2315	0.0000	1.5701	0.0014	0.2298	0.0001	0.2972	0.0002	0.2919	0.0019	0.2377
VG	0.0000	1.2944	0.0001	2.8800	0.0022	0.3409	0.0001	0.3096	0.0002	0.4269	0.0030	0.3091
						6 months						
NIG	0.0000	1.6878	0.0003	0.8600	0.0019	0.0557	0.0003	0.2501	0.0010	0.2608	0.0023	0.0668
VG	0.0000	1.5009	0.0003	0.9554	0.0021	0.0667	0.0003	0.2327	0.0014	0.3616	0.0037	0.1140

The table reports the mean pricing errors for the noisy call prices in terms of $\bar{L}_a$ and $\bar{L}_r$—an average of absolute pricing error L_a and relative pricing error L_r over 1000 iterations—for three different ranges of moneyness: (0.7,0.8), (0.8,0.9), and (0.97,1.03). The noisy option prices are constructed by multiplying a Gamma-distributed noise to the option prices calculated from affine jump-diffusion (15) with Gaussian jumps and exponential jumps, respectively.

errors it is fair to say that the two classes of distributions are comparable.

3.3.2 GARCH Option Pricing Models. Besides the affine jump-diffusion models, we also consider the GARCH(1,1) option pricing model of Heston and Nandi (2000)—see also Duan (1995)—to assess the approximation errors of the VG, NIG, and GST distributions. We consider the GARCH option pricing example as a reasonable alternative—used by practitioners as well as academics. Many other models—such as stochastic volatility models—require quite involved estimation procedures and filtering of latent volatilities. The GARCH option pricing model is arguably on the same level of simplicity as our moment-implied density approach. Many other methods are much more involved and harder to implement.

The GARCH(1,1) process under the risk neutral measure is

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}\sigma_t^2 + \sigma_t\epsilon_t, \quad \epsilon_t \overset{\text{iid}}{\sim} N(0, 1)$$

$$\sigma_t^2 = \omega + a(\epsilon_{t-1} - c\sigma_{t-1})^2 + b\sigma_{t-1}^2. \quad (18)$$

The time-t price of European call option with strike K at maturity T is explicitly given by

$$C_t(K; T, S_t, \sigma_{t+1}^2)$$
$$= \frac{1}{2}\left(S_t - Ke^{-r(T-t)}\right) + \frac{e^{-r(T-t)}}{\pi}$$

$$\times \left(\int_0^\infty Re\left[\frac{K^{-i\phi}f(t,T;i\phi+1)}{i\phi}\right]d\phi \right.$$
$$\left. - K\int_0^\infty Re\left[\frac{K^{-i\phi}f(t,T;i\phi)}{i\phi}\right]d\phi\right), \quad (19)$$

where $f(t; T, \phi) = S_t^\phi \exp(A_t + B_t\sigma_{t+1}^2)$ and

$$A_t = \phi r + A_{t+1} + \omega B_{t+1} - 0.5\ln(1 - 2aB_{t+1})$$

$$B_t = \phi(-0.5 + c) - 0.5c^2 + \frac{0.5(\phi - c)^2}{1 - 2aB_{t+1}} + bB_{t+1} \quad (20)$$

with terminal condition $A_T = B_T = 0$.

The values of parameters are taken from Table 1 of Heston and Nandi (2000): $r = 0$, $\omega = 5.02 * 10^{-6}$, $a = 1.32 * 10^{-6}$, $b = 0.589$, $c = 422$. We simulate 2000 daily prices under the risk neutral measure, with initial state variables $S_0 = 100$ and $\sigma_1^2 = (0.15^2)/252$. We drop the first 1000 observations, so we end up with a sample path of 1000 daily observations $\{(S_t, \sigma_{t+1}^2), t = 1, \ldots, 1000\}$. For each t, we calculate the call price via (19) and put price using the put/call parity, with $T = 22$ (in days) or 1 month and $T = 132$ (in days) or 6 months. We also consider 14 days to maturity which is more relevant to the option pricing exercise in Section 4.2. Moreover, as described in Section 3.3.1, we add Gamma-distributed noise to the theoretical option prices.

Table 3. Mean pricing errors for the noisy call options simulated from the GARCH model

| | (0.7, 0.8) | | (0.8, 0.9) | | (0.97, 1.03) | |
Moneyness	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$
			14 days			
NIG	0.0939	0.7828	0.0111	0.9763	0.4571	0.3997
VG	0.0915	0.7580	0.0112	0.9830	0.4756	0.4028
			22 days			
NIG	0.0495	0.8666	0.0362	31.0302	0.5566	0.4289
VG	0.0497	0.8758	0.0414	36.0376	0.6105	0.4426
			132 days			
NIG	0.0007	34.9586	0.0248	43.6354	1.9885	0.8831
VG	0.0009	45.3229	0.0251	50.2347	1.9902	0.8837

The table reports the mean pricing errors for the noisy call prices in terms of $\bar{L}_a$ and $\bar{L}_r$—an average of absolute pricing error L_a and relative pricing error L_r over 1000 iterations—for three different ranges of moneyness: (0.7,0.8), (0.8,0.9), and (0.97,1.03). The noisy option prices are constructed by multiplying a Gamma-distributed noise to the option prices calculated from (19).
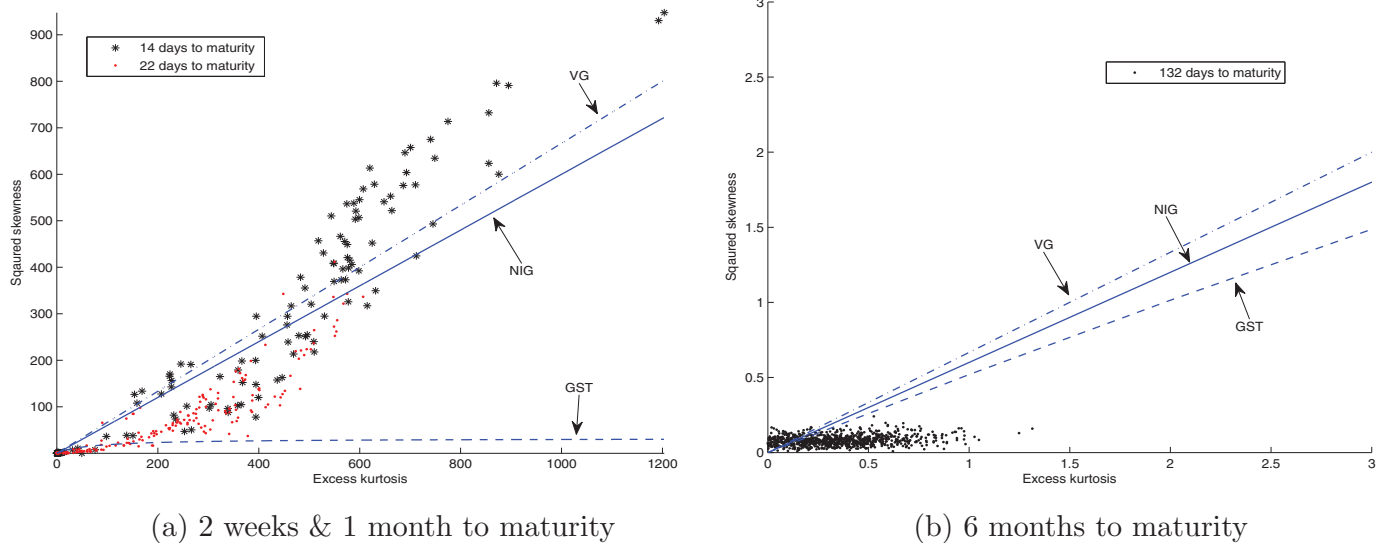
(a) 2 weeks & 1 month to maturity

(b) 6 months to maturity

Figure 5. Skewness and kurtosis of noisy option prices generated by the GARCH model. The figure displays the squared skewness and excess kurtosis calculated from the noisy option prices which are constructed by multiplying a Gamma-distributed noise to the option prices calculated from the GARCH(1,1) option pricing model (18). Superimposed, we draw the feasible domains of the VG, NIG, and GST distributions.

The noisy prices, which violate the arbitrage-free bounds, are removed.

Figure 5 displays the squared skewness and excess kurtosis calculated from the noisy option prices with 14 days and 22 days to maturity (Panel (a)) and 132 days to maturity (Panel (b)). Superimposed are the feasible domains of the VG, NIG, and GST distributions. Since the GST distribution has a more restrictive feasible domain, we consider only the VG, NIG distributions to approximate the option prices implied by the GARCH(1,1). The mean pricing errors are reported in Table 3. The VG and NIG approximations do a poor job in terms of the relative measure for the deep OTM options with time-to-maturity over 1 month. When the time-to-maturity changes from 6 months to 2 weeks, the relative pricing error decreases significantly for all the three moneyness ranges, and the two distributions—VG and NIG—are comparable.

4.   OPTION PRICING

We start from the observation that a good method of option pricing should be able to price contracts in situations where only a few contracts are traded. We, therefore, design experiments where we compute risk neutral moments using a smaller set of option contracts than is available. We then price contracts which are not used to compute the moments via the approximate densities. This means we use a subset of existing market prices to extract risk neutral moments and another set of existing market prices to appraise the accuracy of the approximate option prices. Hence, we will examine, through an experimental design, how data sparseness affects option pricing via VG, NIG, and GST density approximations. We do this in such a way that we can appraise how well our methods perform to price options that are deep out-of-the-money versus options that are not. The former is the most challenging task to achieve for any method—and we show that we do very well. It will be helpful to explain the

empirical investigation first with an illustrative example—which is covered first—followed by a full sample implementation.

4.1   An Illustrative Empirical Case

We start with an illustrative example and then proceed to a full sample formal analysis. To illustrate the design we use SPX options (European options on the S&P 500 index) priced on September 23, 2009 (which is a Wednesday) with 7 days to maturity (i.e., September 30, 2009) to illustrate the procedure. There are 42 call options and 42 put options with 7 days to maturity on 2009/9/23. The data consists of 20 OTM calls and 22 OTM puts. Figure 6(a) plots all the available OTM options. We then focus on the options with moneyness (i.e., the ratio of index level S and strike price K) between 0.8 and 1.2, which are displayed in Figure 6(b). Since the valuation of formulas (8), (9), and (10) require the calls and put to have the same (or similar) distance from strike price to the index level to mitigate estimation error (recall the discussion in footnote 7), we use $|K/S - 1|$ as x-axis in Figure 6(b). To evaluate the option pricing via the VG, NIG, and GST approximations and assess how data sparseness affects pricing accuracy, we consider two strategies: (1) select the call/put pairs with $|K/S - 1|$ closest to 0.03, 0.06, 0.09, 0.12, and 0.15, and (2) select the call/put pairs with $|K/S - 1|$ closest to 0.03, 0.09, and 0.15. Therefore, we pick call/put pairs in Circles 1–5 in Figure 6(b) for Strategy 1, and call/put pairs in Circles 1, 3, and 5 for Strategy 2. For each strategy, we derive the approximating densities—the VG, NIG, and GST distributions, and then price all the "unused" call options.[11]

Table 4 reports the pricing errors measured by L_a and L_r (defined in (17) where C_i^{obs} is the observed market call option price). Denote by $[a, b]$ the range of moneyness of call options

