

# Accounting for Biases in Black-Scholes\*

David Backus,<sup>†</sup> Silverio Foresi,<sup>‡</sup> Kai Li,<sup>§</sup> and Liuren Wu<sup>¶</sup>

First draft: August 3, 1997

This version: November 25, 1997

## Abstract

Prices of currency options commonly differ from the Black-Scholes formula along two dimensions: implied volatilities vary by strike price (volatility smiles) and maturity (implied volatility of at-the-money options increases, on average, with maturity). We account for both using Gram-Charlier expansions to approximate the conditional distribution of the logarithm of the price of the underlying security. In this setting, volatility is approximately a quadratic function of moneyness, a result we use to infer skewness and kurtosis from volatility smiles. Evidence suggests that both kurtosis in currency prices and biases in Black-Scholes option prices decline with maturity.

**JEL Classification Codes:** G12, G13, F31, C14.

**Keywords:** currency options; skewness and kurtosis; Gram-Charlier expansions; implied volatility.

---

\*We welcome comments, including references to related papers we inadvertently overlooked. We thank Jose Campa, Steve Figlewski, Ken Garbade, Cliff Hurvich, Josh Rosenberg, and Raghu Sundaram for helpful comments and suggestions. The most recent version of this paper is available at <http://www.stern.nyu.edu/~dbackus>.

<sup>†</sup>Stern School of Business, New York University and NBER; [dbackus@stern.nyu.edu](mailto:dbackus@stern.nyu.edu).

<sup>‡</sup>Salomon Brothers Inc; [sforesi@sbi.com](mailto:sforesi@sbi.com).

<sup>§</sup>Stern School of Business, New York University; [kli1@stern.nyu.edu](mailto:kli1@stern.nyu.edu).

<sup>¶</sup>Stern School of Business, New York University; [lwu1@stern.nyu.edu](mailto:lwu1@stern.nyu.edu).

# 1 Introduction

Since Black and Scholes (1973) revolutionized the theory of option pricing, increasingly refined applied work has documented a number of anomalies or biases: examples in which observed prices of options differ systematically from the Black-Scholes formula. These observations have motivated, in turn, development of more complex models in which (among other things) jumps and stochastic volatility in the price of the underlying generate theoretical option prices different from Black-Scholes.

For currency options, biases are apparent across different strike prices and maturities. Strike price or “moneyness” biases include familiar smiles, smirks, and skews in which implied volatility varies across strike prices of otherwise similar options. For options on dollar prices of major currencies, the most common pattern is a smile, with implied volatility lower at the money than in or out of the money. Equally interesting, in our view, are biases across maturities. Although these are obscured by variation over time in the term structure of volatility, we document an average pattern of lower at-the-money volatility for shorter options. This maturity bias suggests a time element to departures from Black-Scholes that has largely been overlooked.

We show that both biases are natural consequences of stationary models in which changes in the logarithm of the price of the underlying security — here a currency — have greater kurtosis than normal random variables. In continuous time, departures from normality are typically modeled with jump-diffusions. We choose instead discrete-time Gram-Charlier expansions, an approach pioneered in finance by Jarrow and Rudd (1982). Whatever its source, excess kurtosis is well understood to produce volatility smiles. We derive a particularly simple version of this result in which implied volatility is approximately a quadratic function of moneyness, with coefficients related to the skewness and kurtosis of the price. Less well understood, we think, is how the effect of kurtosis varies with maturity. We show, in two related theoretical environments, that kurtosis biases downward the implied volatility of at-the-money options, but the bias eventually declines with maturity. For long options, the Black-Scholes formula is a good approximation in these models.

We argue that currency and option prices conform broadly with this theory. The excess kurtosis of log-changes in dollar prices of foreign currencies declines rapidly as the time interval is extended. Further, an example of volatility smiles for several maturities at the same date shows clearly that they get less pronounced as the maturity increases. The rate of convergence in this case is slower than we would expect

from models with iid increments, but is potentially consistent with the second-order time-dependence of stochastic-volatility models with fat-tailed innovations. In this respect, the behavior of currency options accords with related work on currency prices (Baillie and Bollerslev 1989 and Drost, Nijman, and Werker 1996, among others).

## 2 Biases in Black-Scholes

We begin with a selective list of systematic differences in observed prices of currency options from the Black-Scholes formula. Much of this dates back to Bodurtha and Courtadon's (1987, pp 33-50) comprehensive assessment of currency options, and even before.

Lists of Black-Scholes biases invariably begin with differences in implied volatilities across strike prices or moneyness. Although there have been striking examples of volatility skews in currency options, with volatility increasing or decreasing monotonically with moneyness, options on dollar exchange rates more commonly exhibit smiles, with minimum volatility roughly at the money. See, for example, Campa, Chang, and Reider (1997), Malz (1996), and Rosenberg (1996). Figure 1 provides an example from the over-the-counter market: one-month deutschemark options at five strike prices on April 3, 1996. This observation was supplied by Jose Campa and comes from data described in Campa, Chang, and Reider (1997). The figure contains a classic example of a volatility smile.

Smiles and other patterns in implied volatilities are often attributed to departures from normality in the logarithm of the price of the underlying. In Table 1, we report sample moments for three currencies of daily log-changes or depreciation rates of the spot exchange rate,

$$x_{t+1} = \log S_{t+1} - \log S_t, \tag{1}$$

where  $S_t$  is the dollar price of one unit of foreign currency at date  $t$  and time is measured in business days. The mean and standard deviation are expressed in annual units. With 260.6 business days per year in our sample, on average, we multiplied the estimated mean by 260.6 and the standard deviation by  $260.6^{1/2}$ .

Of particular interest to us are indicators of skewness and kurtosis, which we define in terms of the cumulants  $\kappa_j$  of a random variable  $x$ . The first cumulant is the mean,  $\kappa_1 = E(x)$ . The second is the variance,  $\kappa_2 = E(x - \kappa_1)^2$ . The third is the

third central moment:  $\kappa_3 = E(x - \kappa_1)^3$ . The fourth is  $\kappa_4 = E(x - \kappa_1)^4 - 3(\kappa_2)^2$ . The standard indicators of skewness and kurtosis are scaled versions of the third and fourth cumulants:

$$\gamma_1 = \frac{\kappa_3}{(\kappa_2)^{3/2}} \quad (\text{skewness}) \quad (2)$$

$$\gamma_2 = \frac{\kappa_4}{(\kappa_2)^2} \quad (\text{kurtosis}). \quad (3)$$

Both are invariant to changes in the scale of  $x$  and equal zero if  $x$  is normal.

The estimates in Table 1 replace expectations with sample moments. Standard deviations are just above 10% annually for the mark and the yen, less than half that for the Canadian dollar. Skewness  $\gamma_1$  is small, but estimated values of  $\gamma_2$  suggest that log-changes in currency prices are not normal: all three currency prices exhibit greater kurtosis than normal random variables. More formal statements to this effect are made by Akgiray and Booth (1988), Baillie and Bollerslev (1989), Drost, Nijman, and Werker (1996), Hsieh (1989), and Jorion (1988).

Less commonly noted are differences in option prices across maturities. Table 2 summarizes two years of daily volatility quotes for at-the-money options offered by a large international bank (Zhu 1997). The quotes come from at-the-money straddles, a combination of a put and a call with the same strike price; see, for example, Campa and Chang (1996, p 730). Most interesting to us is the term structure of at-the-money implied volatility. Although the term structure at different points in time has been increasing, decreasing, or even flat, mean and median volatility exhibit a steady rise with maturity. The latter is pictured in Figure 2, where we see a similar pattern in all three currencies. Average differences in at-the-money volatility across maturities are small relative to the time series variation in volatility, but they are nevertheless a recurring feature of option data, including longer time series from currency options traded on the Philadelphia Exchange and a nonoverlapping over-the-counter dataset summarized by Campa and Chang (1995, Tables I and II). This feature of the data has not been emphasized in earlier work, but was implicit in remarks made by Black (1975, p 64) (“Options with less than three months to maturity tend to be overpriced” by the Black-Scholes formula) and Bodurtha and Courtadon (1987, p 49) (“at-the-money [currency] options with less than 90 days to maturity are overpriced” by their models).

### 3 A Risk-Neutral Framework

Like many of our predecessors, we study option prices in a risk-neutral world. This assumption is not innocuous, but it allows us to focus our attention on the distribution of the underlying spot price.

The first element of our framework is the conditional distribution of future spot exchange rates. The conditional distribution of the one-period change in the log-price of the spot rate is determined by the properties of  $x$  in equation (1). Over  $n$  periods, the change is

$$\begin{aligned}\log S_{t+n} &= \log S_t + \sum_{j=1}^n x_{t+j} \\ &= \log S_t + x_{t+1}^n,\end{aligned}\tag{4}$$

with the obvious definition of  $x_{t+1}^n$ . Then  $S_{t+n} = S_t \exp(x_{t+1}^n)$  and the conditional distribution of  $S_{t+n}$  depends on that of  $x_{t+1}^n$ .

The second element is the valuation of options. The price of a European call option on a currency with strike price  $K$  is

$$C_{nt} = E_t [M_{t,t+n}(S_{t+n} - K)^+],\tag{5}$$

where  $M_{t,t+n}$  is a multi-period stochastic discount factor and  $x^+ \equiv \max(0, x)$ . Put-call parity,

$$P_{nt} = C_{nt} + Ke^{-r_{nt}n} - S_t e^{-r_{nt}^*n},\tag{6}$$

gives us the price  $P_{nt}$  of a put with the same maturity  $n$  and strike price  $K$ .

The pricing relation (5) is quite general, but leaves us with the difficulty of disentangling the effects of  $M$  and  $S$ . We follow tradition and sidestep this issue by assuming that the two are independent. With this “risk-neutral” structure, the call price depends in a clear way on the conditional distribution of  $x_{t+1}^n$ :

$$\begin{aligned}C_{nt} &= e^{-r_{nt}n} E_t (S_{t+n} - K)^+ \\ &= e^{-r_{nt}n} \int_{\log(K/S_t)}^{\infty} (S_t e^x - K) f(x) dx,\end{aligned}\tag{7}$$

where  $f$  is the conditional density of  $x_{t+1}^n$  and  $r_{nt}$  is the continuously-compounded  $n$ -period yield, measured in the same units as  $n$ .

If the  $n$ -period depreciation rate  $x_{t+1}^n$  is conditionally normal, with (say) mean  $\mu_n$  and standard deviation  $\sigma_n$ , then  $S_{t+n}$  is conditionally log-normal and the solution

to (7) is the familiar Black-Scholes (1973) formula, as adapted to currency options by Garman and Kohlhagen (1983):

$$C_{nt} = S_t e^{-r_{nt}^* n} \Phi(d) - K e^{-r_{nt} n} \Phi(d - \sigma_n), \quad (8)$$

where

$$d = \frac{\log(S_t/K) - (r_{nt}^* - r_{nt})n + \sigma_n^2/2}{\sigma_n}, \quad (9)$$

$\Phi$  is the standard normal cumulative distribution function, and  $r_{nt}^*$  is the foreign-currency yield. This follows from the arbitrage condition (interest parity),

$$\log E_t S_{t+n} - \log S_t = (r_{nt} - r_{nt}^*)n, \quad (10)$$

implying

$$\mu_n = (r_{nt} - r_{nt}^*)n - \sigma_n^2/2, \quad (11)$$

and evaluation of the integral defined by (7). Details are provided in Appendix A.1.

Largely for mathematical convenience, we use the variable  $d$  as a measure of moneyness in later sections, with  $d = 0$  indicating a call option at the money,  $d > 0$  in the money, and  $d < 0$  out of the money. This choice differs from standard practice, but the differences are small. The at-the-money straddles used to construct the implied volatilities used in Table 2, for example, define at-the-money by the forward rate,

$$K = S_t e^{(r_{nt} - r_{nt}^*)n},$$

implying  $d = \sigma_n/2$ . For an annual option on the deutschemark,  $d$  is approximately 0.06, which is tiny relative to the range of  $d$  pictured in Figure 1. For shorter options, the choice of at-the-money  $d$  is even smaller.

