

**Finance and Economics Discussion Series  
Divisions of Research & Statistics and Monetary Affairs  
Federal Reserve Board, Washington, D.C.**

**Zero Bound, Option-Implied PDFs, and Term Structure Models**

**Don H. Kim**

**2008-31**

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

# ZERO BOUND, OPTION-IMPLIED PDFS, AND TERM STRUCTURE MODELS

DON H. KIM\*

ABSTRACT. This paper points out that several known ways of modeling non-negative nominal interest rates lead to different implications for the risk-neutral distribution of the short rate that can be checked with options data. In particular, Black's boundary models ("interest rates as options") imply a probability density function (pdf) that contains a Dirac delta function and a cumulative distribution function (cdf) that is nonzero at the zero boundary [ $P(K) \propto K$  behavior for put option price for strike  $K$  near 0], while models like the CIR and positive-definite quadratic-Gaussian (QG) models have a zero cdf at the boundary [ $P(K) \propto K^\alpha$ ,  $\alpha > 1$ ]. Eurodollar futures options data are found to favor Black's boundary models: the CIR/QG models, even multifactor versions, have difficulty capturing option prices accurately not only in low interest rate environments but also in higher interest rate environments, and data in early 2008 provide an almost tangible signature of the Dirac delta function in Black's boundary pdf models. Options data also contradict the prediction of well-known models whose cdf is zero at the zero boundary, namely that the risk-neutral pdf is always positively skewed.

## 1. INTRODUCTION

It is well known, at least since Breeden and Litzenberger (1978), that options at a broad range of strikes can provide information about the whole risk-neutral distribution of the underlying security prices, not just mean and variance. Policy makers and market participants have utilized this fact to deduce the risk-neutral distribution of short-term interest rates from eurodollar futures options or federal funds futures options, and many studies have explored techniques that are useful in this regard and discussed their applications to various settings.<sup>1</sup> However, this literature so far has not made much connection with the vast literature on dynamic term structure models, i.e., relatively little work has been done to investigate what the option-implied risk-neutral distribution tells about

---

*Date:* June 12, 2008.

\* Division of Monetary Affairs, Federal Reserve Board. email: Don.H.Kim@frb.gov. I thank Alain Chaboud, Jim Clouse, Darrell Duffie, Benson Durham, Young-Ho Eom, Tony Rodrigues, Jonathan Wright, and the seminar participants at Yonsei University for helpful comments, and Dion Chiu and Scott Konzem for help with eurodollar futures options data. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Board.

<sup>1</sup>See, for example, the papers in BIS (1999).

the underlying stochastic process of interest rates and whether known term structure models can capture the risk neutral distribution implicit in option prices.<sup>2</sup> The present paper is a contribution toward filling this gap.

We shall be focusing in particular on the the zero boundary behavior of the short rate and its ramifications for term structure models. Perhaps the best known model that recognizes the presence of the zero bound is the CIR model, written down by Cox, Ingersoll, and Ross in their seminal paper (1985). This type of model has the feature that the cumulative distribution function (cdf) of the short rate vanishes as the short rate approaches zero. A different treatment of zero-boundary behavior was proposed by Fischer Black in his last paper (1995), in which the nominal short rate is modeled as  $r_t = \max[x_t, 0]$ , where  $x_t$  is a “shadow-rate” process that can go below zero. As we shall see, this type of model implies a short rate cdf that is nonzero at the zero boundary. Proper modeling of the behavior near the zero boundary (e.g., whether it is better described as CIR-like, Black-like, or something else) is a basic and historically important problem in term structure modeling, but it has significance beyond that, as different boundary behaviors could signify different underlying macroeconomic mechanism of the interest rate determination.

In the case of Japan, studies by Gorovoi and Linetsky (2004) and Ueno, Baba, and Sakurai (2006) using yields data give strong support for Black-type models. However, some might argue that this conclusion does not necessarily carry over to the postwar US economy.<sup>3</sup> Indeed, non-negative interest rate modeling with US data has so far been predominantly in terms of models whose short rate cdf is zero at the zero boundary, e.g., the BGM model (Brace, Gutarek, and Musiela (1997)), the multifactor CIR model, the quadratic-Gaussian (QG) model, and the non-affine model of Ahn and Gao (1999).<sup>4</sup> As the US short rate in the past 50 years has not been low enough to see a significant effect of the zero bound on the yield curve,<sup>5</sup> the distinction between the two types of models is difficult to discern from *yields data*. A key point of the present paper is that *data on out-of-money options* can help, as the models may imply strong enough differences in risk-neutral distributions, especially near the zero bound. Recently (early 2008), short-term interest rates have come down to a fairly low level while uncertainty about interest rates has remained relatively high, thus it is particularly interesting to explore what the

---

<sup>2</sup>Hórdahl (2000) and Hórdahl and Vestin (2005) explore the risk-neutral and physical measure probability density functions of certain term structure models, but they do not apply the models to actual options data.

<sup>3</sup>In the US, interest rates seemed to display a Black-like behavior during the Great Depression of 1930s, which may have been one of the motivations for Black’s paper (1995, see p1376). The analysis of the yield curve during that period is complicated by certain institutional effects; see Cecchetti (1988).

<sup>4</sup>Some results in the literature could be construed as an indication of reflection-like behavior. For example, Goldstein and Kierstead (1997) note that according to Ait-Sahalia (1996) the US short rate in the 1973-95 period had (physical-measure) dynamics that was similar to a Brownian motion with two reflecting barriers.

<sup>5</sup>Bomfim (2003) notes that even when the federal funds target rate was at the historical low of 1% in 2003, the effect of the zero bound was not very tangible in the yield curve.

options data from this episode tell about term structure models. Beyond fundamental theoretical interest, modeling of risk-neutral distribution in such environment is an important practical problem for market participants as well.

This paper makes largely two contributions. The first is to explore the implications of well-known and tractable term structure models that respect the non-negativity condition for nominal interest rates (e.g., the multifactor positive-definite CIR/QG models) for the risk-neutral distribution of the short rate. In particular, this paper derives the asymptotic form of their risk-neutral probability density function (pdf) at vanishing interest rates (relevant for the discussion of the behavior near the zero boundary), which does not seem to have been explored in the literature. This paper also develops a method for fast computation of option prices for the risk-neutral distribution implied by the multifactor CIR model and (positive-definite) QG model, and explores how well these models perform in matching the observed out-of-money eurodollar futures option prices. In addition, this paper points out that known models which have zero cdf at the zero boundary imply a fairly strong restriction on the shape of the distribution, namely that the distribution is always positively skewed, and examines whether this prediction is born out in the options data. The second contribution is to examine the implications of Black-type boundary models for option prices and option-implied pdfs. After deriving the risk-neutral pdf of the simplest Black's boundary model ("Black-Vasicek model"), this paper develops flexible parametric models of risk-neutral pdf that are consistent with Black's boundary behavior and lead to convenient calculation of option prices, and explore their practical performance.

The main findings are as follows. A flexible parametric pdf model with Black's boundary behavior (normal-mixture shadow rate model) is found to perform quite well in matching the option prices across various strikes in a variety of interest rate environment from 1998 to 2008. The model produces implied prices of the 1-year-maturity 0.5%-strike eurodollar futures option in February and March of 2008 that agree fairly well with actual settlement prices. On the other hand, the traditional lognormal-mixture pdf model and 2-factor and 3-factor (positive-definite) CIR/QG models substantially underprice this "low-strike" put option. The CIR/QG models have difficulty matching the option prices accurately not only in low interest rate environments but also in higher interest rate environments, often generating large and biased pricing errors. In addition, options data contradict these models' prediction that the skewness of the pdf is always positive.

The remainder of this paper is organized as follows. Section 2 sets up the stage for the later parts of the paper, discussing empirically important zero boundary behaviors and reviewing the risk-neutral pdf extraction techniques. It also introduces a mathematical object called the Dirac  $\delta$ -function, which seldom appears in finance literature but is needed for the discussion of the pdfs with Black's boundary behavior. Section 3 examines the pdfs and option prices in one-factor models, comparing Black's boundary model with CIR/QG models. Section 4 examines richer pdf models and develops useful techniques

for them; the application of these models to actual options data and the empirical results are discussed in Section 5. Section 6 concludes.

## 2. PRELIMINARY DISCUSSION

**2.1. “Zero-cdf model” vs Black’s model.** Nominal interest rates cannot go below zero. The best known model that respects this condition is the CIR model. This model is nested by the so-called CEV model<sup>6</sup>

$$(1) \quad dr_t = \kappa(\theta - r_t)dt + \sigma r_t^\alpha dB_t,$$

which admits two kinds of boundary behavior: when  $\alpha > 1/2$ , the short rate  $r_t$  does not reach zero, i.e., zero is an *inaccessible boundary* (also called unattainable boundary). When  $\alpha < 1/2$ , the short rate  $r_t$  can reach zero; in this case, zero is an *accessible boundary*. The CIR model ( $\alpha = 1/2$ ) can display both behaviors depending on the parameters. More specifically, if the Feller condition

$$(2) \quad k\theta/\sigma^2 > 1/2$$

is satisfied, the zero boundary is inaccessible, and if not, it is accessible.

In the accessible boundary case ( $\alpha < 1/2$ , or  $\alpha = 1/2$  and  $k\theta/\sigma^2 < 1/2$ ), zero is a regular boundary (according to the terminology of Feller (1952)),<sup>7</sup> and additional conditions at the boundary can be imposed. For example, requiring that  $r_t$  stay at zero forever after it hits the zero boundary amounts to having an absorbing boundary. As the permanent stay of the short rate at zero would not be a promising description of the actual economy, we shall not consider the absorbing boundary scenario further in this paper, and instead focus on the regular “unrestricted boundary behavior”, after Longstaff (1992).<sup>8</sup>

Most of the non-negative interest rate models applied to the US data have been either inaccessible boundary models or regular unrestricted boundary models. The former include the CIR model with  $k\theta/\sigma^2 > 1/2$ , the BGM model, and Ahn and Gao (1999)’s non-affine model. The latter include the CIR model with  $k\theta/\sigma^2 < 1/2$  and (positive-definite) quadratic-Gaussian model (Beaglehole and Tenney (1992)). All these models (i.e., both the inaccessible boundary models and regular unrestricted boundary models) share the feature that the cdf of the conditional distribution of the short rate is zero at the zero boundary. Therefore, in this paper we shall refer to these models collectively as “zero-cdf models”, for brevity.