---

[11]We do not consider call options with moneyness greater than 1.03, because they are not liquid, that is, infrequently traded.
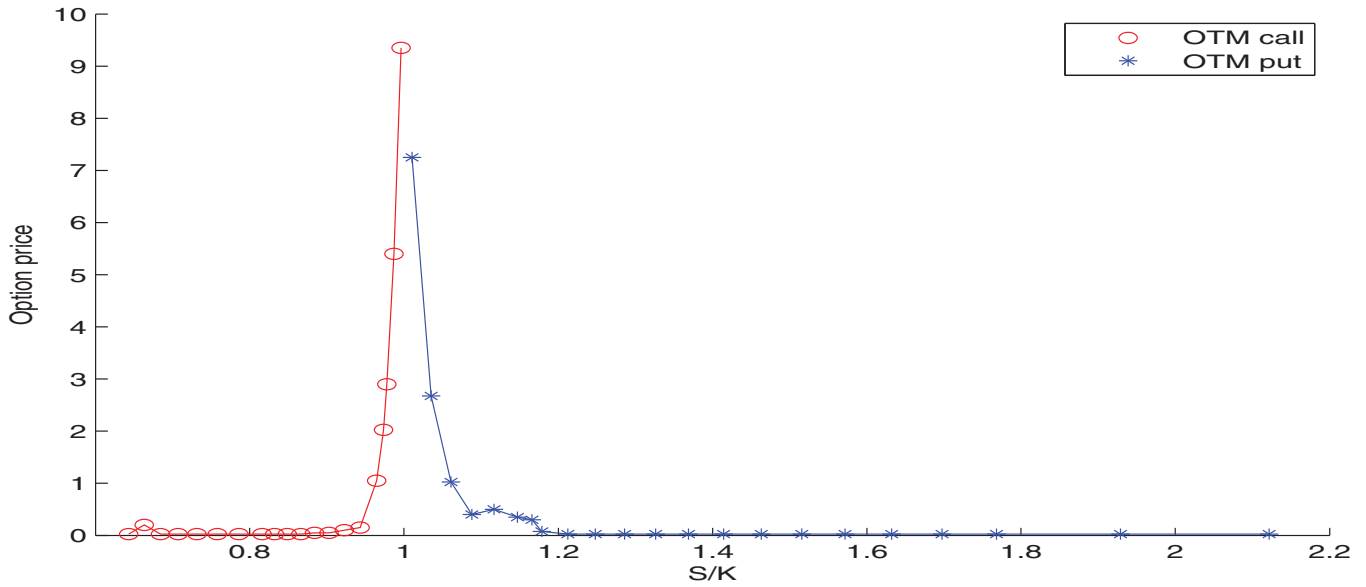
(a) OTM calls/puts
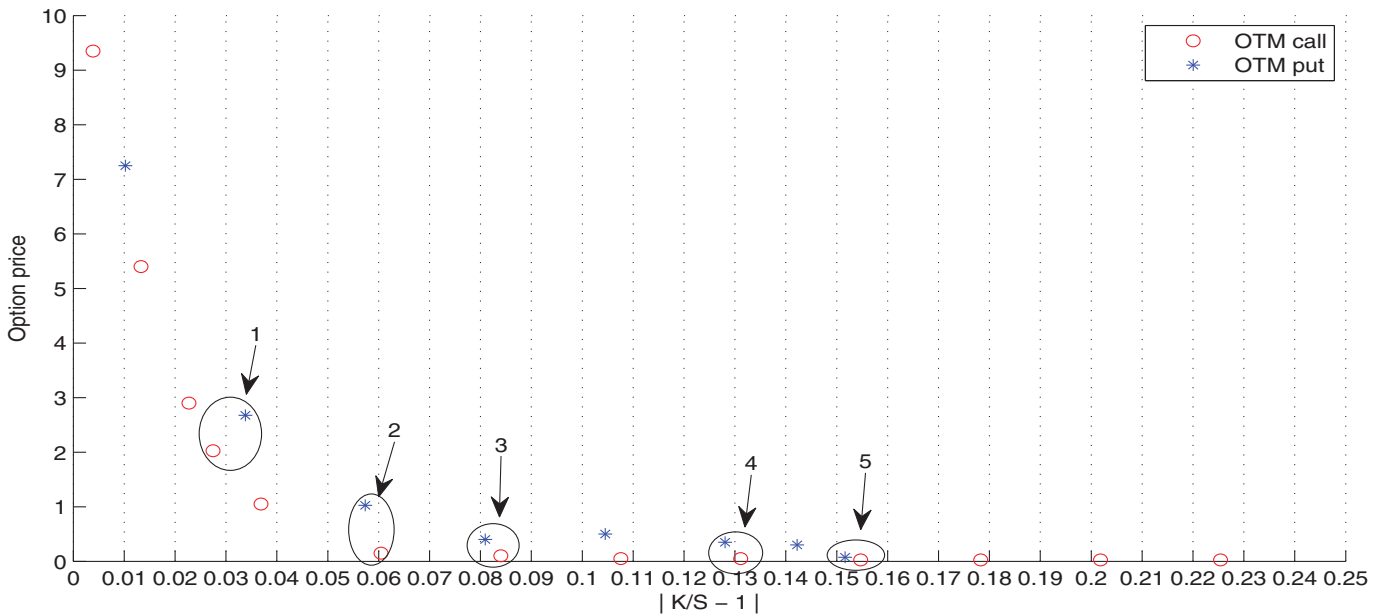
(b) Data used

Figure 6. OTM SPX calls and puts, 2009/9/23, expired on 2009/9/30. The figure plots all the available out-of-the-money S&P 500 index call/put options priced on September 23, 2009 with 7 days to maturity.

that are used for pricing. We divide all the "unused" calls into three groups: Group 1 contains "unused" options with moneyness within $[a, b]$, Group 2 contains "unused" options with moneyness less than a, and Group 3 contains "unused" calls with moneyness greater than b. In particular:

- Strategy 1: $a = 0.8660$ and $b = 0.9733$. There are 2 points in $[a, b]$, 10 points in '$< a$', 4 points in '$> b$', a total of 16 unused points.

- Strategy 2: $a = 0.8660$ and $b = 0.9733$. There are 4 points in $[a, b]$, 10 points in '$< a$', 4 points in '$> b$', a total of 18 unused points.

The three groups are labeled as $[a, b]$, $< a$, and $> b$, respectively, in Table 4. The following observations emerge from examining Table 4:

- The GST approximation is not feasible using Strategy 1, while it is for Strategy 2.

Table 4. Pricing SPX call options on 2009/9/23

	Overall		[a, b]		< a		> b	
	L_a	L_r	L_a	L_r	L_a	L_r	L_a	L_r
				Strategy 1				
NIG	2.6837	0.9449	0.6992	0.9606	0.0675	0.9998	5.3434	0.7816
VG	3.1201	0.9194	0.6105	0.7757	0.0674	0.9893	6.2244	0.7962
GST	NA	NA	NA	NA	NA	NA	NA	NA
				Strategy 2				
NIG	2.2326	0.8451	0.3739	0.6122	0.0673	0.9782	4.7200	0.6685
VG	2.4797	0.8147	0.3193	0.4496	0.0672	0.9597	5.2494	0.6942
GST	2.0950	0.9493	0.5315	1.0000	0.0675	1.0000	4.4110	0.7454

The table reports the pricing errors measured by L_a and L_r (defined in (17) where C_i^{obs} is the observed market call option price) for S&P 500 index call options on September 23, 2009, with 7 days to maturity. Two strategies are considered. For Strategy 1, we use five pairs of call/put options to construct the approximating VG, NIG, and GST distributions, which are further used to price other unused options. For Strategy two, three pairs of call/put options are selected to do the pricing. Denote by $[a, b]$ the range of moneyness of call options that are used for pricing. We divide all the "unused" calls into three groups: Group 1 contains "unused" options with moneyness within $[a, b]$, Group 2 contains "unused" options with moneyness less than a, and Group 3 contains "unused" calls with moneyness greater than b. The three groups are labeled as $[a, b]$, $< a$, and $> b$, respectively.

- The pricing errors are relatively small in Strategy 2. In other words, the accuracy is improved when we use fewer options to derive the NIG, VG, and GST approximations.
- Overall, the VG approximation is best, though Strategy 1 picks NIG and Strategy 2 picks GST for group 3—perhaps in light of a small sample size in that group.

Since Strategy 2 yields a smaller pricing error than Strategy 1 in terms of both absolute and relative measures, we take a closer look at this phenomenon. There are 16 "unused" calls in Strategy 1 and 18 in Strategy 2. We pick up the "unused" call options in common (i.e., 16) and compare their market values with the prices derived from both strategies. The results are shown in Figure 7, where "VG1" and "NIG1" refer to the option prices estimated by VG and NIG, respectively, using Strategy 1, while "VG2" and "NIG2" are for Strategy 2. "Market" means the observed market price. It is apparent that the prices derived from Strategy 2 is much closer to the market prices, especially for the at-the-money (ATM) options (i.e., S/K between 0.97 and 1.03).

Since the pricing errors are smaller in Strategy 2 with fewer observations, we have that the accuracy is improved when fewer options are used to derive the NIG, VG, and GST approximations. This may sound counterintuitive since usually more observations often lead to high accuracy. It is worth recalling that we are dealing with approximate models where more data may either improve or worsen the approximation. Hence, we are not in a regular setting of sampling theory with increased accuracy with more data. This explains why fewer observations may, as it turns out, be better.

4.2 Full Sample Analysis

We examine now all the SPX options from January 4, 1996 to October 30, 2009, which are priced on *Wednesdays*.[12] For each Wednesday and time-to-maturity T_0, we pick the ATM call

option with strike price closest to the index level, and then calculate its Black-Scholes implied volatility (IV). We, therefore, obtain a sample of IVs for time-to-maturity T_0. Based on this sample, we obtain the empirical first and third quartiles, that is, Q_1 and Q_3. Denote by $D(1, T_0)$ all the Wednesday calls whose IV is less than Q_1, $D(2, T_0)$ the Wednesday calls whose IV is between Q_1 and Q_3, and $D(3, T_0)$ the set of Wednesday calls whose IV is larger than Q_3. Based on observations in Section 3.2, we consider T_0 as one week (i.e., 5–14 days), 1 month (i.e., 17–31 days), and 3 months (i.e., 81–94 days). Therefore, we end up with nine different combinations. Table 5 lists, for each combination, the number of days which yield skewness-kurtosis combination within the feasible domain of the VG, NIG, GST distributions using Strategies 1 and 2, respectively.[13] Namely, the triplets in Table 5 are the numbers of days feasible for the VG, NIG, and GST distributions. For instance, the triplet (27, 21, 0) means that the VG distribution can model 27 dates out of 46 (i.e., first quartile yields 46 data points), NIG can model 21 data points while none of the 46 days yields a skewness-kurtosis combination in the feasible domain of the GST distribution. We note that, as expected, the GST is the most restrictive. We, therefore, consider two sample configurations—one where all three distributions are applicable. This is a relatively small sample that arguably gives unfair advantage to the more restrictive GST distribution. Hence, we also examine a sample where only the VG and NIG distributions are feasible. This is a larger and hence more reliable sample. We analyze the pricing errors for days identified in Table 5 starting with the most restrictive sample configurations involving all days where GST, VG, and NIG distributions apply. We focus on the set $D(2, T_0)$ with $T_0 = 17–31$ and 81–94, $D(3, T_0)$ with $T_0 = 5–14$ and 17–31 because they have relatively more observations than the other cells. For days identified in each of the cells, we repeat the analysis described for the single day 2009/9/23, and derive L_a and L_r. We then average L_a and L_r within each cell. These averages, denoted by $\bar{L}_a$ and $\bar{L}_r$ respectively, are reported in Table 9.

To appraise the performance of the VG, NIG, and GST approximations, we use the prices computed from the

---

[12]We pick Wednesday—as this is common in the empirical option pricing literature—since it avoids some of the day-of-the-week effects that occur particularly on Fridays and Mondays.