## 4 Accounting for Moneyness Bias

We turn to prices of options of fixed maturity when log-changes in the spot exchange rate are not normal. In continuous time, non-normality is modeled with “jumps” or “point processes.” In discrete time, the conditional distribution of log-changes can be specified directly, and we have some flexibility over its form. A convenient choice for our purposes is a Gram-Charlier expansion, in which a normal density is augmented with additional terms capturing the effects of skewness and kurtosis. Johnson, Kotz, and Balakrishnan (1994, pp 25-30) and Kolassa (1994, ch 3) describe the underlying statistical theory. This approach was introduced to finance by Jarrow

and Rudd (1982) and has since been applied by Abken, Madan, and Ramamurtie (1996), Brenner and Eom (1997), Knight and Satchell (1997), Longstaff (1995), and Madan and Milne (1994). Our application differs from Jarrow and Rudd's in approximating the conditional distribution of the logarithm of the price rather than the price itself. This relatively subtle difference leads to a particularly simple expression for volatility smiles in terms of higher moments of the log-price of the underlying.

A Gram-Charlier expansion generates an approximate density function for a standardized random variable that differs from the standard normal in having potentially nonzero skewness and excess kurtosis. In our application, let the  $n$ -period log-change in the spot rate  $x_{t+1}^n$  [equation (4)] have conditional mean  $\mu_n$  and standard deviation  $\sigma_n$ , and define the standardized variable

$$w = \frac{x_{t+1}^n - \mu_n}{\sigma_n}. \quad (12)$$

A Gram-Charlier expansion defines an approximate density for  $w$  by

$$f(w) = \varphi(w) - \gamma_{1n} \frac{1}{3!} D^3 \varphi(w) + \gamma_{2n} \frac{1}{4!} D^4 \varphi(w), \quad (13)$$

where  $\varphi(w) = (2\pi)^{-1/2} \exp(-w^2/2)$  is the standard normal density and  $D^j$  denotes the  $j$ th derivative of what follows. Equation (13) is typically viewed as an approximation to an arbitrary density with nonzero higher moments, but for moderate values of  $\gamma_{1n}$  and  $\gamma_{2n}$  it is a legitimate density in its own right. In such cases it provides a more parsimonious representation of a distribution with skewness and kurtosis than the normal mixtures implied by jump-diffusions. Most important to us is that departures from normality are closely related to measures of higher moments.

Some of the properties of the Gram-Charlier expansion  $f$  are apparent from its cumulant generating function (the log of its moment generating function),

$$\begin{aligned} \psi(s; w) &= \log E(e^{sw}) \\ &= s^2/2! + \gamma_{1n} s^3/3! + \gamma_{2n} s^4/4!. \end{aligned} \quad (14)$$

In  $w$  is normal, this power series is zero after the quadratic term. The Gram-Charlier expansion extends the series for two additional terms. Since the derivatives of  $\psi$  define the cumulants, the coefficients are

$$\left. \frac{\partial^j \psi(s; w)}{\partial s^j} \right|_{s=0} = \kappa_{jn}.$$

For a standardized random variable the standard deviation is one, so  $\kappa_{3n} = \gamma_{1n}$  and  $\kappa_{4n} = \gamma_{2n}$ , the standard indicators of skewness and kurtosis. And since  $x^n$  is a linear translation of  $w$ , these indicators apply to it, too. We leave the maturity  $n$  explicit for later use.

We use the relation between the density and its higher moments to derive expressions relating a call option's price, implied volatility, and "delta" to the same moments. Call prices are given by

**Proposition 1** *Consider currency options in an environment characterized by the risk-neutral pricing relation (7), with the evolution of the spot rate governed by equations (4), (12), and the Gram-Charlier expansion (13). Call prices in this setting are approximately*

$$C_{nt} \cong S_t e^{-r_{nt}^* n} \Phi(d) - K e^{-r_{nt} n} \Phi(d - \sigma_n) + S_t e^{-r_{nt}^* n} \varphi(d) \sigma_n \left[ \frac{\gamma_{1n}}{3!} (2\sigma_n - d) - \frac{\gamma_{2n}}{4!} (1 - d^2 + 3d\sigma_n - 3\sigma_n^2) \right], \quad (15)$$

where  $d$  is defined by (9).

A proof is given in Appendix A.2. As with our approach to the Black-Scholes formula, it relies on a combination of an arbitrage condition and the patience to evaluate the integral defined by (7). The approximation refers to our elimination of terms involving powers three and higher of  $\sigma_n$ , which are negligible in practice.

Equation (15) expresses the call price as Black-Scholes plus terms involving the skewness ( $\gamma_{1n}$ ) and kurtosis ( $\gamma_{2n}$ ) of the logarithm of the underlying spot exchange rate. Kurtosis lowers prices of at-the-money options ( $d = 0$ ) relative to Black-Scholes, but raises prices of far in- and out-of-the-money options ( $|d| > 1$ ). The combination stems from the effect of kurtosis on the density: it increases the probabilities of both small and large realizations. For at-the-money options, the former dominates, leading to lower prices than Black-Scholes. For out-of-the-money options, the latter dominates and the price exceeds its Black-Scholes value. Figure 3 illustrates the differences between Black-Scholes call prices (the solid line) and those of Proposition 1 with positive skewness (dashed line) and kurtosis (dash-dot line).

The picture is more familiar if we translate the bias into volatility units. Define implied volatility  $v_n$  as the value of  $\sigma_n$  that equates the Black-Scholes formula (8) to the observed price given values of the other parameters. We refer to the relation between implied volatility and moneyness as the implied volatility smile. If the call price is in fact generated by (15), then the implied volatility smile is given by

**Proposition 2** *Let the price of a call option be the Gram-Charlier formula (15). Then the implied volatility smile is*

$$v_n(d) \cong \sigma_n \left[ 1 - \frac{\gamma_{1n}}{3!}d - \frac{\gamma_{2n}}{4!}(1 - d^2) \right], \quad (16)$$

with  $d$  defined by (9).

A proof is given in Appendix A.2.

Although the form of (16) is new, its content is reminiscent of Merton's (1976, Section 4) discussion of the qualitative features of volatility smiles in jump-diffusion models. The most attractive feature of our version is its simplicity: it expresses option price biases in terms of the widely-used implied volatility and relates them to higher moments in the distribution of the underlying. The equation is an approximation for several reasons: the Gram-Charlier expansion, a linear approximation of call prices in terms of volatility, and the elimination of terms involving powers of two and higher in  $\sigma_n$ . Of these, only the first and last seem to have much effect. Figure 4 compares the volatility smiles implied by excess kurtosis under our approximation (16) (dashed line) and under an exact inversion of the call price formula (15). For  $|d| < 1.5$  the differences are small, but for larger values the true implied volatility is smaller than that implied by Proposition 2. Figure 5 illustrates the differences implied by equation (16) between positive kurtosis (solid line) and a combination of positive kurtosis and negative skewness (dashed line) that mimics the volatility smirks and skews documented in other markets.

As an example of how equation (16) might be used in practice, consider the volatility smile for one-month deutschemark options pictured in Figure 1. A least squares fit of a quadratic function to the smile implies moments of  $\sigma_n = 0.0892/12^{1/2}$ ,  $\gamma_{1n} = -0.168$ , and  $\gamma_{2n} = 1.591$ . In this way, the volatility smile gives us estimates of the true volatility and higher moments of the spot exchange rate process. Just as the Black-Scholes formula made volatility visible from option prices, Proposition 2 makes skewness and kurtosis visible from the shape of volatility smiles.

Before turning to maturity bias, we take a short digression to consider the effects of skewness and kurtosis on the delta of an option:

**Corollary 1** *The delta of the Gram-Charlier call price (15) is*

$$e^{r_{nt}^*n} \Delta(d) \equiv \frac{\partial C_{nt}}{\partial (S_t e^{-r_{nt}^*n})}$$

$$\begin{aligned}
&= \Phi(d) - \frac{\gamma_{1n}}{3!}\varphi(d)(1 - d^2 + 3\sigma_n - 2\sigma_n^2) \\
&\quad + \frac{\gamma_{2n}}{4!}\varphi(d)[3d(1 + 2\sigma_n^2) + 4d^2\sigma_n - d^3 - 4\sigma_n + 3\sigma_n^3], \quad (17)
\end{aligned}$$

with  $d$  defined by (9).

The proof consists of differentiating (15).

The delta, like the call price and implied volatility, is the conventional Black-Scholes result [here  $e^{-r_{nt}^*n}\Phi(d)$ ] plus additional terms involving skewness and kurtosis. Examples are pictured in Figure 6. For at-the-money options, the delta simplifies considerably. Setting  $d = 0$  and eliminating terms involving powers of two and higher of  $\sigma_n$ ,

$$\Delta(0) \cong e^{-r_{nt}^*n} \left[ \Phi(0) - \frac{1}{3!}\varphi(0)(\gamma_{1n} + \gamma_{2n}\sigma_n) \right].$$

Since  $\sigma_n$  is generally small, kurtosis has little effect on the delta at the money, despite having a substantial effect on the price. Skewness, on the other hand, has little effect on the at-the-money price, but can have a sizable effect on the delta.

Equation (17) is a reminder that the first-order approximations often used in hedging and risk management of options can differ substantially from the Black-Scholes benchmark when the log-price of the underlying exhibits skewness or kurtosis. We will see shortly that skewness and kurtosis are most evident in short-dated options, where the difficulties of hedging are widely thought to be most severe.

## 5 The Quality of the Approximation

Propositions 1 and 2 base approximations of call prices and implied volatility smiles, respectively, on a Gram-Charlier approximation to the conditional density of the log-price of the underlying. In this section, we examine the accuracy of the Gram-Charlier approximation to the smile, equation (16) of Proposition 2, when the distribution is generated by a jump-diffusion (Merton 1976). This distribution has some support in the empirical literature (Akgiray and Booth 1988, for example), and is capable of generating a wide range of non-normal behavior. We show that with what we regard as reasonable parameter values, higher moments inferred from option prices using (16) are similar to the true moments of the process for the underlying. This does not constitute a general endorsement of the Gram-Charlier

approximation, but gives us some reason to believe that estimates of higher moments from option prices are useful indicators of the same properties of the underlying.

The jump-diffusion process defines the density of the  $n$ -period log-price change as a countable mixture of normals:

$$f(x) = \sum_{j=0}^{\infty} p_j \varphi(x; \nu_0 + j\nu_1, \tau_0^2 + j\tau_1^2), \quad (18)$$

where  $p_j = e^{-\lambda_n} \lambda_n^j / j!$  is the (poisson) probability of  $j$  jumps,  $\lambda_n$  is the jump intensity (the expected number of jumps in  $n$  periods),  $(\nu_0, \tau_0^2)$  are the mean and standard deviation of the diffusion,  $(\nu_1, \tau_1^2)$  are the mean and standard deviation of the jump, and  $\varphi(x; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp[-(x - \mu)^2 / 2\sigma^2]$ . The jump-diffusion density (18) exhibits greater kurtosis than the normal and, if  $\nu_1$  is nonzero, nonzero skewness as well. The first four cumulants are

$$\begin{aligned} \kappa_1 &= \nu_0 + \lambda_n \nu_1 \\ \kappa_2 &= \tau_0^2 + \lambda_n (\tau_1^2 + \nu_1^2) \\ \kappa_3 &= \lambda_n (\nu_1^3 + 3\nu_1 \tau_1^2) \\ \kappa_4 &= \lambda_n (3\tau_1^4 + 6\nu_1^2 \tau_1^2 + \nu_1^4). \end{aligned}$$

If  $\nu_1 = 0$ , skewness is zero and kurtosis depends on the intensity ( $\lambda_n$ ) and variance ( $\tau_1^2$ ) of jumps. When  $\nu_1 \neq 0$ , its sign carries over to the third cumulant.