---

<sup>6</sup>This model (in the physical measure) was empirically investigated by Chan et al (1992), who obtained an  $\alpha$  estimate of 1.5.

<sup>7</sup>See, also Karlin and Taylor (1981, Chap 15).

<sup>8</sup>Longstaff (1992) calls the bond pricing for the CIR model without further conditions “the unrestricted equilibrium”. One could consider yet other types of boundary condition, such as those leading to “sticky boundary” or “elastic boundary”, but these yield much less tractable models, and I am not aware of any studies in the finance literature that have explicitly investigated these possibilities.

Let us now discuss Black's boundary models. As the short rate in these models can stay at the zero boundary for an extended period, they have the feature that a finite probability mass is concentrated at the point  $r = 0$ ; this means that the probability density function of the conditional distribution,  $f(r)$ , is infinite at  $r = 0$ . Mathematically,

$$(3) \quad \text{Prob}(r = 0) = \lim_{\epsilon \rightarrow 0^+} \int_{-\epsilon}^{\epsilon} f(r) dr = w > 0, \quad \Rightarrow f(r = 0) = \infty.$$

A natural way to describe such a distribution is to use the Dirac  $\delta$ -function, which is defined as<sup>9</sup>

$$(4) \quad \delta(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{if } x = 0, \end{cases} \quad \int_{-\infty}^{\infty} \delta(x) dx = 1.$$

It has the property  $\int_{-\infty}^{\infty} h(x) \delta(x - a) dx = h(a)$ , where  $h(\cdot)$  is an arbitrary function, and it can be thought of as the pdf of a normal distribution (or any other distribution) with a vanishing variance, i.e.,

$$(5) \quad \delta(x - a) = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2\pi} \epsilon} \exp\left(-\frac{(x - a)^2}{2\epsilon^2}\right).$$

Using the  $\delta$ -function, the pdf of Black's boundary models can be expressed as<sup>10</sup>

$$(6) \quad f(r) = w \delta(r) + \tilde{f}(r)$$

where  $\tilde{f}(r)$  is an improper probability density function with support at  $[0, \infty]$  (i.e.,  $\int_{-\infty}^{\infty} \tilde{f}(r) dr = \int_0^{\infty} \tilde{f}(r) dr = 1 - w < 1$ ).

The cumulative distribution function of the short rate,  $F(r) = \int_{-\infty}^r f(s) ds$ , takes the form

$$(7) \quad F(r) = w + \int_0^r \tilde{f}(s) ds$$

(for  $r > 0$ ). Note that the presence of the  $\delta$ -function in the pdf translates to a nonzero cdf at the zero boundary (more precisely,  $\lim_{r \rightarrow 0^+} F(r) \neq 0$ ).

**2.2. Option-implied risk-neutral pdfs.** Consider a eurodollar futures option with strike  $K$  and maturity  $\tau$ .<sup>11</sup> Let  $\tilde{r}$  denote the futures rate at the maturity of the option,

<sup>9</sup>See, for example, Dirac (1958, p58). One place the  $\delta$ -function appears in the finance literature is in the specification of initial conditions for the Kolmogorov (Fokker-Planck) equations; see, e.g., Goldstein and Kierstead (1997).

<sup>10</sup>Note, incidentally, that the pdf of the absorbing boundary models also contains a delta function.

<sup>11</sup>In the case of eurodollar futures options, the futures price is quoted as  $100 - [\text{futures rate}]$ . Therefore, a put option on *futures contract (futures price)* is a call option on *futures rate*. In this paper, "put option" will always mean put option on the futures rate, and "strike" is in terms of futures rate (not futures price).

and let  $f_\tau(\tilde{r})$  be the risk-neutral pdf of  $\tilde{r}$ . The put option price  $P_\tau(K)$  and the call option price  $C_\tau(K)$  can be expressed in terms of  $f_\tau$  as

$$(8) \quad \begin{aligned} P_\tau(K) &= e^{-r f \tau} \int_{-\infty}^K (K - \tilde{r}) f_\tau(\tilde{r}) d\tilde{r}, \\ C_\tau(K) &= e^{-r f \tau} \int_K^{\infty} (\tilde{r} - K) f_\tau(\tilde{r}) d\tilde{r}, \end{aligned}$$

where the prefactor  $e^{-r f \tau}$  denotes the price of a zero-coupon bond maturing  $\tau$  periods hence.<sup>12</sup> In the case of the eurodollar futures options, the futures rate at the option maturity is the 3-month LIBOR. For simplicity, we shall treat this rate as the short rate of interest, and drop the “tilde” symbol from  $\tilde{r}$ .

Taking successive derivatives of both sides of eq. (8) gives a simple formula for the pdf,

$$(9) \quad f(K) = e^{r f \tau} P''(K) = e^{r f \tau} C''(K),$$

where the double prime notation denotes second derivatives; the maturity index  $\tau$  has been dropped for notational simplicity. In practice, options are available at discrete values of  $K$  (say,  $K_1^p, K_2^p, \dots$  for put options and  $K_1^c, K_2^c, \dots$  for call options), and observed prices could contain some errors. Taking numerical second derivatives (with an interpolated function) to obtain  $f(r)$  typically leads to unrobust results.<sup>13</sup> Therefore, a common practice is to assume a flexible functional form for  $f(r; \gamma)$ , where  $\gamma$  collectively denotes the parameters of  $f$ , and solve the least squares problem

$$(10) \quad \hat{\gamma} = \operatorname{argmin} \zeta(\gamma),$$

$$\zeta(\gamma) = \sum_{i=1}^{n_p} (P^{ob}(K_i^p) - P(K_i^p))^2 + \sum_{i=1}^{n_c} (C^{ob}(K_i^c) - C(K_i^c))^2 + (F^{ob} - F)^2$$

where the superscript *ob* denotes the observed values. Including the futures pricing error helps to pin down the mean of the distribution; note that the model-implied current futures value  $F$  is  $E(r) = \int_{-\infty}^{\infty} r f(r; \gamma) dr$ . This least squares problem amounts to minimizing the pricing errors in absolute terms. Because options far out of money tend to have very low prices, this metric in effect substantially discounts the information in such options, as compared to near-the-money options. One could also consider minimizing the pricing errors in relative terms (i.e., replacing  $(P^{ob} - P)^2$  in eq. (10) by  $(P - P^{ob})^2 / P^{ob2}$ , etc.), but this is not advisable, as the farther out-of-money options

---

<sup>12</sup>More precisely, we have  $P_\tau(K) = E[\exp(-\int_0^\tau r_s ds) \max[0, K - \tilde{r}]]$ , where  $E$  denotes the risk-neutral expectation. Making the approximation  $E[\exp(-\int_0^\tau r_s ds) \max[0, K - \tilde{r}]] \approx E[\exp(-\int_0^\tau r_s ds)] E[\max[0, K - \tilde{r}]]$  (which is common in this literature) gives eq. (8).

<sup>13</sup>To overcome this problem, Shimko (1993) proposed to fit a smooth functional form for the Black-Scholes implied volatility and then take derivatives of the Black-Scholes formula, but this method does not tell much about the tail parts of the pdf.

prices are expected to be less reliable, due to their lower liquidity and more pronounced market microstructural effects (minimum tick sizes, etc.).

There are many possibilities for parametrizing  $f(r; \gamma)$ . The general wisdom is that a flexible parametric form with 4 to 6 parameters is sufficient to match observed option prices at various strikes.<sup>14</sup> A parametrization substantially richer than this can often lead to strange implications at minimal improvements in the fit of option prices (overfitting). Perhaps the most popular parametric form is the “mixture of lognormals”:<sup>15</sup>

$$(11) \quad f(r) = \beta \frac{1}{\sqrt{2\pi} \sigma_1 r} \exp\left(-\frac{(\log r - \mu_1)^2}{2\sigma_1^2}\right) + (1 - \beta) \frac{1}{\sqrt{2\pi} \sigma_2 r} \exp\left(-\frac{(\log r - \mu_2)^2}{2\sigma_2^2}\right).$$

Note that if  $\beta = 1$ , eq. (11) becomes just the lognormal distribution, which is an implicit assumption behind the so-called Black implied-volatility formula (Black (1976)), which is often used to quote cap/floor and swaption prices. Parametric forms like  $f(r)$  in eq. (11) cannot be derived from known classes of term structure models. However, certain parametric forms *can* be derived from a specific class of term structure models. In Section 4.1, we shall discuss such an example, where an implicit parametric form for  $f(r)$  is derived from the multifactor CIR and quadratic-Gaussian models.

It is also worth noting that despite the typically good performance, forms like (11) do have a limitation that is particularly relevant to our problem (pdf modeling in the proximity of the zero bound). Because the lognormal distribution has the feature that  $f(r) \rightarrow 0$  as  $r \rightarrow 0$ , the form (11) can have difficulty when the actual distribution contains a  $\delta$ -function at  $r = 0$ . As noted earlier (eq. (5)), the delta function  $\delta(r)$  can be approximated by a lognormal distribution with mean and variance close to zero, which can be achieved by having a small  $\sigma_1$  and a large and negative  $\mu_1$  in eq. (11), but in such a two-component mixture model this may incur a substantial cost in the fit of other aspects of the pdf. Section 4.2 develops parametric forms for  $f(r)$  that can describe Black’s boundary behavior more easily and naturally.

To have a manageable scope, this paper will be focusing only on the processes and distributions in the *risk-neutral* measure (which determines *pricing*), and not the physical (real world) measure.<sup>16</sup> Therefore, henceforth we shall often drop the adjective “risk-neutral” for brevity.