[13]We removed days which have fewer than five calls or five puts.
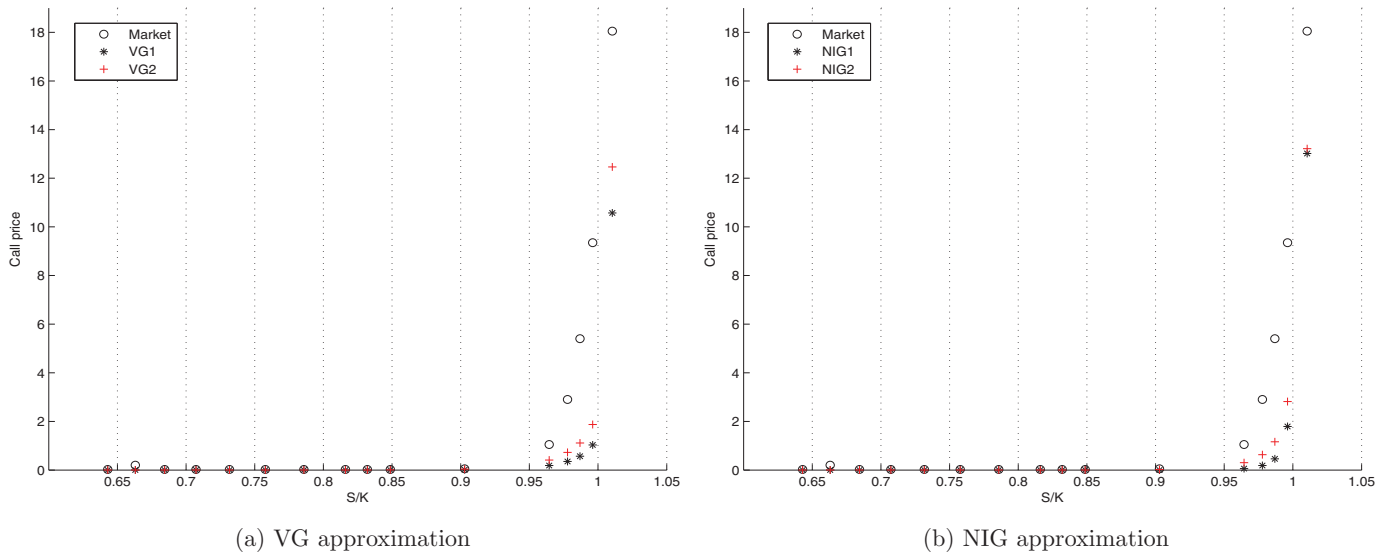
(a) VG approximation

(b) NIG approximation

Figure 7. Compare Strategy 1 with Strategy 2 for SPX call options on September 23, 2009. The figure compares the market prices of 16 call options (on September 23, 2009 with 7 days to maturity) with the prices derived from Strategy 1 and Strategy 2. "VG1" and "NIG1" refer to the option prices estimated by VG and NIG, respectively, using Strategy 1, while "VG2" and "NIG2" are for Strategy 2. "Market" means the observed market price.

Black-Scholes model and the GARCH(1,1) option pricing model ([18]) as benchmarks, which will be denoted BLS and GARCH, respectively, in the tables. The volatility that enters the Black-Scholes formula is the Black-Scholes implied volatility of the previous-day ATM call option, which has the shortest time-to-maturity and strike price closest to the index level.[14] To obtain the GARCH option pricing ([19]), we first estimate model parameters by considering the following GARCH(1,1) under the physical measure

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r + \tilde{\lambda}\sigma_t^2 + \sigma_t\epsilon_t, \quad \epsilon_t \sim N(0, 1)$$

$$\sigma_t^2 = \omega + a(\epsilon_{t-1} - \tilde{c}\sigma_{t-1})^2 + b\sigma_{t-1}^2. \quad (21)$$

We use the S&P 500 index over the full sample—from January 4, 1996 to October 30, 2009 to estimate the parameters $(a, b, \tilde{c}, \omega, \tilde{\lambda})$. The maximum likelihood estimators are $(\hat{a}, \hat{b}, \hat{\tilde{c}}, \hat{\omega}, \hat{\tilde{\lambda}}) = (1.72e - 05, 0.4895, 164.1, 7.5e - 05, 1.4685)$, and hence $\hat{c} = \hat{\tilde{c}} + \hat{\tilde{\lambda}} + .5 = 166.0685$.

What do we learn from Table 9? For median volatility regimes $(D(2, .))$ we note that option pricing based on Black-Scholes implieds perform best in absolute terms (L_a) but not in relative terms. This observation applies to both strategies and across moneyness. In terms of relative pricing errors the picture is quite different. Namely, in terms of relative pricing errors, we find that either the VG approximation or the GST one is the best. It should also be noted, however, that the NIG and VG approximations typically yield similar relative pricing errors. The GST appears to perform better for ATM options $> b$. It is also worth noting that Strategies 1 and 2 yield comparable pricing errors. This is rather impressive as it means that we can use less data (i.e., contracts) and still find similar pricing errors.

In turbulent times $(D(3, .))$ we note that BLS starts to perform poorly both in absolute and relative pricing error terms. Overall—all maturities pooled—we find GST to perform best in absolute terms, but the VG approximation is again clearly dominant, with NIG quite similar in terms of performance. It is clear that the distributional approximations make a big difference compared to Black-Scholes implied option pricing as the pricing errors are substantially larger for BLS. In terms of relative pricing errors, the differences between BLS and distributional approximations are particularly important.

What happens when we drop the GST distribution, which is most restrictive in terms of feasibility? The results—where we remove the GST distribution—appear in Table 10. Given the different sample configuration, we look at a larger set of combinations, namely, we look at the sets $D(1, T_0)$ with $T_0 = 5–14$ and $17–31$, $D(2, T_0)$ with $T_0 = 5–14$, $17–31$ and $81–94$, and finally $D(3, T_0)$ with $T_0 = 5–14$ and $17–31$. Hence, we not only looking at a larger dataset within each cell (see Table 5) but also a broader set of maturity/moneyness/volatility-state combinations. We start with the cases that are common

Table 5. Number of days that yield skewness-kurtosis combination within the feasible domains of the VG, NIG, GST distributions, SPX options (1996–2009)

T_0	Total	Strategy	$D(1, T_0)$	$D(2, T_0)$	$D(3, T_0)$
$5 \sim 14$	184	1	(27, 21, 0)	(66, 55, 3)	(41, 39, 19)
		2	(34, 28, 2)	(78, 76, 32)	(43, 40, 32)
$17 \sim 31$	352	1	(34, 22, 1)	(114, 88, 15)	(60, 48, 25)
		2	(65, 55, 2)	(142, 126, 40)	(62, 53, 31)
$81 \sim 94$	255	1	(16, 7, 0)	(52, 36, 17)	(5, 4, 4)
		2	(22, 10, 0)	(43, 30, 14)	(4, 4, 3)

The table lists, for each of nine combinations, the number of days which yield the risk neutral skewness-kurtosis combination within the feasible domains of the VG, NIG, GST distributions using Strategies 1 and 2, respectively. The skewness and kurtosis are extracted from S&P 500 index options from January 4, 1996 to October 30, 2009.

---

[14]This is considered the most heavily traded and accurately priced contract; see Bates ([1995]) and Chernov and Ghysels ([2000]).

between Tables 9 and 10. We note that whenever GST is best when feasible the NIG distribution fills the void. However, since NIG and VG are again often close—it is also the case that VG is equally appealing. Overall, we find quantitatively the same results as in the more constraint sample where GST is feasible. Namely, Strategies 1 and 2 appear similar, the absolute pricing errors are dominated by BLS in the low volatility regime—as before—but in terms of relative pricing errors there are clear gains across all moneynesses and maturities. Moreover, once we move to high volatility states BLS no longer performs well. Finally, it is also worth noting that the GARCH option pricing model reported in Tables 9 and 10 performs very poorly, both in absolute and relative terms. Hence, the methods we propose are clearly superior to the GARCH option pricing model.

5. FROM THE RISK NEUTRAL MEASURE TO THE PHYSICAL MEASURE

In this section, we discuss the relation between the risk neutral measure and the physical measure in terms of the first four moments.

Let $f_t(x; \tau)$ be the conditional density function of $R_t(\tau) = \ln(S_{t+\tau}) - \ln(S_t)$ at time t under the risk neutral measure Q, while $\tilde{f}_t(x; \tau)$ the conditional density under the physical measure P. Consider a power utility function with coefficient of relative risk aversion θ_1, and scaling factor θ_0. The pricing kernel is then

$$m_t(R_t(\tau), \theta) = \theta_0 e^{-\theta_1 R_t(\tau)}. \tag{22}$$

The physical density, risk neutral density, and utility function are related in the following way:

$$\tilde{f}_t(x; \tau) = \frac{e^{x\theta_1} f_t(x; \tau)}{\int e^{y\theta_1} f_t(y; \tau) dy} \tag{23}$$

and θ_1 should be such that $\int e^{y\theta_1} f_t(y; \tau) dy$ exists.[15] The following proposition links the cumulants under the risk neutral measure and the physical measure.

Proposition 5.1. Suppose that there exist $0 < r_q, r_p < \infty$ such that $\int e^{y\lambda} f_t(y; \tau) dy < \infty$ for $|\lambda| < r_q$ and $\int e^{y\lambda} \tilde{f}_t(y; \tau) dy < \infty$ for $|\lambda| < r_p$. Let κ_n be the nth cumulant of $R_t(\tau)$ conditional on information up to time t under the risk neutral measure, and $\tilde{\kappa}_k$ under the physical measure. Then for $l \geq 1$,

$$\tilde{\kappa}_l = \kappa_l + \sum_{n=1}^{\infty} \kappa_{n+l} \frac{\theta_1^n}{n!}. \tag{24}$$

The proof can be found in Rompolis and Tzavalis (2010). To a leading order, we have $\tilde{\kappa}_l \approx \kappa_l + \theta_1 \kappa_{l+1}$. Moreover, $\kappa_l = \tilde{\kappa}_l - \theta_1 \tilde{\kappa}_{l+1} + \sum_{n=2}^{\infty} (n-1) \kappa_{n+l} \frac{\theta_1^n}{n!} \approx \tilde{\kappa}_l - \theta_1 \tilde{\kappa}_{l+1}$, for $l \geq 1$. Bakshi, Kapadia, and Madan (2003) obtained similar results for the central moments of return.

Let $\widetilde{\text{Mean}}(t, \tau)$ $\widetilde{\text{var}}(t, \tau)$, $\widetilde{\text{Skew}}(t, \tau)$, and $\widetilde{\text{EKurt}}(t, \tau)$ be the time t conditional mean, variance, skewness, and excess kurtosis of $\ln(S_{t+\tau})$ under the physical measure. It follows from Proposition 5.1 that,

$$\widetilde{\text{Mean}}(t, \tau) \approx \text{Mean}(t, \tau) + \theta_1 \kappa_2 \tag{25}$$

$$\widetilde{\text{var}}(t, \tau) \approx \text{var}(t, \tau) + \theta_1 \kappa_3 \tag{26}$$

$$\widetilde{\text{Skew}}(t, \tau) \approx \text{Skew}(t, \tau) + \theta_1 \big(\kappa_2^{-3/2}\kappa_4 - 1.5\kappa_2^{-5/2}\kappa_3^2\big)$$
$$= \text{Skew}(t, \tau) + \theta_1(\text{EKurt}(t, \tau)$$
$$- 1.5\text{Skew}(t, \tau)^2)\text{Var}(t, \tau)^{1/2} \tag{27}$$

$$\widetilde{\text{EKurt}}(t, \tau) \approx \text{EKurt}(t, \tau) + \theta_1 \big(\kappa_2^{-2}\kappa_5 - 2\kappa_2^{-3}\kappa_3\kappa_4\big)$$
$$= \text{EKurt}(t, \tau) + \theta_1 \big(\kappa_2^{-2}\kappa_5 - 2\kappa_2^{1/2}$$
$$\times \text{EKurt}(t, \tau)\text{Skew}(t, \tau)\big), \tag{28}$$

where κ_5 needs the pricing of $E_t^Q(e^{-r\tau} R_t(\tau)^5)$, which is

$$E_t^Q(e^{-r\tau} R_t(\tau)^5)$$
$$= \int_{S_t}^{\infty} \frac{20(\ln(K/S_t))^3 - 5(\ln(K/S_t))^4}{K^2} C(t, \tau; K) dK$$
$$+ \int_0^{S_t} \frac{20(\ln(K/S_t))^3 - 5(\ln(K/S_t))^4}{K^2} P(t, \tau; K) dK$$

Equations (25)–(28) allow us to calculate the first four moments under the physical measure.

Because the transformation formulas (25)–(28) are derived under the assumption that $\tilde{\kappa}_l \approx \kappa_l + \theta_1\kappa_{l+1}$, we report in Table 6 the sample mean and standard deviation of the risk neutral cumulants $\kappa_2, \kappa_3, \kappa_4$, and κ_5 for the S&P 500 index over the full sample—from January 4, 1996, through October 30, 2009.[16] On average, the higher-order cumulants are of order 10^{-4} for time-to-maturity up to 1 month. Therefore, it is reasonable to believe that the transformation formulas (25)–(28) provide a good approximation if θ_1 is relatively small, say, less than 10.