This departure from normality results in option prices that differ from Black-Scholes. The price of a European call with strike price  $K$  is

$$C_{nt} = e^{-r_{tn}} \sum_{j=0}^{\infty} p_j W_j, \quad (19)$$

where

$$\begin{aligned} W_j &= S_t e^{(\nu_0 + j\nu_1) + (\tau_0^2 + j\tau_1^2)/2} \Phi(d_j) - K \Phi(d_j - [\tau_0^2 + j\tau_1^2]^{1/2}) \\ d_j &= \frac{\log(S_t/K) + (\nu_0 + j\nu_1) + (\tau_0^2 + j\tau_1^2)}{(\tau_0^2 + j\tau_1^2)^{1/2}}. \end{aligned}$$

The result follows from repeated application of equation (30) in Appendix A.1. Similar formulas are reported by Bates (1996b, equation 9), Jarrow and Rudd (1982, equation 17), and Merton (1976, equation 18)

Table 3 summarizes the accuracy of Gram-Charlier-based approximations to skewness and kurtosis in the jump-diffusion model. For a variety of choices of parameter values, we compute call prices for the jump-diffusion model [equation (19)],

implied volatilities [inverting (8)], and Gram-Charlier estimates of higher moments [a least squares fit of equation (16)]. In each case, the moneyness vector  $d$  corresponds to  $\Phi(d) = 0.1, 0.2, \dots, 0.9$ . The jump-intensity parameter is chosen to correspond to estimates in the literature. Other parameters are chosen to reproduce plausible values of domestic and foreign interest rates, the standard deviation  $(\kappa_2)^{1/2}$ , skewness  $\gamma_1$ , and kurtosis  $\gamma_2$ . Given interest rates, the non-jump mean  $\nu_0$  is adjusted to obey

$$(r_t - r_t^*)n = \nu_0 + \tau_0^2/2 + \lambda_n \left( e^{\nu_1 + \tau_1^2/2} - 1 \right),$$

an application of the arbitrage condition (10). Our benchmark values are  $r_t = r_t^* = 0$ ,  $\sigma = (\kappa_2)^{1/2} = 0.1$  (annualized),  $\gamma_1 = 0$ , and  $\gamma_2 = 1$ . Estimates of the jump-intensity parameter are reported by Akgiray and Booth (1988, Table 2), Bates (1996a, Table 2), and Jorion (1988, Table 3). Annualized values range between 1 and 50, with a median of about 10 (just under one jump per month, on average), which we use as our starting point.

With benchmark parameter values (Panel A), the Gram-Charlier approximation of the volatility smile implies skewness and kurtosis estimates of  $-0.006$  and  $0.898$ . Both values are close to the moments of the jump distribution (0 and 1, respectively). In that sense, they suggest that Proposition 2 is a passable approximation. In fact, some of the error comes not from the Gram-Charlier approximation, but from the additional approximations made in moving from Proposition 1 to Proposition 2. A least squares fit of the more accurate equation (15) generates estimates of  $(\gamma_1, \gamma_2) = (-0.013, 0.917)$ .

In the remaining panels of Table 3, we examine the sensitivity of these estimates to the choice of parameters. In Panel B, we consider different values of jump intensity. With larger values, the approximation gets better (with  $\lambda = 15$ ,  $\gamma_2$  rises to 1.046), but with smaller values it gets worse ( $\gamma_2 = 0.748$  when  $\lambda = 5$ ). Values above 15 are incompatible with  $\gamma_2 = 1$ . Values below 5 imply larger, less frequent jumps, and the Gram-Charlier estimates of kurtosis are well below those of the jump-diffusion model.

In Panel C, we examine volatility. Neither smaller nor larger values has an appreciable effect on estimates of higher moments. In Panel D, we examine options with longer maturities. Three-month options allow more time for jumps, and therefore require greater jump variance  $\tau_1^2$  to generate the same amount of kurtosis. The result is a slight overestimate of kurtosis. For the six-month option we set  $\gamma_2 = 0.5$  (the model is incapable of reproducing  $\gamma_2 = 1$  with the benchmark value

of  $\lambda$ ). The Gram-Charlier estimate is 0.540. In all of these examples, the errors in Gram-Charlier estimates of skewness and kurtosis are small.

Panel E suggests that the accuracy of the approximation changes little when we vary the amount of kurtosis. Skewness, however, poses some difficulties (Panel F). The Gram-Charlier approximation systemically underestimates the (absolute) amount of skewness, and for larger values underestimates kurtosis as well. This difficulty highlights a difference between the Gram-Charlier and jump models. In the Gram-Charlier approximation, skewness has (approximately) a linear effect on implied volatility. In the jump-diffusion model, skewness alone produces concave volatility smiles. The approximation therefore reduces its estimate of  $\gamma_2$  to compensate.

In short, the Gram-Charlier approximation works well in some cases, less well in others. Others here refers to examples with low jump intensity or substantial skewness. In our opinion, the approximation is reasonably good in the range of parameter values indicated by currency data, but logic strongly suggests that there can be no general defense of the approximation. The Gram-Charlier expansion arises from a Taylor series approximation to the cumulant generating function, and examples are commonplace of functions for which a short Taylor series is an extremely poor approximation. For similar reasons, there must certainly exist examples in which Gram-Charlier approximations to call prices are poor. Whether these examples are realistic is impossible to say without knowing what they are. We take some comfort, however, in the ability of the model to approximate estimates of kurtosis from jump-diffusions when plausible parameter values are used.

## 6 Accounting for Maturity Bias 1

The second bias concerns maturity: average at-the-money implied volatility is typically smaller at short maturities than long ones (Table 2 and Figure 2). To study this issue, we must specify a stochastic process for log-price changes: a description of the conditional distribution of price changes over periods of arbitrary length  $n$ . We consider two such processes. In this section, we examine iid innovations. The iid structure is not realistic, but allows an especially simple characterization of the effects of maturity. In the next section, we explore the possibility of time-dependence induced by stochastic volatility. In both cases, departures from normality and Black-Scholes go to zero with maturity. Since kurtosis lowers at-the-money volatility, we have a potential explanation for the maturity bias.

Suppose, then, that  $n$ -period log-changes in the exchange rate are composed of iid one-period components with finite moments of all order. Specifically, let daily depreciation rates  $x$  be

$$\begin{aligned}\log S_{t+1} - \log S_t &= x_{t+1} \\ &= \mu_{t+1} + \sigma \varepsilon_{t+1},\end{aligned}$$

where  $\{\varepsilon_t\}$  are iid with mean zero and variance one. Over  $n$  periods, the log-change in the spot rate is

$$\begin{aligned}\log S_{t+n} - \log S_t &= x_{t+1}^n \\ &= \mu_n + \sigma \sum_{j=1}^n \varepsilon_{t+j},\end{aligned}\tag{20}$$

with the obvious definitions of  $x_t^n$  and  $\mu_n$ .

Now consider the behavior of cumulants and moments over different time intervals  $n$ . An arbitrary  $\varepsilon$  with finite cumulants  $\kappa_j$  has cumulant generating function

$$\psi(s; \varepsilon) = \sum_{j=1}^{\infty} \frac{\kappa_j s^j}{j!}.$$

Recall that the cumulant generating function of the sum of independent random variables is the sum of the generating functions of the individual random variables. Then  $\varepsilon_{t+1}^n \equiv \sum_{j=1}^n \varepsilon_{t+j}$  has cumulant generating function

$$\psi(s; \varepsilon_{t+1}^n) = \sum_{j=1}^{\infty} \frac{n\kappa_j s^j}{j!}$$

and hence has cumulants  $n\kappa_j$ . As a result, the standard deviation increases in the familiar way with the square root of maturity:  $\sigma_n = n^{1/2}\sigma$ . In addition, the skewness and kurtosis indicators  $\gamma_{1n}$  and  $\gamma_{2n}$  defined by (2,3) decline with  $n$ :

$$\gamma_{1n} = \frac{n\kappa_3}{(n\kappa_2)^{3/2}} = \frac{\gamma_{11}}{n^{1/2}}\tag{21}$$

$$\gamma_{2n} = \frac{n\kappa_4}{(n\kappa_2)^2} = \frac{\gamma_{21}}{n}\tag{22}$$

Similar expressions are reported by Das and Sundaram (1996) for jump-diffusions. Thus departures from normality, as indicated by higher moments, decline with maturity.

The Gram-Charlier expansion is a special case of this setup with cumulants equal to zero after the fourth. In this setting:

**Proposition 3** *Suppose  $n$ -period log-changes in the exchange rate (20) are composed of standardized iid one-period innovations  $\varepsilon_t$ , whose density is given by a Gram-Charlier expansion (13) with third and fourth cumulants  $\kappa_3$  and  $\kappa_4$ . As the maturity  $n$  of a call option approaches infinity, skewness and kurtosis approach zero and call prices approach the Black-Scholes formula.*

The proposition follows from the formulas for skewness and kurtosis, equations (21) and (22), and the relations between higher moments and option prices, (15) and (16). ■

Proposition 3 suggests an explanation of the maturity bias: the tendency for average or median at-the-money volatility to rise with maturity. The effects of kurtosis at different maturities are pictured in Figure 7. For one-month options, kurtosis generates the smile we saw in Figure 4. For three-month options and the iid structure of this section, kurtosis declines by a factor of 3. This results in a less sharply curved smile and a general reduction in differences from the Black-Scholes formula. For at-the-money options, implied volatility is higher for the longer option. We might thus expect to see that at-the-money volatility rises with maturity, as we saw (on average) in the data.

Proposition 3 and the moment properties (21,22) go beyond these qualitative properties and provide explicit predictions of how prices of currencies and options behave across maturities. Each of these properties can be compared with the evidence. Consider currencies. In equation (22), kurtosis  $\gamma_{2n}$  is proportional to  $1/n$ . Figure 8 is an attempt to compare this prediction with the data. We computed  $\gamma_{2n}$  for different values of  $n$  for the ten years of data used in Table 1. (To put them on the same scale, the estimates of  $\gamma_{2n}$  for each currency are divided by  $\gamma_{21}$ .) We see in the figure that kurtosis declines rapidly with  $n$ . Moreover, the pattern of decline is not much different from the solid line  $1/n$ , which is what the iid model predicts. The yen is very close to this line. The Canadian dollar and the mark are similar, but decline less rapidly for small  $n$  than the benchmark line. On the whole, the data are roughly in line with equation (22), but exhibit less rapid convergence.

A more limited look at option prices conveys a similar message. Equations (16) and (22) tell us that Proposition 3 should show up as volatility smiles with progressively less slope and curvature as we increase maturity. Exchange-traded options generally do not have enough depth across the moneyness spectrum to allow us to estimate smiles precisely, much less compare them across maturities, but the Campa, Chang, and Reider (1997) data includes maturities between one day and

eighteen months. Strike prices in this data set are chosen to set the Black-Scholes  $\Delta$ ,  $e^{-r_n^* t} \Phi(d)$ , equal to 0.1, 0.25, 0.75, and 0.9, with an additional at-the-money strike corresponding to  $d = \sigma_n/2$  (see the discussion at the end of Section 3). Smiles on April 3, 1996 (our only observation, and the first one in Campa, Chang, and Reider's data) are graphed in Figure 9 for maturities between one day and one year. As the theory suggests, the smiles get increasingly flatter as the maturity rises. It's essential here that the smiles are expressed in units of moneyness that are comparable across maturities. We use  $d$ , as suggested by Proposition 2, but the important thing is that the natural spread in the distribution over time (the increase in  $\sigma_n$  with  $n$ ) be counteracted. Once this is done, the smiles flatten out with maturity as Proposition 3 suggests.

More concretely, the smiles in Figure 9 can be used to infer moments of  $x^n$ , as we described in Section 4. The results of this exercise are reported in Table 4. The dramatic rise in the level of the smiles (and the difference between Figures 7 and 9) is attributed to a rising term structure of volatility. More relevant to Proposition 3: both skewness and kurtosis decline with maturity. The primary discrepancy with the theory of this section is the rate of convergence: estimates of kurtosis inferred from option prices (Table 4) decline substantially less quickly than the  $1/n$  rate of equation (22).

## 7 Accounting for Maturity Bias 2

Despite its lack of realism, the iid model in the last section provides a relatively good qualitative explanation for both moneyness and maturity biases in Black-Scholes. Quantitatively, it suggests that higher moments decline rapidly with maturity; volatility smiles suggest that the rate of decline is less rapid. In this section, we consider time-dependence in one-period depreciation rates  $x$ . Since depreciation rates exhibit little evidence of autocorrelation, we are led to consider dependence through second moments, or stochastic volatility. We show that such models can generate slower convergence of higher moments to zero over the range of maturities observed in option markets.

Stochastic volatility models are motivated by clear evidence of predictable variation in conditional variance. This is apparent in the dynamics of implied volatilities (Table 2, for example) and in estimates of GARCH and related models. Notable applications to currency options include Bates (1996a), Melino and Turnbull (1990), and Taylor and Xu (1994).

