NBER WORKING PAPER SERIES

TAMING THE SKEW: HIGHER-ORDER MOMENTS IN MODELING ASSET PRICE PROCESSES IN FINANCE

Sanjiv Ranjan Das
Rangarajan K. Sundaram

Working Paper 5976

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1997

We benefited from discussions with Menachem Brenner, John Campbell, George Chacko, Silverio Foresi, Blake LeBaron and S.R.S. Varadhan. We are also grateful to seminar participants at the NBER AP Lunch seminar and UCLA for their comments. The second author would also like to thank the National Science Foundation for research support under grant 94-10485. This paper is part of NBER's research program in Asset Pricing. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Taming the Skew: Higher-Order Moments in
Modeling Asset Price Processes in Finance
Sanjiv Ranjan Das and Rangarajan K. Sundaram
NBER Working Paper No. 5976
March 1997
Asset Pricing

ABSTRACT

It is widely acknowledged that many financial markets exhibit a considerably greater degree of kurtosis (and sometimes also skewness) than is consistent with the Geometric Brownian Motion model of Black and Scholes (1973). Among the many alternative models that have been proposed in this context, two have become especially popular in recent years: models of jump-diffusions, and models of stochastic volatility. This paper explores the statistical properties of these models with a view to identifying simple criteria for judging the consistency of either model with data from a given market; our specific focus is on the patterns of skewness and kurtosis that arise in each case as the length of the interval of observations changes. We find that, regardless of the precise parameterization employed, these patterns are strikingly similar *within* each class of models, enabling a simple consistency test along the desired lines. As an added bonus, we find that for most parameterizations, the set of possible patterns differs sharply *across* the two models, so that data from a given market will typically not be consistent with both models. However, there exist exceptional parameter configurations under which skewness and kurtosis in the two models exhibit remarkably similar behavior from a qualitative standpoint. The results herein will be useful to empiricists, theorists and practitioners looking for parsimonious models of asset prices.

Sanjiv Ranjan Das
Graduate School of Business
Harvard University
Soldiers Field Road
Boston, MA 02163
and NBER
sdas@hbs.edu

Rangarajan K. Sundaram
Department of Finance
Stern School of Business
New York University
New York, NY 10012
rsundara@stern.nyu.edu

1 Introduction and Summary

The Geometric Brownian Motion (GBM) model of Black and Scholes (1973) remains the preeminent model in Finance for describing the evolution of asset prices in equity, stock index, and currency markets. Under this process, the (continuously compounded) returns $Z(h)$ from holding the asset over any period of length h are presumed to be distributed normally with a mean and variance that are linear in h:

$$Z(h) \sim N[\mu h, \sigma^2 h]. \tag{1}$$

Since any normal distribution is symmetric about its mean and has zero excess kurtosis, expression (1) carries two related—but distinct—implications about the *higher-order* moments of the returns $Z(h)$:

1. The skewness and the excess kurtosis of $Z(h)$ are both zero.
2. The skewness and excess kurtosis of $Z(h)$ do not depend on h.

Of course, since conditional and unconditional returns distributions coincide under a GBM (witness (1)), these implications apply equally to the respective conditional and unconditional moments.

It is now widely acknowledged that these implications are routinely violated in a number of financial markets. At the unconditional level, observed price histories in many markets indicate the presence of non-zero skewness and a strong degree of leptokurtosis (positive excess kurtosis) in the returns distributions. Moreover, far from being independent of the length h of the interval of observation, the extent of unconditional skewness and leptokurtosis tends to depend in a non-trivial manner on h. For example, using data from 1974–1985, Jorion (1988) reports that:

- Excess kurtosis in the \$/DM exchange rate is 3.29 when weekly observations are used, but falls to 1.56 when monthly observations are employed.
- Excess kurtosis in the distribution of the value-weighted CRSP index is 2.922 under weekly observations, but is only 0.89 under monthly observations.

Jorion (1988) and Bates (1996) contain references to a number of other studies that have reported similar results. These studies indicate that the extent of violation differs across markets, but there are also some strong common patterns; for instance, the degree of unconditional leptokurtosis typically declines as the interval of observation increases.[1]

From a conditional standpoint also, there is much indirect evidence from options markets that the implications of normality fail to hold. It is well known, for instance, that implied volatilities obtained from option prices under the Black–Scholes model (1) are typically higher for out-of-the-money options than for otherwise identical at-the-money options (the so-called "volatility skew"). It is also typically the case that the differences in implied volatilities become less pronounced as options with a greater time-to-maturity are considered. Finally, the implied volatility curve is often quite asymmetric, with implied volatility estimates for out-of-the-money and in-the-money options differing markedly. (See Figure 1 for an illustration of all three points using S&P 500 index options.) The existence of the volatility skew is, as Bates (1996) and others have observed, strongly suggestive of the presence of leptokurtosis in the conditional returns distribution.[2] The reduction in the skew at longer maturities indicates that the degree of conditional leptokurtosis declines as the horizon increases. Finally, the asymmetric nature of the implied volatility curve implies the presence of skewness in the conditional returns distribution.

Many alternative continuous-time models have been proposed in the Finance literature to account for these systematic departures from the assumptions of the GBM model. Among these, two have become especially popular in recent years. The first, the class of *jump-diffusion* models, augments the return distribution of the GBM with a Poisson-driven "jump" process. Models of this sort were initially proposed by Merton (1976); more recently, they have been studied by Jarrow and Rosenfeld (1984), Ball and Torous (1985), Ahn (1992), Amin (1993), and Das and Foresi (1996), among others. The second alternative, the class of *stochastic volatility* models. also involves a generalization of the GBM process, this time by allowing the volatility σ of the return process to itself evolve stochastically over time. Recent papers adopting such an approach include Wiggins (1987), Melino and Turnbull (1990), Stein and Stein (1991), Amin and Ng (1993), and Heston (1993). In addition, Jorion (1988)

[1] The evidence on skewness is more mixed. While skewness also appears to be monotonically related to h in many markets, it is a decreasing function of h in some cases, and an increasing function of h in others.

[2] Evidence of conditional leptokurtosis also comes from the ARCH/GARCH literature in the form of fat-tailed residuals. See Bollerslev, Chou, and Kroner (1992) for a survey of this literature.

and Bates (1996) have each studied models of jump-diffusions combined with stochastic volatility.[3]

At a purely intuitive level, it is not difficult to see how jump-diffusion and stochastic volatility models could each lead to return distributions that exhibit skewness and leptokurtosis. Skewness in jump-diffusions should arise from the distribution of jump-sizes; for instance, if the jump size is always positive, the return distribution should be positively skewed. Moreover, the presence of jumps in the returns process creates outliers which add fatness to the tails of the distribution; thus, returns under a jump-diffusion should always be leptokurtic. In stochastic volatility models, skewness should arise if there is non-zero correlation between the stochastic processes driving changes in the returns and the volatility, respectively; for instance, positive correlation between these processes should result in positive skewness, since high returns will be accompanied by high volatility, and low returns by low volatility. Leptokurtosis should obtain from the changing values of volatility (the "volatility of volatility" as it is sometimes called).

Unfortunately, there are two problems with this scenario. First, while these intuitive arguments appear compelling, the literature does not, to the best of our knowledge, contain a formal confirmation that either class of models can, in fact, generate non-zero skewness and/or leptokurtosis in the returns distribution. Second, on a more important note, even the generation of non-zero skewness and leptokurtosis by the posited price process is not, by itself, sufficient. As we have indicated above, the empirical literature also points to the existence of definite *patterns* concerning the manner in which skewness and excess kurtosis vary with the length h of the interval of observation. The question naturally arises: to what extent are these patterns consistent with jump-diffusions or stochastic volatility models? More generally:

> How do the moments of the returns $Z(h)$ in (a) jump-diffusion models, and (b) stochastic volatility models, depend on the length h of the holding period?

This question has received surprisingly little attention in the Finance literature. Yet, there are obvious advantages to be gained by identifying general patterns that can—and cannot—

[3]Strictly speaking, Jorion's paper involves a jump-diffusion combined with an ARCH model, rather than with a stochastic volatility model. While conditional variance in an ARCH model is a *nonstochastic* function of past innovations, the work of Nelson (1990) indicates that, at least in some cases, ARCH models may be considered as discrete-time approximations of a continuous-time stochastic volatility model.

be generated under each process. Most importantly, there is the possibility that the exercise will result in simple criteria for judging the appropriateness of either model in a given setting. If it were to turn out, for example, that the degree of excess kurtosis in stochastic volatility models necessarily *increases* with h, then such models cannot be profitably used to represent price processes in markets in which excess kurtosis *declines* as h increases. Second, on a related note, there is the possibility that identifying all possible patterns under each of the two classes of models will provide us with criteria for distinguishing between them using only data on asset (or option) prices. Finally, such an exercise would make full use of the fact that data on prices is available at a variety of intervals (daily, weekly, monthly). In particular, the "aliasing" problem that the two models' implications could be very similar for certain intervals of observation but not for others does not arise.[4]

This paper undertakes a systematic investigation of the behavior of the returns $Z(h)$ as a function of h under both jump-diffusions and stochastic volatility models. The structure of this paper, and a summary of our main findings, are described in the subsections following. Subsection 1.3 relates our paper to some others in the literature.

1.1 Structure of this Paper

Our analysis begins in section 3 with an examination of the properties of returns under a jump-diffusion process. The model of the jump process that we employ is more general than is commonly used in the literature in that we place no restrictions on the distribution of the jump size. Since the returns process in a jump-diffusion has increments that are stationary and independent, conditional and unconditional distributions coincide in this case. Our analysis in this section does not, therefore, attempt to distinguish between conditional and unconditional moments.

Sections 4 and 5 turn to the study of stochastic volatility models. Following Heston (1993) and Bates (1996), we represent volatility as evolving according to a mean-reverting square-root process.[5] In addition, to capture the possibility that changes in volatility could

[4]An advantage of working in continuous-time is that it is possible to gauge the impact of varying h continuously. In discrete-time, h would have to be an integer multiple of the minimum time interval.

[5]In part, our choice of this process was guided by the many advantages it possesses over alternatives (see Bates (1996) for details). However, our analysis of the stochastic volatility model appears robust to this specification. Our calculations show that the same qualitative results obtain if the square-root process for the variance of instantaneous returns is replaced with an Ornstein–Uhlenbeck process for the standard

depend on the level of prices, we allow for arbitrary correlation between the Wiener processes driving the returns and volatility processes, respectively. This correlation factor, which we denote by ρ plays a central role in many of our results (see especially result 3(b) below). Unlike models of jump-diffusions, the conditional and unconditional returns of a stochastic volatility process will not, in general, coincide. Section 4 investigates the behavior of skewness and kurtosis in the *conditional* returns of stochastic volatility models, while Section 5 does likewise for the *unconditional* returns.

Section 6 concludes. The appendices contain proofs of results omitted in the main body of the text.