### 3. ONE-FACTOR MODELS

**3.1. Black-Vasicek model.** Let us now derive the pdf and option prices for the simplest Black’s boundary model, which has the shadow rate described by the one-factor

<sup>14</sup>See, e.g., Jackwerth (2004)’s survey and the papers in BIS (1999).

<sup>15</sup>McManus (1999, included in BIS (1999)) has compared with performance of the lognormal-mixture pdf with other parametric forms in the context of eurodollar futures options.

<sup>16</sup>Nonetheless, the risk-neutral and physical processes take the same (or similar) form in many term structure model specifications, thus many of the qualitative findings in the present paper also apply to the physical measure as well.



Vasicek process:

$$(12) \quad r_t = \max[0, x_t],$$

$$(13) \quad dx_t = \kappa(\theta - x_t)dt + \sigma dB_t.$$

This model, which we shall refer to as the ‘‘Black-Vasicek model,’’ has been theoretically and empirically studied by Gorovoi and Linetsky (2004) and Ueno, Baba, and Sakurai (2006) and others, but its risk-neutral pdf has not been discussed in the literature, to my knowledge.<sup>17</sup>

The conditional distribution of this model, i.e., the transition density  $f(r_{t+\tau}|\mathcal{F}_t)$ , can be easily obtained if one thinks in terms of the shadow rate process  $x_t$  rather than the  $r_t$  process. Note that if  $x_{t+\tau} > 0$ , we have  $r_{t+\tau} = x_{t+\tau}$ ; hence in this case, the distribution of  $r_{t+\tau}$  equals that of  $x_{t+\tau}$ . On the other hand, if  $x_{t+\tau} \leq 0$ , we have  $r_{t+\tau} = 0$ . Thus all scenarios in which  $x_{t+\tau}$  is non-positive maps to a single point  $r_{t+\tau} = 0$ . Therefore, we have

$$(14) \quad f(r_{t+\tau}|x_t) = f^S(x_{t+\tau}|x_t)I(x_{t+\tau}) + w(x_t)\delta(x_{t+\tau}),$$

where  $f^S(x_{t+\tau}|x_t)$  denotes the transition density of the  $x_t$  process,  $w(x_t)$  is the weight of  $f^S(x_{t+\tau}|x_t)$  in the nonpositive region, i.e.,

$$(15) \quad w(x_t) = \int_{-\infty}^0 f^S(x_{t+\tau}|x_t) dx_{t+\tau},$$

and  $I(x)$  is the indicator function ( $I(x) = 1$  for  $x \geq 0$  and  $I(x) = 0$  for  $x < 0$ ). It may be of interest to note that the distribution in eq. (14) can be thought of as a special case of the ‘‘mixture’’ distribution, consisting of a truncated normal distribution  $\frac{1}{1-w(x_t)}f^S(x_{t+\tau}|x_t)I(x_{t+\tau})$  and an infinitely sharp distribution  $\delta(x_{t+\tau})$  with weights  $1 - w(x_t)$  and  $w(x_t)$ , respectively.

The transition density function  $f^S(x_{t+\tau}|x_t)$  for the Vasicek process is well-known: Since the stochastic differential equation (13) has the solution

$$(16) \quad x_{t+\tau} = e^{-\kappa\tau}x_t + (1 - e^{-\kappa\tau})\theta + \int_0^\tau e^{-\kappa(\tau-s)}\sigma dB_s$$

the random variable  $x = x_{t+\tau}$  is conditionally normally distributed<sup>18</sup>

$$(17) \quad x \sim N(\tilde{\mu}, \tilde{\sigma}^2),$$

with conditional mean and variance given by

$$(18) \quad \tilde{\mu} = e^{-\kappa\tau}x_t + (1 - e^{-\kappa\tau})\theta,$$

$$\tilde{\sigma}^2 = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa\tau}).$$

---

<sup>17</sup>Gorovoi and Linetsky (2004) derive an expression for the option prices in the Black-Vasicek model using eigenfunction expansions, but they do not discuss pdfs.

<sup>18</sup>In this paper, the symbol  $\sim$  means ‘‘is distributed as’’.

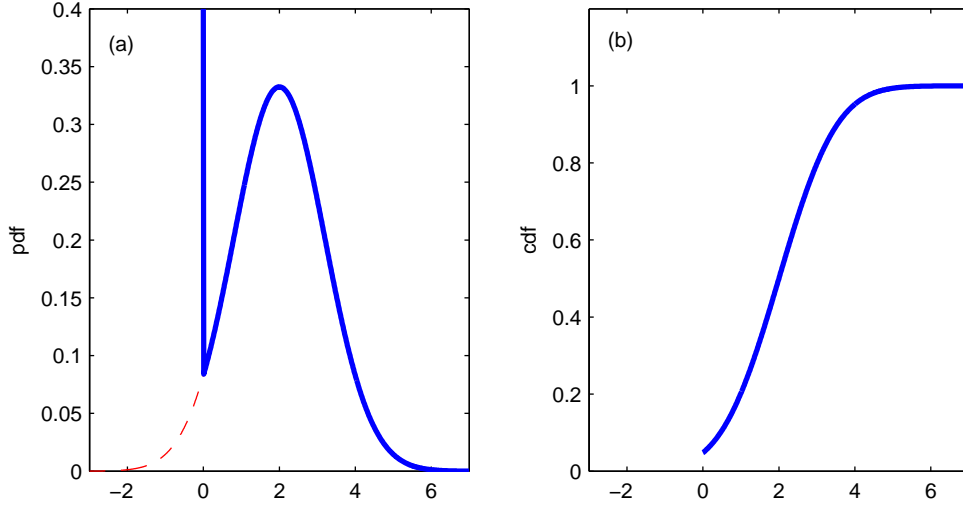


Figure 1: Black-Vasicek model pdf (a) and cdf (b). The vertical line at 0 in Figure (a) pictures a delta function.

Thus, suppressing time indices and simplifying notations, we have the conditional distribution of the short rate given by

$$(19) \quad f(r) = w \delta(r) + \frac{1}{\tilde{\sigma}} \phi\left(\frac{r - \tilde{\mu}}{\tilde{\sigma}}\right) \mathbf{I}(r),$$

where  $w = \Phi(-\tilde{\mu}/\tilde{\sigma})$ , and  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the standard normal pdf and cdf, respectively (i.e.,  $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ ,  $\Phi(x) = \int_{-\infty}^x dt e^{-t^2/2}/\sqrt{2\pi}$ ). Integrating eq. (19) gives the cdf,<sup>19</sup>

$$(20) \quad F(r) = [w + \Phi((r - \tilde{\mu})/\tilde{\sigma}) - \Phi(-\tilde{\mu}/\tilde{\sigma})] \mathbf{I}(r) = \Phi((r - \tilde{\mu})/\tilde{\sigma}) \mathbf{I}(r).$$

The pdf and cdf of the short rate in the Black-Vasicek model are illustrated in Figure 1a,b. The dashed line in Figure 1b is the part of the shadow rate distribution that collapses onto the delta function at  $r = 0$ , denoted by a vertical line at zero in Figure (a). Note also that the cdf  $F(r)$  approaches a finite number as  $r \rightarrow 0^+$  (Figure (b)).

The prices of  $\tau$ -maturity put/call options with strike  $K$  are straightforward to evaluate. Substituting eq. (19) into eq. (8), we have

$$(21) \quad P(K) = e^{-r_f \tau} [w K + (K - \tilde{\mu})(\Phi(k) - \Phi(-\tilde{\mu}/\tilde{\sigma})) + \tilde{\sigma}(\phi(k) - \phi(-\tilde{\mu}/\tilde{\sigma}))],$$

$$(22) \quad C(K) = e^{-r_f \tau} [\tilde{\sigma} \phi(k) - (K - \tilde{\mu})(1 - \Phi(k))],$$

where  $k \equiv (K - \tilde{\mu})/\tilde{\sigma}$ .

<sup>19</sup>Note that the indicator function  $\mathbf{I}(r)$  and the Dirac delta function  $\delta(r)$  are related by the identity  $\frac{d\mathbf{I}(r)}{dr} = \delta(r)$ . Thus, taking the derivative of eq. (20), we recover the expression (19).

**3.2. Two zero-cdf models.** Let us now consider two well-known and tractable positive interest rate models: the CIR model and the QG model. The CIR model is given by

$$(23) \quad dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t.$$

It is well known (CIR, 1985) that the random variable  $r \equiv r_{t+\tau}$  in this model is conditionally distributed (at time  $t$ ) as

$$(24) \quad r \sim b\chi_{\nu,\lambda}^2,$$

where  $\chi_{\nu,\lambda}^2$  denotes the noncentral chi-squared random variable with  $\nu$  degrees of freedom and noncentrality parameter  $\lambda$ , and

$$(25) \quad \begin{aligned} b &= \frac{\sigma^2(1 - e^{-\kappa\tau})}{4\kappa} \\ \nu &= \frac{4\kappa\theta}{\sigma^2} \\ \lambda &= \frac{4\kappa}{\sigma^2(1 - e^{-\kappa\tau})}e^{-\kappa\tau}r_t. \end{aligned}$$

Consider now the positive-definite 1-factor QG model

$$(26) \quad \begin{aligned} r_t &= x_t^2, \\ dx_t &= \kappa(\theta - x_t)dt + \sigma dB_t. \end{aligned}$$

Since the random variable  $x = x_{t+\tau}$  is conditionally normally distributed (eq. (17)), it can be seen immediately that the distribution of random variable  $r = r_{t+\tau}$  is given by eq. (24), i.e., *the same form* as the CIR model, with

$$(27) \quad \begin{aligned} b &= \tilde{\sigma}^2 \\ \nu &= 1 \\ \lambda &= \tilde{\mu}^2/\tilde{\sigma}^2, \end{aligned}$$

where  $\tilde{\mu}$  and  $\tilde{\sigma}$  were given in eq. (18).