Next we will discuss in detail the VG, NIG, and GST distributions under the physical measure, and the estimation of the relative risk aversion coefficient θ_1.

5.1 The GH Family of Distributions: The Risk Neutral Measure or the Physical Measure?

Suppose that the conditional distribution of $R_t(\tau)$ given information up to time t is approximated by $\text{GH}(\alpha_t, \beta_t, \mu_t, b_t, p_t)$ under the physical measure P. Consider the moment-generating function of $\text{GH}(\alpha, \beta, \mu, b, p)$. For $|\beta + z| < \alpha$,

$$M_{\text{GH}}(z; \alpha, \beta, \mu, b, p) = \int_{\mathbb{R}} e^{zx} f_{\text{GH}}(x) dx$$
$$= e^{\mu z} \frac{(\alpha^2 - \beta^2)^{p/2}}{(\alpha^2 - (\beta + z)^2)^{p/2}} \frac{K_p(b\sqrt{\alpha^2 - (\beta + z)^2})}{K_p(b\sqrt{\alpha^2 - \beta^2})}, \tag{29}$$

---

[15]See Aït-Sahalia and Lo (2000b), Jackwerth (2000), Rosenberg and Engle (2002), Liu et al. (2007), among others.

[16]The results are similar in magnitude to the estimates of κ_3 and κ_4 reported in Rompolis and Tzavalis (2010) who looked at S&P 500 index returns and options over 22 trading days from January 1996, to December 2007.

Table 6. Summary statistics for SPX from January 1996 to October 2009

	Skewness	Excess kurtosis	κ_2	κ_3	κ_4	κ_5
		5~14 days to maturity				
Mean	−1.5131	22.8377	0.0015	−0.0002	0.0001	−0.0001
SD	1.7441	50.4268	0.0020	0.0005	0.0005	0.0008
		17~31 days to maturity				
Mean	−1.3942	7.3299	0.0037	−0.0006	0.0003	−0.0003
SD	0.8565	13.2251	0.0050	0.0018	0.0014	0.0016
		81~94 days to maturity				
Mean	−1.2633	3.0027	0.0129	−0.0031	0.0016	−0.0011
SD	0.6652	8.9794	0.0144	0.0076	0.0065	0.0070
		171~199 days to maturity				
Mean	−1.1723	1.5461	0.0258	−0.0069	0.0027	−0.0009
SD	0.5268	3.6149	0.0239	0.0138	0.0103	0.0076

The table reports the sample mean (mean) and standard deviation (SD) of the risk neutral skewness, the risk neutral excess kurtosis, and the risk neutral cumulants κ_2, κ_3, κ_4, and κ_5 for S&P 500 index over the full sample (January 4, 1996 – October 30, 2009). The risk neutral moments and cumulants are extracted from option prices via (8), (9), and (10).

where $f_{GH}(x)$ is defined in (1). The pricing kernel (22) satisfies the following arbitrage free conditions

$$E_t^P[m_t(R_t(\tau), \theta_t)] = e^{-r\tau}, \quad E_t^P\left[m_t(R_t(\tau), \theta_t)e^{R_t(\tau)}\right] = 1$$

(30)

if and only if

$$M_{GH}(-\theta_1; \alpha_t, \beta_t, \mu_t, b_t, p_t) = \theta_0^{-1}e^{-r\tau},$$

$$M_{GH}(1 - \theta_1; \alpha_t, \beta_t, \mu_t, b_t, p_t) = \theta_0^{-1}$$

(31)

given that $|\beta_t - \theta_1| < \alpha_t$ and $|\beta_t + 1 - \theta_1| < \alpha_t$.[17] It follows that $\log M_{GH}(1; \alpha_t, \beta_t - \theta_1, \mu_t, b_t, p_t) = r\tau$. Therefore, $R_t(\tau)$ is $GH(\alpha_t, \beta_t - \theta_1, \mu_t, b_t, p_t)$ under the risk neutral measure. On the other hand, if the conditional distribution of $R_t(\tau)$ belongs to the GH family under the risk neutral measure, then after exponential tilting the real-world distribution of $R_t(\tau)$ is within the GH family as well again in the absence of arbitrage. In particular, under the arbitrage free assumption, $R_t(\tau)$ is NIG (or VG) distributed under the risk neutral measure if and only if it is NIG (or VG) distributed under the physical measure, but it does not apply to the GST distribution.

Next we will consider estimating θ_1 when the approximating densities are the VG and NIG. We use the maximum likelihood estimation outlined in Liu et al. (2007). Let $\{t_i^*, 1 \leq i \leq n\}$ be the expiration times of the options and let t_i be the time when risk neutral densities are formed for options that expire at time t_i^*. $\{t_i^*, t_i\}$ should satisfy $t_i < t_i^* \leq t_{i+1}$ for $1 \leq i \leq n - 1$ with $t_n^* = t_{n+1}$ so that the densities do not overlap. With $\tau = t_i^* - t_i$, θ_1 is estimated by maximizing the following log-likelihood function

$$\log L(R_{t_1}(\tau), \ldots, R_{t_n}(\tau)|\theta_1) = \sum_{i=1}^n \log \tilde{f}_{t_i}(R_{t_i}(\tau)|\theta_1, \hat{\Psi}_{t_i}),$$

(32)

where $\tilde{f}_{t_i}(R_{t_i}(\tau)|\theta_1, \hat{\Psi}_{t_i})$ is the real-world distribution of $R_{t_i}(\tau)$ conditional on information up to t_i, and $\hat{\Psi}_{t_i}$ represents the

vector of estimated parameters in the associated risk neutral density. To be specific, if $R_{t_i}(\tau)$ is NIG$(\alpha_{t_i}, \beta_{t_i}, \mu_{t_i}, b_{t_i})$ distributed under the risk neutral measure, then $\tilde{f}_{t_i}(\cdot)$ is the density function of NIG$(\alpha_{t_i}, \beta_{t_i} + \theta_1, \mu_{t_i}, b_{t_i})$ with $\Psi_{t_i} = (\alpha_{t_i}, \beta_{t_i}, \mu_{t_i}, b_{t_i})$. $\hat{\Psi}_{t_i}$ is obtained by matching the moments as described in Section 3.1. Similarly, if $R_{t_i}(\tau)$ is modeled by VG$(\alpha_{t_i}, \beta_{t_i}, \mu_{t_i}, p_{t_i})$ under the risk neutral measure, then it is VG$(\alpha_{t_i}, \beta_{t_i} + \theta_1, \mu_{t_i}, p_{t_i})$ distributed under the physical measure with $\Psi_{t_i} = (\alpha_{t_i}, \beta_{t_i}, \mu_{t_i}, p_{t_i})$. Table 7 reports the estimates of θ_1 using the S&P 500 index options from January 4, 1996, to October 30, 2009, with $\tau = 7, 14, 22$ days.

5.2 Value-at-Risk Forecasting

Last but certainly not least, we consider value-at-risk (VaR) forecasting. We use the notation introduced in Section 3.1. $-R_t(\tau)$ can be viewed as the loss from time t to $t + \tau$. The τ−period-ahead VaR forecast at level $100(1 - \alpha)\%$ is defined as $\text{VaR}_t(\tau, \alpha) = \inf_y\{y : P(-R_t(\tau) \geq y|\mathcal{F}_t) \leq \alpha\}$. We consider VaR evaluated under the risk neutral measure Q as well as the physical measure P, which will be denoted by $\text{VaR}_t^Q(\tau, \alpha)$ and $\text{VaR}_t^P(\tau, \alpha)$, respectively. Let $\widehat{\text{VaR}}_t^Q(\tau, \alpha)$ (or $\widehat{\text{VaR}}_t^P(\tau, \alpha)$) be the estimate of $\text{VaR}_t^Q(\tau, \alpha)$ (or $\text{VaR}_t^P(\tau, \alpha)$). The real-world estimates are obtained using θ_1 reported in Table 7. Therefore, only the VG and NIG distributions are considered in this exercise.

Table 7. The estimates of θ_1 using SPX options from January 1996 to October 2009

Time to maturity	VG	NIG
7 days	2.8953	3.0914
14 days	0.7484	0.2421
22 days	0.0140	0.0198

The table reports the estimates of the relative risk aversion coefficient θ_1 using S&P 500 index options from January 4, 1996, to October 30, 2009, with 7 days, 14 days, 22 days to maturity.

---

[17]See, for instance, Singleton (2006) and León, Mencía, and Sentana (2009).

(a) VG approximation

(b) NIG approximation

Figure 8. The 95% VaR forecast for SPX over 7 days. The figure plots the losses $\{-R_t(7)\}$ (the curves) of S&P 500 index from January 1996 to October 2009, and the 95% VaR forecasts under the risk neutral measure (i.e., stars for $\widehat{\text{VaR}}_t^Q(7, 0.05)$) and the physical measure (i.e., dots for $\widehat{\text{VaR}}_t^P(7, 0.05)$). The forecasts are derived using the VG and the NIG densities. The real world forecasts are obtained using θ_1 reported in Table 7. The vertical line divides the series into two subperiods: (1) prior to August 2008 and (2) the financial crisis period.
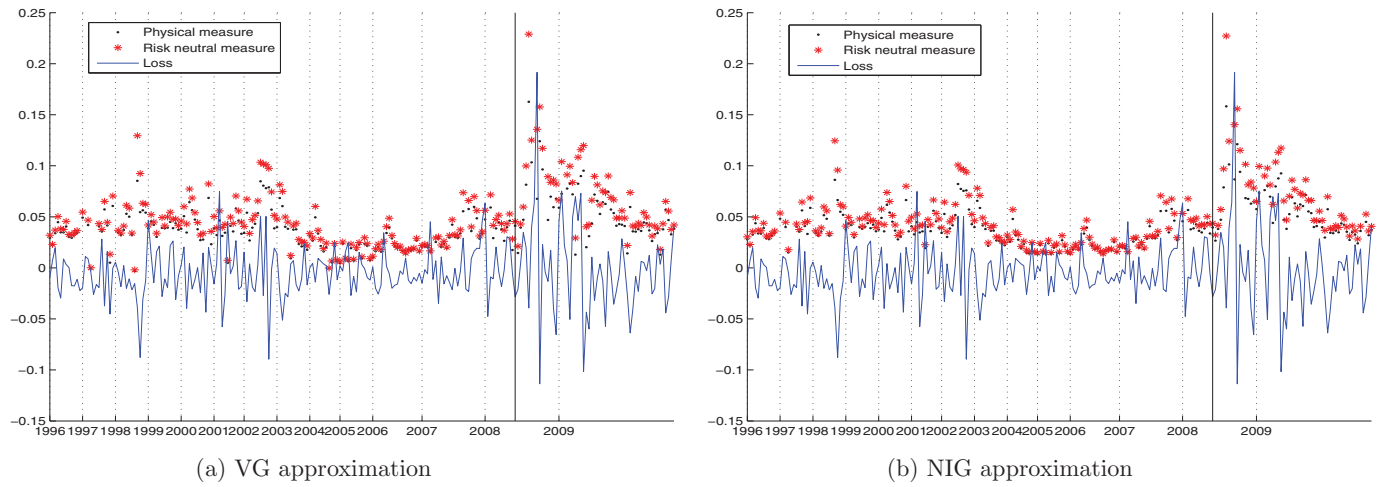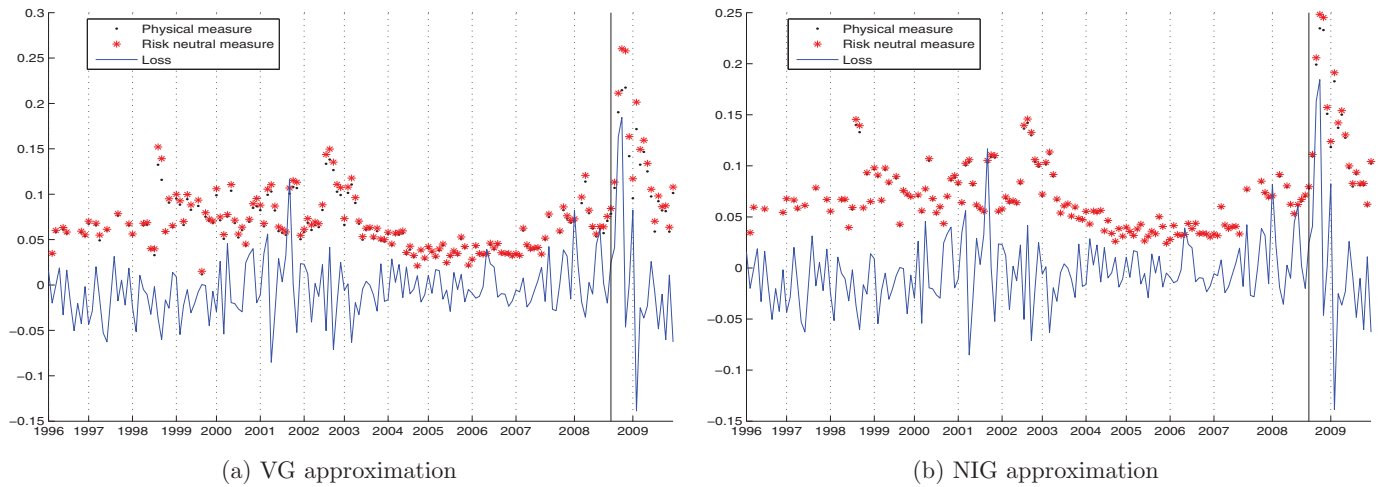