1.2 Main Results: A Summary

We provide closed-form expressions for the first four moments of the conditional and unconditional returns distributions under both jump-diffusions and stochastic volatility. Confirming the intuition presented above, we show that that excess kurtosis is always strictly positive under both models, but that skewness can be positive, negative, or zero. However, the manner in which skewness and kurtosis depend on h is very different in the two cases, with the differences being starker for conditional returns than for unconditional returns. We summarize the results by model below. Recall that conditional and unconditional returns distributions coincide in a jump-diffusion model, but not in a stochastic volatility model.

1. Concerning jump-diffusions:

   (a) There are three possible patterns of skewness: (i) skewness is zero everywhere, (ii) skewness is positive and decreasing in h, or (iii) skewness is negative and increasing in h. In the last two cases, skewness must be proportional to $1/\sqrt{h}$.

   (b) Excess kurtosis is strictly positive and decreasing in h. Indeed, excess kurtosis must be proportional to $1/h$.

2. Concerning conditional returns under stochastic volatility:

   (a) There are three possible patterns of skewness: (i) skewness is zero everywhere, (ii) skewness is positive and tent-shaped in h, increasing from zero towards a

---

deviation of instantaneous returns.

maximum, and then decreasing asymptotically back to zero, or (iii) skewness is negative and "v"-shaped in h, decreasing from zero to a minimum, and then increasing asymptotically back to zero again. For plausible values of the remaining parameters, the maximum in case (ii) and minimum in case (iii) are attained only when h is of the order of at least several months.

(b) Excess kurtosis is strictly positive and is tent-shaped in h, increasing from zero towards a maximum, and then decreasing asymptotically to zero again. For plausible values of the remaining parameters, the maximum value of excess kurtosis is attained only when h is of the order of at least several months.

3. Concerning unconditional returns under stochastic volatility:

   (a) There are three possible patterns of skewness, viz., the same as those identified in 2(a). It is again the case that, for plausible values of the remaining parameters, the extremum values of skewness are attained only when h is of the order of at least several months.

   (b) Excess kurtosis is strictly positive, but its behavior as h varies depends on the value of the correlation factor ρ. If $|\rho|$ is sufficiently high, kurtosis will be tent-shaped; if not, it will be decreasing everywhere in h. (See Section 5 for further details.)

In comparing *conditional* returns under the two classes of models, a sharp contrast arises from the fact that the extremum values of conditional skewness and kurtosis in a stochastic volatility model are attained only at relatively large values of h. If the extrema had occurred at a very small value of h (of the order of one day, for example), then the patterns of conditional skewness and kurtosis in stochastic volatility and jump-diffusion models would have been indistinguishable. As it stands, however, Results 1 and 2 imply that for at least small to moderate values of h, the two classes of models exhibit opposite patterns of conditional kurtosis; and that if skewness is non-zero, they also exhibit opposite patterns of conditional skewness. This offers the researcher a potential way to identify the "more" appropriate model in a given setting.

The situation is less rosy where *unconditional* returns are concerned. The surprising and unintuitive result 3(b), which does not seem to have been anticipated in the literature, implies that there are plausible parameter configurations under which kurtosis in a stochastic volatility model behaves in qualitatively the same way as kurtosis under a jump-diffusion. This implies in turn that in the absence of skewness, there may be no way to distinguish between the two classes of models using only unconditional (asset price) data. Too much should not be made of this, however. From a theoretical standpoint, the only way the stochastic volatility model will exhibit both zero skewness and decreasing kurtosis in the unconditional returns is when $\rho = 0$.[6] For all other values of ρ, unconditional skewness is non-zero, providing a differentiation criterion. Secondly, from an empirical standpoint, many markets do exhibit at least some degree of skewness in their unconditional returns distributions. Finally, regardless of the behavior of unconditional returns, the *conditional* returns can always be used to distinguish between the models.

This paper's presentation of closed-form solutions for the first four moments of the stochastic processes studied has an important by-product: it becomes possible to exploit these closed-forms to implement method of moments estimation. This is a logical follow-up to the results presented here, and one with which we are involved currently.

Finally, it is an obvious point, but one that must be stressed, that the set of patterns we identify as being possible under jump-diffusions and stochastic volatility models is by no means exhaustive. It is entirely possible that a given market may exhibit skewness and/or kurtosis patterns that do not fit either model. In this case, neither model is appropriate in the strict sense of the word, so our results cannot provide a guideline to the researcher on how to proceed.

1.3 Comments on the Related Literature

The "standard" approach in the literature to comparing jump-diffusion and stochastic volatility models is an econometric one: the parameters of the two processes are estimated using data at a fixed frequency, and the process that fits the data better is identified. We provide a brief description here of two papers that use this procedure, Jorion (1988) and Bates (1996).

Jorion (1988) undertakes an empirical study of stock prices and $/DM exchange rates using both weekly and monthly data. Using maximum likelihood methods, he estimates the parameters of four models: a GBM model, a jump-diffusion model (with normally-distributed jump sizes), an ARCH model, and a combined jump-diffusion/ARCH model. He

[6] It is interesting, therefore, to note that some of studies of stochastic volatility models in the literature have, for purposes of analytic tractability, restricted attention to the case $\rho = 0$.

finds that with monthly data, there is not much difference between the models, but that when weekly data is used the GBM is rejected against all three alternatives. The combined jump-diffusion/ARCH model performs best, with the jump-diffusion being, a priori, more probable than the ARCH model. Lastly, concerning possible model misspecification, Jorion finds that with monthly data, all the models fit the observations well, but with weekly data, only the jump-diffusion does not seem incorrectly specified.

Maximum likelihood comparison of jump diffusion and stochastic volatility models is often hampered by the fact that the transition probabilities for the likelihood function are not typically available in closed form. Bates (1996) develops a jump-diffusion/stochastic volatility model, and derives the transition density function as well as option pricing functions for this model. Bates undertakes cross-sectional estimation using weekly data on options prices, and, like Jorion, finds that the jump-diffusion submodel explains the data better than the stochastic volatility submodel.

The approach taken by Jorion and Bates to comparing jump-diffusions and stochastic volatility models is, of course, much more sophisticated than using the simple rules of thumb we develop here. However, we think there are some advantages to using our results to form at least a preliminary judgment about the consistency of the different models with the given setting. For one thing, the econometric approach is computationally non-trivial and does not always lend itself to easy practical implementation. In Bates' setting, for instance, obtaining maximum likelihood estimates involves first solving a partial differential equation to obtain the characteristic function of the probability density, then performing numerical Fourier inversion of the characteristic function to recover the density, and finally using further numerical work to implement maximum likelihood estimation. Secondly, the econometric approach does not take explicit advantage of the fact that data is available at different frequencies. This gives rise to the "aliasing" problem that the alternative models may not be dissimilar at the chosen frequency, but could be very different at others. For instance, we mentioned above the finding of Jorion (1988) that the different models he considers are not very different when monthly data is employed, but diverge significantly under weekly data. In contrast, the properties of the alternative models that we develop in this paper rely precisely on the fact that the models' implications diverge at different data frequencies.

A second branch of the literature that is related to our paper concerns the issues that arise when estimating continuous-time models with discrete-time data (e.g., Melino (1994). Especially relevant in this context is the recent paper of Ait-Sahalia (1996), which asks whether it is possible, using discrete data to identify whether the "true" continuous-time model is Markovian, and, if so, whether it has continuous sample paths. There is an obvious similarity between these questions, and those motivating the analysis in this paper. However, there are important differences. For one thing, our primary purpose is to identify whether observed patterns of skewness and kurtosis in different markets are consistent with either class of models, rather than to identify a "differentiation" criterion between the two models (the latter is simply a useful by-product of our analysis). Secondly, Ait-Sahalia's focus is on interest-rate processes, while we examine processes that are more commonly used to represent stocks, stock indices, and exchange rates. Lastly, and importantly, the test proposed by Ait-Sahalia to distinguish between continuous and discontinuous sample-path processes requires the processes to be Markovian and univariate. This rules out our setting: the returns and volatility processes in the stochastic volatility model are jointly Markovian, but not, of course, univariate; the returns process on its own is not Markovian.

Some Additional Remarks

The results of this paper also clarify and extend observations made in the literature concerning the properties of models of jump-diffusions and/or stochastic volatility. For example, Bates (1996, p.73) asserts that

> Stochastic volatility implies a direct relationship between the magnitude of conditional skewness and excess kurtosis, whereas jumps imply a strong inverse relationship.

The same claim is also made in Das and Foresi (1996). As our results show, this statement is only partially true, since both conditional skewness and conditional kurtosis in stochastic volatility models are tent-shaped, and not monotone.

Secondly, Merton (1976) and Jorion (1988) provide a closed-form expression for the variance of a jump-diffusion process in the special case where the jump-size is normally distributed. Proposition 3.1 shows that their expressions are actually incorrect. For details, see the remark following the statement of Proposition 3.1.

Finally, we noted above that a "volatility skew" arises in most markets when the Black-Scholes model is used to recover implied volatilities; and that the extent of the skew typically declines as option maturity increases. Put in terms of the model's pricing implications, this means that the Black-Scholes model tends to underprice out-of-the-money options relative to options that are at-the-money, but that the extent of the pricing bias decreases as the horizon increases. It has been suggested (see, e.g., Bodurtha and Courtadon (1987), or Jorion (1988)) that both the existence of the bias and the way the bias changes with the horizon are consistent with a jump-diffusion process for asset prices. Intuitively, the argument goes, the possibility of jumps is very important for short-maturity options that are out-of-the-money (since a single jump could take the option into-the-money), but becomes progressively less important as maturity increases.

It appears to us that this argument is incomplete. Underpricing, in our mind, is a consequence of ignoring the conditional leptokurtosis in asset prices, in particular, of assigning too small a likelihood to "extreme" events. Thus, a decrease in the pricing bias at longer maturities occurs *because* the degree of conditional leptokurtosis declines as the horizon h increases. Now, to the extent that jump-diffusions can generate leptokurtosis, they can provide an explanation of the *existence* of the bias. However, unless it can also be shown that conditional excess kurtosis in a jump-diffusion decreases with increasing horizons, it cannot be claimed that they explain the *pattern* of bias changes. Our paper demonstrates that conditional leptokurtosis in a jump-diffusion model does, in fact, vary in the required fashion, completing the argument.

2 Geometric Brownian Motion: A Brief Review

Let S_0 denote the initial (time-0) price of the asset. Under Geometric Brownian Motion (GBM), the time-t price S_t evolves according to

$$S_t \;=\; S_0 \exp\{\alpha t + \sigma W_t\}, \tag{2}$$

where α and σ are given constants, and W_t is a standard Brownian motion process.