The generic stochastic volatility model starts with log-price changes of the form

$$x_{t+1} = \mu_{t+1} + z_t^{1/2} \varepsilon_{t+1}, \quad (23)$$

where  $\varepsilon_{t+1}$  is conditionally independent of  $z_t$  and  $\{\varepsilon_t\}$  is a sequence of iid draws with zero mean and unit variance. The new element,  $z_t$ , is the conditional variance. Popular processes for  $z$  include the square-root model used by Bates (1996a),

$$z_{t+1} = (1 - \beta)\omega + \beta z_t + \alpha z_t^{1/2} \eta_{t+1}, \quad (24)$$

the logarithmic model of Kim and Shephard (1994),

$$\log z_{t+1} = (1 - \beta)\omega + \beta \log z_t + \alpha \eta_{t+1}, \quad (25)$$

and the GARCH(1,1) model estimated by Baillie and Bollerslev (1989),

$$\begin{aligned} z_{t+1} &= (1 - \beta)\omega + (\beta - \alpha)z_t + \alpha z_t \varepsilon_{t+1}^2 \\ &= (1 - \beta)\omega + \beta z_t + \alpha z_t \eta_{t+1}, \end{aligned} \quad (26)$$

with  $\eta_t = \varepsilon_t^2 - 1$ . In all of these models,  $\{\eta_t\}$  is iid with zero mean and unit variance. Although these models are not equivalent, their properties are similar in many applications.

Whatever the process, variation in  $z$  generates kurtosis in  $x$  and  $x^n$ . With no particular structure on  $z$  but existence of unconditional mean and variance, the unconditional kurtosis in  $x$  is

$$\gamma_2(x) = 3c(z) + \gamma_2(\varepsilon)[1 + c(z)], \quad (27)$$

where  $c(z) = \text{Var}(z)/E(z)^2$  is the squared coefficient of variation of  $z$  (the ratio of the standard deviation to the mean). See Appendix A.3. Equation (27) quantifies two ways of generating excess kurtosis in the unconditional distribution of  $x$ : kurtosis in the innovations  $\varepsilon$  [represented by  $\gamma_2(\varepsilon)$ ] and stochastic volatility [represented by  $c(z)$ ]. Note, too, that the two effects complement each other: the total is greater than the sum of the parts.

Similar relations hold conditionally and over multiple periods if  $\{\eta_t\}$  and  $\{\varepsilon_t\}$  are independent. This assumption is violated in the GARCH model, since  $\eta_t = \varepsilon_t^2 - 1$ , but we think the gain in simplicity makes the extra structure worthwhile. Then conditional kurtosis is

$$\gamma_2(x^n) = 3 \frac{\text{Var}_t \sum_{j=1}^n z_{t+j-1}}{\left(E_t \sum_{j=1}^n z_{t+j-1}\right)^2} + \gamma_2(\varepsilon) \frac{E_t \sum_{j=1}^n z_{t+j-1}^2}{\left(E_t \sum_{j=1}^n z_{t+j-1}\right)^2}, \quad (28)$$

which is derived in Appendix A.3. As with unconditional kurtosis, the interaction of fat-tailed innovations and stochastic volatility results in greater kurtosis than the sum of the two individually.

Further progress requires us to be more specific about the process for  $z$ . In the interest of transparency, we study a linear volatility model,

$$z_{t+1} = (1 - \beta)\omega + \beta z_t + \alpha \eta_{t+1}, \quad (29)$$

with  $\{\eta_t\}$  independent of  $\{\varepsilon_t\}$ . When  $0 < \beta < 1$ , the components of conditional kurtosis (28) are

$$\begin{aligned} z_{t+n} &= \omega + \beta^n(z_t - \omega) + \alpha \sum_{j=1}^n \beta^{n-j} \eta_{t+j} \\ \sum_{j=1}^n z_{t+j-1} &= n\omega + \left(\frac{1 - \beta^n}{1 - \beta}\right)(z_t - \omega) + \alpha \sum_{j=1}^{n-1} \left(\frac{1 - \beta^{n-j}}{1 - \beta}\right) \eta_{t+j} \\ E_t \sum_{j=1}^n z_{t+j-1} &= n\omega + \left(\frac{1 - \beta^n}{1 - \beta}\right)(z_t - \omega) = O(n) \\ E_t \sum_{j=1}^n z_{t+j-1}^2 &= \sum_{j=1}^n [\omega + \beta^{j-1}(z_t - \omega)]^2 + \alpha^2 \sum_{j=1}^n \left(\frac{1 - \beta^{2(j-1)}}{1 - \beta^2}\right) = O(n) \\ \text{Var}_t \sum_{j=1}^n z_{t+j-1} &= \alpha^2 \sum_{j=1}^n \left(\frac{1 - \beta^j}{1 - \beta}\right)^2 = O(n). \end{aligned}$$

(The notation  $f(n) = O(n^k)$  means that  $f(n)/n^k$  has a finite nonzero limit. We say that  $f(n)$  is of order  $n^k$ .) When  $\beta = 1$ , the expressions become

$$\begin{aligned} E_t \sum_{j=1}^n z_{t+j-1} &= n z_t = O(n) \\ E_t \sum_{j=1}^n z_{t+j-1}^2 &= n z_t^2 + \alpha^2 n(n-1)/2 = O(n^2) \\ \text{Var}_t \sum_{j=1}^n z_{t+j-1} &= \alpha^2 n(2n-1)(n-1)/6 = O(n^3). \end{aligned}$$

This limiting case provides an upper bound to the effects of stochastic volatility on conditional kurtosis.

We can now assess the qualitative features of conditional kurtosis. When  $\beta = 1$ , conditional kurtosis is

$$\gamma_2(x^n) = 3 \frac{O(n^3)}{O(n^2)} + \gamma_2(\varepsilon) \frac{O(n^2)}{O(n^2)}.$$

The second term has a finite limit, but the first grows without bound with  $n$ . In sharp contrast to the previous section, there is no tendency for higher moments to decline with maturity. When  $0 < \beta < 1$ , however, the qualitative features of the last section return. Conditional kurtosis is

$$\gamma_2(x^n) = 3 \frac{O(n)}{O(n^2)} + \gamma_2(\varepsilon) \frac{O(n)}{O(n^2)},$$

so both terms eventually converge to zero.

Consider, now, the rate at which higher moments converge in this model. In the stationary case, the model combines two mechanisms studied extensively by Das and Sundaram (1996): non-normal innovations and stochastic volatility. Without stochastic volatility (that is, with  $\alpha = 0$ , which we think of as analogous to a pure jump model), kurtosis follows the pattern we documented in the last section:  $\gamma_2(x^n) = \gamma_2(\varepsilon)/n$ . Without non-normal innovations (that is, with  $\gamma_2(\varepsilon) = 0$ ), kurtosis is hump-shaped, approaching zero at very short and very long maturities. As Ait-Sahalia and Lo (1997, p 3) note, neither corresponds to observed option prices: the former declines too rapidly to zero, the latter is too small at short maturities.

A numerical example suggests that a combination of non-normal innovations and stochastic volatility gives a much better account of the behavior of conditional kurtosis across maturities than either on its own. This complements work by Baillie and Bollerslev (1989), Drost, Nijman, and Werker (1996), Hsieh (1989), and Jorion (1988), who proposed similar models and showed that they accounted for many of the observed properties of currency prices. The parameter values are  $\beta = 0.9834$ ,  $\omega = 0.1116^2$  (annualized),  $\alpha = 0.0094$  (annualized), and  $\gamma_2(\varepsilon) = 2.912$ , and the state variable is set equal to its mean ( $z_t = \omega$ ). The relation between kurtosis and maturity is pictured in Figure 10. For maturities between one day and eighteen months, kurtosis declines from 2.9 to 0.5, as in the option data summarized in Table 4. Note that the kurtosis in this example is substantially greater than we get from “jumps” [nonzero  $\gamma_2(\varepsilon)$ ] or stochastic volatility (nonzero  $\alpha$ ) alone.

The tendency for kurtosis to decline with maturity in this model is a consequence of a stronger result: the central limit theorem. As Diebold (1988) showed in similar environments, average changes in log-prices are normal over long enough time intervals. As a result, departures from the Black-Scholes formula decline with the maturity of the option. This statement doesn’t apply to all theoretical environments (the unit root volatility model is a counterexample), but it appears to be a reasonable approximation for prices of currencies and currency options.

## 8 Final Thoughts

We have continued a line of research initiated by Jarrow and Rudd (1982) of using Gram-Charlier expansions to explore the impact of departures from log-normality in the underlying on the prices of options. We add two things: (i) an extremely simple relation between the shape of implied volatility smiles and the skewness and kurtosis of the underlying log-price process and (ii) an examination of how smiles and higher moments vary with maturity. The latter includes the suggestion that biases in the popular Black-Scholes formula should be greatest for short options and disappear entirely at long enough maturities. We argue that a model with both fat-tailed innovations and stochastic volatility can account for the relatively slow decline in kurtosis with maturity that we infer from option prices.

These conclusions point us in a number of directions. One is to use option prices to study the dynamics of conditional higher moments in asset prices. Studies based on the price alone have been hindered by the difficulty of estimating such moments from small numbers of realizations. (Hansen 1994 remains the only study we know to attempt this.) Option prices give us another source of information and raise the hope that we will be able to document the time series properties of skewness and kurtosis, as earlier work has done with conditional variances. Another direction is the behavior of exotic options. Appendix A.4 extends the theory to digital options, but the more difficult barrier options remain. A third direction concerns the nature of option markets and data. We have modeled option prices in a competitive, frictionless world, but perhaps bid/ask spreads and noncompetitive pricing account for some of the differences between observed option prices and the Black-Scholes formula.

## A Derivations and Proofs

### A.1 Black-Scholes Formula

We adapt Rubinstein's (1976) approach to the Black-Scholes formula. Let the  $n$ -period log-price change  $x_{t+1}^n$  be normal with mean  $\mu_n$  and variance  $\sigma_n^2$ , so the density function is  $f(x) = (2\pi\sigma_n)^{-1/2} \exp[-(x - \mu_n)^2/2\sigma_n^2]$ . The integral in (8) has two terms. The first is

$$\int_{\log(K/S_t)}^{\infty} S_t e^x f(x) dx = S_t e^{\mu_n + \sigma_n^2} \Phi(d)$$

with

$$d = \frac{\log(S_t/K) + \mu_n + \sigma_n^2}{\sigma_n}.$$

The second is

$$\int_{\log(K/S_t)}^{\infty} K f(x) dx = K \Phi(d - \sigma_n).$$

One version of the Black-Scholes formula is

$$\text{BS} = e^{-rnt^n} \left[ S_t e^{\mu_n + \sigma_n^2} \Phi(d) - K \Phi(d - \sigma_n) \right], \quad (30)$$

which we note for later use. The conventional Black-Scholes formula, equation (8), results from using the arbitrage condition (11) to eliminate  $\mu_n$ .