(a) VG approximation

(b) NIG approximation

Figure 9. The 95% VaR forecast for SPX over 14 days. The figure plots the losses $\{-R_t(14)\}$ (the curves) of S&P 500 index from January 1996 to October 2009, and the 95% VaR forecasts under the risk neutral measure (i.e., stars for $\widehat{\text{VaR}}_t^Q(14, 0.05)$) and the physical measure (i.e., dots for $\widehat{\text{VaR}}_t^P(14, 0.05)$). The forecasts are derived using the VG and the NIG densities. The real world forecasts are obtained using θ_1 reported in Table 7. The vertical line divides the series into two subperiods: (1) prior to August 2008 and (2) the financial crisis period.

We consider again the SPX options from January 4, 1996 to October 30, 2009, with 7 days and 14 days to maturity. Hence, we do as if we hold the market portfolio or something similar to that, and compute its VaR forecast. Figures 8 and 9 plot the losses $\{-R_t(\tau)\}$ over the life time of options, that is, $\tau = 7$ and 14. The vertical line divides the series into two subperiods: (1) Prior to August 2008 and (2) the Financial crisis period.[18] Superimposed are the out-of-sample VaR forecasts at the 95% level under the risk neutral measure (stars) and the physical measure (dots). The estimates are derived using the VG and the NIG densities. There are 229 observations for $\tau = 7$. Among them, 216 observations

can be modeled by the VG density while 213 by the NIG density. For $\tau = 14$, 155 out of 172 can be approximated by the VG and 149 out of 172 by the NIG density.

To better understand the difference between the VG and NIG densities and their goodness of fit, we consider the VaR failure indicator which is defined as $I_t(\tau, \alpha) = I_{\{-R_t(\tau) > \widehat{\text{VaR}}_t^P(\tau, \alpha)\}}$, that is, $I_t(\tau, \alpha) = 1$ if $R_t(\tau) < -\widehat{\text{VaR}}_t^P(\tau, \alpha)$ and 0 otherwise. Figure 10 plots the 95% VaR forecasts under the physical measure (i.e., $\widehat{\text{VaR}}_t^P(\tau, 0.05)$) using both VG and NIG densities against the losses $\{-R_t(\tau)\}$ over the lift time of the options for $\tau = 7$ and 14. Denote by T the number of VaR forecasts. Then $\sum_{t=1}^{T} I_t(\tau, \alpha)$ is the total number of exceedances of the $100(1 - \alpha)$% VaR forecasts, and $T^{-1} \sum_{t=1}^{T} I_t(\tau, \alpha)$ is referred

---

[18]We also remove days which have only two pairs of contracts.

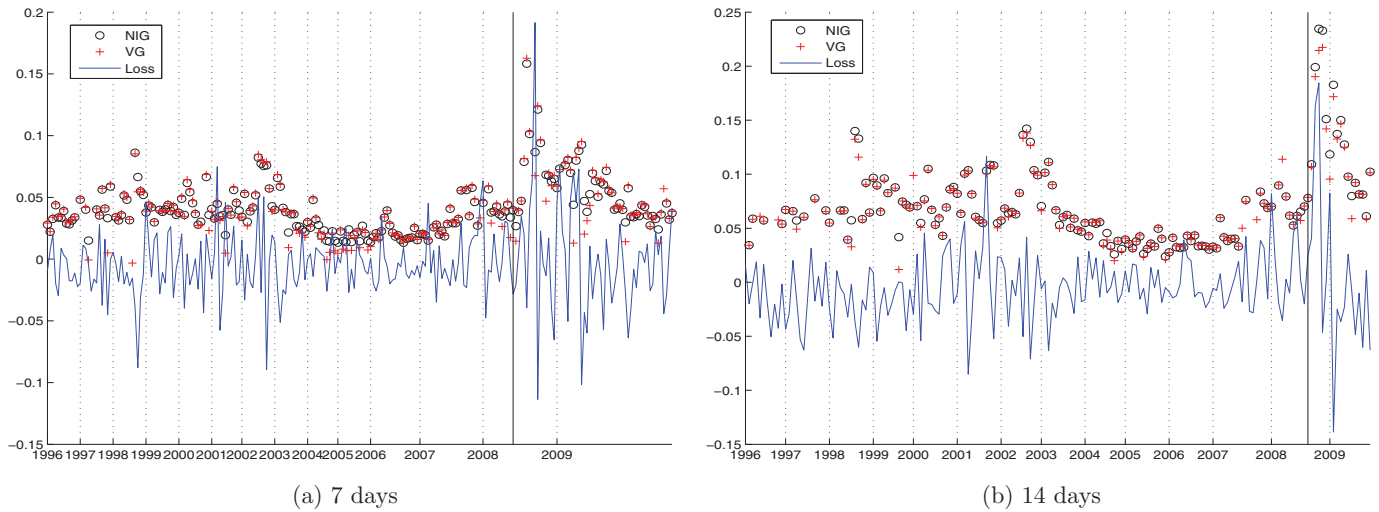(a) 7 days   (b) 14 days

Figure 10. The 95% VaR forecasts under the physical measure for SPX (Jan. 1996–Oct. 2009). The figure shows the 95% VaR forecasts under the physical measure (i.e., $\widehat{\mathrm{VaR}}_t^P(\tau, 0.05)$) using both VG and NIG densities against the losses $\{-R_t(\tau)\}$ (the curves) over the lift time of the S&P 500 index call options for $\tau = 7$ and 14. The real world forecasts are obtained using θ_1 reported in Table 7. The vertical line divides the series into two subperiods: (1) prior to August 2008 and (2) the financial crisis period.

to as the empirical risk level. We report in Table 8 the empirical risk level (see the columns with label "Emp. Risk") for the full sample and the two subsamples: prior to the financial crisis and during the crisis. For the VG approximation, we report the empirical risk levels across all the observations which are within the VG feasible domain—see the rows labeled "VG(216)" and "VG(155)," as well as the empirical risk levels for observations which are common for both VG and NIG distributions—see the rows labeled "VG(213)" and "VG(149)." The latter allows us to compare the two densities. For nominal risk level 0.025 and $\tau = 7$, the two densities yield the same empirical risk level over the full sample and they over-estimate the risk. For the subsample "Prior to August 2008," the VG provides a more accurate estimate of the risk. For the "Financial Crisis Period," the NIG

underestimates the risk, but it provides a more accurate estimate than the VG approximation. When the nominal risk level increases to 0.05, both over-estimate the risk in the three samples, and overall the NIG provides a more accurate estimate. We also consider the nominal risk level 0.05 and $\tau = 14$. The two densities yield the same results and they underestimate the risk.

To backtest the accuracy of VaR forecasts, we consider the unconditional coverage test of Christoffersen (1998), which tests for the null hypothesis of $E^P(I_t(\tau, \alpha)) = \alpha$ against a two-sided alternative. The column labeled "Uncond" in Table 8 reports the p-values of unconditional coverage test. The test is not applicable when $\alpha = 0.05$ and $\tau = 14$ during the financial crisis period. This is because the exceedance of VaR forecast is 0. The p-values, when applicable, indicate that there is no significant

Table 8. VaR Backtesting

	Full sample		Prior to August 2008		Financial crisis period	
	Emp. Risk	Uncond	Emp. risk	Uncond	Emp. risk	Uncond
$\alpha = 0.025, \tau = 7$						
VG(216)	0.0324	0.5046	0.0312	0.6258	0.0357	0.6291
VG(213)	0.0282	0.7715	0.0253	0.9797	0.0364	0.6128
NIG(213)	0.0282	0.7715	0.0316	0.6071	0.0182	0.7337
$\alpha = 0.05, \tau = 7$						
VG(216)	0.0694	0.2144	0.0688	0.3019	0.0714	0.4881
VG(213)	0.0657	0.3140	0.0633	0.4608	0.0727	0.4676
NIG(213)	0.0610	0.4745	0.0570	0.6942	0.0727	0.4676
$\alpha = 0.05, \tau = 14$						
VG(155)	0.0194	0.0467	0.0219	0.0909	0	–
VG(149)	0.0201	0.0584	0.0229	0.1128	0	–
NIG(149)	0.0201	0.0584	0.0229	0.1128	0	–

The table reports the empirical risk levels (see the columns with label "Emp. risk"), and the p-values of the unconditional coverage test of Christoffersen (1998) (see the columns with label "Uncond"). The empirical risk level is defined as the total exceedances divided by the number of forecasts. The unconditional coverage test is not applicable when $\alpha = 0.05$ and $\tau = 14$ during the financial crisis period, and hence the p-values are not reported. For the VG approximation, we report the empirical risk levels and p-values across all the observations which are within the VG feasible domain—see the rows labeled "VG(216)" and "VG(155)," as well as for observations which are common for both VG and NIG distributions—see the rows labeled "VG(213)" and "VG(149)."

Table 9. Pricing S&P 500 index call options (1996–2009)—GST, VG, and NIG

	Strategy 1								Strategy 2							
	Overall		[a, b]		< a		> b		Overall		[a, b]		< a		> b	
	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$
$D(2, T_0), T_0 = 17\text{–}31$																
BLS	4.96	1.92	5.06	3.19	0.38	2.34	6.46	0.23	4.26	2.50	4.09	3.25	0.50	3.60	5.57	0.24
GARCH	70.28	106.06	71.48	77.80	27.71	227.45	83.66	3.42	66.50	106.11	66.00	103.62	27.03	217.96	78.79	3.79
NIG	7.05	0.64	3.52	0.71	0.19	0.92	9.45	0.40	5.06	0.63	2.67	0.70	0.18	0.94	7.28	0.35
VG	7.59	0.61	3.37	0.60	0.17	0.84	10.28	0.44	5.49	0.58	2.54	0.56	0.17	0.86	8.05	0.40
GST	6.27	0.70	4.10	0.94	0.20	1.00	8.18	0.35	4.61	0.73	3.11	0.92	0.19	1.00	6.34	0.32
$D(2, T_0), T_0 = 81\text{–}94$																
BLS	5.14	1.68	5.79	0.47	3.04	4.42	5.64	0.10	5.84	1.86	6.58	0.58	3.57	5.11	6.19	0.11
GARCH	146.49	92.06	156.76	15.20	111.12	234.51	158.86	3.24	151.12	93.61	160.81	17.33	111.23	241.89	163.84	3.26
NIG	9.62	0.55	9.45	0.58	0.78	0.89	11.99	0.25	9.29	0.58	9.09	0.59	0.83	0.90	11.88	0.24
VG	10.01	0.51	10.04	0.58	0.64	0.74	12.41	0.26	9.63	0.54	9.59	0.59	0.70	0.78	12.21	0.25
GST	9.25	0.58	9.30	0.60	0.89	0.98	11.39	0.23	9.13	0.60	9.02	0.62	0.93	0.97	11.60	0.23
$D(3, T_0), T_0 = 5\text{–}14$																
BLS	4.27	6.03	5.93	5.63	1.10	5.57	8.44	0.40	4.07	4.97	5.03	5.19	0.73	4.02	7.37	0.43
GARCH	18.54	42.34	22.58	57.40	3.26	40.49	32.48	1.98	20.93	52.91	23.50	66.99	3.67	55.97	34.75	2.22
NIG	3.43	0.81	2.29	0.67	0.19	0.94	7.92	0.43	3.17	0.78	1.83	0.66	0.19	0.94	7.03	0.39
VG	3.73	0.77	2.23	0.60	0.18	0.89	8.61	0.46	3.42	0.75	1.77	0.57	0.18	0.89	7.63	0.42
GST	3.13	0.91	2.59	0.92	0.20	1.00	7.05	0.40	2.87	0.89	2.10	0.90	0.20	1.00	6.21	0.35
$D(3, T_0), T_0 = 17\text{–}31$																
BLS	6.92	7.97	9.77	2.51	2.81	11.77	10.34	0.32	6.04	6.45	8.05	2.92	2.03	9.28	8.99	0.30
GARCH	41.64	60.18	54.31	16.09	17.24	86.49	60.87	1.95	43.62	70.38	54.73	27.31	16.89	105.13	62.85	2.12
NIG	4.70	0.79	4.88	0.68	0.29	0.96	8.72	0.28	4.23	0.77	4.05	0.65	0.25	0.97	7.63	0.26
VG	4.94	0.74	4.97	0.63	0.26	0.89	9.26	0.30	4.46	0.72	4.16	0.58	0.23	0.91	8.10	0.28
GST	4.50	0.83	5.03	0.75	0.31	1.00	8.08	0.26	4.03	0.81	4.16	0.74	0.27	1.00	7.08	0.24