Let $h > 0$ denote the length of time between observations of the price, and let $Z(h)$ denote the (continuously compounded) return from holding the asset over the time interval $[t, t+h]$. Then, by definition:

$$Z(h) \;=\; \log\left(\frac{S_{t+h}}{S_t}\right). \tag{3}$$

From (2), $Z(h)$ is given by

$$Z(h) = \alpha h + \sigma(W_{t+h} - W_t). \tag{4}$$

Since (W_τ) is a Wiener process, the difference $(W_{t+h} - W_t)$ is normally distributed with a mean of zero and a variance of h. It follows from (4), therefore, that the returns $Z(h)$ are themselves normally distributed with mean equal to αh and variance equal to $\sigma^2 h$:

$$Z(h) \sim N\left(\alpha h, \sigma^2 h\right). \tag{5}$$

Now recall that if X is any random variable with mean m and variance $s^2 > 0$, then the skewness and kurtosis of X are defined as

$$\text{Skewness}\,(X) = \frac{E[(X-m)^3]}{s^3} \quad \text{and} \quad \text{Kurtosis}\,(X) = \frac{E[(X-m)^4]}{s^4}, \tag{6}$$

and that the *excess kurtosis* of X is given by Excess Kurtosis $(X) = [\text{Kurtosis}(X) - 3]$.

It is well known that any normal distribution has zero skewness and a kurtosis of 3. Therefore, expression (5) carries two strong implications about the returns $Z(h)$:

1. The returns have zero skewness and zero excess kurtosis.

2. The skewness and kurtosis of the returns do not depend on h.

As mentioned in the Introduction, however, both conditions are routinely violated in a number of financial markets. In the following sections, we examine two of the proposals—models of *jump-diffusion*, and models of *stochastic volatility*—that have been made to overcome these problems.

3 Jump–Diffusion Models

A jump–diffusion model is obtained by augmenting the return process in a GBM with a Poisson jump process. Given any date $t \geq 0$, and a holding-period of length $h > 0$, the returns $Z_t(h)$ over the period $[t, t+h]$ in such a model are given by

$$Z_t(h) = \begin{cases} x, & \text{if } K = 0 \\ x + y_1 + \cdots + y_K, & \text{if } K \geq 1. \end{cases} \tag{7}$$

where:

1. $x \sim N(\alpha h, \sigma^2 h)$.

2. $y_1, y_2, \ldots$ is an i.i.d. sequence with common distribution G.

3. K is distributed Poisson with parameter λh, with $\lambda \geq 0$. That is, for $k = 0, 1, 2, \ldots$ we have

$$\text{Prob}\,(K = k) \;=\; \left(\frac{e^{-\lambda h}(\lambda h)^k}{k!}\right).$$

Observe that $Z_t(h)$ does not depend on t in any way. Therefore, conditional and unconditional returns distributions coincide in a jump-diffusion model. To simplify exposition in the sequel, we will drop the subscript t, and simply use $Z(h)$ to denote the returns over an interval of length h.

It is easy to see that Geometric Brownian Motion is a special case of jump-diffusion processes that arises when either (a) $\lambda = 0$, or (b) G is degenerate at zero.[7] All of our results in this section are valid for this special case also. However, since our focus is primarily on non-trivial jump-diffusion models, we will suppose henceforth that

Assumption 1 G does not place unit mass on the point $y = 0$. Moreover, $\lambda > 0$.

Our second assumption concerns the moments of the distribution G. Let ν_j denote the j-th moment of G:

$$\nu_j \;=\; \int_{-\infty}^{+\infty} y^j dG(y).$$

Assumption 2 ν_j exists and is finite for $j = 1, 2, 3, 4$.

Assumption 2 is a mild condition that is met under virtually all the distributions used in practice; it is also a necessary requirement in the context of any sensible discussion of higher-order moments.

We do not place any other restrictions on the distribution G. In particular, G could be a continuous distribution (for example, normal, as is commonly assumed), a discrete

[7]When $\lambda = 0$, we have $K = 0$ with probability one, so from (7), the jump-diffusion is simply a model of Geometric Brownian Motion. Similarly, if G is degenerate at zero, then $y_1 = \cdots = y_k = \cdots = 0$ with probability one, so from (7), the jump-diffusion is *identical* to the underlying geometric Brownian motion process.

distribution (for example, Bernoulli), or some combination of the two. Similarly, G may itself be a symmetric distribution with zero skewness and excess kurtosis, or it may possess non-zero skewness and/or excess kurtosis.

Note that under Assumptions 1 and 2, ν_4 must always be strictly positive.[8] No corresponding restrictions are implied, however, on the sign of ν_3, which could be positive, negative or zero. For instance, if G is the distribution

$$y = \begin{cases} +1, & \text{with probability } p \\ -1, & \text{with probability } (1-p) \end{cases}$$

then

$$\nu_3 \begin{cases} < 0, & \text{if } p < \frac{1}{2} \\ = 0, & \text{if } p = \frac{1}{2} \\ > 0, & \text{if } p > \frac{1}{2} \end{cases}$$

The following is the central result of this section:

Proposition 3.1 *The variance, skewness, and kurtosis of the jump-diffusion process (7) are given by:*

$$\text{Variance}[Z(h)] = E\left[(Z(h) - E[Z(h)])^2\right] = h(\sigma^2 + \lambda\nu_2). \tag{8}$$

$$\text{Skewness}[Z(h)] = \frac{E[(Z(h) - E[Z(h)])^3]}{[Var(Z(h))]^{3/2}} = \frac{1}{\sqrt{h}}\left[\frac{\lambda\nu_3}{(\sigma^2 + \lambda\nu_2)^{3/2}}\right]. \tag{9}$$

$$\text{Kurtosis}[Z(h)] = \frac{E[(Z(h) - E[Z(h)])^4]}{[Var(Z(h))]^2} = 3 + \frac{1}{h}\left[\frac{\lambda\nu_4}{(\sigma^2 + \lambda\nu_2)^2}\right]. \tag{10}$$

Therefore:

1. *Concerning Skewness:*

   (a) *If $\nu_3 = 0$, then the skewness of the jump-diffusion is also zero.*

   (b) *If $\nu_3 > 0$, the skewness is positive and inversely proportional to $\sqrt{h}$.*

   (c) *If $\nu_3 < 0$, the skewness is negative and inversely proportional to $\sqrt{h}$.*

[8] By its definition, ν_4 is always non-negative. Assumption 1 guarantees that it must be non-zero, while Assumption 2 ensures it is finite.

2. *The kurtosis of the jump-diffusion process (7) is always strictly positive and is inversely proportional to h.*

Remark Under the assumption that $G \sim N(\theta, \delta^2)$, Merton (1976, Section 4) and Jorion (1988, Section 5) state that the variance of $Z(h)$ is given by $(\sigma^2 + \lambda\delta^2)$. Proposition 3.1 shows that this assertion is incorrect, and that the correct expression for the variance in their models should be $(\sigma^2 + \lambda[\delta^2 + \theta^2])$. □

Proof See Appendix A. □

The content of Proposition 3.1 is illustrated through the following examples. The first example considers a "degenerate" case, where G places unit mass on some point. The second example applies Proposition 3.1 to the more interesting situation where G is a Bernoulli distribution; The third and last example examines the most widely-used distribution in practice, the normal distribution.

Example 3.2 Let G be the distribution that places unit mass on the point m, where $m \neq 0$. Then, $\nu_j = m^j$, $j = 1, 2, 3, 4$. In this case, therefore, the skewness and kurtosis of the jump-diffusion are given by

$$\text{Skewness}\,(Z(h)) = \frac{1}{\sqrt{h}}\left[\frac{\lambda m^3}{(\sigma^2 + m^2\lambda)^{3/2}}\right]$$

$$\text{Kurtosis}\,(Z(h)) = 3 + \frac{1}{h}\left[\frac{\lambda m^4}{(\sigma^2 + \lambda m^2)^2}\right]$$

Therefore, the jump-diffusion has positive skewness if $m > 0$ and negative skewness if $m < 0$. Of course, in either case it has positive excess kurtosis. □

Example 3.3 Let G be the Bernoulli distribution $y = \pm 1$ with probability $\frac{1}{2}$. Then, we have $\nu_1 = \nu_3 = 0$, and $\nu_2 = \nu_4 = 1$. From (9) and (10), we now obtain:

$$\text{Skewness}\,(Z(h)) = 0.$$

$$\text{Kurtosis}\,(Z(h)) = 3 + \frac{1}{h}\left[\frac{\lambda}{(\sigma^2 + \lambda)^2}\right].$$

□

Example 3.4 Let G be a normal distribution with mean θ and variance δ^2. Then, we have: $\nu_1 = \theta$, $\nu_2 = \delta^2 + \theta^2$, $\nu_3 = \theta^3 + 3\theta\delta^2$, and $\nu_4 = \theta^4 + 6\theta^2\delta^2 + 3\delta^4$. From (9)–(10), therefore:

$$\begin{aligned}
\text{Skewness}\,(Z(h)) &= \frac{1}{\sqrt{h}}\left[\frac{\lambda(\theta^3 + 3\theta\delta^2)}{(\sigma^2 + \lambda\delta^2 + \lambda\theta^2)^{3/2}}\right]. \\
\text{Kurtosis}\,(Z(h)) &= 3 + \frac{1}{h}\left[\frac{\lambda(\theta^4 + 6\theta^2\delta^2 + 3\delta^4)}{(\sigma^2 + \lambda\delta^2 + \lambda\theta^2)^2}\right].
\end{aligned}$$

Note that the skewness of the jump-diffusion is positive if $\theta > 0$ and negative if $\theta < 0$. In the special case where $\theta = 0$, these expressions simplify considerably:

$$\begin{aligned}
\text{Skewness}\,(Z(h)) &= 0. \\
\text{Kurtosis}\,(Z(h)) &= 3 + \frac{1}{h}\left[\frac{3\lambda\delta^4}{(\sigma^2 + \lambda\delta^2)^2}\right].
\end{aligned}$$

□

4 Conditional Returns in Stochastic Volatility Models

Stochastic volatility models generalize Geometric Brownian Motion by allowing the volatility of the return process to itself evolve stochastically over time. From (2), the cumulative returns $x_t = \ln(S_t/S_0)$ under a GBM obey $x_t = \exp\{\alpha t + \sigma W_t\}$; expressing this in stochastic differential form, we have

$$dx_t = \alpha\, dt + \sigma\, dW_t, \tag{11}$$

In a stochastic volatility model, the instantaneous variance $V = \sigma^2$ is no longer required to be constant, but is allowed to change with time. That is, (11) is replaced by the joint process

$$dx_t = \alpha\, dt + \sqrt{V_t}\, dW_t \tag{12}$$

$$dV_t = \xi(t, V_t)\, dt + \beta(t, V_t)\, dB_t \tag{13}$$

where W_t and B_t are standard Brownian motion processes. For the purposes of analysis, we must choose a functional form for (13). Following Heston (1993) and Bates (1996), we shall assume that volatility follows a mean-reverting square-root process given by

$$dV_t = \kappa(\theta - V_t)\, dt + \eta\sqrt{V_t}\, dB_t, \tag{14}$$

where κ, θ, and η are all strictly positive constants. This specification has many advantages, one of which is that V_t is guaranteed to remain strictly positive. (See Bates (1996) for a more detailed discussion.) Finally, we adopt a general specification for the relationship between the Wiener processes W_t and B_t:

$$dB_t \, dW_t = \rho \, dt. \tag{15}$$

Expression (15) captures the possibility that increases in volatility could be related to the level of asset prices. Expressions (12), (14), and (15) complete the description of the model we shall analyze in this section and the next.