### A.2 Propositions 1 and 2

**Proposition 1.** The proof has the same steps as the derivation of Black-Scholes, but requires more work in evaluating integrals. The Gram-Charlier density (13) is

$$f(w) = \left( 1 - \frac{\gamma_{1n}}{3!} D^3 + \frac{\gamma_{2n}}{4!} D^4 \right) \varphi(w),$$

where  $w = (x^n - \mu_n)/\sigma_n$  and  $\varphi(w) = (2\pi)^{-1/2} \exp(-w^2/2)$ . Application of (7) thus involves the integral

$$\begin{aligned} \int_{w^*}^{\infty} (S_t e^{\mu_n + \sigma_n w} - K) f(w) dw &= \int_{w^*}^{\infty} (S_t e^{\mu_n + \sigma_n w} - K) \varphi(w) dw \\ &\quad - \frac{\gamma_{1n}}{3!} \int_{w^*}^{\infty} (S_t e^{\mu_n + \sigma_n w} - K) \varphi'''(w) dw \\ &\quad + \frac{\gamma_{2n}}{4!} \int_{w^*}^{\infty} (S_t e^{\mu_n + \sigma_n w} - K) \varphi''''(w) dw \\ &= I_1 - \frac{\gamma_{1n}}{3!} I_2 + \frac{\gamma_{2n}}{4!} I_3, \end{aligned}$$

with  $w^* = (\log(K/S_t) - \mu_n)/\sigma_n$ . The first piece has the same form as Black-Scholes:

$$e^{-r_{nt}n} I_1 = \text{BS}.$$

See (30). The second piece we evaluate by repeated application of integration by parts:

$$I_2 = -\sigma_n K \varphi(w^*)(w^* + \sigma_n) - \sigma_n^3 e^{r_{nt}n} \text{BS} - \sigma_n^3 K \Phi(-w^*).$$

This uses a property of derivatives of the normal density:  $\lim_{x \rightarrow \infty} e^x \varphi^{(n)}(x) = 0$ . The third piece is

$$I_3 = \sigma_n K \varphi(w^*) \left[ (w^*)^2 - 1 + w^* + \sigma_n^2 \right] + \sigma_n^4 e^{r_{nt}n} \text{BS} + \sigma_n^4 K \Phi(-w^*).$$

The call price is therefore

$$\begin{aligned} C_{nt} &= e^{-r_{nt}n} \left( I_1 - \frac{\gamma_{1n}}{3!} I_2 + \frac{\gamma_{2n}}{4!} I_3 \right) \\ &= \text{BS} \left( 1 + \frac{\gamma_{1n}}{3!} + \frac{\gamma_{2n}}{4!} \right) \\ &\quad + \frac{\gamma_{1n}}{3!} \left[ e^{-r_{nt}n} \sigma_n K \varphi(w^*)(w^* + \sigma_n) + e^{-r_{nt}n} \sigma_n^3 K \Phi(-w^*) \right] \\ &\quad + \frac{\gamma_{2n}}{4!} \left[ e^{-r_{nt}n} \sigma_n K \varphi(w^*) \left[ (w^*)^2 - 1 + w^* + \sigma_n^2 \right] + e^{-r_{nt}n} \sigma_n^4 K \Phi(-w^*) \right]. \end{aligned}$$

This formula is exact given the Gram-Charlier expansion. Equation (15) of the proposition results from (i) substitution of  $w^* = \sigma_n - d$ , (ii) application of the arbitrage condition,

$$\mu_n = (r_{nt} - r_{nt}^*)n - \sigma_n^2/2 - \sigma_n^3 \gamma_{1n}/3! - \sigma_n^4 \gamma_{2n}/4!, \quad (31)$$

(iii) substitution using the identity

$$S_t e^{-r_{nt}^*n} \varphi(d) = K e^{-r_n n} \varphi(d - \sigma_n), \quad (32)$$

and (iv) elimination of terms involving  $\sigma_n^3$  and  $\sigma_n^4$ , which are extremely small for options of common maturities. ■

**Proposition 2.** Consider the Black-Scholes formula as a function of implied volatility  $v_n$ . A linear approximation around the point  $v_n = \sigma_n$  is

$$\begin{aligned} C_{nt} &= S_t e^{-r_{nt}^*n} \Phi[d(v)] - K e^{-r_n n} \Phi[d(v) - v] \\ &\cong S_t e^{-r_{nt}^*n} \Phi[d(\sigma_n)] - K e^{-r_n n} \Phi[d(\sigma_n) - \sigma_n] + S_t e^{-r_{nt}^*n} \varphi(d)(v - \sigma_n). \end{aligned}$$

We equate this to the Gram-Charlier price, equation (15), apply (32), and eliminate terms of power two and higher in  $\sigma_n$ . The result is equation (16). ■

### A.3 Stochastic Volatility

We derive moments of  $n$ -period log-changes in the exchange rate for an arbitrary stochastic volatility model, as described in Section 7.

For  $n = 1$ , unconditional moments and cumulants follow from the derivatives of the moment generating function. Taking expectations with respect to  $\varepsilon$  first,

$$\begin{aligned}\phi(s; x) &= E \exp[s\mu_{t+1} + sz_t^{1/2}\varepsilon_{t+1}] \\ &= E_z \exp[s\mu_{t+1} + \frac{s^2}{2!}z_t + \frac{s^3}{3!}z_t^{3/2}\gamma_1(\varepsilon) + \frac{s^4}{4!}z_t^2\gamma_2(\varepsilon) + \dots],\end{aligned}$$

where  $E_z$  denotes the expectation with respect to  $z$ . The moments of  $x$  come from derivatives evaluated at  $s = 0$ . They imply cumulants (see Section 2)

$$\begin{aligned}\kappa_1(x) &= \mu_{t+1} \\ \kappa_2(x) &= E(z) \\ \kappa_3(x) &= \gamma_1(\varepsilon)E(z^{3/2}) \\ \kappa_4(x) &= 3\text{Var}(z) + \gamma_2(\varepsilon)E(z^2).\end{aligned}$$

The standard indicators of skewness and kurtosis are

$$\begin{aligned}\gamma_1(x) &= \gamma_1(\varepsilon)\frac{E(z^{3/2})}{E(z)^{3/2}} \\ \gamma_2(x) &= 3\frac{\text{Var}(z)}{E(z)^2} + \gamma_2(\varepsilon)\frac{E(z^2)}{E(z)^2},\end{aligned}$$

as stated in equation (27).

Conditional higher moments follow from similar methods. If  $\{z_t\}$  and  $\{\varepsilon_t\}$  are independent, as we assumed, the moment generating function is

$$\begin{aligned}\phi(s; x^n) &= E \exp[s \sum_{j=1}^n (\mu_{t+j} + z_{t+j-1}^{1/2}\varepsilon_{t+j})] \\ &= E_z \exp[\sum_{j=1}^n (s\mu_{t+1} + \frac{s^2}{2!}z_{t+j-1} + \frac{s^3}{3!}z_{t+j-1}^{3/2}\gamma_1(\varepsilon) + \frac{s^4}{4!}z_{t+j-1}^2\gamma_2(\varepsilon) + \dots)],\end{aligned}$$

where  $E_z$  now denotes the conditional expectation with respect to the  $z$ 's. Conditional cumulants are

$$\kappa_1(x^n) = \sum_{j=1}^n \mu_{t+j} = \mu_n$$

$$\begin{aligned}
\kappa_2(x^n) &= E_t \sum_{j=1}^n z_{t+j-1} \\
\kappa_3(x^n) &= \gamma_1(\varepsilon) E_t \sum_{j=1}^n z_{t+j-1}^{3/2} \\
\kappa_4(x^n) &= 3 \text{Var}_t \sum_{j=1}^n z_{t+j-1} + \gamma_2(\varepsilon) E_t \sum_{j=1}^n z_{t+j-1}^2.
\end{aligned}$$

This leads to equation (28).

#### A.4 Digital Options

The Gram-Charlier expansion leads to relatively simple expressions for digital options. In the risk-neutral framework of Section 3, the price is

$$D_{nt} = e^{-r_{nt}n} \int_{\log(K/S_t)}^{\infty} f(x) dx.$$

If  $x^n$  is normal (the Black-Scholes case), the price is

$$D_{nt} = e^{-r_{nt}n} \Phi(d - \sigma_n). \quad (33)$$

For the Gram-Charlier expansion, the price is

$$\begin{aligned}
D_{nt} &= e^{-r_{nt}n} \Phi(d - \sigma_n) \\
&\quad - e^{-r_{nt}n} \varphi(d - \sigma_n) \left[ \frac{\gamma_{1n}}{3!} [1 - (d - \sigma_n)^2] + \frac{\gamma_{2n}}{4!} [(d - \sigma_n)^3 - 3(d - \sigma_n)] \right].
\end{aligned}$$

If we use (33) to compute implied volatility, the analog to Proposition 2 is

$$v_n(d) = \sigma_n - \left[ \frac{\gamma_{1n}}{3!} [1 - (d - \sigma_n)^2] + \frac{\gamma_{2n}}{4!} [(d - \sigma_n)^3 - 3(d - \sigma_n)] \right].$$

Note that the skewness and kurtosis terms aren't multiplied by  $\sigma_n$ , as they are in equation (16). For that reason, they can have larger effects.

## References

- Abken, Peter, Dilip Madan, and Sailesh Ramamurtie, 1996, "Estimation of risk-neutral and statistical densities by hermite polynomial approximation: With an application to eurodollar futures options," manuscript, Federal Reserve Bank of Atlanta, June.
- Ait-Sahalia, Yacine, and Andrew Lo, 1997, "Nonparametric estimation of state-price densities implicit in financial asset prices," manuscript, University of Chicago, June.
- Akgiray, Vedat, and Geoffrey Booth, 1988, "Mixed diffusion-jump process modeling of exchange rate movements," *Review of Economics and Statistics* 70, 631-637.
- Bates, David, 1996a, "Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options," *Review of Financial Studies* 9, 69-107.
- Bates, David, 1996b, "Dollar jump fears, 1984-1992: Distributional abnormalities implicit in currency futures data," *Journal of International Money and Finance* 15, 65-93.
- Baillie, Richard, and Tim Bollerslev, 1989, "The message in daily exchange rates: A conditional variance tale," *Journal of Business and Economic Statistics* 7, 297-305.
- Black, Fischer, and Myron Scholes, 1973, "The pricing of options and corporate liabilities," *Journal of Political Economy* 81, 637-654.
- Black, Fischer, 1975, "Fact and fantasy in the use of options," *Financial Analysts Journal* 31 (July-August), 36-72.
- Bodurtha, James, and Georges Courtadon, 1987, *The Pricing of Foreign Currency Options*, Monograph Series in Finance and Economics 1987-4/5, New York: Salomon Center.
- Brenner, Menachem, and Young-Ho Eom, 1997, "No-arbitrage option pricing: New evidence on the validity of the martingale property," manuscript, New York University and Federal Reserve Bank of New York.
- Campa, Jose, and Kevin Chang, 1995, "Testing the expectations hypothesis on the term structure of volatilities," *Journal of Finance* 50, 529-547.

- Campa, Jose, and Kevin Chang, 1996, "Arbitrage-based tests of target-zone credibility: Evidence from ERM cross-rate options," *American Economic Review* 86, 726-740.
- Campa, Jose, Kevin Chang, and Robert Reider, 1997, "Implied exchange rate distributions: Evidence from OTC option markets," manuscript, New York University, August.
- Das, Sanjiv, and Rangarajan Sundaram, 1997, "Taming the skew: Higher-order moments in modeling asset price processes in finance," NBER Working Paper No. 5976.
- Diebold, Francis, 1988, *Empirical Modeling of Exchange Rate Dynamics*, New York: Springer-Verlag.
- Drost, Feike, Theo Nijman, and Bas Werker, 1996, "Estimation and testing in models containing both jumps and stochastic volatility," forthcoming, *Journal of Business and Economic Statistics*.
- Garman, Mark, and Steven Kohlhagen, 1983, "Foreign currency option values," *Journal of International Money and Finance* 2, 231-237.
- Hansen, Bruce, 1994, "Autoregressive conditional density estimation," *International Economic Review* 35, 705-730.
- Hsieh, David, 1989, "Modeling heteroskedasticity in daily foreign exchange rates," *Journal of Business and Economic Statistics* 7, 307-317.
- Jarrow, Robert, and Andrew Rudd, 1982, "Approximate option valuation for arbitrary stochastic processes," *Journal of Financial Economics* 10, 347-369.
- Johnson, Norman, Samuel Kotz, and N. Balakrishnan, 1994, *Continuous Univariate Distributions, Volume 1, Second Edition*, New York: Wiley.
- Jorion, Philippe, 1988, "On jump processes in the foreign exchange and stock markets," *Review of Financial Studies* 1, 427-445.
- Kim, Sangjoon, and Neil Shephard, 1994, "Stochastic volatility: Optimal likelihood inference and comparison with ARCH models," manuscript, Nuffield College, Oxford, March.
- Knight, John, and Stephen Satchell, 1997, "Pricing derivatives written on assets with arbitrary skewness and kurtosis," manuscript, Trinity College, Cambridge, November.