The pdf of the  $r \sim b\chi_{\nu,\lambda}^2$  distribution (CIR/QG models) is given by

$$(28) \quad f(r) = \frac{1}{b}g\left(\frac{r}{b}\right)I(r),$$

with

$$(29) \quad g(x) = \frac{1}{2}e^{-(\lambda+x)/2} \left(\frac{x}{\lambda}\right)^{(\nu-2)/4} I_{(\nu-2)/2}(\sqrt{\lambda x}),$$

where  $I_a(\cdot)$  is the modified Bessel function of the first kind of order  $a$ .<sup>20</sup> The option prices and cdf are not available in simple form for this distribution, but it is straightforward to compute them by one-dimensional numerical integrations with  $f(r)$ .

---

<sup>20</sup>See, for example, Johnson, Kotz, and Balakrishnan (1994).

**3.3. Comparing the model implications.** To get a feel for the qualitative difference between Black's boundary models and zero-cdf (CIR/QG) models, it is instructive to compare the behavior of the pdf  $f(r)$  in the  $r \rightarrow 0$  limit.

For the CIR/QG model, expanding the expression (29) in powers of  $r$ , using the well-known series expansion of the modified Bessel function of the first kind

$$(30) \quad I_a(z) = \left(\frac{1}{2}z\right)^a \sum_{j=0}^{\infty} \frac{(z^2/4)^j}{\Gamma(j+1)\Gamma(a+j+1)},$$

gives

$$(31) \quad f(r) = \frac{e^{-\lambda/2}}{(2b)^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} r^{\frac{\nu}{2}-1} + \frac{e^{-\lambda/2}(\lambda-\nu)}{2(2b)^{\frac{\nu}{2}+1}\Gamma(\frac{\nu}{2}+1)} r^{\frac{\nu}{2}} + O(r^{\frac{\nu}{2}+1}),$$

where the notation  $O(r^n)$  denotes terms of order  $n$  or higher.<sup>21</sup> Thus, the cdf is given by

$$(32) \quad F(r) = \frac{e^{-\lambda/2}}{(2b)^{\frac{\nu}{2}}\frac{\nu}{2}\Gamma(\frac{\nu}{2})} r^{\frac{\nu}{2}} + \frac{e^{-\lambda/2}(\lambda-\nu)}{2(2b)^{\frac{\nu}{2}+1}(\frac{\nu}{2}+1)\Gamma(\frac{\nu}{2}+1)} r^{\frac{\nu}{2}+1} + O(r^{\frac{\nu}{2}+2}).$$

From eq. (31) it can be seen that depending on  $\nu$ , we can expect qualitatively different behavior of  $f(r)$ :

$$(33) \quad \lim_{r \rightarrow 0} f(r) = \begin{cases} \infty, & \nu < 2 \\ 0, & \nu > 2. \end{cases}$$

Recall that, for the CIR model (eq. (23)),  $\nu$  is given by eq. (25), therefore, the condition  $\nu > 2$  is the same as Feller condition ( $\kappa\theta > \frac{1}{2}\sigma^2$ ). It makes sense that accessible boundary case (Feller condition violated) shows more probability density in the low  $r$  region than the inaccessible boundary case.

Note also that since  $\nu = 1$  for the QG model, we have  $\lim_{r \rightarrow 0} f(r) = \infty$  for the QG model. As the zero boundary is accessible in the QG model (since  $x_t$  in eq. (26) is unrestricted, it can reach zero), it shows a qualitatively similar small- $r$  behavior as the CIR model with violated Feller condition. It should be noted, however, that an accessible zero boundary does not always imply  $\lim_{r \rightarrow 0} f(r) = \infty$ , as we shall see with the case of the 3-factor QG model (Sec. 4.1). Note also that even in the case  $\lim_{r \rightarrow 0} f(r) = \infty$  (one-factor CIR model with violated Feller condition and the QG model), the cdf is zero at  $r = 0$ , since  $F(r) \propto r^{\nu/2}$  in the  $r \rightarrow 0$  limit.

The small- $r$  behavior of the pdf in eq. (31) implies that the put option price for small strike  $K$  is given by

$$(34) \quad P(K) = e^{-r_f\tau} \int_0^K (K-r)f(r)dr = e^{-r_f\tau} \frac{e^{-\lambda/2}}{(2b)^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} \frac{4}{\nu(\nu+2)} K^{\frac{\nu}{2}+1} + O(K^{\frac{\nu}{2}+1}).$$

---

<sup>21</sup>This result can be also derived using the Laplace transform technique, as shown in Appendix A.

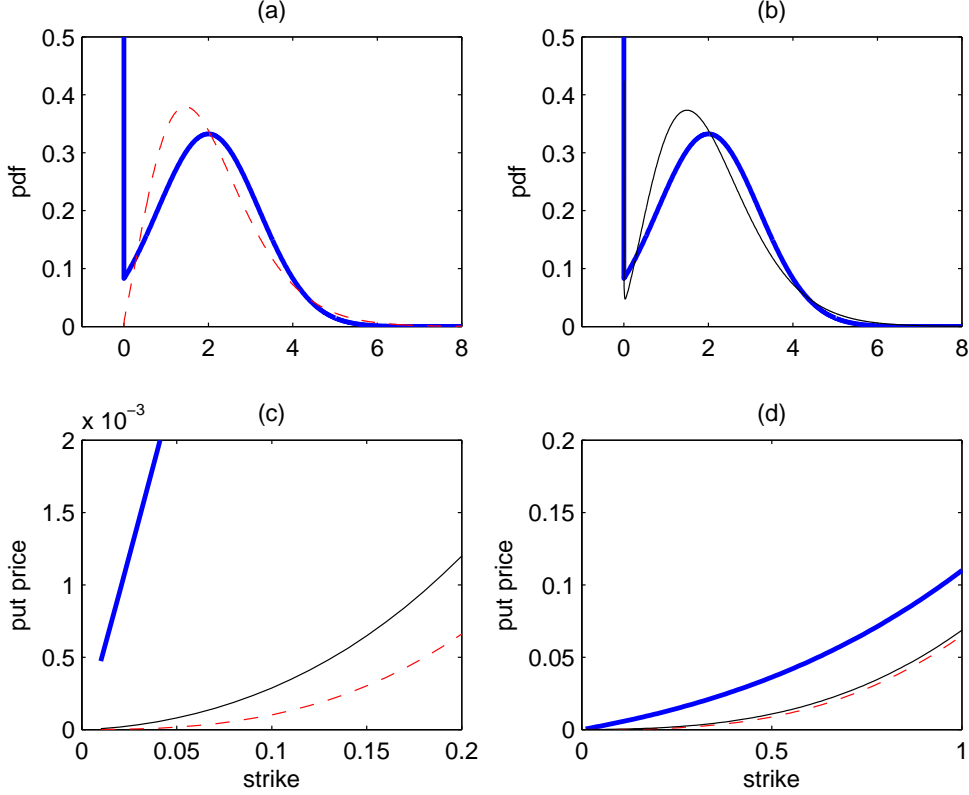


Figure 2: (a) The Black-Vasicek pdf (thick solid line) and inaccessible boundary CIR/QG model pdf (thin dashed line). (b) Accessible boundary model pdf (thin solid line). (c) Put option prices for the Black-Vasicek model (thick solid line), accessible boundary CIR/QG model (thin solid line), and inaccessible boundary model (thin dashed line). (d) Same as (c).

These asymptotic behaviors of CIR/QG models can be compared with the those of the Black-Vasicek model:

$$\begin{aligned}
 (35) \quad f(r) &= w \delta(r) + \frac{\phi(-\tilde{\mu}/\tilde{\sigma})}{\tilde{\sigma}} + O(r^1) \\
 F(r) &= w + \frac{\phi(-\tilde{\mu}/\tilde{\sigma})}{\tilde{\sigma}} r + O(r^2) \\
 P(K) &= e^{-rf\tau} w K + O(K^2).
 \end{aligned}$$

(for  $r, K > 0$ ), which are easily derived from the formulae in Section 3.1.

Figure 2 graphically illustrates the qualitative difference between Black's boundary models and CIR/QG models. Figure 2a plots, in thick solid line, the pdf of the Black-Vasicek model with parameters in eq. (19) given by  $\tilde{\mu} = 2, \tilde{\sigma} = 1.2$ , which imply  $w = 0.0478$ . It also shows, in thin dashed line, the pdf of the  $r \sim b\chi_{\nu,\lambda}^2$  model with  $\nu = 3$  (corresponding to the CIR model satisfying the Feller condition), whose  $b$  and  $\lambda$  were

determined such that the model matches the mean and variance of the Black-Vasicek model;<sup>22</sup> requiring the means and variances of the distributions be similar amounts to having comparable at-the-money option prices. Figure 2b shows the pdf of the  $r \sim b\chi_{\nu,\lambda}^2$  model with  $\nu = 1$  (corresponding to the CIR model violating the Feller condition and the QG model) in thin solid line. As discussed with eq. (31), this pdf has a singularity of the form  $r^{-1/2}$ , but it has a local minimum at a point very close to zero,<sup>23</sup> and above that it looks similar to the pdf of the  $r \sim b\chi_{\nu,\lambda}^2$  distribution with  $\nu = 3$  (i.e., Figure 2a).

Despite the rough similarity in the shape of the pdf of the Black-Vasicek model and the  $r \sim b\chi_{\nu=1,\lambda}^2$  model, the model-implied put prices are drastically different, with the Black-Vasicek model producing far larger numbers, as shown in Figure 2c. This is because the  $r^{-1/2}$  singularity in the  $r \sim b\chi_{\nu=1,\lambda}^2$  model is a much weaker singularity than the  $\delta$ -function singularity in the Black-Vasicek model; to put it simply, *Black's boundary model has a much larger probability weight at or near zero than the CIR/QG model*. Note also that in the low interest rate region the Black-Vasicek model shows a linear dependence on  $K$ , as predicted by eq. (35).