As earlier, let $Z_t(h) = x_{t+h} - x_t$ denote the returns from holding the asset between times t and $t + h$. It is immediate from (12) that these returns will depend on the path taken by the volatility process V_τ over this interval. Moreover, it is also immediate from (14) that this path will depend on the initial value V_t of the volatility at time t. Thus, the returns $Z_t(h)$ depend on time–t information through V_t.[9] It follows that conditional and unconditional returns distributions will not necessarily coincide in a stochastic volatility model. In this section, we examine the properties of the conditional returns distribution; the unconditional distribution is the subject of the section following.

For notational simplicity, let $x_t = x$ and $V_t = v$. Denote by

$$F(x, v, h, s) = E[\exp\{isx_{t+h}\} \mid x, v]$$

the characteristic function of x_{t+h} conditional on time-t information. A standard argument (see, e.g., Duffie (1996)) establishes F may be obtained as the solution to the Kolmogorov Backward Equation

$$\alpha \frac{\partial F}{\partial x} + \frac{1}{2} v \frac{\partial^2 F}{\partial x^2} + \kappa(\theta - v)\frac{\partial F}{\partial v} + \frac{1}{2}\eta^2 v \frac{\partial^2 F}{\partial v^2} - \frac{\partial F}{\partial h} + \rho\eta v \frac{\partial^2 F}{\partial x \partial v} \; = \; 0, \tag{16}$$

subject to the initial condition

$$F(x, v, 0; s) \; = \; e^{isx}, \quad \text{for all } x, v, s. \tag{17}$$

Proposition 4.1 *The solution of (16)–(17) has the form*

$$F(x, v, h; s) \; = \; C(h) \exp[isx + A(h) + vB(h)]. \tag{18}$$

[9]Note, however, that the distribution of $Z_t(h)$ does not depend on x_t.

Letting $\gamma = \kappa - \rho\eta i s$ and $\phi = \sqrt{\gamma^2 + \eta^2 s^2}$, the terms $A(\cdot)$, $B(\cdot)$, and $C(\cdot)$ are given by

$$
\begin{aligned}
A(h) &= i\alpha s h \\
B(h) &= \frac{-s^2\left[e^{\phi h} - 1\right]}{(\phi + \gamma)\left[e^{\phi h} - 1\right] + 2\phi} \\
C(h) &= \left[\frac{2\phi\left(e^{(\phi+\gamma)\frac{h}{2}}\right)}{(\phi + \gamma)\left[e^{\phi h} - 1\right] + 2\phi}\right]^{\frac{2\kappa\theta}{\eta^2}}
\end{aligned}
$$

Proof See Appendix B. □

It is an easy matter now to obtain the characteristic function $F^*(h, v, s)$ of $Z_t(h)$. Indeed, since $Z_t(h) = x_{t+h} - x$, the function F^* is related to the function F of Proposition 4.1 through $F^*(v, h, s) = F(x, v, h, s)e^{-isx}$. Substituting for $F(x, v, h, s)$ from (18), we obtain

$$F^*(v, h, s) = C(h)\exp\{A(h) + vB(h)\}, \tag{19}$$

where $A(\cdot)$, $B(\cdot)$, and $C(\cdot)$ are as given in Proposition 4.1.

Expression (19) may now be used to obtain the moments of $Z_t(h)$ in the usual way: if we let $\xi_j = E[(Z_t(h))^j]$, then we have

$$\xi_j = \frac{1}{i^j} \cdot \frac{\partial^j F^*}{\partial s^j}(v, h; s)\bigg|_{s=0}$$

Grinding through the computations, we finally obtain the desired expressions:

$$\text{Skewness}\,(Z_t(h)) = \left(\frac{3\eta\rho e^{\frac{1}{2}\kappa h}}{\sqrt{\kappa}}\right)\left[\frac{\theta(2 - 2e^{\kappa h} + \kappa h + \kappa h e^{\kappa h}) - v(1 + \kappa h - e^{\kappa h})}{(\theta[1 - e^{\kappa h} + \kappa h e^{\kappa h}] + v[e^{\kappa h} - 1])^{3/2}}\right] \tag{20}$$

$$\text{Kurtosis}\,(Z_t(h)) = 3\left[1 + \eta^2\left(\frac{\theta A_1 - vA_2}{B}\right)\right] \tag{21}$$

where, letting $y = \kappa h$,

$$A_1 = [1 + 4e^y - 5e^{2y} + 4ye^y + 2ye^{2y}] + 4\rho^2[6e^y - 6e^{2y} + 4ye^y + 2ye^{2y} + y^2e^y] \tag{22}$$

$$A_2 = 2[1 - e^{2y} + 2ye^y] + 8\rho^2[2e^y - 2e^{2y} + 2ye^y + y^2e^y] \tag{23}$$

$$B = 2\kappa[\theta(1 - e^y + ye^y) + v(e^y - 1)]^2 \tag{24}$$

Expressions (20)-(24) may now be used to identify properties of the skewness and kurtosis of the conditional returns $Z_t(h)$. We begin with skewness first and then turn to kurtosis.

4.1 Skewness of the Conditional Returns

For notational simplicity, let $S_t(h)$ denote the skewness of $Z_t(h)$. When we wish to emphasize the dependence of $S_t(h)$ on any of the parameters of the model, we will write $S_t(h;\rho)$, etc.

Proposition 4.2 *$S_t(h)$ has the following properties:*

1. *$S_t(h)$ is positive if $\rho > 0$, zero if $\rho = 0$, and negative if $\rho < 0$.*

2. *$S_t(h;\rho) = -S_t(h;-\rho)$ for all ρ.*

3. $\lim_{h\downarrow 0} S_t(h) = \lim_{h\uparrow\infty} S_t(h) = 0$.

Proof See Appendix C. □

Properties 1 and 3 of Proposition 4.2 imply that when $\rho \neq 0$, skewness *cannot* be a monotone function of h as it was in the case of jump-diffusions. So how does skewness behave with changing h? To see the answer, consider the derivative $S_t'(h)$. Letting $a = v/\theta$, it can be seen by direct calculation that h satisfies $S_t'(h) = 0$ if and only if $y = \kappa h$ satisfies

$$\begin{aligned}(1-a)(4-3a) + (1-a)^2 y - (8 - 8a + 3a^2)e^y & \\ - 2a(2-a)ye^y + (4-a)e^{2y} - ye^{2y} - 2(1-a)y^2 e^y &= 0.\end{aligned} \tag{25}$$

Equation (25) has a unique solution for each value of a. (Solutions for select values of a are given in Table 1 below.) Combined with Properties 1 and 3, this uniqueness implies that

- If $\rho > 0$, skewness increases from zero to a maximum, and then decreases asymptotically back to zero.

- If $\rho < 0$, skewness decreases from zero to a minimum, and then increases asymptotically back to zero.

Figures 2–5 illustrate this behavior of conditional skewness for various parameter configurations.

From a practical standpoint, this raises the obvious question: what of the value h at which skewness is maximized (minimized if $\rho < 0$)? This question is of especial importance in deriving criteria to distinguish between jump-diffusions and stochastic volatility models. To wit, if the maximizing value h were to turn out to be very small, the increasing portion of

Table 1: Solutions of Equation (25) as a varies

This table presents the values of y that solve (25) for a range of values of a.

a	0.33	0.50	1.00	2.00	3.00
y	1.41	1.66	2.15	2.73	3.11

Table 2: Values of h that Maximize Conditional Skewness

This table presents the values of h that maximize $S_t(h)$ for a given range of values of a and κ.

κ	a	h (years)	κ	a	h (years)
0.10	3.00	31.11	5.00	3.00	0.62
0.10	2.00	27.31	5.00	2.00	0.55
0.10	1.00	21.49	5.00	1.00	0.43
0.10	0.50	16.57	5.00	0.50	0.33
0.10	0.33	14.06	5.00	0.33	0.28
1.00	3.00	3.11	50.00	3.00	0.06
1.00	2.00	2.73	50.00	2.00	0.05
1.00	1.00	2.15	50.00	1.00	0.04
1.00	0.50	1.66	50.00	0.50	0.03
1.00	0.33	1.41	50.00	0.33	0.03

the skewness curve would become empirically irrelevant, and the conditional skewness curve under stochastic volatility would essentially resemble the one under jump-diffusions. Now, it is immediate from the discussion in the last paragraph that the value of h that maximizes $S_t(h)$ is precisely the value of h for which $y = \kappa h$ satisfies (25). In particular, therefore, the maximizing value of h depends only on two parameters, κ which is the rate of mean-reversion in the volatility process, and a, which is the ratio of current volatility v to long-term mean volatility θ. Table 2 summarizes the way in which this maximizing value changes with κ and a.

Plausible values for κ (i.e., ones typically found empirically) are in a neighborhood of unity. If the current value of the volatility v is also not too far off its long-term mean level θ (i.e., if a and a^{-1} are both not very large), then it is seen from Table 2 that the maximizing h is of the order of several months, and sometimes even years. From an empirical standpoint therefore, the increasing part of $S_t(h)$ is very relevant, creating a sharp contrast with conditional skewness in jump-diffusion models.

4.2 Kurtosis of the Conditional Returns

For notational simplicity, let $K_t(h)$ denote the kurtosis of $Z_t(h)$. As with skewness, when we wish to emphasize the dependence of kurtosis on a particular parameter, we shall write $K_t(h;\rho)$, etc.

Proposition 4.3 *$K_t(h)$ has the following properties:*

1. *Excess kurtosis is strictly positive everwhere: $K_t(h) > 3$ for all $h > 0$.*
2. *The degree of kurtosis only depends on the absolute value of ρ: $K_t(h;\rho) = K_t(h;-\rho)$.*
3. $\lim_{h\downarrow 0} K_t(h) = \lim_{h\uparrow\infty} K_t(h) = 0$.

Proof See Appendix D. □

Property 3 implies, once again, that kurtosis cannot be a monotone function of h as it was for jump-diffusions. Indeed, exactly the same procedure as used in the case of skewness above shows that

- Kurtosis is tent-shaped as a function of h, increasing from zero to a maximum, and then decreasing asymptotically back to zero again.

We omit the details here, since the expressions are significantly more complicated than in the earlier case. Figures 2–5 illustrate this behavior of $K(h)$.

Table 3 shows how the value of h that maximizes $K_t(h)$ depends on the three relevant parameters, κ, ρ, and $a = v/\theta$. It can be seen that the maximizing value of h increases with ρ and a, and decreases with κ. More importantly, as with skewness, it is again the case that for plausible parameter values, this maximizing value of h is quite large (of the order of several months, and even years). Thus, the increasing portion of the kurtosis curve is significant from an empirical standpoint, creating a nice contrast with the behavior of conditional kurtosis in jump-diffusions.

Table 3: Values of h that Maximize Conditional Kurtosis

This table presents the values of h that maximize $K_t(h)$ for a range of values of the parameters κ, ρ, and $a = v/\theta$.