Table 10. Pricing S&P 500 index call options (1996–2009)—VG and NIG

| | Strategy 1 | | | | | | | | Strategy 2 | | | | | | | |
| | Overall | | $[a,b]$ | | $<a$ | | $>b$ | | Overall | | $[a,b]$ | | $<a$ | | $>b$ | |
	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$	$\bar{L}_a$	$\bar{L}_r$
$D(1,T_0), T_0 = 5\text{--}14$																
BLS	2.38	3.43	0.71	5.36	0.19	1.00	2.92	1.67	2.05	2.94	0.56	4.31	0.16	1.00	2.63	1.46
GARCH	43.94	199.44	34.94	341.49	2.41	20.88	50.02	37.49	41.92	219.08	32.67	349.91	2.26	24.91	48.44	34.24
NIG	2.49	0.70	0.15	0.77	0.19	0.99	3.18	0.62	2.27	0.69	0.15	0.78	0.16	0.99	2.98	0.59
VG	8.39	0.75	0.13	0.67	0.18	0.98	10.70	0.75	6.41	0.73	0.13	0.75	0.16	0.98	8.38	0.68
$D(1,T_0), T_0 = 17\text{--}31$																
BLS	4.09	4.09	2.74	6.40	0.16	0.95	4.78	0.78	3.53	3.41	2.30	5.22	0.17	0.97	4.25	0.67
GARCH	87.84	164.36	81.49	233.08	19.91	269.34	94.06	13.40	86.04	223.84	78.86	331.64	22.10	230.47	93.44	13.42
NIG	4.29	0.75	1.01	0.93	0.16	0.99	5.36	0.61	3.38	0.70	0.74	0.86	0.17	0.99	4.45	0.54
VG	5.54	0.70	0.88	0.77	0.16	0.98	6.97	0.63	4.15	0.62	0.60	0.66	0.17	0.98	5.47	0.55
$D(2,T_0), T_0 = 5\text{--}14$																
BLS	2.15	2.33	1.33	3.42	0.15	1.21	3.28	0.43	2.16	3.13	1.24	4.39	0.15	1.17	3.26	0.46
GARCH	34.22	101.08	30.39	151.94	3.46	74.06	44.75	6.02	34.07	122.32	29.10	183.35	3.47	69.83	44.26	6.09
NIG	4.01	0.79	0.64	0.85	0.14	0.98	6.16	0.60	3.30	0.70	0.48	0.72	0.14	0.96	5.08	0.52
VG	5.25	0.74	0.55	0.69	0.14	0.96	7.94	0.64	3.82	0.66	0.39	0.59	0.14	0.94	5.86	0.54
$D(2,T_0), T_0 = 17\text{--}31$																
BLS	4.65	4.01	4.87	5.26	0.53	4.95	6.24	0.29	5.12	5.88	5.03	7.87	0.49	4.54	6.68	0.34
GARCH	65.79	120.82	68.99	106.46	25.12	233.32	80.05	4.05	67.92	159.95	68.63	168.39	26.13	268.21	81.95	4.29
NIG	5.86	0.76	3.55	0.90	0.18	0.98	8.47	0.44	4.59	0.72	2.61	0.83	0.18	0.98	6.83	0.36
VG	6.27	0.72	3.28	0.78	0.18	0.94	9.23	0.48	4.96	0.65	2.41	0.67	0.18	0.94	7.54	0.41
$D(2,T_0), T_0 = 81\text{--}94$																
BLS	6.71	3.69	8.05	0.80	4.36	8.29	7.77	0.15	6.44	3.18	7.26	0.73	4.20	8.33	7.16	0.13
GARCH	140.85	134.46	154.56	17.07	109.35	306.92	155.50	3.25	140.84	120.73	150.84	19.16	105.04	323.08	153.86	3.34
NIG	8.10	0.60	8.77	0.60	0.71	0.95	10.19	0.21	7.85	0.59	8.17	0.63	0.65	0.95	9.67	0.21
VG	8.38	0.56	9.20	0.59	0.62	0.85	10.50	0.22	8.10	0.56	8.53	0.62	0.56	0.85	9.93	0.22
$D(3,T_0), T_0 = 5\text{--}14$																
BLS	3.86	4.73	4.74	5.19	0.69	4.13	7.57	0.44	3.85	4.75	4.73	5.11	0.68	4.04	7.31	0.43
GARCH	20.05	47.67	23.92	65.87	3.51	52.71	34.58	2.20	20.63	53.37	23.01	67.00	3.56	55.82	34.73	2.27
NIG	3.68	0.85	2.09	0.79	0.19	0.97	7.95	0.46	3.06	0.80	1.69	0.71	0.19	0.95	6.79	0.39
VG	3.95	0.81	1.98	0.69	0.18	0.94	8.57	0.49	3.30	0.76	1.61	0.61	0.18	0.91	7.38	0.43
$D(3,T_0), T_0 = 17\text{--}31$																
BLS	7.00	8.96	9.12	4.40	2.29	11.67	10.82	0.36	7.00	8.54	9.04	5.74	2.01	11.27	10.22	0.36
GARCH	46.03	78.94	58.56	35.02	19.30	116.75	66.04	2.17	48.04	91.28	57.62	46.76	20.31	141.16	67.32	2.25
NIG	4.92	0.81	4.82	0.76	0.26	0.98	8.74	0.30	4.18	0.80	3.91	0.73	0.25	0.98	7.50	0.26
VG	5.18	0.77	4.83	0.69	0.24	0.93	9.32	0.32	4.38	0.75	3.94	0.65	0.23	0.93	7.95	0.28

departure from adequacy except for the VG at $\alpha = 0.05$ and $\tau = 14$ which yields a marginally significant p-value—considering the 5% significance level.

6. CONCLUDING REMARKS

We revisit in this article the generalized hyperbolic family of distributions, which covers a wide range of distributions commonly used in the literature to model the financial risk. We specifically study the VG, NIG, and GST distributions, whose densities have closed-form expression in terms of the first four moments. We study the properties of the NIG, VG, and GST distributions in terms of tail behavior and feasible domain. Among them, the VG distribution admits the largest possible combinations of skewness and (excess) kurtosis while the GST is the most restrictive.

In the context of risk management we analyze option pricing—assuming an unknown risk neutral density which is approximated by the class of distributions we study as well as Gram-Charlier and Edgeworth expansions. We show, through numerical and empirical evidence, that the VG and NIG distributions are roughly similar as candidate approximating risk neutral densities for option pricing applications. The GST—with a more restrictive feasible domain—also performs well when applied.

We also find that the NIG and VG approximations work extremely well in terms of option pricing in particular during high volatility periods, compared to the industry standard of using previous period ATM implied Black-Scholes volatilities. Moreover, we find that with only a small set of quoted contracts, we can extrapolate and price very well options not used in the computation of risk neutral moments and hence options not used to compute the approximations. When compared to GARCH option pricing model, the methods proposed in this article are remarkably accurate and clearly dominating. Unfortunately, we also find from the tail density test that the distributional approximations do not do so well as far as VaR goes.

APPENDICES

A. PROOFS

Proof of Proposition 2.2. Note that

$$m_n = \frac{2^{\lceil \frac{n}{2} \rceil} \bar{\gamma}^p b^{2\lceil \frac{n}{2} \rceil} \beta^{\tilde{n}}}{\sqrt{\pi} K_p(\bar{\gamma}) \bar{\alpha}^{p+\lceil \frac{n}{2} \rceil}} \sum_{k=0}^{\infty} \frac{2^k \bar{\beta}^{2k} \Gamma(k + \lceil \frac{n}{2} \rceil + \frac{1}{2})}{\bar{\alpha}^k (2k + \tilde{n})!} K_{p+k+\lceil \frac{n}{2} \rceil}(\bar{\alpha}),$$

(A.1)

where $\tilde{n} \equiv n(\bmod 2)$ $\bar{\alpha} = b\alpha$, $\bar{\beta} = b\beta$, $\bar{\gamma} = \sqrt{\bar{\alpha}^2 - \bar{\beta}^2}$ (see Barndorff-Nielsen and Stelzer 2005). Expression (2) is derived from (A.1) together with the fact that

$$K_v(z) = \frac{z^v}{x^v} \sum_{k=0}^{\infty} \frac{1}{2^k k!} \frac{y^{2k}}{x^k} K_{v+k}(x),$$

$$z = \sqrt{x^2 - y^2}, x > 0, y > 0, v \in \mathbb{R}.$$

$$\sum_{k=0}^{\infty} \frac{\bar{\beta}^{2k}}{\bar{\alpha}^k} \frac{(2k + 3)}{2^k k!} K_{p+k+2}(\bar{\alpha}) = K_{p+3}(\bar{\gamma}) \frac{\bar{\alpha}^{p+2} \bar{\beta}^2}{\bar{\gamma}^{p+3}}$$

$$+ K_{p+2}(\bar{\gamma}) \frac{3\bar{\alpha}^{p+2}}{\bar{\gamma}^{p+2}},$$

$$\sum_{k=0}^{\infty} \frac{\bar{\beta}^{2k}}{\bar{\alpha}^k} \frac{(2k + 3)(2k + 1)}{2^k k!} K_{p+k+2}(\bar{\alpha})$$

$$= K_{p+4}(\bar{\gamma}) \frac{\bar{\beta}^4 \bar{\alpha}^{p+2}}{\bar{\gamma}^{p+4}} + K_{p+3}(\bar{\gamma}) \frac{6\bar{\beta}^2 \bar{\alpha}^{p+2}}{\bar{\gamma}^{p+3}} + K_{p+2}(\bar{\gamma}) \frac{3\bar{\alpha}^{p+2}}{\bar{\gamma}^{p+2}}.$$

$\square$

Proof of Proposition 2.4. (1) Use the fact that $\lim_{b \to 0} \frac{b^k K_{p+k}(\bar{\gamma})}{K_p(\bar{\gamma})} = \frac{2^k \Gamma(p+k)}{\gamma^k \Gamma(p)}$ for $k \in \mathbb{Z}^+$ and $p > 0$, and the dominant convergence theorem.