Although the  $r, K \rightarrow 0$  behavior of the pdfs and put option prices derived above for the one-factor Black-Vasicek and CIR/QG models are useful for getting a firm handle on the zero boundary behavior of these models, these predictions might not be cleanly testable empirically, as options might not exist (trade) at very low strikes, and even if they existed their quality may be in doubt. Furthermore, certain market realities (to be discussed at the end of Sec. 5.2) may blur the clean asymptotic behaviors. Still, even if one moves onto regions where the asymptotic behaviors like (34) and (35) are no longer accurate, the basic intuition would still carry over, and one would expect Black's boundary model to produce higher put option prices for small  $K$ 's than zero-cdf models (with comparable mean and variance) when the rates are low enough (or the distribution is wide enough) that the weight of the  $\delta$ -function piece in Black's boundary model is non-negligible. This is illustrated in Figure 2d (which plots the same objects as Figure 2c but on a larger scale). It is also interesting to note that the put option prices for the  $r \sim b\chi_{\nu=3,\lambda}^2$  model (inaccessible boundary) and the  $r \sim b\chi_{\nu=1,\lambda}^2$  model (accessible boundary) are quite similar, beyond the very small  $K$  region.

#### 4. RICHER PARAMETRIC FORMS

The one-factor models considered in the previous section, though instructive, may be too parsimonious to capture option prices accurately for a broad range of strikes. Let us now therefore consider richer versions of zero-cdf models and Black's boundary models.

---

<sup>22</sup>Since the mean and variance of the  $r \sim b\chi_{\nu,\lambda}^2$  distribution are given by  $b(\nu + \lambda)$  and  $2b^2(\nu + 2\lambda)$ , respectively, for a given  $\nu$  ( $= 1, 3$  here) one can easily solve for  $b$  and  $\lambda$ .

<sup>23</sup>The approximate location of this point can be determined by setting the derivative of eq. (31) to zero and solving for  $r$ .

**4.1. Multifactor zero-cdf models.** A non-negative, multifactor version of the CIR model can be constructed by adding up independent CIR factors:

$$(36) \quad \begin{aligned} r_t &= x_{1t} + x_{2t} + \cdots + x_{nt}, \\ dx_{it} &= \kappa_i(\theta_i - x_{it})dt + \sigma_i\sqrt{x_{it}}dB_{it} \end{aligned}$$

where the Brownian motions  $B_{1t}, B_{2t}, \dots$  are independent of each other. This ‘‘multifactor CIR’’ model has been used in many settings, e.g., Duffie and Singleton (1997) and Feldhütter and Lando (2007).<sup>24</sup>

Because  $x_i$ ’s are independent, the results about the 1-factor CIR model carries over easily; it is straightforward to show that the distribution of  $r = r_{t+\tau}$  is given by

$$(37) \quad r \sim b_1\chi_{\nu_1, \lambda_1}^2 + b_2\chi_{\nu_2, \lambda_2}^2 + \cdots + b_n\chi_{\nu_n, \lambda_n}^2, \quad (b_i > 0),$$

where  $\chi_{\nu_i, \lambda_i}^2$ ’s are independent noncentral chi-squared random variables with  $\nu_i$  degrees of freedom and noncentrality parameter  $\lambda_i$ , and

$$(38) \quad \begin{aligned} b_i &= \frac{\sigma_i^2(1 - e^{-\kappa_i\tau})}{4\kappa_i} \\ \nu_i &= \frac{4\kappa_i\theta_i}{\sigma_i^2} \\ \lambda_i &= \frac{4\kappa_i e^{-\kappa_i\tau}}{\sigma_i^2(1 - e^{-\kappa_i\tau})}x_{it}. \end{aligned}$$

Let us now consider the general  $n$ -factor positive-definite QG model,

$$(39) \quad \begin{aligned} r_t &= (\mathbf{x}_t - \alpha)^T \Psi (\mathbf{x}_t - \alpha), \\ d\mathbf{x}_t &= \mathcal{K}(\theta - \mathbf{x}_t)dt + \Sigma dB_t, \end{aligned}$$

where the superscript T denotes matrix transpose;  $\mathbf{x}_t$  is an  $n$ -dimensional vector  $\mathbf{x}_t = [x_{1t}, \dots, x_{nt}]^T$ ;  $\mathcal{K}$  and  $\Sigma$  are  $n \times n$  constant matrices;  $\alpha$  and  $\theta$  are  $n$ -dimensional constant vectors, and  $B_t$  is an  $n$ -dimensional vector of standard Brownian motions. In order to insure the non-negativity of  $r_t$ , we require  $\Psi$  to be a symmetric positive-definite matrix.

It is straightforward to show that the conditional distribution of the short rate in the positive-definite QG model also takes the form (37),<sup>25</sup> with

$$(40) \quad \begin{aligned} b_i &= d_i \quad (> 0) \\ \nu_i &= 1 \\ \lambda_i &= ([\mathcal{P}^T C^{-1}(\tilde{\mu} - \alpha)]_i)^2, \end{aligned}$$

<sup>24</sup>The multifactor CIR model belongs to Dai and Singleton (2000)’s  $A_{j=n}(n)$  affine model classification. Other kinds of multifactor affine models (the  $A_{j<n}(n)$  models) are not considered in this paper, as they do not respect the non-negativity condition for nominal interest rates. The concluding section of this paper, however, does make some comments relevant to these models.

<sup>25</sup>For a derivation, see Ahn, Dittmar, and Gallant (2002, Appendix B).

where  $C$ ,  $\tilde{\mu}$  are given by

$$(41) \quad \tilde{\mu} = e^{-\mathcal{K}\tau} \mathbf{x}_t + (I - e^{-\mathcal{K}\tau})\theta,$$

$$(42) \quad C = \text{Chol}(\tilde{\Omega}), \quad (\text{i.e., } \tilde{\Omega} = C C^T),$$

$$(43) \quad \tilde{\Omega} = \int_0^\tau e^{-\mathcal{K}(\tau-s)} \Sigma \Sigma^T e^{-\mathcal{K}^T(\tau-s)} ds,$$

(the notation  $e^X$  denoting the matrix exponential), and  $d_i$  and  $\mathcal{P}$  are from the diagonalization of the symmetric positive-definite matrix  $C^T \Psi C$ :

$$(44) \quad C^T \Psi C = \mathcal{P} D \mathcal{P}^{-1} = \mathcal{P} D \mathcal{P}^T,$$

where  $D$  is a diagonal matrix with diagonal elements  $d_1, d_2, \dots, d_n$  (the eigenvalues of the  $C^T \Psi C$  matrix). Note that the dependence of  $\lambda_i$  on  $\mathbf{x}_t$  has been suppressed in the notation.

Models like the multifactor CIR model have been criticized for the restrictive nature of the factor correlation structure. However, the positive-definite QG models have been widely believed to accommodate a rich factor correlation structure, due to the ability to specify  $\mathcal{K}$  and  $\Sigma$  flexibly; see, e.g., Ahn, Dittmar, and Gallant (2002). Therefore, it bears emphasizing that not only the multifactor CIR model but also the positive-definite QG model has a short rate distribution given by a positive linear combination of independent noncentral  $\chi^2$  random variables, and that *this holds no matter how flexible the model is and no matter how many factors it has*.

The pdf  $f(r)$  and cdf  $F(r)$  of the distribution for the multifactor CIR/QG model (eq. (37)) do not have simple closed-form expressions in general, but the characteristic function  $\hat{f}(w) \equiv E(e^{i\omega r})$  has a simple expression

$$(45) \quad \hat{f}(w) = \prod_{j=1}^n (1 - 2b_j i\omega)^{-\nu_j/2} \exp\left(\frac{\lambda_j b_j i\omega}{1 - 2b_j i\omega}\right),$$

due to the independence of the noncentral  $\chi^2$  variables in eq. (37). From this,  $f(r)$  and  $F(r)$  can be computed by evaluating the well-known Fourier-inversion integrals (see, e.g., Stuart and Ord (1994, Chap. 4)),

$$(46) \quad f(r) = \frac{1}{\pi} \int_0^\infty \text{Re}(e^{-i\omega r} \hat{f}(\omega)) d\omega$$

$$(47) \quad F(r) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Im}(e^{i\omega r} \hat{f}(-\omega))}{\omega} d\omega.$$

The option prices  $P(K)$  and  $C(K)$  for this model (integrals (8)) are also not available in closed-form, but they can be expressed in terms of certain cumulative distribution functions. For the computation of pdfs from options data (eq. (10)), it is desirable to have a formula that allows fast computation of option prices. Appendix B derives such a formula, using the celebrated saddle-point formula of Lugannani and Rice (1980).



Let us now consider the small- $r$  behavior of  $f(r)$ . Appendix A shows that

$$(48) \quad f(r) = \frac{e^{-\Sigma_i \lambda_i / 2}}{\prod_i (2b_i)^{\nu_i / 2}} \left( \frac{1}{\Gamma(\frac{1}{2} \Sigma_i \nu_i)} r^{\frac{1}{2} \Sigma_i \nu_i - 1} + \frac{\Sigma_i (\lambda_i - \nu_i) / 4b_i}{\Gamma(\frac{1}{2} \Sigma_i \nu_i + 1)} r^{\frac{1}{2} \Sigma_i \nu_i} + \dots \right).$$

This form has an interesting dependence on the number of factors: the exponent in the leading term increases with number of factors. For the QG model, we have

$$(49) \quad f(r \rightarrow 0) = \begin{cases} \infty, & n = 1, \\ \text{finite}, & n = 2, \\ 0, & n = 3. \end{cases}$$

This pattern is due to the fact that when there are more factors it is less likely for  $r$  to approach 0 since all of the factors have to become small. Therefore, higher-factor QG models are even more different from Black's boundary models insofar as the zero-boundary behaviors are concerned.

To get further insights into the distribution described by eq. (37), it is useful to examine its cumulants (hence moments), which is straightforward to calculate from the cumulant generating function ( $\log(\hat{f}(-i\omega))$ ). The mean, variance, and the third central moment of this distribution are easily derived from the cumulants of the noncentral  $\chi^2$  variables (see, e.g., Johnson, Kotz, Balakrishnan (1994, p447)):

$$(50) \quad \begin{aligned} E(r) &= \Sigma_{i=1}^n b_i (\nu_i + \lambda_i) \\ Var(r) &= \Sigma_{i=1}^n 2 b_i^2 (\nu_i + 2\lambda_i) \\ E[(r - E(r))^3] &= \Sigma_{i=1}^n 8 b_i^3 (\nu_i + 3\lambda_i). \end{aligned}$$

It can be seen that because  $b_i, \nu_i, \lambda_i > 0$  for all  $i$ , the skewness  $E[(r - E(r))^3] / Var(r)^{3/2}$  is always positive.