κ	a	ρ	h (years)	κ	a	ρ	h (years)
0.10	2.00	0.00	24.92	5.00	2.00	0.00	0.4984
0.10	2.00	0.25	25.47	5.00	2.00	0.25	0.5095
0.10	2.00	0.50	26.13	5.00	2.00	0.50	0.5226
0.10	2.00	0.75	26.47	5.00	2.00	0.75	0.5293
0.10	1.00	0.00	18.93	5.00	1.00	0.00	0.3785
0.10	1.00	0.25	19.18	5.00	1.00	0.25	0.3826
0.10	1.00	0.50	19.52	5.00	1.00	0.50	0.3904
0.10	1.00	0.75	19.72	5.00	1.00	0.75	0.3944
0.10	0.50	0.00	13.94	5.00	0.50	0.00	0.2788
0.10	0.50	0.25	13.94	5.00	0.50	0.25	0.2788
0.10	0.50	0.50	13.94	5.00	0.50	0.50	0.2788
0.10	0.50	0.75	13.94	5.00	0.50	0.75	0.2788
1.00	2.00	0.00	2.49	50.00	2.00	0.00	0.0498
1.00	2.00	0.25	2.55	50.00	2.00	0.25	0.0510
1.00	2.00	0.50	2.61	50.00	2.00	0.50	0.0523
1.00	2.00	0.75	2.65	50.00	2.00	0.75	0.0529
1.00	1.00	0.00	1.89	50.00	1.00	0.00	0.0379
1.00	1.00	0.25	1.91	50.00	1.00	0.25	0.0383
1.00	1.00	0.50	1.95	50.00	1.00	0.50	0.0390
1.00	1.00	0.75	1.97	50.00	1.00	0.75	0.0390
1.00	0.50	0.00	1.39	50.00	0.50	0.00	0.0279
1.00	0.50	0.25	1.39	50.00	0.50	0.25	0.0279
1.00	0.50	0.50	1.39	50.00	0.50	0.50	0.0279
1.00	0.50	0.75	1.39	50.00	0.50	0.75	0.0279

5 Unconditional Returns in a Stochastic Volatility Process

In the previous section, we showed that conditional on $V_t = v$, the characteristic function $F^*(v, h, s)$ of $Z_t(h)$ has the form

$$D(h, s) \exp\{vB(h, s)\}, \tag{26}$$

where D is defined from the functions A and C of Proposition 4.1 by

$$D(h, s) = C(h, s) \exp\{A(h, s)\}.$$

Now, it is well known that the square-root process (14) followed by volatility has a stationary density that is gamma, and is given by

$$g(v) = \frac{1}{\Gamma(\nu)} \omega^{\nu} v^{\nu-1} e^{-\omega v} \tag{27}$$

where

$$\omega = \frac{2\kappa}{\eta^2} \quad \text{and} \quad \nu = \frac{2\kappa\theta}{\eta^2}. \tag{28}$$

Combining (26) and (27), the characteristic function of the *unconditional* returns $Z(h)$, denoted say $G(h, s)$, is given by

$$G(h, s) = \int F^*(v, h, s) g(v) dv = D \int \exp\{vB\} g(v) dv. \tag{29}$$

But this last integral on the right-hand side is simply the moment generating function $E(e^{tv})$ of the gamma distribution evaluated at $t = B$. It is well-known that this moment-generating function has the form

$$\left(\frac{\omega}{\omega - B}\right)^{\nu}.$$

Therefore, the characteristic function $G(h, s)$ of the unconditional returns has the form

$$G(h, s) = D(h, s) \left(\frac{\omega}{\omega - B(h, s)}\right)^{\nu}. \tag{30}$$

Successive differentiation of the function G now delivers all the moments of the unconditional returns. Letting $y = \kappa h$, working through the calculations yields the following:

$$\text{Skewness}\,(Z(h)) = 3\left(\frac{\rho\eta}{\sqrt{\kappa\theta}}\right)\left[\frac{1 - e^y + ye^y}{y^{3/2}e^y}\right] \tag{31}$$

$$\text{Kurtosis}\,(Z(h)) = 3\left[1 + \frac{\eta^2}{\kappa\theta y^2 e^y}\left(1 - e^y + ye^y + 4\rho^2[2 - 2e^y + y + ye^y]\right)\right]. \tag{32}$$

Expressions (31) and (32) may now be used to derive properties of unconditional skewness and kurtosis in the stochastic volatility model. We begin with unconditional skewness.

5.1 Skewness of the Unconditional Returns

For notational simplicity, let $S(h)$ denote the skewness of the unconditional returns $Z(h)$. As earlier, when we wish to emphasize the dependence of $S(h)$ on a parameter, we will write $S(h;\rho)$, etc.

Proposition 5.1 *$S(h)$ has the following properties:*

1. *$S(h)$ is positive if $\rho > 0$, zero if $\rho = 0$, and negative if $\rho < 0$.*
2. *$S(h;\rho) = -S(h;-\rho)$.*
3. *$\lim_{h\downarrow 0} S(h) = \lim_{h\uparrow 0} S(h) = 0$.*

Proof See Appendix E. □

Properties 1 and 3 of Proposition 5.1 show that unconditional skewness cannot be monotone in h if $\rho \neq 0$. The natural question, therefore, is: how does skewness behave for $h > 0$? The key, once again, lies in the behavior of the derivative $S'(h)$. A simple computation shows that when $\rho \neq 0$, it is the case that $S'(h) = 0$ for some $h > 0$ if and only if $y = \kappa h$ satisfies

$$3e^y - ye^y - 2y - 3 = 0. \tag{33}$$

There is only one non-trivial (i.e., non-zero) solution to (33), which is given approximately by $y = 2.15$. The uniqueness of the solution implies that

- If $\rho > 0$, then skewness increases from zero to a maximum and then decreases asymptotically to zero. The maximum occurs approximately at $\kappa h = 2.15$.

- If $\rho < 0$, then skewness decreases from zero to a minimum, and then increases aymptotically to zero. The minimum occurs approximately at $\kappa h = 2.15$.

A remarkable feature of these results, that deserves emphasis, is that the point where skewness reaches its extremum values depends on only a single parameter of the model, namely κ. The only part played by ρ is in determining whether this extremum is a maximum (if $\rho > 0$) or a minimum (if $\rho < 0$).

As mentioned above, reasonable values of κ (i.e., those typically found empirically) are in a neighborhood of unity. Even for a value such as $\kappa = 5$, however, the value of h that maximizes unconditional skewness is of the order of $2.15/5 = 0.43$ years, or almost five months. Therefore, for a substantial interval of values of h, unconditional skewness in the stochastic volatility model is increasing in h, creating a sharp contrast with the behavior of unconditional skewness in jump-diffusions.

5.2 Kurtosis of the Unconditional Returns

Let $K(h)$ denote the kurtosis of the conditional returns given h. When we wish to emphasize the dependence of $K(h)$ on the other parameters, we will write, for example, $K(h;\rho)$.

Proposition 5.2 *$K(h)$ has the following properties:*

1. *Excess kurtosis is strictly positive everywhere: $K(h) > 0$ at all $h > 0$.*

2. *Excess kurtosis only depends on the absolute value of ρ: $K(h;\rho) = K(h;-\rho)$.*

3. $\lim_{h\downarrow 0} K(h) = 3(1+b)$ *where $b = \eta^2/2\kappa\theta$, and* $\lim_{h\uparrow\infty} K(h) = 0$.

Remark Note that, in contrast to the behavior of conditional excess kurtosis, excess kurtosis of the unconditional distribution does not approach zero as $h \to 0$. We are unable to find an intuitive explanation for this behavior. □

Proof See Appendix F. □

Since excess kurtosis has a strictly positive limit as $h \to 0$, but goes to zero as $h \to \infty$, Proposition 5.2 does *not* rule out the possibility that excess kurtosis could be monotone in

h. To analyze further the behavior of kurtosis as a function of h, we examine its deriative $K'(h)$. Letting $y = \kappa h$, this derivative is seen to be

$$\frac{3\eta^2}{\theta}\left[\frac{(2e^y - 2 - y - ye^y) + 4\rho^2(4e^y - 4 - 3y - y^2 - ye^y)}{y^3 e^y}\right]. \tag{34}$$

In anaylyzing this expression, it is helpful to first consider the polar cases $\rho = 0$ and $\rho = \pm 1$. We begin with the case $\rho = 0$. In this case, (34) reduces to

$$\frac{3\eta^2}{\theta}\left(\frac{2e^y - 2 - y - ye^y}{y^3 e^y}\right). \tag{35}$$

It is easily shown that the term in parenthesis on the right-hand side of (35) is strictly negative at all $y > 0$.[10] As a consequence, $K'(h) < 0$ at all $h > 0$, so *excess kurtosis is strictly decreasing for* $h > 0$.

At the other polar extreme is the case $\rho = \pm 1$. In this case, (34) becomes

$$\frac{3\eta^2}{\theta}\left(\frac{18e^y - 18 - 13y - 5ye^y - 4y^2}{y^3 e^y}\right). \tag{36}$$

It can be shown that there is a unique value of $y > 0$ (given approximately by $y = 1.83$) at which $(18e^y - 18 - 13y - 5ye^y - 4y^2) = 0$. Since $y = \kappa h$, it is the case that excess kurtosis is now strictly increasing in h for $h < 1.83/\kappa$, and is strictly decreasing in h for $h > 1.83/\kappa$.

As is apparent from the foregoing, the behavior of excess kurtosis depends on the value of ρ. The nature of this dependence can be understood by considering the full form (34) of the derivative $K'(h)$. The sign of this derivative at $h > 0$ is the same as the sign of the numerator of the term in parenthesis. Denoting this numerator by $g(y)$, we may rewrite g as

$$g(y) = 2(1 + 8\rho^2)e^y - 2(1 + 8\rho^2) - (1 + 12\rho^2)y - (1 + 4\rho^2)ye^y - 4\rho^2 y^2. \tag{37}$$

Some tedious calculation shows that

$$g(0) = g'(0) = g''(0) = 0, \tag{38}$$

[10] To see this, note that the denominator $y^3 e^y$ is strictly positive for $y > 0$. It suffices, therefore, to show that the function $f(y) = 2e^y - 2 - y - ye^y$ is negative at all $y > 0$. A few simple calculations show that $f(0) = f'(0) = 0$ and that $f''(y) = -ye^y < 0$ for all y ¿ 0. It follows immediately that $f(y)$ is strictly positive, and even strictly increasing, for $y > 0$.

and that

$$g'''(y) = (4\rho^2 - 1)e^y - (1 + 4\rho^2)ye^y. \tag{39}$$

If $4\rho^2 < 1$ (equivalently, if $|\rho| < \frac{1}{2}$), then both terms on the RHS of (39) are negative for all $y > 0$, so $g'''(y) < 0$ for all $y > 0$. Combined with (38), this implies $g(y)$ is negative at all $y > 0$. Since the other terms in (34) are all strictly positive at $y > 0$, it follows that (34) is negative at all $y > 0$. We have shown, therefore, that *for* $|\rho| < \frac{1}{2}$, *excess kurtosis is itself strictly decreasing for all* $y > 0$.