(2) Note that (4) implies

$$S = \frac{2\rho(3 + R)}{\sqrt{V}\alpha(1 - R^2)}, \quad K = \frac{6(1 + 6R + R^2)}{V\alpha^2(1 - R)^2(1 + R)},$$

where $\rho = \frac{\beta}{\alpha}$ ($|\rho| < 1$) and $R = \rho^2 < 1$. Therefore,

$$\frac{3S^2}{2K} = \frac{R(3 + R)^2}{(1 + 6R + R^2)(1 + R)}.$$

(A.2)

Define $C = \frac{3S^2}{2K}$ and $C \in (0, 1)$. Note that $f(R) = \frac{R(3+R)^2}{(1+6R+R^2)(1+R)}$ is continuous and strictly increasing on $(0, 1)$ with range $(0, 1)$. Given S, K and $3S^2 < 2K$, there exists a unique $R \in (0, 1)$ satisfying (A.2), or $(C - 1)R^3 + (7C - 6)R^2 + (7C - 9)R + C = 0$. It follows immediately that

$$\alpha = \frac{2\rho(3 + R)}{\sqrt{V}S(1 - R^2)} = \frac{2\sqrt{R}(3 + R)}{\sqrt{V}|S|(1 - R^2)},$$

$$\beta = \rho\alpha = \frac{2R(3 + R)}{\sqrt{V}S(1 - R^2)}$$

$$p = \frac{V(1 - R)^2\alpha^2}{2(1 + R)} = \frac{2R(3 + R)^2}{S^2(1 + R)^3},$$

$$\mu = M - \frac{2\beta p}{\alpha^2 - \beta^2} = M - \frac{2\sqrt{V}R(3 + R)}{S(1 + R)^2}.$$

$\square$

Proof of Proposition 2.5. (1) It follows from the fact that $\lim_{\gamma \to 0} \frac{K_{p+k}(\bar{\gamma})}{\gamma^k K_p(\bar{\gamma})} = \frac{b^k \Gamma(-p-k)}{2^k \Gamma(-p)}$ for $k \in \mathbb{Z}^+$ and $p < 0$, and an application of the dominated convergence theorem.

(2) Define $\rho = \frac{2b^2\beta^2}{(v-2)(v-4)}$. $0 < \rho < b^2\beta^2/12$. Expression (5) implies that

$$V = \frac{(1+\rho)b^2}{v-2}$$

$$S^2 = \left(3 + \frac{4(v-4)\rho}{v-6}\right)^2 \frac{2\rho}{(v-4)(1+\rho)^3}$$

$$K = \left[\frac{2(5v-22)(v-4)\rho^2}{(v-6)(v-8)} + \frac{8(v-4)\rho}{v-6} + 1\right]$$
$$\times \frac{6}{(1+\rho)^2(v-4)}$$

and thus $b^2 = V(v-2)/(1+\rho)$, and $\beta^2 = \rho(1+\rho)(v-4)/(2V)$, where $\rho > 0$ and $v > 8$ are solutions of the following system of equations

$$S^2 = \left(3 + \frac{4(v-4)\rho}{v-6}\right)^2 \frac{2\rho}{(v-4)(1+\rho)^3} \quad \text{(A.3)}$$

$$K = \left[\frac{2(5v-22)(v-4)\rho^2}{(v-6)(v-8)} + \frac{8(v-4)\rho}{v-6} + 1\right]$$
$$\times \frac{6}{(1+\rho)^2(v-4)}$$

or

$$0 = 2\rho[3(v-6) + 4(v-4)\rho]^2 - S^2(v-4)(v-6)^2(1+\rho)^3$$

$$0 = 12(v-4)(5v-22)\rho^2 + 48(v-4)(v-8)\rho$$
$$+ 6(v-6)(v-8) - K(v-4)(v-6)(v-8)(1+\rho)^2.$$

Note that (A.3) may not have solutions with arbitrary combination of K and S^2. We next need to justify that the necessary and sufficient (N&S) condition under which (A.3) has one and only one solution is (6). Let $x = \rho$ and $y = v - 8$. Both are positive. Fix y, S, K, and define

$$f_1(x; y, K) \equiv x^2 \left[\frac{12(5y+18)}{(y+2)y} - K\right] + 2x \left[\frac{24}{y+2} - K\right]$$
$$+ \frac{6}{y+4} - K \equiv Ax^2 + 2Bx + C \quad \text{(A.4)}$$

$$f_2(x; y, S) \equiv x^3 \left[\frac{32(y+4)}{(y+2)^2} - S^2\right] + 3x^2 \left[\frac{16}{y+2} - S^2\right]$$
$$+ 3x \left[\frac{6}{y+4} - S^2\right] - S^2$$
$$\equiv Dx^3 + 3Ex^2 + 3Fx - S^2. \quad \text{(A.5)}$$

Since

$$\frac{12(5y+18)}{(y+2)y} > \frac{32(y+4)}{(y+2)^2} > \frac{24}{y+2} > \frac{16}{y+2} > \frac{6}{y+4} > 0,$$

the N&S condition that (A.4) has one and only one positive root is $\frac{12(5y+18)}{(y+2)y} > K > \frac{6}{y+4}$. The positive root is $x_1(y; K) = \frac{\sqrt{B^2-AC}-B}{A}$ and $y_L < y < y_U$ where $y_L = \max(0, 6/K - 4)$ and $y_U = \sqrt{1 + 156/K + 900/K^2} - 1 + 30/K$. Define

$$g(y; S, K) = f_2(x_1(y; K); y, S), \quad \text{for } y_L < y < y_U. \quad \text{(A.6)}$$

Lemma A.3 implies that g has root(s) in (y_L, y_U) if and only if

$$\lim_{y \to y_U^-} D > 0 \quad \text{or} \quad \frac{32(y_U+4)}{(y_U+2)^2} > S^2. \quad \text{(A.7)}$$

The root is unique as well. Therefore, the N&S condition that one can estimate the GST parameters via the first four moments is (6). □

To complete the proof of Proposition 2.5, we need to justify the following lemmas:

Lemma A.1. $x_1(y; K)$ is increasing in y, and $\lim_{y \to y_L^+} x_1(y; K) = 0$, $\lim_{y \to y_U^-} x_1(y; K) = +\infty$.

Proof. (1) Note that $x_1' = -\frac{A'x_1^2 + 2B'x_1 + C'}{2\sqrt{B^2-AC}} > 0$. x_1 is increasing in y.
(2) If $K < 3/2$, then $y_L = 6/K - 4 > 0$. Note that $\lim_{y \to y_L^+} A > 0$, $\lim_{y \to y_L^+} B > 0$, $\lim_{y \to y_L^+} C \nearrow 0$. We have $\lim_{y \to y_L^+} x_1 = 0$. If $K \geq 3/2$, then $y_L = 0$. Similarly we have $\lim_{y \to y_L^+} x_1 \searrow 0$.
(3) Note that $\lim_{y \to y_U^-} A \searrow 0$, $\lim_{y \to y_U^-} B < 0$, and $\lim_{y \to y_U^-} C < 0$. $\lim_{y \to y_U^-} x_1 = \lim_{y \to y_U^-} \frac{-C}{\sqrt{B^2-AC}+B} = +\infty$. □

Lemma A.2. For $g(y; S, K)$ defined in (A.6), if there exists $y_0 > 0$ such that $g(y_0) = 0$, then $g'(y_0) > 0$.

Proof. Define $a = A + K$, $b = B + K$, $c = C + K(= F + S^2)$, $d = D + S^2$, and $e = E + S^2$. Then

$$K = \frac{ax_1^2 + 2bx_1 + c}{(x_1+1)^2}, \quad S^2 = \frac{dx_1^3 + 3ex_1^2 + 3cx_1}{(x_1+1)^3}\bigg|_{y=y_0}.$$
□

Therefore,

$$\frac{dx_1}{dy} = -\frac{A'x_1^2 + 2B'x_1 + C'}{2(Ax_1+B)} = -\frac{(a'x_1^2 + 2b'x_1 + c')(1+x_1)}{2((a-b)x_1 + (b-c))},$$

$$\frac{dg}{dy}\bigg|_{y=y_0} = D'x_1^3 + 3E'x_1^2 + 3F'x_1$$
$$+ 3\left(Dx_1^2 + 2Ex_1 + F\right)\frac{dx_1}{dy}\bigg|_{y=y_0}$$
$$= (d'x_1^2 + 3e'x_1 + 3c')x_1$$
$$- 3\left((d-e)x_1^2 + 2(e-c)x_1 + c\right)$$
$$\times \frac{(a'x_1^2 + 2b'x_1 + c')x_1}{2((a-b)x_1^2 + (b-c)x_1)}\bigg|_{y=y_0}$$
$$= -(a'x_1^2 + 2b'x_1 + c')x_1$$
$$\times \left(\frac{3}{2}\frac{(d-e)x_1^2 + 2(e-c)x_1 + c}{(a-b)x_1^2 + (b-c)x_1}\right.$$
$$\left.- \frac{d'x_1^2 + 3e'x_1 + 3c'}{a'x_1^2 + 2b'x_1 + c'}\right)\bigg|_{y=y_0}.$$

With some algebra, one can show that

$$\frac{3}{2}\frac{(d-e)x_1^2 + 2(e-c)x_1 + c}{(a-b)x_1^2 + (b-c)x_1} - \frac{d'x_1^2 + 3e'x_1 + 3c'}{a'x_1^2 + 2b'x_1 + c'} > 0$$

for all $y > 0$, and hence $g'(y_0) > 0$.

Lemma A.3. Consider $g(y; S, K)$ defined in (A.6). If $\lim_{y \to y_U^-} D > 0$, g has one and only one root in (y_L, y_U). If $\lim_{y \to y_U^-} D \le 0$, $g(y) < 0$ for all $y \in (y_L, y_U)$.

Proof. If $\lim_{y \to y_U^-} D > 0$, then $D(y) > 0$ for all $y \in (y_L, y_U)$. It follows from Lemma A.1 that $\lim_{y \to y_L^+} g = -S^2 < 0$ and $\lim_{y \to y_U^-} g = +\infty$. Therefore $g(y)$ has one and only one root in (y_L, y_U) due to intermediate value theorem and Lemma A.2. □

Next we will show that $g(y) < 0$ for all $y \in (y_L, y_U)$ if $\lim_{y \to y_U^-} D \le 0$. Define $y_D = \inf\{y \in (y_L, y_U) : D(y) \le 0\}$. If $y_D = y_L$, then $F < E < D \le 0$ on (y_L, y_U) and hence $g(y) < 0$ for $y \in (y_L, y_U)$. Suppose $y_D \in (y_L, y_U)$. We have $D(y) > 0$ for $y \in (y_L, y_D)$. Note that $\lim_{y \to y_D} D = 0$, $\lim_{y \to y_D} E < 0$, $\lim_{y \to y_D} F < 0$, and $\lim_{y \to y_D} x_1$ is finite and positive. Thus $\lim_{y \to y_D} g < 0$. Note also that $\lim_{y \to y_L^+} g = -S^2 < 0$. It follows from Lemma A.2 that $g(y) < 0$ for $y \in (y_L, y_D]$ and hence $g(y) < 0$ for $y \in (y_L, y_U)$.

B.　FAST FOURIER TRANSFORM FOR DENSITY FUNCTION AND CALL PRICE

We use fast Fourier transform of Carr and Madan (1999) to numerically evaluate conditional density function and call option prices derived from Model (15). The conditional density $f_t(y; T, x_t) = \frac{1}{\pi} \int_0^\infty e^{-iuy} \Psi_t(iu; T, x_t) du$ which is inverse Fourier transform of characteristic function, is approximated by

$$f_t(y_l; T, x_t) \approx \sum_{j=1}^N e^{-(l-1)(j-1)\frac{2\pi}{N}i} \tilde{x}(j),$$

$$\tilde{x}(j) = \frac{\eta}{\pi} \delta_j e^{ibu_j} \Psi_t(iu_j; T, x_t), \qquad (B.8)$$

where $y_l = -b + \lambda(l-1)$ with $b = N\lambda/2$ and $\lambda = 2\pi/(\eta N)$ for $l = 1, 2, \ldots, N$. $u_j = \eta(j-1)$, and $\delta_j = 1/2$ when $j = 1$ and 1 otherwise. For call price, consider the fact that

$$C_t(K; T, x_t) \equiv E((e^{Y_T} - K)^+ | \mathcal{F}_t)$$

$$= \frac{e^{-\xi \log(K)}}{\pi} \int_0^\infty e^{-iu \log(K)} \phi(u) du,$$

$$\phi(u) = \frac{e^{-r(T-t)} \Psi_t(iu + (1+\xi); T, x_t)}{\xi^2 + \xi - u^2 + iu(2\xi + 1)},$$

where ξ is a dampening coefficient. Therefore, $C_t(K; T, x_t)$ is approximated by

$$C_t(K_l; T, x_t) \approx e^{-\xi \log(K_l)} \sum_{j=1}^N e^{-(l-1)(j-1)\frac{2\pi}{N}i} \tilde{x}(j),$$

$$\tilde{x}(j) = \delta_j \frac{\eta}{\pi} e^{iu_j b} \phi(u_j), \qquad (B.9)$$

where $\log(K_l) = -b + \lambda(l-1)$ with $b = N\lambda/2$ and $\lambda = 2\pi/(\eta N)$ for $l = 1, 2, \ldots, N$.