More generally speaking, one cannot create a negatively skewed distribution by forming a positive linear combination of independent, positive-skewed random variables. Therefore, a model like

$$(51) \quad r_t = b_1 x_{1t} + b_2 x_{2t} + \dots + b_n x_{nt}, \quad (b_i > 0),$$

where  $x_{it}$  is Ahn and Gao (1999)'s one-factor process for the short rate, would again have the feature that the short rate distribution is always positively skewed.<sup>26</sup>

**4.2. Flexible pdfs with Black's boundary behavior.** This section develops flexible parametric forms for the risk-neutral pdf that are consistent with Black's boundary behavior. The earlier discussion with the Black-Vasicek model points to the way: write

---

<sup>26</sup>I have not seriously attempted to prove that the conditional distribution of the short rate in Ahn and Gao (1999)'s model [ $r_t = 1/x_t$ ,  $dx_t = \tilde{\kappa}(\tilde{\theta} - x_t)dt + \tilde{\sigma}\sqrt{x_t}dB_t$ , ( $\tilde{\kappa}, \tilde{\theta} > 0$ ,  $2\tilde{\kappa}\tilde{\theta} > \tilde{\sigma}^2$ )] is positively skewed for all feasible values of the parameters, but I believe so, as the right tail of the distribution is thicker and longer than the left tail. (It can be shown that  $f(r) \propto r^{-a}e^{-b/r}$  ( $a > 7/4$ ,  $b > 0$ ) for  $r \rightarrow 0$ , and  $f(r) \propto r^{-a}$  ( $a > 2$ ) for  $r \rightarrow \infty$ .)

down a flexible form for the shadow rate distribution  $f^S$  that has some weight below zero, and take

$$(52) \quad f(r) = f^S(r)\mathbf{I}(r - \underline{r}) + w\delta(r - \underline{r}), \quad w = \int_{-\infty}^{\underline{r}} dx f^S(x).$$

Here we have a slightly more general expression in which the lower boundary  $\underline{r}$  is not necessarily zero.

The distribution  $f^S$  can be modeled with similar degree of parsimony as the pdfs that are in use in the extant literature. A natural candidate for  $f^S$  is the ‘‘mixture of normals’’ form, i.e.,

$$(53) \quad f^S(r) = \beta_1 \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(r - \mu_1)^2}{2\sigma_1^2}\right) + \beta_2 \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(r - \mu_2)^2}{2\sigma_2^2}\right),$$

where  $\beta_2 = 1 - \beta_1$ . Another possibility is the Gram-Charlier/Edgeworth expansion

$$(54) \quad f^S(r) = \frac{1}{\sigma} \phi\left(\frac{r - \mu}{\sigma}\right) \left(1 + \frac{\gamma_3}{6} H_3\left(\frac{r - \mu}{\sigma}\right) + \frac{\gamma_4}{24} H_4\left(\frac{r - \mu}{\sigma}\right) + \dots\right),$$

where  $H_j$  are Hermite polynomials ( $H_3(x) = x^3 - 3x$ ,  $H_4(x) = x^4 - 6x^2 + 3$ , etc.). This expansion differs from the Gram-Charlier/Edgeworth expansion commonly used in the option-implied pdf literature by the fact that it uses unit *normal* variable, rather than unit *lognormal* variable. Note that in both eq. (53) and (54),  $f^S$  is allowed to have some probability weight below  $r = 0$ .

These forms, which are richer than that of the Black-Vasicek model (eq. (19)), can be viewed as accommodating a more complicated process for the shadow rate, e.g., a multifactor model with stochastic volatility.

The formula for  $P(K)$ ,  $C(K)$  and  $F$  for the mixture-of-normals shadow rate model (eq. (53)) and the Gram-Charlier/Edgeworth shadow rate model (eq. (54)) are provided in Appendix C.

## 5. EMPIRICAL EVIDENCE

**5.1. Empirical strategy.** To examine how well zero-cdf models perform in matching option prices, I have implemented two versions of pdf models, which we shall refer to as the QG3 model and the CIR2 model:

$$(55) \quad \text{QG3 : } r \sim b_1 \chi_{1,\lambda_1}^2 + b_2 \chi_{1,\lambda_2}^2 + b_3 \chi_{1,\lambda_3}^2,$$

$$(56) \quad \text{CIR2 : } r \sim b_1 \chi_{\nu_1,\lambda_1}^2 + b_2 \chi_{\nu_2,\lambda_2}^2,$$

( $b_i, \nu_i, \lambda_i > 0$ ). For Black’s boundary model, I have implemented the model in which the shadow rate distribution is that of the mixture of normals, i.e., eq. (53), which we shall refer to as the B-MN model.<sup>27</sup> For comparison with a popular benchmark pdf, I have

<sup>27</sup>I have also explored empirically the Black’s boundary model with Gram-Charlier/Edgeworth shadow rate distribution (eq. (54)), and obtained results that are similar to the B-MN model. For brevity, only the B-MN model results will be reported in this paper.

model	pdf parameters	description
QG3	$b_1, b_2, b_3, \lambda_1, \lambda_2, \lambda_3$	3-factor QG(CIR) models.
CIR2	$b_1, b_2, \nu_1, \nu_2, \lambda_1, \lambda_2$	2-factor CIR(QG) models.
B-MN	$\beta_1, \mu_1, \mu_2, \sigma_1, \sigma_2$	Black's boundary; mixture of normals.
MLN	$\beta_1, \mu_1, \mu_2, \sigma_1, \sigma_2$	Mixture of lognormals.

Table 1: Summary of the models. The pdf parameters for the QG3 and CIR2 models are defined in eqs. (55) and (56). The pdf parameters for the MLN and B-MN models are defined in eqs. (11) and (53), respectively.

also computed the pdf based on the mixture of lognormals, i.e., eq. (11), which we shall denote as the MLN model. These four models are summarized in Table 1.

The QG3 model covers all possible 3-factor positive-definite QG models, but it also includes a special case of the 3-factor CIR model ( $\kappa_i \theta_i = \frac{1}{4} \sigma_i^2$ ). The CIR2 model covers all possible cases of the 2-factor CIR models, but it also nests the 2-factor QG model. This model also nests the Longstaff-Schwartz (1992) model, which was regarded by at least Hórdahl (2000) as a promising model for describing the short rate distribution. In terms of the number of parameters, the QG3 and CIR2 models are richer than the B-MN and MLN models by one parameter.

It is also worth noting that, because the QG3 and CIR2 models originate from specific term structure models, the pdf parameters  $b_i, \nu_i, \lambda_i$  ( $\nu_i = 1$  for QG3) are linked to the elementary parameters of the term structure model and the state variables  $x_{it}$ .<sup>28</sup> Therefore the pdf parameters can be obtained from an estimated term structure model. However, there are many ways of estimating term structure models,<sup>29</sup> which could lead to different estimated elementary parameters and state variables, and in turn, different pdfs. Therefore, in this paper we shall simply treat the pdf parameters for the QG3 model and the CIR2 model as free parameters to be determined in the optimization problem (10), in the same way as the pdf parameters for the B-MN and MLN models are determined. This way can be viewed as giving the QG3 and CIR2 model a maximum chance to fit options data.<sup>30</sup>

Lastly, note that the estimation of multifactor CIR and QG models in the literature have often included a constant term in the short rate, i.e.,  $r_t = c + x_{1t} + \dots + x_{nt}$  and

<sup>28</sup>Recall, from eqs. (38) and (40), that  $\lambda_i$  in the multifactor CIR/QG models depend on  $x_{it}$ .

<sup>29</sup>For example, one could use just the current yield curve information (daily recalibration) or use historical yields data to estimate the model in a time-consistent way. The estimation technique itself could be of many variety (e.g., MLE, GMM, etc.). One could also opt or opt not to use options data in the estimation.

<sup>30</sup>This is analogous to the daily recalibration of term structure models that is often done in finance industry. Note that after setting the left hand side of eq. (38) or eq. (40) to the pdf parameters determined this way, one can search for the elementary parameters and state variables that best fit the yield curve, but the yield curve fit may be poorer than the unconstrained case.

$r_t = c + (\mathbf{x}_t - \alpha)^T \Psi (\mathbf{x}_t - \alpha)$ . Empirically, Duffie and Singleton (1997) obtained a negative  $c$  for the multifactor CIR model, but in that case it is no longer a non-negative interest rate model. Ahn, Dittmar, and Gallant (2002) obtained a positive  $c$  in their estimation of the QG model; in this case, there is no probability weight in the region  $[0, c]$ . In any case, the inclusion of the constant term does not change major qualitative behaviors of these models, including the prediction that the pdfs are always positively skewed.

**5.2. Data.** For the empirical investigation, we shall confine attention to the eurodollar futures option with maturity of about 1 year. The 1-year maturity was chosen so as to have a horizon that is reasonably short but at the same time long enough to have a sizeable width of the distribution.

To get a sense of the performance of the models in a variety of contingencies, I have computed the pdfs for the above four models for every last business day of the month from May 1998 to March 2008. Besides these monthly data, I also have computed the pdfs from a second dataset consisting of daily data from Feb 4, 2008 to April 28, 2008. As we shall see, this period is particularly interesting as regards the zero boundary behavior question.

The options data used in this paper are from CME (Chicago Mercantile Exchange). The prefactor  $e^{-r_f \tau}$  in eq. (8) was computed using the 1-year LIBOR. The CME eurodollar futures options are American options. I have converted their price into corresponding European price by subtracting the early exercise premium computed based on Baroni-Adesi and Whaley (1987)'s formula; these corrections tended to be fairly small.