To sum up the behavior of the kurtosis of unconditional returns:

- If $|\rho| < \frac{1}{2}$, then excess kurtosis is a strictly decreasing frunction of h.
- If $|\rho|$ is sufficiently large, however, then excess kurtosis is tent-shaped in h, increasing to a maximum and then decreasing asymptotically to zero.

For small enough values of $|\rho|$, then, there is no qualitiative difference between the behavior of unconditional kurtosis in jump-diffusion models and that in stochastic volatility models. Note, however, that unconditional skewness in jump-diffusions and stochastic volatility models always exhibit different patterns of behavior, provided only skewness is non-zero, i.e., $\rho \neq 0$. Thus, except for the knife-edge case of $\rho = 0$, it is possible to tell apart the two classes of models using the pattern of unconditional skewness, and maybe also the pattern of unconditional kurtosis.

6 Conclusion

This paper explores the features of skewness and kurtosis when continuous time jump diffusion and stochastic volatility models are used for asset pricing. Determining the "true" return generating process for financial assets with discrete data is complicated by the aliasing problem: it may not possible to distinguish between the alternative classes of models if data at a fixed interval is used.

We show in this paper that using data at different intervals provides one way for generating criteria to distinguish between jump diffusions and stochastic volatility models. Data may be examined either unconditionally by simply looking at the time series of asset prices, or conditionally from options prices. We show that:

- With conditional data, skewness and kurtosis of a stochastic volatility model are each increasing as a function of the sampling interval h for at least small to moderate values of h; in a jump-diffusion, each is a strictly decreasing function of h.

- When using unconditional data, the pattern of skewness differs across the two models in exactly the same way it did for conditional data. However, unconditional kurtosis could have the same pattern in both models, and may not be useful as a discriminating device.

Finally, the moments derived in this paper may be employed in method of moments estimation procedures where the moments may exploit data at different intervals. To the best of our knowledge, this has not been undertaken so far in the literature.

A Proof of Proposition 3.1

We will begin by proving the following result on the moments of $Z_t(h)$. The result is stated in the form of a lemma for ease of future reference.

Lemma A.1 *In the jump-diffusion model, the first four moments of $Z_t(h)$ are given by the following expressions:*

$$E[Z_t(h)] = h[\alpha + \nu_1\lambda] \tag{40}$$

$$E\left[(Z_t(h))^2\right] = h[\sigma^2 + \nu_2\lambda] + h^2[\alpha + \nu_1\lambda]^2 \tag{41}$$

$$E\left[(Z_t(h))^3\right] = h[\nu_3\lambda] + 3h^2[\alpha + \nu_1\lambda][\sigma^2 + \nu_2\lambda] + h^3[\alpha + \nu_1\lambda]^3 \tag{42}$$

$$E\left[(Z_t(h))^4\right] = \left\{ \begin{array}{l} h[\nu_4\lambda] + h^2[3(\sigma^2 + \nu_2\lambda)^2 + 4\nu_3\lambda(\alpha + \nu_1\lambda)] \\ +6h^3[\alpha + \nu_1\lambda]^2[\sigma^2 + \nu_2\lambda] + h^4[\alpha + \nu_1\lambda]^4. \end{array} \right. \tag{43}$$

Proof We use the method of direct calculation. Some notational simplification is useful in this process. First, define the sequnce (w_k) by $w_0 = 0$ and for $k = 1, 2, 3, \ldots$,

$$w_k = y_1 + \cdots + y_k.$$

Note that in this notation, we can write the returns $Z_t(h)$ over the period $[t, t+h]$ as

$$Z_t(h) = x + w_K,$$

where, of course, $x \sim N(\alpha h, \sigma^2 h)$. Secondly, let p_k denote the probability that $K = k$:

$$p_k = \frac{e^{-\lambda h}(\lambda h)^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

Finally, throughout this proof, we will suppress dependence of the holding-period returns on t and h, and simply write Z for $Z_t(h)$.

Two preliminary observations are important. First, since K follows a Poisson distribution with parameter λh, we have the following expressions for the moments of K:

$$E[K] = \lambda h \tag{44}$$

$$E[K^2] = \lambda h + (\lambda h)^2 \tag{45}$$

$$E[K^3] = \lambda h + 3(\lambda h)^2 + (\lambda h)^3 \tag{46}$$

$$E[K^4] = \lambda h + 7(\lambda h)^2 + 6(\lambda h)^3 + (\lambda h)^4. \tag{47}$$

Secondly, for any fixed value of $k \geq 1$, the independence of the variables $y_1, y_2, \ldots$ implies[11] the following expressions for the moments of w_k:[12]

$$E[w_k] = k\nu_1 \tag{48}$$

$$E[w_k^2] = k[\nu_2 - \nu_1^2] + k^2\nu_1^2 \tag{49}$$

$$E[w_k^3] = \begin{cases} \nu_3, & \text{if } k = 1 \\ k[\nu_3 - 3\nu_2\nu_1 + 2\nu_1^3] + 3k^2[\nu_2\nu_1 - \nu_1^3] + k^3\nu_1^3, & \text{if } k \geq 2 \end{cases} \tag{50}$$

$$E[w_k^4] = \begin{cases} \nu_4, & \text{if } k = 1 \\ 2\nu_4 + 8\nu_3\nu_1 + 6\nu_2^2, & \text{if } k = 2 \\ k[\nu_4 - 4\nu_3\nu_1 - 3\nu_2^2 + 12\nu_2\nu_1^2 - 6\nu_1^4] & \\ +k^2[4\nu_3\nu_1 + 3\nu_2^2 - 18\nu_2\nu_1^2 + 11\nu_1^4] & \\ +6k^3[\nu_2\nu_1^2 - \nu_1^4] + k^4\nu_1^4, & \text{if } k \geq 3 \end{cases} \tag{51}$$

Now, note that for $j = 1, 2, 3, 4$, it is the case that

$$E[w_K^j] = \sum_{k=0}^{\infty}(p_k w_k^j). \tag{52}$$

Using (52) and the definition of p_k, it can be shown that

$$E[w_K] = h[\nu_1\lambda]. \tag{53}$$

$$E[w_K^2] = h[\nu_2\lambda] + h^2[\nu_1^2\lambda^2]. \tag{54}$$

$$E[w_K^3] = h[\nu_3\lambda] + h^2[3\nu_1\nu_2\lambda^2] + h^3[\nu_1^3\lambda^3]. \tag{55}$$

$$E[w_K^4] = h[\nu_4\lambda] + h^2[4\nu_1\nu_3\lambda^2 + 3\nu_2^2\lambda^2] + h^3[6\nu_1^2\nu_2\lambda^3] + h^4[\nu_1^4\lambda^4]. \tag{56}$$

Now combining (44)—(56), we finally obtain:

$$\begin{aligned} E[Z] &= E[x + w_K] \\ &= E[x] + E[w_K] \\ &= \alpha h + \nu_1 \cdot (\lambda h) \\ &= (\alpha + \nu_1\lambda)h. \end{aligned}$$

[11]Details omitted here and elsewhere in the proof are available on request from the authors.

[12]All moments of w_0 are, of course, equal to zero.

$$
\begin{aligned}
E[Z^2] &= E[(x+w_K)^2] \\
&= E[x^2 + 2xw_K + w_K^2] \\
&= E[x^2] + 2E[x]E[w_K] + E[w_K^2] \\
&= [\sigma^2 h + \alpha^2 h^2] + 2[\alpha h \nu_1 \lambda h] + [(\nu_2 - \nu_1^2)\lambda h + \nu_1^2(\lambda h + (\lambda h)^2)] \\
&= h[\sigma^2 + \nu_2\lambda] + h^2[\alpha + \nu_1\lambda]^2
\end{aligned}
$$

$$
\begin{aligned}
E[Z^3] &= E[(x+w_K)^3] \\
&= E[x^3 + 3x^2 w_K + 3xw_K^2 + w_K^3] \\
&= E[x^3] + 3E[x^2]E[w_K] + 3E[x]E[w_K^2] + E[w_K^3] \\
&= \begin{cases} [\alpha^3 h^3 + 3\alpha\sigma^2 h^2] + 3[\alpha^2 \nu_1 \lambda h^3 + \sigma^2 \nu_1 \lambda h^2] \\ +3[\alpha\nu_1^2\lambda^2 h^3 + \alpha\nu_2\lambda h^2] + [\nu_1^3\lambda^3 h^3 + 3\nu_1\nu_2\lambda^2 h^2 + \nu_3\lambda h] \end{cases} \\
&= h[\lambda\nu_3] + 3h^2[\alpha + \nu_1\lambda][\sigma^2 + \nu_2\lambda] + h^3[\alpha + \nu_1\lambda]^3
\end{aligned}
$$

$$
\begin{aligned}
E[Z^4] &= E[(x+w_K)^4] \\
&= E[x^4 + 4x^3 w_K + 6x^2 w_K^2 + 4xw_K^3 + w_K^4] \\
&= E[x^4] + 4E[x^3]E[w_K] + 6E[x^2]E[w_K^2] + 4E[x]E[w_K^3] + E[w_K^4] \\
&= \begin{cases} [3\sigma^4 h^2 + 6\alpha^2\sigma^2 h^3 + \alpha^4 h^4] \\ +4[3\alpha\sigma^2\nu_1\lambda h^3 + \alpha^3\nu_1\lambda h^4] \\ +6[\sigma^2\nu_2\lambda h^2 + \sigma^2\nu_1^2\lambda^2 h^3 + \alpha^2\nu_2\lambda h^3 + \alpha^2\nu_1^2\lambda^2 h^4] \\ +4[\alpha\nu_3\lambda h^2 + 3\alpha\nu_1\nu_2\lambda^2 h^3 + \alpha\nu_1^3\lambda^3 h^4] \\ +[\lambda\nu_4 h + 4\nu_1\nu_3\lambda^2 h^2 + 3\nu_2^2\lambda^2 h^2 + 6\nu_1^2\nu_2\lambda^3 h^3 + \nu_1^4\lambda^4 h^4] \end{cases} \\
&= \begin{cases} h[\lambda\nu_4] + h^2[3(\sigma^2 + \nu_2\lambda)^2 + 4\nu_3\lambda(\alpha + \nu_1\lambda)] \\ +6h^3[\alpha + \nu_1\lambda]^2[\sigma^2 + \nu_2\lambda] + h^4[\alpha + \nu_1\lambda]^4 \end{cases}
\end{aligned}
$$