Considering the criterion stated in Carr and Madan (1999), we take in numerical calibration $N = 2^{13}$, $\eta = 0.25$ for both (B.8) and (B.9), and $\xi = 4$ in (B.9).

REFERENCES

Aas, K., and Haff, I. (2006), "The Generalized Hyperbolic Skew Student's t-Distribution," *Journal of Financial Econometrics*, 4, 275–309. [88,90]

Aas, K., Haff, I. H., Dimakos, X. K., and Center, N. (2006), "Risk Estimation Using the Multivariate Normal Inverse Gaussian Distribution," *Journal of Risk*, 8, 39–60. [88]

Aït-Sahalia, Y., and Lo, A. (2000a), "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94, 9–51. [89]

——— (2000b), "Nonparametric Risk Management and Implied Risk Aversion," *Journal of Econometrics*, 94, 9–51. [102]

Avramidis, A., and L'Ecuyer, P. (2006), "Efficient Monte Carlo and quasi-Monte Carlo Option Pricing Under the Variance-Gamma Model," *Management Science*, 52, 1930–1994. [88]

Bakshi, G., Kapadia, N. and Madan, D. (2003), "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *The Review of Financial Studies*, 16, 101–143. [88,92,94,102]

Bali, T., and Theodossiou, P. (2007), "A Conditional-SGT-VaR Approach With Alternative GARCH Models," *Annals of Operations Research*, 151, 241–267. [88]

Bams, D., Lehnert, T., and Wolff, C. (2005), "An Evaluation Framework for Alternative VaR-Models," *Journal of International Money and Finance*, 24, 944–958. [88]

Barndorff-Nielsen, O. (1977), "Exponentially Decreasing Distributions for the Logarithm of Particle Size," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 353, 401–419. [89]

——— (1997a), "Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling," *Scandinavian Journal of Statistics*, 24, 1–13. [88]

——— (1997b), "Processes of Normal Inverse Gaussian Type," *Finance and Stochastics*, 2, 41–68. [88]

Barndorff-Nielsen, O., and Shephard, N. (2001), "Normal Modified Stable Processes," *Theory of Probability and Mathematical Statistics*, 65, 1–20. [90]

Barndorff-Nielsen, O., and Stelzer, R. (2005), "Absolute Moments of Generalized Hyperbolic Distributions and Approximate Scaling of Normal Inverse Gaussian Levy Processes," *Scandinavian Journal of Statistics*, 32, 617–637. [90,108]

Barton, D., and Dennis, K. (1952), "The Conditions Under Which Gram-Charlier and Edgeworth Curves are Positive Definite and Unimodal," *Biometrika*, 39, 425–427. [93]

Bates, D. (1995), "Testing Option Pricing Models," Discussion paper, National Bureau of Economic Research. [101]

Battalio, R., and Schultz, P. (2006), "Options and the Bubble," *The Journal of Finance*, 61, 2071–2102. [92]

Bauwens, L., and Laurent, S. (2005), "A New Class of Multivariate Skew Densities, With Application to Generalized Autoregressive Conditional Heteroscedasticity Models," *Journal of Business and Economic Statistics*, 23, 346–354. [88]

Bibby, B., and Sørensen, M. (2003), "Hyperbolic Processes in Finance," in *Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev, Philadelphia, PA: Elsevier Science, pp. 211–248. [88,90]

Bondarenko, O. (2003), "Estimation of Risk-Neutral Densities Using Positive Convolution Approximation," *Journal of Econometrics*, 116, 85–112. [96]

Carr, P., Geman, H., Madan, D., and Yor, M. (2002), "The Fine Structure of Asset Returns: An Empirical Investigation*," *The Journal of Business*, 75, 305–332. [88]

Carr, P., and Madan, D. (1999), "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance*, 2, 61–73. [95,110]

Chen, Y., Hardle, W., and Jeong, S. (2008), "Nonparametric Risk Management With Generalized Hyperbolic Distributions," *Journal of the American Statistical Association*, 103, 910–923. [88]

Chernov, M., and Ghysels, E. (2000), "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 407–458. [101]

Chihara, T. (1989), "Hamburger Moment Problems and Orthogonal Polynomials," *Transactions of the American Mathematical Society*, 315, 189–203. [89]

Christoffersen, P. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841–862. [105]

Conrad, J., Dittmar, R., and Ghysels, E. (2013), "Ex Ante Skewness and Expected Stock Returns," *Journal of Finance*, 68, 85–124. [92]

Dennis, P., and Mayhew, S. (2002), "Risk-Neutral Skewness: Evidence From Stock Options," *Journal of Financial and Quantitative Analysis*, 37, 471–493. [92,93]

Duan, J.-C. (1995), "The GARCH Option Pricing Model," *Mathematical Finance*, 5, 13–32. [97]

Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376. [95]

Eberlein, E. (2001a), "Application of Generalized Hyperbolic Lévy Motions to Finance," in *Lévy Processes: Theory and Applications*, eds. O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, Basel, Switzerland: Birkhäuser Verlag, pp. 319–337. [88]

——— (2001b), "Application of Generalized Hyperbolic Levy Motions to Finance," *Levy Processes: Theory and Applications*, Boston, MA: Birkhauser, pp. 319–336. [88]

Eberlein, E., and Hammerstein, E. (2004), "Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes," in *Seminar on Stochastic Analysis, Random Fields, and Applications IV: Centro Stefano Franscini, Ascona, May 2002*, p. 221. [88,90]

Eberlein, E., and Keller, U. (1995a), "Hyperbolic Distributions in Finance," *Bernoulli*, 1, 281–299. [88]

——— (1995b), "Hyperbolic Distributions in Finance," *Bernoulli*, 281–299. [89]

Eberlein, E., Keller, U., and Prause, K. (1998), "New Insights Into Smile, Mispricing, and Value at Risk: The Hyperbolic Model," *Journal of Business*, 71, 371–405. [88]

Eberlein, E., and Prause, K. (2002), "The Generalized Hyperbolic Model: Financial Derivatives and Risk Measures," in *Mathematical Finance—Bachelier Finance Congress 2000*, eds. H. Geman, D. Madan, S. Pliska, and T. Vorst, p. 245–267. [88]

Eriksson, A., Forsberg, L., and Ghysels, E. (2004), "Approximating the Probability Distribution of Functions of Random Variables: A New Approach," Discussion paper CIRANO. [91]

Eriksson, A., Ghysels, E., and Wang, F. (2009), "The Normal Inverse Gaussian Distribution and the Pricing of Derivatives," *The Journal of Derivatives*, 16, 23–37. [88,93]

Figlewski, S. (2007), "The Implied Risk Neutral Density for the US Market Portfolio," Discussion Paper NYU. [88]

Frecka, T., and Hopwood, W. (1983), "The Effects of Outliers on the Cross-Sectional Distributional Properties of Financial Ratios," *The Accounting Review*, 58, 115–128. [88]

Haas, M., and Pigorsch, C. (2007), "Financial Economics, Fat-Tailed Distributions," *Encyclopedia of Complexity and Systems Science*, New York: Springer, pp. 3404–3435. [90]

Heston, S. L., and Nandi, S. (2000), "A Closed-Form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625. [97]

Hirsa, A., and Madan, D. (2004), "Pricing American Options Under Variance Gamma," *Journal of Computational Finance*, 7, 63–80. [88]

Jackwerth, J. (2000), "Recovering Risk Aversion From Option Prices and Realized Returns," *Review of Financial Studies*, 13, 433–451. [102]

Kalemanova, A. (2007), "The Normal Inverse Gaussian Distribution for Synthetic CDO Pricing," *The Journal of Derivatives*, 14, 80–93. [88]

Keller-Ressel, M. (2011), "Moment Explosions and Long-Term Behavior of Affine Stochastic Volatility Models," *Mathematical Finance*, 21, 73–98. [96]

Konikov, M., and Madan, D. (2002), "Option Pricing Using Variance Gamma Markov Chains," *Review of Derivatives Research*, 5, 81–115. [88]

Kuester, K., Mittnik, S., and Paolella, M. (2006), "Value-at-Risk Prediction: A Comparison of Alternative Strategies," *Journal of Financial Econometrics*, 4, 53–89. [88]

Lee, R. (2004), "Option Pricing by Transform Methods: Extensions, Unification and Error Control," *Journal of Computational Finance*, 7, 51–86. [95]

León, A., Mencía, J., and Sentana, E. (2009), "Parametric Properties of Semi-Nonparametric Distributions, With Applications to Option Valuation," *Journal of Business and Economic Statistics*, 27, 176–192. [89,103]

Liu, X., Shackleton, M., Taylor, S., and Xu, X. (2007), "Closed-Form Transformations From Risk-Neutral to Real-World Distributions," *Journal of Banking & Finance*, 31, 1501–1520. [102,103]

Madan, D., Carr, P., and Chang, E. (1998), "The Variance Gamma Process and Option Pricing," *European Finance Review*, 2, 79–105. [88]

Madan, D., and Milne, F. (1994), "Contingent Claims Valued and Hedged by Pricing and Investing in a Basis," *Mathematical Finance*, 4, 223–245. [93]

Madan, D., and Seneta, E. (1990), "The Variance Gamma (VG) Model for Share Market Returns," *Journal of Business*, 63, 511. [88]

Moosbrucker, T. (2006), "Pricing CDOs With Correlated Variance Gamma Distributions," Research Report, Department of Banking, University of Cologne. [88]

Prause, K. (1999), *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, Ph.D. dissertation, University of Freiburg, Freiburg (Breisgau). [88,90]

Ribeiro, C., and Webber, N. (2004), "Valuing Path-Dependent Options in the Variance-Gamma Model by Monte Carlo With a Gamma Bridge," *Journal of Computational Finance*, 7, 81–100. [88]

Rompolis, L., and Tzavalis, E. (2010), "Risk Premium Effects on Implied Volatility Regressions," *Journal of Financial Research*, 33, 125–151. [102]

Rosenberg, J., and Engle, R. (2002), "Empirical Pricing Kernels," *Journal of Financial Economics*, 64, 341–372. [102]

Rosenberg, J., and Schuermann, T. (2006), "A General Approach to Integrated Risk Management With Skewed, Fat-Tailed Risks," *Journal of Financial Economics*, 79, 569–614. [88]

Rubinstein, M. (1998), "Edgeworth Binomial Trees," *Journal of Derivatives*, 5, 20–27. [93]

Rydberg, T. (1997), "The Normal Inverse Gaussian Levy Process: Simulation and Approximation," *Communications in Statistics: Stochastic Models*, 13, 887–910. [88]

Rydberg, T. (1999), "Generalized Hyperbolic Diffusion Processes With Applications in Finance," *Mathematical Finance*, 9, 183–201. [88]

Seneta, E. (2004), "Fitting the Variance-Gamma Model to Financial Data," *Journal of Applied Probability*, 41, 177–187. [88]

Singleton, K. (2006), *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*, Princeton University Press. [103]

Theodossiou, P. (1998), "Financial Data and the Skewed Generalized T Distribution," *Management Science*, 44, 1650–1661. [88]

Venter, J., and de Jongh, P. (2002), "Risk Estimation Using the Normal Inverse Gaussian Distribution," *Journal of Risk*, 4, 1–24. [88]

Wang, S. (2000), "A Class of Distortion Operators for Pricing Financial and Insurance Risks," *Journal of Risk and Insurance*, 15–36. [88]

Widder, D. (1946), *The Laplace Transform*, Princeton Mathematical series, 6, Princeton, NJ: Princeton University Press. [89]