- Kolassa, John, 1994, *Series Approximation Methods in Statistics*, New York: Springer-Verlag.
- Longstaff, Francis, 1995, "Option pricing and the martingale restriction," *Review of Financial Studies* 8, 1091-1124.
- Madan, Dilip, and Frank Milne, 1994, "Contingent claims valued and hedged by pricing and investing in a basis," *Mathematical Finance* 4, 223-245.
- Malz, Allan, 1996, "Option-based estimates of the probability distribution of exchange rates and currency excess returns," manuscript, Federal Reserve Bank of New York, March.
- Melino, Angelo, and Stuart Turnbull, 1990, "Pricing foreign currency options with stochastic volatility," *Journal of Econometrics* 45, 239-265.
- Merton, Robert, 1976, "Option pricing when underlying stock prices are discontinuous," *Journal of Financial Economics* 3, 125-144.
- Rosenberg, Joshua, 1996 "Pricing multivariate contingent claims using estimated risk-neutral density functions," manuscript, New York University, November.
- Rubinstein, Mark, 1976, "The valuation of uncertain income streams and the pricing of options," *Bell Journal of Economics* 7, 407-425.
- Taylor, Stephen, and Xinzhong Xu, 1994, "The magnitude of implied volatility smiles: Theory and empirical evidence for exchange rates," *Review of Futures Markets* 13, 355-380.
- Zhu, Yingzi, 1997, *Three Essays in Mathematical Finance*, Doctoral dissertation submitted to the Department of Mathematics, New York University, May.

**Table 1**  
**Properties of Daily Exchange Rate Changes**

Statistic	Canadian Dollar	German Mark	Japanese Yen
Mean (Annualized)	0.002	0.045	0.062
Std Deviation (Annualized)	0.045	0.117	0.116
Skewness	-0.171	-0.090	0.351
Kurtosis	3.558	2.204	6.936

Statistics pertain to daily depreciation rates,  $x_{t+1} = \log S_{t+1} - \log S_t$ , computed from dollar prices  $S$  of other currencies. The data are from Datastream and cover business days between January 3, 1986 and May 16, 1996 (2704 observations). Means and standard deviations have been multiplied by 260.6 and  $260.6^{1/2}$ , respectively, to convert them to annual units, where 260.6 is the average number of business days in the years 1986-95.

**Table 2**  
**Properties of At-the-Money Volatilities**

Statistic/Maturity	Canadian Dollar	German Mark	Japanese Yen
Mean			
1	5.16	10.52	10.83
2	5.22	10.72	11.18
3	5.27	10.83	11.44
6	5.41	11.00	11.84
12	5.62	11.11	12.10
Median			
1	5.05	9.90	9.98
2	5.30	10.24	10.83
3	5.55	10.50	11.30
6	5.80	11.00	12.07
12	5.98	11.30	12.47
Standard Deviation			
1	1.64	2.94	3.21
2	1.50	2.62	2.88
3	1.40	2.39	2.68
6	1.17	2.05	2.31
12	0.99	1.81	2.00
Autocorrelation			
1	0.948	0.984	0.978
2	0.948	0.988	0.985
3	0.943	0.991	0.988
6	0.922	0.994	0.991
12	0.894	0.995	0.993

Volatilities are quotes from a large international bank, expressed as annual percentages, January 2, 1995, to January 6, 1997 (510 daily observations). See Zhu (1997).

**Table 3**  
**Gram-Charlier Approximations to Jump-Diffusions**

Example	Parameters		Estimates		
	$\tau_0$	$\tau_1$	$\sigma$	$\gamma_1$	$\gamma_2$
A. Benchmark	0.0688	0.0230	0.1000	-0.006	0.898
B. Jump Intensity					
Low ( $\lambda = 5$ )	0.0792	0.0273	0.0997	-0.005	0.748
High ( $\lambda = 15$ )	0.0505	0.0193	0.1002	-0.007	1.046
C. Volatility					
Low ( $\sigma = 0.05$ )	0.0344	0.0155	0.0500	-0.003	0.898
High ( $\sigma = 0.2$ )	0.1375	0.0459	0.1999	-0.012	0.898
D. Maturity					
Three Months	0.0295	0.0302	0.1004	-0.012	1.119
Six Months ( $\gamma_1 = 0.5$ )	0.0295	0.0302	0.1001	-0.009	0.540
E. Kurtosis					
Low ( $\gamma_2 = 0.5$ )	0.0792	0.0193	0.0999	-0.003	0.439
High ( $\gamma_2 = 1.5$ )	0.0595	0.0254	0.1001	-0.009	1.401
F. Skewness					
Low ( $\gamma_1 = -0.1$ )	0.0688	0.0229	0.0997	-0.006	0.897
Medium ( $\gamma_1 = -0.2$ )	0.0688	0.0227	0.0988	-0.006	0.897
High ( $\gamma_1 = -0.5$ )	0.0682	0.0206	0.0896	-0.004	0.801

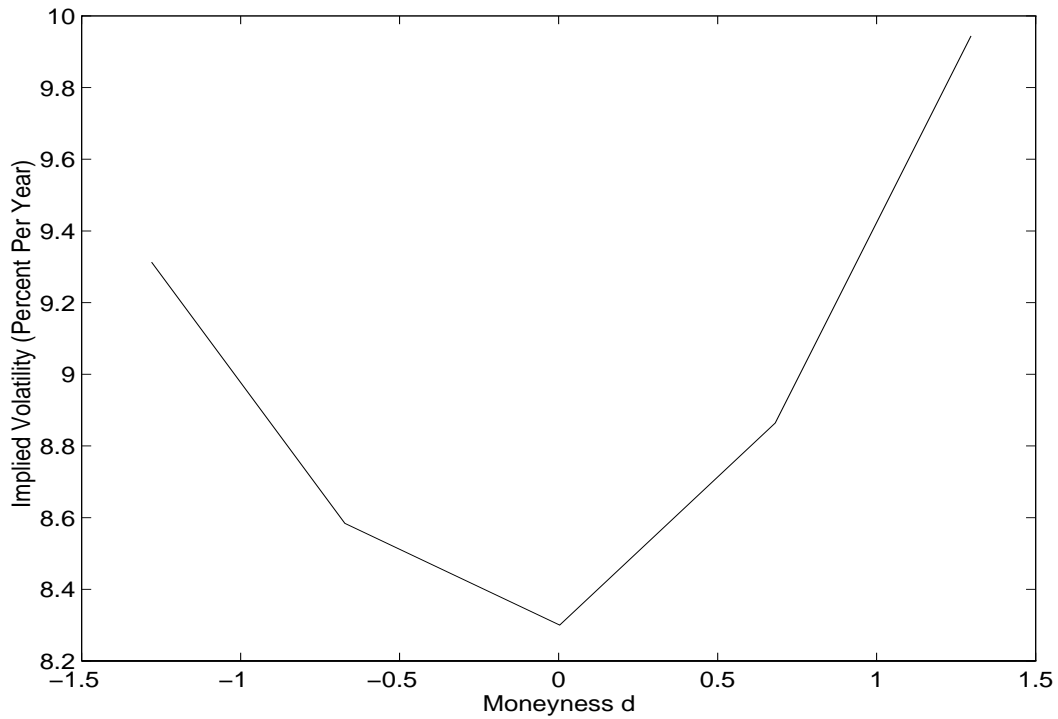
The table reports parameters of jump-diffusion models (annualized standard deviations,  $\tau_0$  and  $\tau_1$ , of non-jump and jump components) and estimates of skewness and kurtosis ( $\gamma_1$  and  $\gamma_2$ ) implied by Gram-Charlier approximations to them based on equation (16). Errors in the approximation are indicated by differences between estimates of skewness and kurtosis and their values in the corresponding jump model. For the benchmark case,  $\gamma_1 = 0$ ,  $\gamma_2 = 1$ , annualized volatility  $\sigma = 0.1$ , annualized jump intensity  $\lambda = 10$ , interest rates are zero, and the maturity of the option is one month. Exceptions are noted in the first column.

**Table 4**  
**Conditional Moments Implied by Volatility Smiles**

Maturity	Standard Deviation	Skewness	Kurtosis
Overnight	0.0849	-0.272	2.912
1 Month	0.0892	-0.168	1.591
2 Months	0.0973	-0.088	1.218
3 Months	0.1018	-0.064	1.033
6 Months	0.1087	-0.038	0.752
9 Months	0.1111	-0.022	0.621
12 Months	0.1116	0.002	0.545
18 Months	0.1117	-0.006	0.447

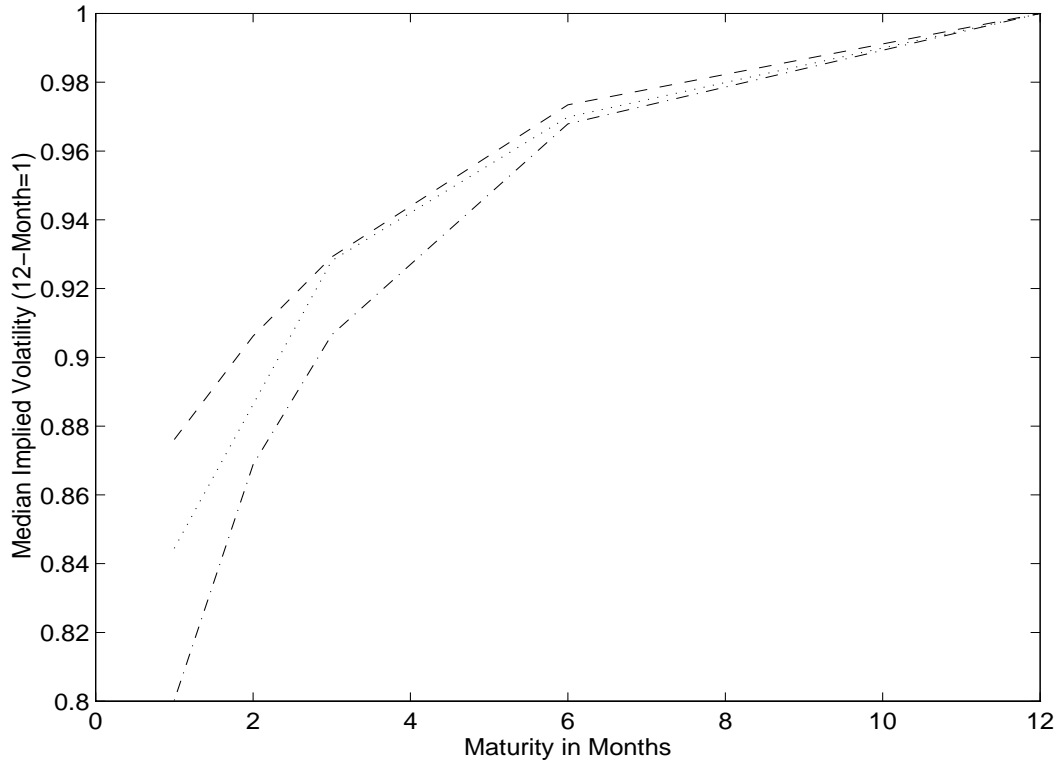
Entries are moments inferred from the volatility smiles in Figure 9, as outlined in Section 4. The standard deviation  $\sigma_n$  is reported in annual units. Skewness and kurtosis are the standard measures,  $\gamma_{1n}$  and  $\gamma_{2n}$ .