This completes the proof of the lemma. □

Direct calculation now shows that the central moments of Z are given by:

$$\text{Variance}[Z] = E[(Z - E[Z])^2] = h[\sigma^2 + \nu_2\lambda].$$

$$\text{Skewness}\,[Z] = \frac{E[(Z-E[Z])^3]}{(\text{Var}[Z])^{3/2}} = \frac{1}{\sqrt{h}}\left[\frac{\lambda\nu_3}{(\sigma^2+\nu_2\lambda)^{3/2}}\right].$$

$$\text{Kurtosis}\,[Z] = \frac{E[(Z-E[Z])^4]}{(\text{Var}[Z])^2} = 3+\frac{1}{h}\left[\frac{\lambda\nu_4}{(\sigma^2+\nu_2\lambda)^2}\right].$$

Proposition 3.1 is proved. □

B Proof of Proposition 4.1

We guess a solution for the characteristic function of the following form:

$$F(x,v,h;s) = \exp[isx + D(h) + vB(h)].$$

Taking derivatives of this guess and substituting them into equation (16), we obtain:

$$\alpha is - \frac{1}{2}vs^2 + \kappa(\theta - v)B + \frac{1}{2}\eta^2B^2v - D_h - vB_h + \rho\eta isBv = 0$$

where the subscripts denote derivatives. Collecting terms in v the equation is rewritten as:

$$\begin{aligned} 0 &= v\left[-\frac{1}{2}s^2 - \kappa B + \frac{1}{2}\eta^2B^2 - B_h + \rho\eta isB\right] \\ &\quad + [\alpha is + \kappa\theta B - D_h] \end{aligned} \tag{57}$$

which is solved subject to the condition that $F(x,v,h,s) = e^{isx}$. Each line of the equation above provides an ordinary differential equation which is solved separately. The initial conditions devolve naturally for each ODE from the initial condition for the PDE. The first ODE is solved subject to the initial condition $B(0) = 0$, and the second ODE is solved subject to the condition $D(0) = 0$. The solution is as given in equation (18), where $D(h) = A(h) + \ln[C(h)]$. The proof is available on request. An analogous derivation may be referenced in Heston (1993) as well. □

C Proof of Proposition 4.2

Property 1 of Proposition 4.2 will be established if we can show that the last term in parenthesis on the RHS of (20) is positive. We will first show that the denominator is positive, and then that the numerator is positive.

To show that the denominator is positive, it suffices to show that

$$\theta(1 - e^y + ye^y) + v(e^y - 1)] > 0 \quad \text{for all } y > 0. \tag{58}$$

Consider first the term $1 - e^y + ye^y$. At $y = 0$, this term is equal to zero. Its derivative at any $y > 0$ is given by the strictly positive quantity ye^y. Therefore, this term is strictly positive at all $y > 0$. The term $(e^y - 1)$ is obviously also strictly positive at all $y > 0$. Since θ and v are positive, (58) is established.

Turning to the numerator, it evidently suffices to show that that all $y > 0$, we have

$$2 - 2e^y + y + ye^y > 0 \quad \text{and} \quad 1 + y - e^y < 0. \tag{59}$$

For notational ease, let $h(y) = 2 - 2e^y + y + ye^y$. It is easily checked that $h(0) = h'(0) = 0$, and that $h''(y) = ye^y > 0$ for all $y > 0$. It follows immediately that $h(y)$ is positive for all $y > 0$, establishing the first part of (59). The other part is trivial since, by definition, $e^y = \sum_{n=0}^{\infty}(y^n/n!) > 1 + y$ if $y > 0$. This completes the proof of Property 1.

Property 2 is immediate from the definition of $S(h)$ in (20). Finally, to see property 3, let

$$g(y) = e^{\frac{1}{2}y}\left[\frac{\theta(2 - 2e^y + y + ye^y) - v(1 + y - e^y)}{(\theta(1 - e^y + ye^y) + v(e^y - 1))^{3/2}}\right].$$

The expression on the right-hand side of this equation has the form $0/0$ at $y = 0$, and ∞/∞ as $y \uparrow \infty$. However, using L'Hopital's rule repeatedly, it can be seen that $\lim_{y\downarrow 0} g(y) = 0$ and that $\lim_{y\uparrow\infty} g(y) = 0$. From (20) it is immediate that

$$\lim_{h\downarrow 0} S(h) = \frac{3\eta\rho}{\sqrt{k}} \cdot \lim_{y\downarrow 0} g(y) \quad \& \quad \lim_{h\uparrow\infty} S(h) = \frac{3\eta\rho}{\sqrt{k}} \cdot \lim_{y\uparrow\infty} g(y).$$

It follows that $S(h)$ goes to zero as $h \downarrow 0$ and $h \uparrow \infty$, completing the proof of the proposition. □

D Proof of Proposition 4.3

To prove Property 1, it suffices to show that $A_1 > 0$ and $A_2 < 0$. Now, A_1 is of the form $f(y) + 4\rho^2 g(y)$, where

$$f(y) = 1 + 4e^y - 5e^{2y} + 4ye^y + 2ye^{2y},$$

and

$$g(y) = 6e^y - 6e^{2y} + 4ye^y + 2ye^{2y} + y^2e^y.$$

A simple computation shows that $f(0) = 0$ and $f'(y) = 4e^y(2 - 2e^y + y + ye^y)$. The term in parenthesis is strictly positive for $y > 0$, as we showed in the proof of Proposition 4.2 (see equation (59). Therefore, $f'(y) > 0$ for all $y > 0$, and so $f(y) > 0$ for all $y > 0$.

The argument for $g(y)$ is similar. It can be checked that $g(0) = g'(0) = g''(0) = g'''(0) = 0$, and that $g^{(4)}(y) > 0$ for all $y > 0$. Therefore, $g(y) > 0$ for all $y > 0$.

Since f and g are both strictly positive functions, A_1 is strictly positive. That $A_2 < 0$ is checked similarly. The details are omitted. This proves Property 1.

Property 2 is immediate from (21) since $K(h)$ only depends on ρ^2. This leaves Property 3. Let

$$h(y) = [\theta(1 - e^y + ye^y) + v(e^y - 1)]^2.$$

Then, for $y = \kappa h$, we have

$$K(h) \;=\; 3\left[1 + \frac{\eta^2}{2k}\left(\frac{\theta A_1 - vA_2}{h(y)}\right)\right].$$

Now, $A_1/h(y) = [f(y)/h(y)] + 4\rho^2[g(y)/h(y)]$, where f and g were defined above. Repeated use of L'Hopital's rule shows that the limits as $y \downarrow 0$ and $y \uparrow \infty$ of $[f(y)/h(y)]$ and $[g(y)/h(y)]$ are all zero. Thus, $A_1/h(y)$ tends to zero as y tends to zero or becomes unbounded. A similar argument shows that $A_2/h(y)$ also tends to zero as y tends to zero or to infinity. Since κ is fixed, $y \downarrow 0$ if and only if $h \downarrow 0$, and $y \uparrow \infty$ only if $h \uparrow \infty$, so we finally obtain

$$\lim_{h \downarrow 0} K(h) = \lim_{h \uparrow \infty} K(h) = 0.$$

This completes the proof of the proposition. □

E Proof of Proposition 5.1

For ease of reference, we reproduce the form of skewness here (recall that $y = \kappa h$):

$$3\left(\frac{\rho\eta}{\sqrt{\kappa\theta}}\right)\left[\frac{1 - e^y + ye^y}{y^{3/2}e^y}\right]. \tag{60}$$

Property 1 of Proposition 5.1 will be proved if we can show that the term $(1 - e^y + ye^y)$ is strictly positive for all $y > 0$. But we have already done this in the proof of Proposition 4.2 above (see equation (58)).

Property 2 of Proposition 5.1 is immediate from the form of skewness (60). This leaves Property 3. Now, the limits as $h \downarrow 0$ and $h \uparrow \infty$ of $S(h)$ are clearly determined entirely by the limits as $y \downarrow 0$ and $y \uparrow \infty$ of the function $h(y)$ defined by

$$h(y) \;=\; \frac{1 - e^y + ye^y}{y^{3/2}e^y}. \tag{61}$$

Indeed, we need only show that $h(y)$ tends to zero as y tends to 0 or $+\infty$.

As $y \downarrow 0$, both the numerator and the denominator of the right-hand side of (61) approach zero. However, repeated application of L'Hopital's rule shows that $\lim_{y\downarrow 0} h(y) = 0$, establishing one part of the desired result.

To see the other part, note that $h(y)$ can be rewritten as

$$h(y) = \frac{1}{y^{3/2}e^y} - \frac{1}{y^{3/2}} + \frac{1}{y^{1/2}}. \tag{62}$$

Since each term on the right-hand side of (62) tends to zero as $y \uparrow \infty$, it is clearly the case that $h(y)$ also goes to zero as $y \uparrow \infty$. This completes the proof of Proposition 5.1

F Proof of Proposition 5.2

For notational ease, define

$$g(y) \;=\; \frac{(1+8\rho^2) - (1+8\rho^2)e^y + (1+4\rho^2)ye^y + 4\rho^2 y}{y^2e^y}.$$

Then, we have

$$K(h) \;=\; 3\left[1 + \left(\frac{\eta^2}{\kappa\theta}\right)g(\kappa h)\right].$$

To prove Property 1 now, it suffices to show that $g(y) > 0$ at all $y > 0$. A simple calculation shows that we have

$$g(0) = g'(0) = 0,$$

and that

$$g''(y) = e^y + (1 + 4\rho^2)ye^y > 0 \quad \text{for all } y > 0.$$

It follows that $g(y)$ is strictly positive (as is $g'(y)$) for $y > 0$, proving Property 1.

Property 2 is an immediate consequence of the fact that $K(h)$ only depends on ρ^2. Finally, to see Property 3, note that

$$\lim_{h \downarrow 0} K(h) = 3\left[1 + \frac{\eta^2}{\kappa\theta}(\lim_{y \downarrow 0} g(y))\right], \tag{63}$$

and

$$\lim_{h \uparrow \infty} K(h) = 3\left[1 + \frac{\eta^2}{\kappa\theta}(\lim_{y \uparrow \infty} g(y))\right]. \tag{64}$$

Repeated applications of L'Hopital's rule to g establishes that

$$\lim_{y \downarrow 0} g(y) = \frac{1}{2} \quad \text{and} \quad \lim_{y \uparrow \infty} g(y) = 0. \tag{65}$$

Substituting (65) into (63) and (64) compeltes the proof of Property 3. □

References

[1] Ahn, C.M. (1992) Option Pricing when Jump Risk is Systematic, *Mathematical Finance* 2(4), 299-308.

[2] Amin, K.I. (1993) Jump Diffusion Option Valuation in Discrete Time, *Journal of Finance* 48(5), 1833–1863.

[3] Amin, K.I. and V. Ng (1993) Option Valuation with Systematic Stochastic Volatility, *Journal of Finance* 48(3), 881–910.

[4] Ball, C. and W. Torous (1985) On Jumps in Common Stock Prices and Their Impact on Call Option Pricing, *Journal of Finance* 40, 155–173.

[5] Bates, D.S. (1996) Jumps and Stocahstic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, *Review of Financial Studies* 9(1), 69–107.

[6] Black, F. and M. Scholes (1973) The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81(3), 637–654.