The eurodollar futures options markets have fixed maturity dates (identical to maturity dates for the eurodollar futures contracts, 4 days a year); for each date in the sample, I chose the option maturity that is closest to 365.25 days. This creates a saw-tooth behavior of the time series of option maturities (oscillating between approximately  $\frac{7}{8}$  and  $\frac{9}{8}$  year) for the monthly sample.

The options are typically available at strikes with multiples of 25 basis points, but sometimes there are options with strikes at odd multiples of 12.5 basis points, which are dropped as they tend to be not as liquid as the even multiples. Also, in order to avoid illiquid options, only the options with strikes at which there is a nonzero open interest are used. Thus the number of call and put options used in the pdf computations varies over time.

Figure 3 shows the minimum put strike and the maximum call strike that were available on each day of the monthly sample. Also shown is the put strike rate that corresponds to the current futures rate minus  $1.5\sigma^*$ , and the call strike rate that corresponds to the futures rate plus  $1.5\sigma^*$ , where the proxy  $\sigma^*$  for the standard deviation of the pdf is computed as

$$(57) \quad \sigma^* = \sqrt{\tau} \sigma^B F,$$

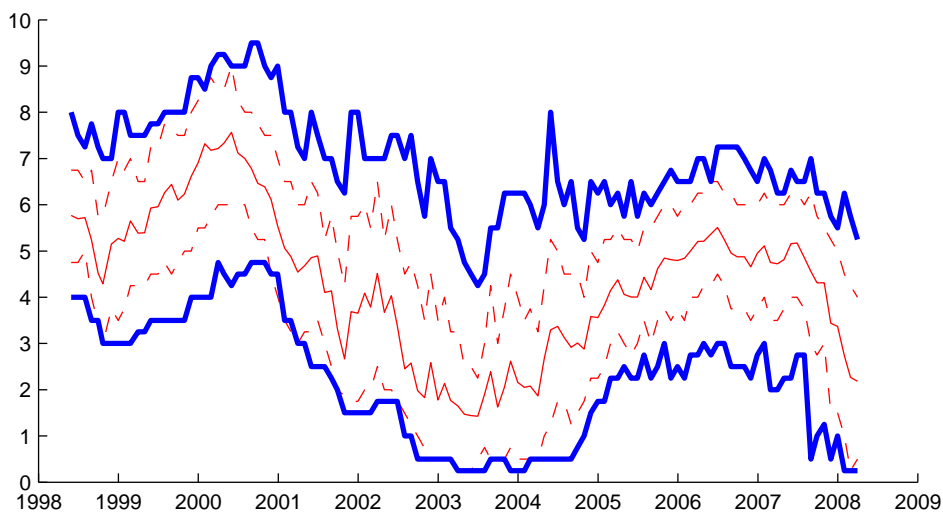


Figure 3: Available option strikes for 1-year maturity options. Maximum call strike and minimum put strike are shown in thick solid lines. The current futures rate  $F$  is shown in thin solid lines. The thin dashed lines are  $F \pm 1.5\sigma^*$ .

where  $\sigma^B$  is the Black implied volatility (Black (1976)) and  $\tau$  is the option maturity.<sup>31</sup> As can be seen, these numbers (thin dashed line) tend to lie inside the  $[\min(\text{put strike}), \max(\text{call strike})]$  band, indicating that typically a fairly wide range of strikes are available, spanning at least  $-1.5\sigma^*$  to  $1.5\sigma^*$  units around the mean of the distribution. I have used only the out-of-money puts and calls (puts at or below the current futures rate and calls above the current futures rate). Thus at each multiples of 25bp (between the minimum put strike and maximum call strike), there is only one option (either put or call), or none (when there is no open interest).

Table 2 provides some statistics regarding the liquidity of options proxied by open interest, for strikes  $0.25\sigma^*$ ,  $0.5\sigma^*$ ,  $\sigma^*$ ,  $1.5\sigma^*$  away from the mean.<sup>32</sup> Options about  $\frac{1}{2}$  standard deviation from the current futures rate are the most active, but the whole strike range (at least up to  $1\frac{1}{2}$  standard deviation) is fairly active; e.g., the median open interest in the  $1.5\sigma^*$  put option is more than a third of that of the  $0.5\sigma^*$  option.

Lastly, as a caveat, note that since the  $r$  here is taken to be the 3-month LIBOR, the “zero bound” (more precisely, the lower bound) is not exactly at zero, as the 3-month LIBOR contains a premium for credit/liquidity risk, besides some differential between

<sup>31</sup>Though simple, this number agrees fairly well with the standard deviation computed from the B-MN model.

<sup>32</sup>Since options are available at discrete strikes in multiples of 0.25%, the strikes closest to the respective numbers were chosen. If there is no strike within 25 basis points of the number, it is not counted. This will be the convention also when the pricing errors for strikes  $\pm 1.5\sigma^*$  away from the mean are examined in Sec. 5.3.

dist. from F	put				call			
	mean	25%	50%	75%	mean	25%	50%	75%
0.25 $\sigma^*$	48664	16320	32749	58906	28072	10106	18352	36747
0.5 $\sigma^*$	49449	15436	37047	57529	32491	9994	20434	37852
$\sigma^*$	45561	4625	25362	99226	27141	9868	19236	34176
1.5 $\sigma^*$	31512	4738	13262	37920	19890	3726	10412	26950

Table 2: Mean open interest, and 25-,50-, and 75-percentile open interest.

the 3-month rate and the instantaneous rate. This means that the proper location of the delta function for the Black's boundary model pdf ( $\underline{r}$ ) should be a small but nonzero number. Furthermore, the delta function itself would be smudged out a bit, in other words,

$$(58) \quad \tilde{\delta}(x - \underline{r}) = \frac{1}{\sqrt{2\pi}\epsilon} \exp\left(-\frac{(x - \underline{r})^2}{2\epsilon^2}\right)$$

where  $\epsilon$  is a small number (but not  $\epsilon \rightarrow 0$  as in eq. (5)), would be a more accurate description. However, as long as  $\epsilon$  is small enough, the delta function  $\delta(x - \underline{r})$  should provide a good approximation of  $\tilde{\delta}(x - \underline{r})$ .

It is difficult to get a precise sense of  $\underline{r}$  for the 3-month LIBOR, but the recent Japanese experience can provide some guidance: even at the height of Bank of Japan's Quantitative Easing Program, the 3-month yen TIBOR rate (an analogue of the US dollar LIBOR) did not go below 8 basis points. For simplicity, the main computations in this paper assumed that the lower bound  $\underline{r}$  is at 0, but I have also computed the pdf for the B-MN model with  $\underline{r} = 0.15\%$  to check the robustness of the results. Intuitively one would expect that the difference between these two settings would be immaterial for option strikes that are substantially larger than the effective lower bound (i.e.,  $K \gg \underline{r}$ ), but at a very low strike (e.g., 25 basis points), there may be a more visible difference between  $\underline{r} = 0$  vs  $\underline{r} = 0.15\%$ .

**5.3. Option pricing performance.** To get a sense of the overall fit across strikes, it is useful to look at the root-mean-square errors, i.e.,

$$(59) \quad \text{RMSE} = (\hat{\zeta}/(1 + n_p + n_c))^{1/2},$$

where  $\hat{\zeta}$  is the minimized value of  $\zeta$  in eq. (10).

Figure 4 shows the time series of RMSEs based on the four models (B-MN, MLN, CIR2, QG3) for the 1998-2008 period. As can be seen in the top panel, the B-MN model (Black's boundary model) and the MLN model (conventional benchmark) performed similarly well, the MLN model producing a somewhat smaller RMSE in 2003-2004 and larger RMSE in 2007-2008. Note that during most of the data period, with the exceptions around 2003 and 2008, the distributions were sufficiently away from the zero bound (as

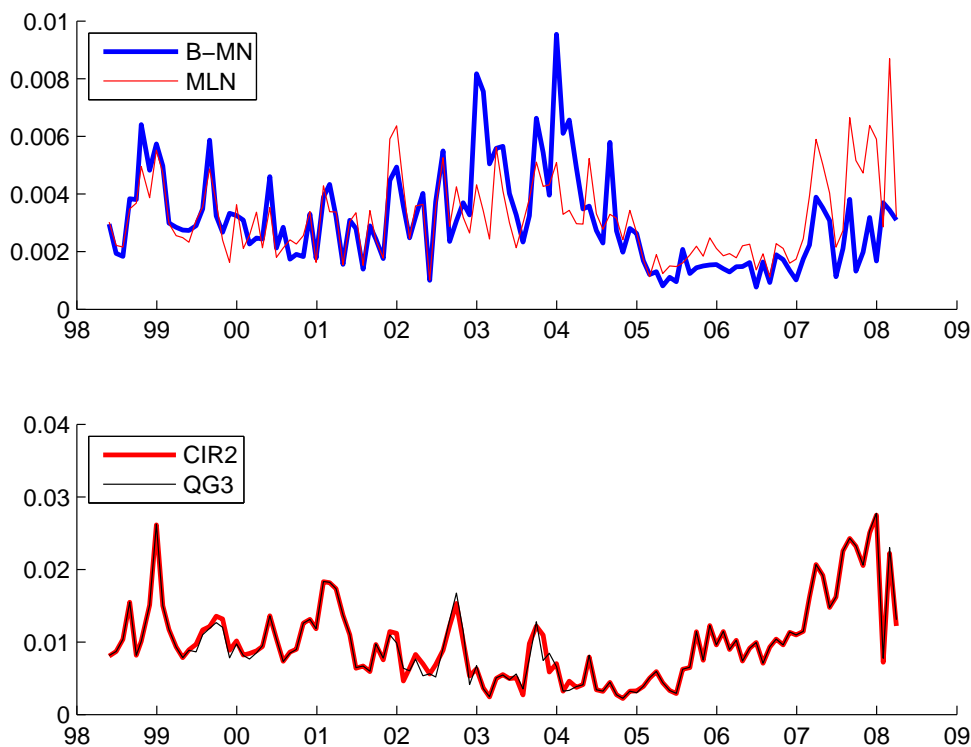


Figure 4: Root mean square option pricing errors. Top panel: B-MN and MLN models. Bottom panel: QG3 and CIR2 models.

can be seen in Figure 3), thus the B-MN model was simply the “mixture of normals” model (i.e.,  $f = f^S$ ); the results indicate that there is little practical difference between the normal-mixture model and the lognormal-mixture model most of the times. On the other hand, the QG3 model and CIR2 model led to notably larger RMSEs. (Note the difference in the y-axis scales in top panel and bottom panel). Between the two models, the RMSE numbers were very similar, practically lying on top of each other.