**Figure 1**  
An Example of a Volatility Smile



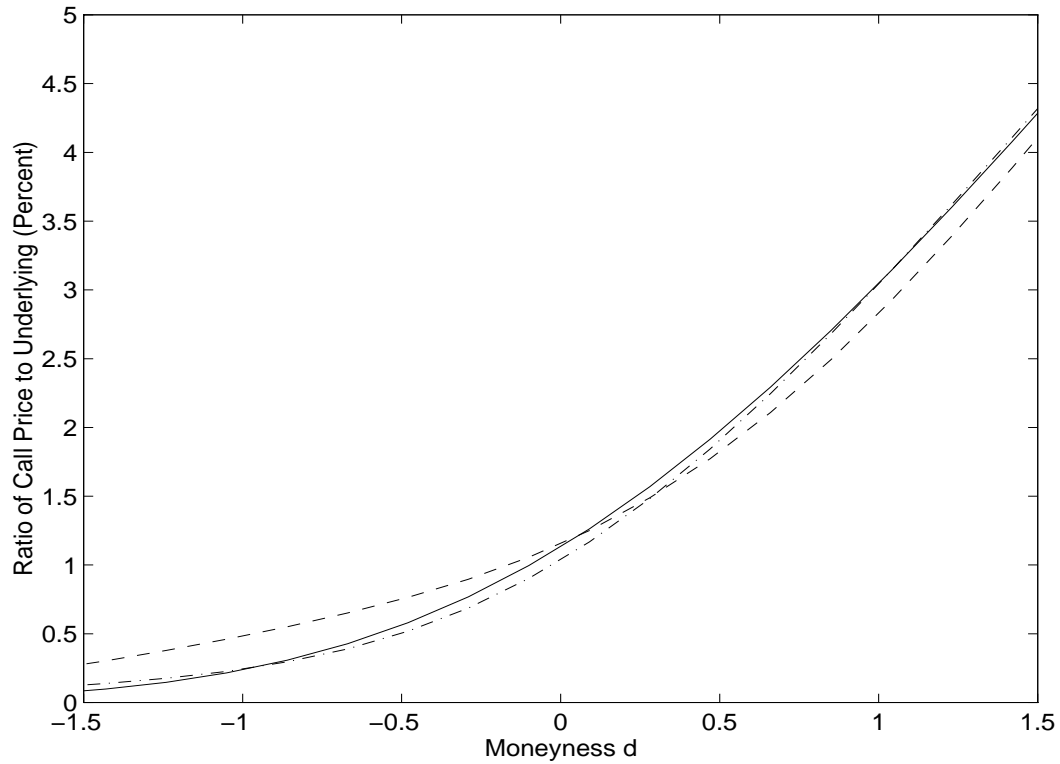
The line plots implied volatility against moneyness for one-month over-the-counter Deutschemark options on April 3, 1996. Moneyness  $d$  is defined by equation (9).

**Figure 2**  
**Median Implied Volatility by Maturity**



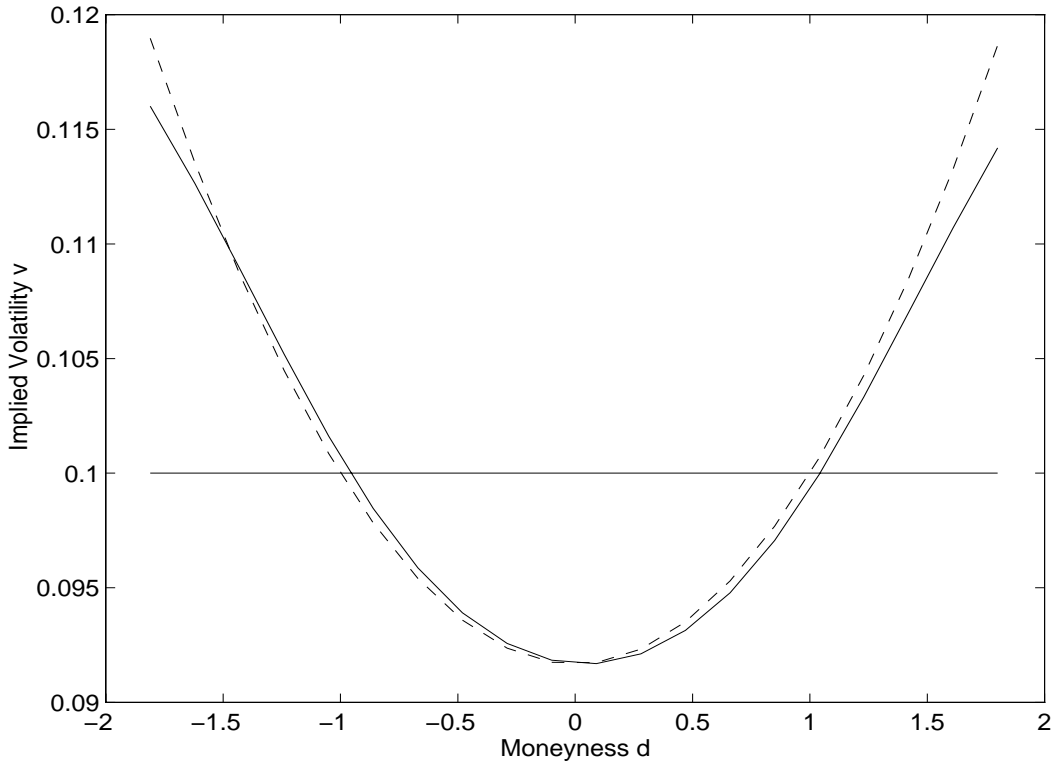
Lines represent median implied volatility quotes for at-the-money options for the Canadian dollar (dotted line), German mark (dashed line), and Japanese yen (dash-dotted line). The lines are scaled to aid comparison: For each currency, median volatility has been divided by its value at 12 months.

**Figure 3**  
**Black-Scholes and Gram-Charlier Call Prices**



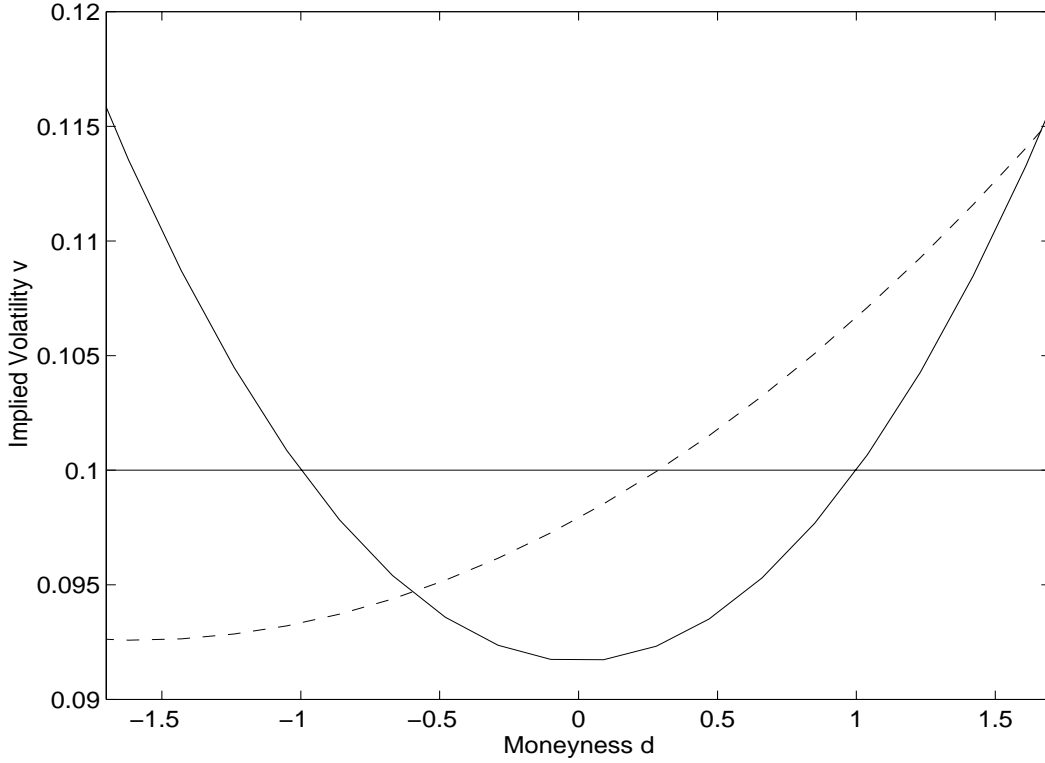
Each line represents the relation between the call price and “moneyness”  $d$  for an underlying distribution of log-changes in the spot exchange rate. The solid line is the Black-Scholes formula, equation (8). The dashed and dash-dotted lines incorporate, respectively, skewness ( $\gamma_{1n} = 2$ ) and kurtosis ( $\gamma_{2n} = 2$ ) into the Gram-Charlier call price formula, equation (15). In each case, we have set  $n = 260.6/12$  (one month),  $\sigma_n = 0.10 * (1/12)^{1/2}$  (0.10 annually), and  $r_{nt} = r_{nt}^* = 0$ .

**Figure 4**  
**Exact and Approximate Volatility Smiles for Gram-Charlier Call Prices**



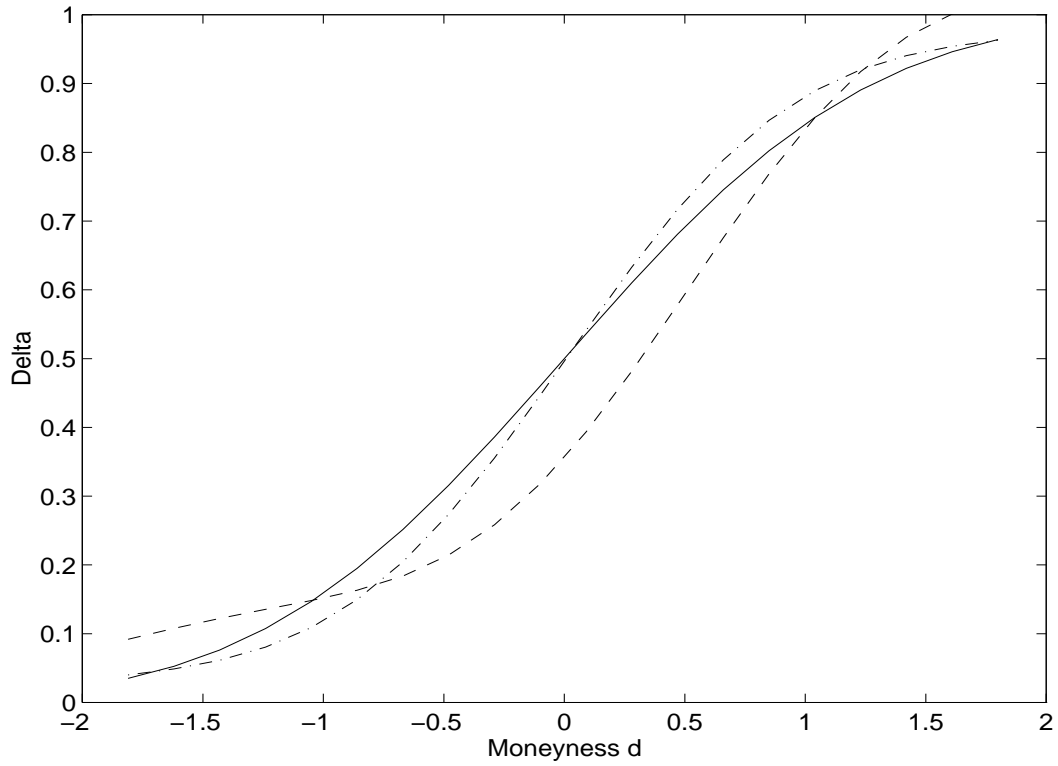
Each line represents the relation between annualized volatility and “moneyness”  $d$  of an option. The solid horizontal line represents the standard deviation  $\sigma_n$ , which is constant. The dashed parabola represents approximate implied volatilities from the Gram-Charlier formula in equation (16) of Proposition 2. The solid curve represents exact implied volatilities based on the Gram-Charlier call price formula, equation (15). Both use  $\gamma_{1n} = 0$ ,  $\gamma_{2n} = 2$ ,  $n = 260.6/12$  (one month),  $\sigma_n = 0.10 * (1/12)^{1/2}$  (0.10 annually), and  $r_{nt} = r_{nt}^* = 0$ .