[7] Bodurtha, J. and G. Courtadon (1987) Tests of the American Option Pricing Model in the Foreign Currency Options Market, *Journal of Financial and Quantitative Analysis* 22, 153–167.

[8] Bollerslev, T., R.Y. Chou, and K.F. Kroner (1992) ARCH Modelling in Finance, *Journal of Econometrics* 52, 5–59.

[9] Das, S.R. and S. Foresi (1996) Exact Prices for Bond and Option Prices with Systematic Jump Risk, *Review of Derivatives Research* 1(1), 7-24.

[10] Duffie, D. (1996) *Dynamic Asset Pricing Theory*, Second Edition, Princeton University Press, Princeton, NJ.

[11] Heston, S.L. (1993) A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6(2), 327–343.

[12] Jarrow, R. and E. Rosenfeld (1984) Jump Risks and the Intertemporal Capital Asset Pricing Model, *Journal of Business* 57, 337–351.

[13] Jorion, P. (1988) On Jump Processes in the Foreign Exchange and Stock Markets, *Review of Financial Studies* 1(4), 427–445.

[14] Karlin, H. and S. Taylor (1975) *A Second Course in Stochastic Processes* (Second Edition), Academic Press, San Diego, CA.

[15] Melino, A. and S.M. Turnbull (1990) Pricing Foreign Currency Options with Stochastic Volatility, *Journal of Econometrics* 45, 239–265.

[16] Merton, R.C. (1976) Option Pricing when the Underlying Process for Stock Returns is Discontinuous, *Journal of Financial Economics* 3, 124–144.

[17] Nelson, D. (1990) ARCH models as Diffusion Approximations, *Journal of Econometrics* 45, 7–38.

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

[19] Wiggins, J.B. (1987) Option Values under Stochastic Volatility: Theory and Empirical Estimates, *Journal of Financial Economics* 19(2), 351–372.

Figure 1: **Implied Volatility Smiles**

This plot depicts implied volatilities for S&P 500 index options obtained using the Black-Scholes pricing formula. Treasury–bill yields of the appropriate maturity were used to obtain the applicable risk free rate. The option values were obtained from the *Wall Street Journal* of 8th November 1996, and pertain to prices prevailing at close of business on November 7, 1996. The S&P 500 Index closed at 727.65 on that day.

Of the three plots in the figure, the uppermost plot shows implied volatility estimates for November maturity options, the middle line that for December maturity options, and the lowest plot pertains to January options.

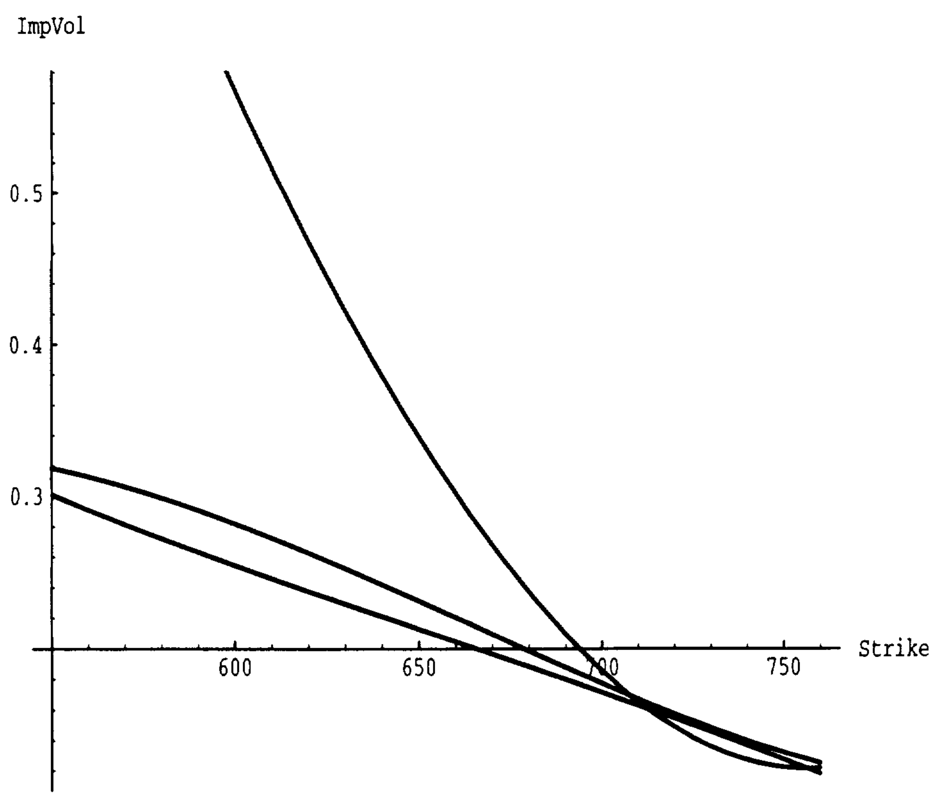

Figure 2: **Conditional Skewness & Kurtosis as κ and h vary**

This plot provides the levels of skewness and kurtosis when both mean reversion (κ) and the data frequency (h) are varied. The base case parameters are: $\alpha = 0.1$, $\kappa = 1.0$, $\theta = 0.01$, $\eta = 0.2$, $\rho = 0.25$ and initial values: $x = 0.1, V = 0.01$.

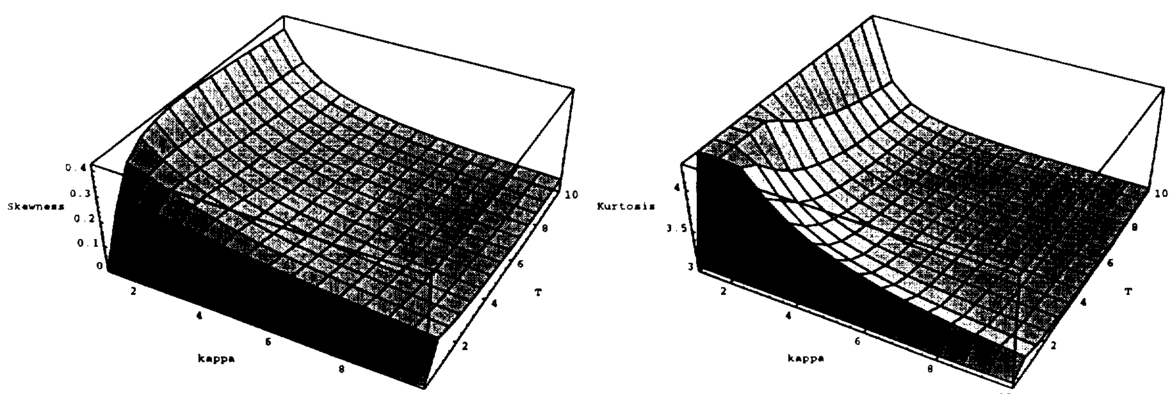

Figure 3: **Conditional Skewness & Kurtosis as θ and h vary**

This plot provides the levels of skewness and kurtosis when both mean volatility (θ) and the data frequency (h) are varied. The base case parameters are: $\alpha = 0.1$, $\kappa = 1.0$, $\theta = 0.01$, $\eta = 0.2$,$\rho = 0.25$ and initial values: $x = 0.1, V = 0.01$.

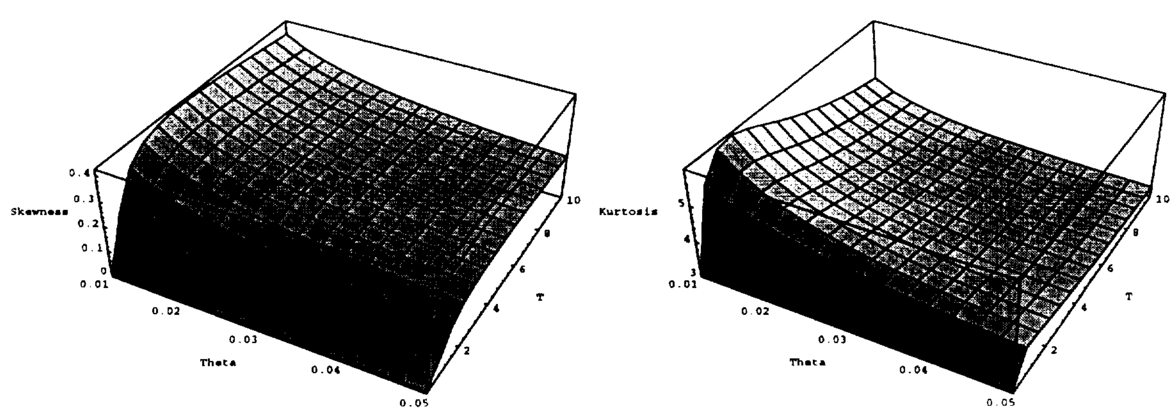

Figure 4: **Conditional Skewness & Kurtosis as η and h vary**

This plot provides the levels of skewness and kurtosis when both volatility variance (η^2) and the data frequency (h) are varied. The base case parameters are: $\alpha = 0.1$, $\kappa = 1.0$, $\theta = 0.01$, $\eta = 0.2$,$\rho = 0.25$ and initial values: $x = 0.1, V = 0.01$.

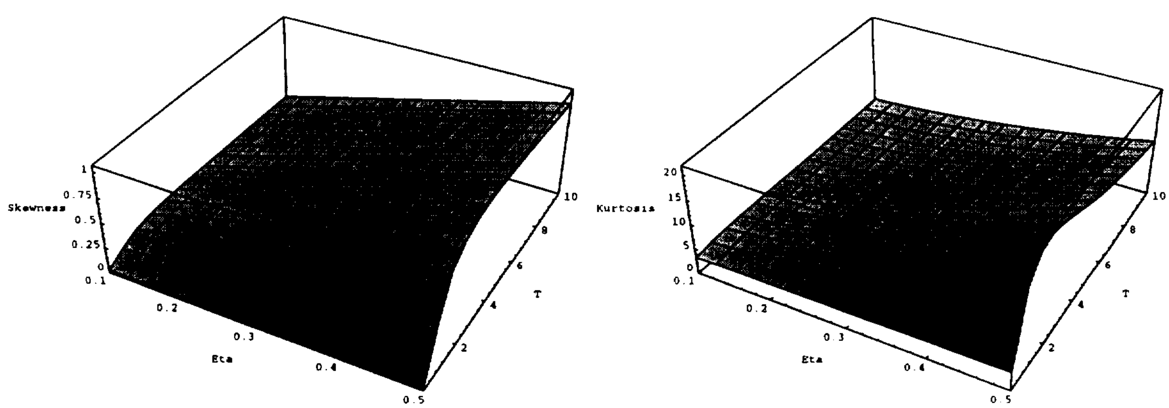

Figure 5: **Conditional Skewness & Kurtosis as ρ and h vary**

This plot provides the levels of skewness and kurtosis when both correlation (ρ) and the data frequency (h) are varied. The base case parameters are: $\alpha = 0.1$, $\kappa = 1.0$, $\theta = 0.01$, $\eta = 0.2$,$\rho = 0.25$ and initial values: $x = 0.1, V = 0.01$.