**Figure 5**  
**Effects of Skewness and Kurtosis on Gram-Charlier Call Prices**



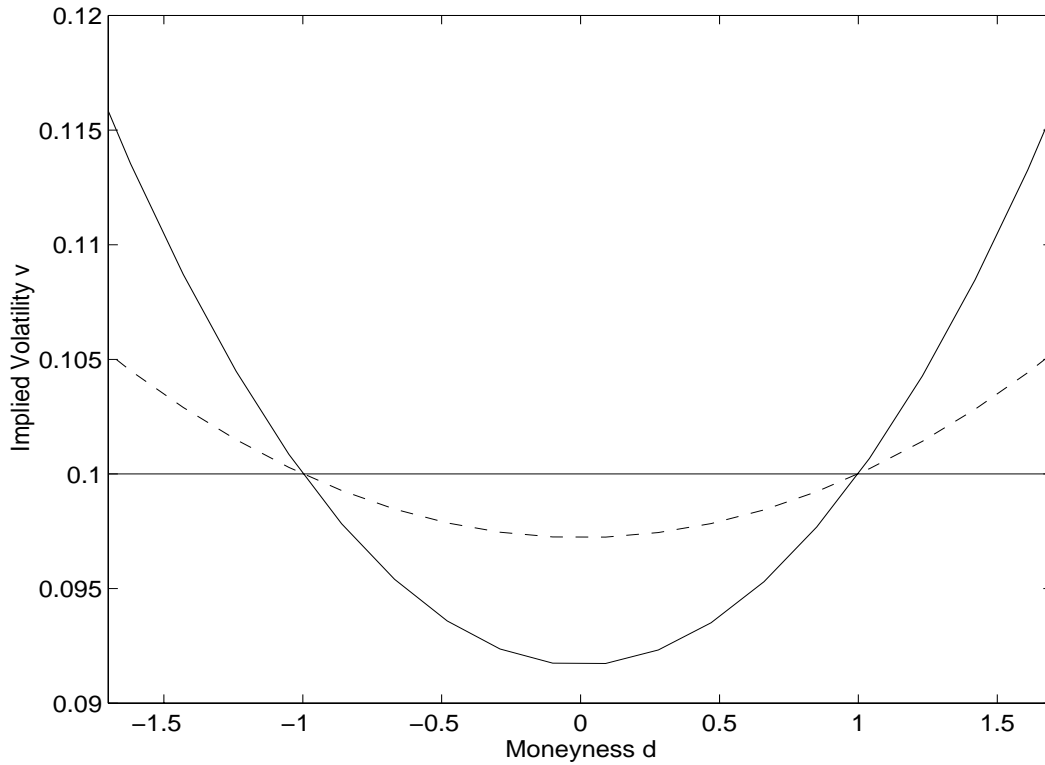
The lines graph approximate annualized implied volatility against moneyness  $d$  for different values of higher moments in the distribution of log-changes in the spot exchange rate. The horizontal solid line represents Black-Scholes. The smile (curved solid line) has positive kurtosis but no skewness ( $\gamma_{1n} = 0, \gamma_{2n} = 2$ ). The skew (dashed line) has negative skewness and positive kurtosis ( $\gamma_{1n} = -0.4, \gamma_{2n} = 0.5$ ). Both are based on the approximation (16). In each case,  $n = 260.6/12$  (one month),  $\sigma_n = 0.10 * (1/12)^{1/2}$  (0.10 annually), and  $r_{nt} = r_{nt}^* = 0$ .

**Figure 6**  
**Deltas for Black-Scholes and Gram-Charlier Call Prices**



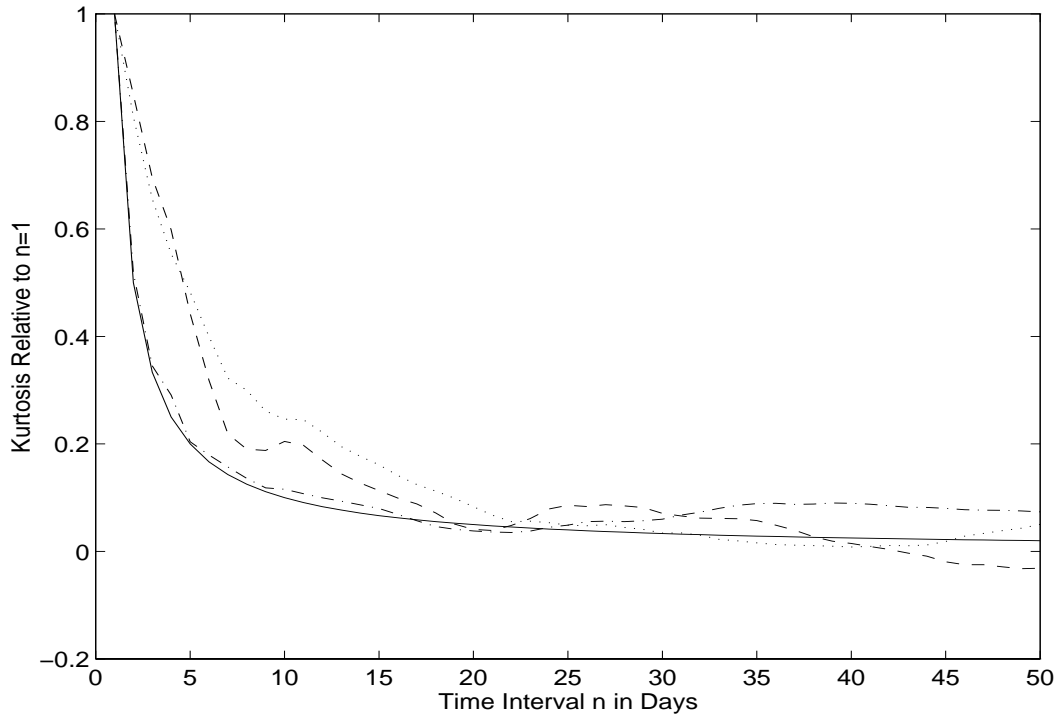
Each line represents the relation between the  $\Delta$  and “moneyness”  $d$  for an underlying distribution of log-changes in the spot exchange rate. The solid line is Black-Scholes:  $\Delta(d) = \Phi(d)$ . The dashed and dash-dotted lines incorporate, respectively, skewness ( $\gamma_{1n} = 2$ ) and kurtosis ( $\gamma_{2n} = 2$ ) into the Gram-Charlier  $\Delta$ , equation (17). Other parameter values are the same as Figure 4.

**Figure 7**  
**Gram-Charlier Volatility Smiles for Two Maturities**



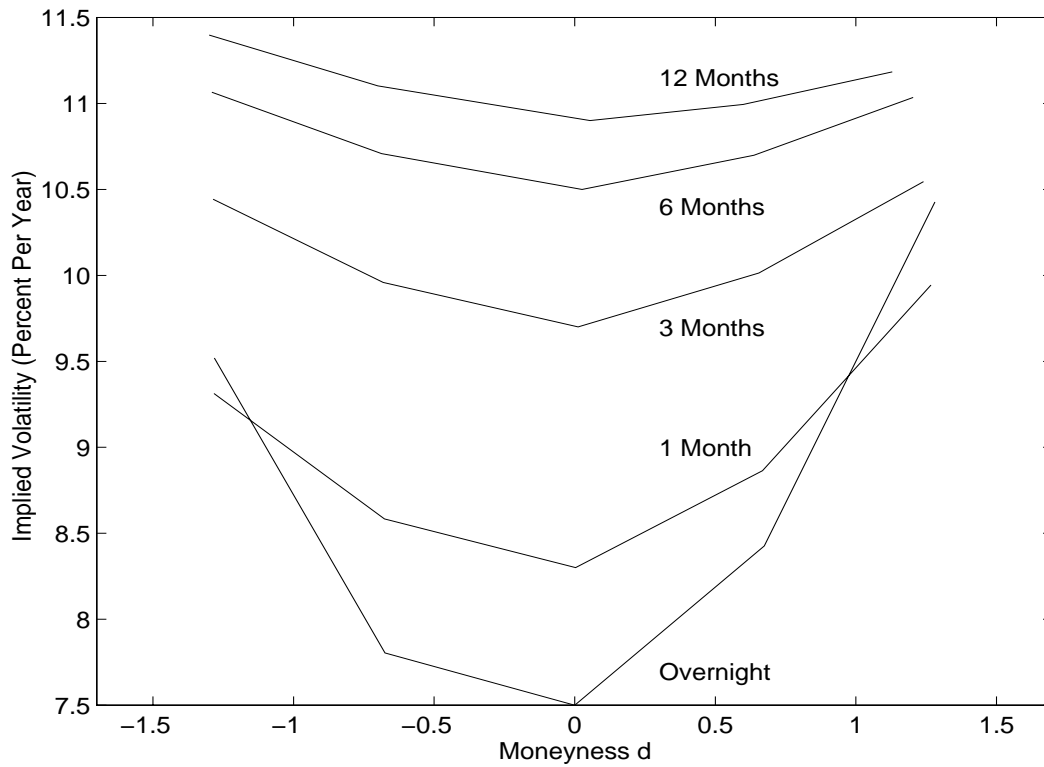
The lines represent volatility smiles for two different maturities. The horizontal solid line represents Black-Scholes. The curved lines have positive kurtosis ( $\gamma_{2n} = 2 \times 260.6 / (12n)$ ) for  $n = 260.6/12$  (one month, solid line) and  $n = 3 \times 260.6/12$  (three months, dashed line). Both are based on the approximation (16). Other parameters are  $\sigma_n = 0.10 * (1/12)^{1/2}$  (0.10 annually), and  $r_{nt} = r_{nt}^* = 0$ .

**Figure 8**  
**Kurtosis in Currency Prices Over Different Time Horizons**



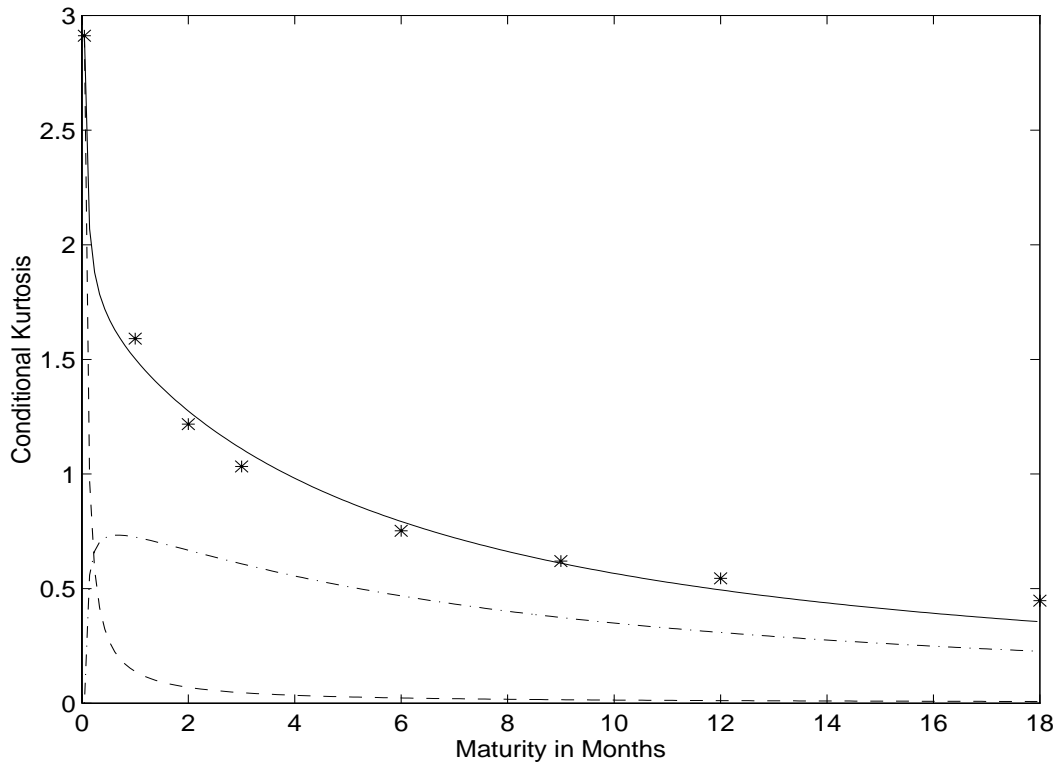
Lines represent estimates of kurtosis ( $\gamma_{2n}$ ) for time intervals of  $n$  days relative to  $n = 1$ . The solid line is  $1/n$  and serves as a theoretical benchmark. The other lines are estimates for the Canadian dollar (dotted line), German mark (dashed line), and Japanese yen (dash-dotted line), all for the same time period covered by Table 1.

**Figure 9**  
**Volatility Smiles for Deutschemark Options on April 3, 1996**



Lines describe volatility smiles for over-the-counter Deutschemark options on April 3, 1996. The maturities of the options are noted in the figure.

**Figure 10**  
**Conditional Kurtosis with Stochastic Volatility**



The figure describes conditional kurtosis v. maturity in the data and in the stochastic volatility model of Section 7. Asterisks represent the data: the values of implied kurtosis reported in Table 4. The solid line is the stochastic volatility model with non-normal innovations. The dashed line represents the same model without stochastic volatility:  $\alpha = 0$  (essentially a pure jump model). The dash-dotted line represents the model with normal innovations:  $\gamma_2(\varepsilon) = 0$  (a gaussian stochastic volatility model).