It is also useful to examine the individual (strike-specific) pricing errors,  $P^{model} - P^{ob}$  and  $C^{model} - C^{ob}$ . Table 3 displays summary statistics for the pricing errors for put options with strikes  $0.25\sigma^*$  and  $1.5\sigma^*$  units below  $F$  and call options with strikes  $0.25\sigma^*$  and  $1.5\sigma^*$  units above  $F$ . It can be seen that the errors for the QG3 and CIR2 models are not only large, but they also deviate systematically from zero. For example, even the 75-percentile value of the pricing error for the put option at the strike of  $F - 1.5\sigma^*$  is negative; this negative bias means that the QG3 and CIR2 models tend to systematically underprice the farther-out-of-the-money put options. On the other hand, they tend to overprice the near-the-money call options.

		$0.25\sigma^*$				$1.5\sigma^*$			
		mean	25%	50%	75%	mean	25%	50%	75%
put	B-MN	-0.0014	-0.0030	-0.0010	0.0009	-0.0003	-0.0019	-0.0001	0.0012
	MLN	-0.0008	-0.0026	-0.0007	0.0017	-0.0017	-0.0032	-0.0009	0.0013
	CIR2	0.0031	-0.0007	0.0031	0.0076	-0.0155	-0.0198	-0.0140	-0.0074
	QG3	0.0028	-0.0011	0.0027	0.0077	-0.0157	-0.0198	-0.0141	-0.0083
call	B-MN	-0.0014	-0.0036	-0.0011	0.0006	0.0011	-0.0008	0.0010	0.0028
	MLN	-0.0011	-0.0027	-0.0010	0.0007	0.0009	-0.0010	0.0008	0.0027
	CIR2	0.0043	0.0009	0.0044	0.0080	0.0005	-0.0056	0.0011	0.0056
	QG3	0.0041	0.0008	0.0041	0.0079	0.0011	-0.0050	0.0010	0.0058

Table 3: Statistics (mean, 25-, 50-, and 75-percentile values) about put option pricing errors at strikes  $F - 0.25\sigma^*$  and  $F - 1.5\sigma^*$  (top half of the table), and call option pricing errors at strikes  $F + 0.25\sigma^*$  and  $F + 1.5\sigma^*$  (bottom half of the table).

In sum, the pdfs based on zero-cdf term structure models (CIR2 and QG3 models) perform notably worse than the MLN and B-MN models despite having one more parameter, suggesting that the kind of restrictions on the shape of the risk-neutral distribution implied by these models are not well supported by data.

**5.4. Is the pdf always positively skewed?** As pointed out in Section 4.1, well-known zero-cdf term structure models, including the multifactor CIR and positive-definite QG models, imply that the skewness of the risk-neutral distribution of the short rate is always positive. To check this prediction, we can examine the model-implied skewness  $E[(r - E(r))^3]/E[(r - E(r))^2]^{3/2}$  from the B-MN and MLN pdfs, shown in Figure 5a. The sharp vertical lines that occur in the MLN measure are a small number of occasions in which the MLN model produced very large skewness values that went out of bounds of the figure, even though option pricing errors were small. Ignoring these pathological cases, the B-MN and MLN models tell fairly similar story. In particular, the risk neutral distribution was negative during an extended period of time in recent years (2006-07). Negatively-skewed pdfs occurred also around 1998 and 2001.

A test of skewness that does not need an assumption about the functional form of the pdf can be constructed as follows. Note that a negatively skewed distribution has  $mean < median$ , while a positively skewed distribution has  $mean > median$ .<sup>33</sup> Since the cumulative distribution function  $F(r)$  at the median is  $1/2$  by definition, one can check whether the skewness is always positive by examining whether the condition

$$(60) \quad F(F) > 1/2$$

<sup>33</sup>This is the idea behind the Pearson measure of skewness =  $3 \text{ [mean-median]}/[\text{standard deviation}]$ . This definition of skewness is not identical to the more common definition  $[\text{third central moment}]/[\text{standard deviation}]^{3/2}$ . However, the two definitions are qualitatively and quantitatively similar for unimodal distributions with moderate asymmetry.



holds, where we have used the fact that the mean of the risk-neutral distribution is the current futures rate. Since

$$(61) \quad F(K) = e^{rf\tau} P'(K) = 1 + e^{rf\tau} C'(K),$$

we can check the condition (60) by evaluating the first derivatives  $P'$  and  $C'$  at  $F$  using the piecewise-cubic Hermite interpolation scheme for  $P(K)$  and  $C(K)$ . Figure 5 shows  $F(F)$ 's based on puts and calls. It can be seen that  $F(F)$  can move below  $1/2$ ; indeed, where this occurs is quite close to where the skewness measures intersect the 0-line in Figure (a).

An example of negatively skewed pdf (April 30, 2007) is shown in Figure 6. The B-MN and MLN models produce fairly similar pdf (except for a small “wiggles” in the MLN pdf), and both clearly show a notable negative skew. As discussed earlier, since zero-cdf models like the QG3 model cannot produce a negatively skewed distribution, the estimated pdf from this model is a nearly symmetric distribution instead. Not

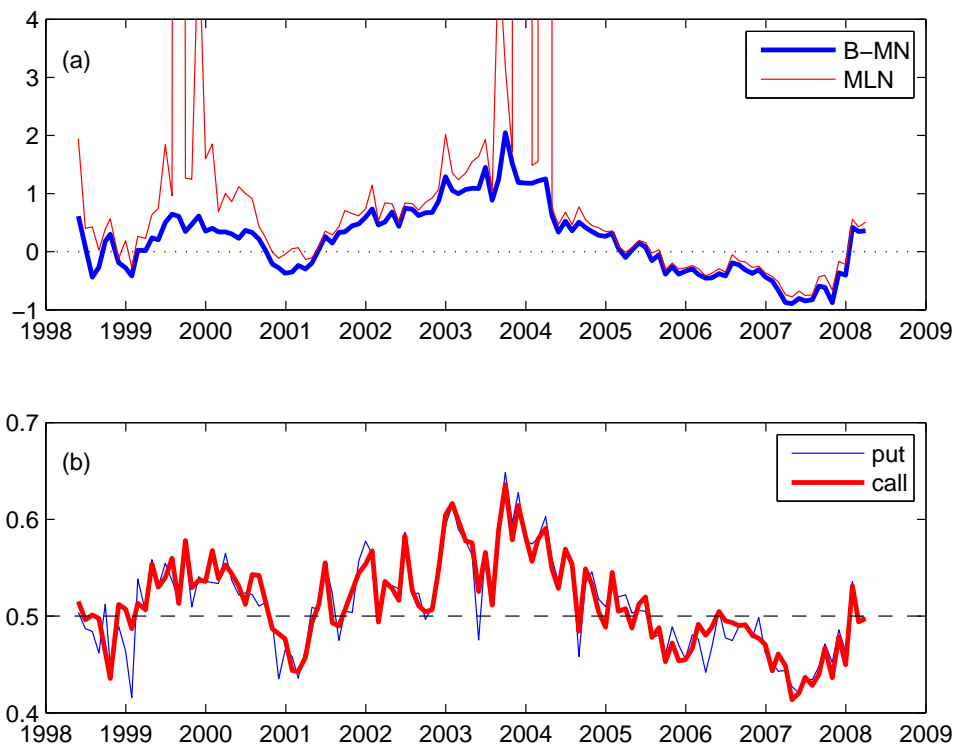


Figure 5: (a) Model-implied skewness from the B-MN and MLN pdfs. (b) Alternative test of skewness based on the Pearson measure.

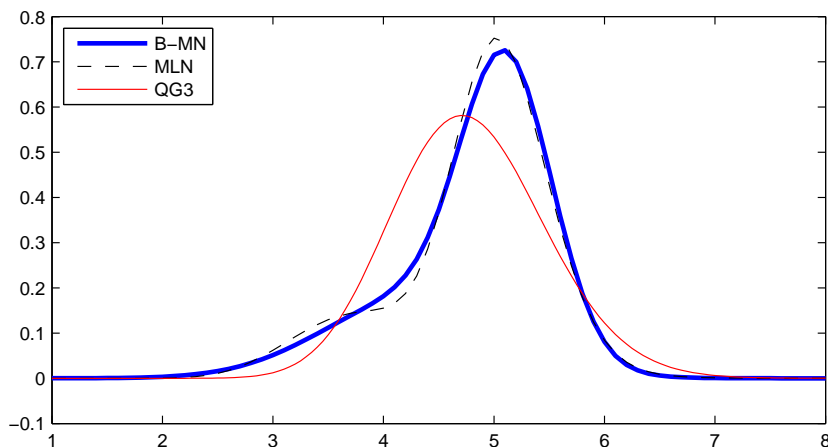


Figure 6: Pdfs on April 30, 2007

surprisingly, pricing errors of the QG3 and CIR2 models tend to be particularly large during the times of negatively skewed pdfs (as can be seen in Figure 4b).

**5.5. Financial market turmoil of 2008.** In early 2008, the turmoil in financial markets that had begun in mid-2007 intensified, posing considerable systemic risk and raising much concern about the broader economy. This led to a substantial lowering of the federal funds target rate (and of the market's expectation of the future target rate). In March, the 1-year-ahead eurodollar futures rate had dropped to levels as low as 1.9%. Amid heightened uncertainty about the path of the interest rate, eurodollar futures options with very low strikes, in particular 0.25% and 0.5%, had begun to trade. This section explores whether the above models, if any, can explain the prices observed for these options.

It befits to start with a cautionary remark that the liquidity in these put options was significantly lower than near-the-money options. The 1-year-maturity 0.25%-strike option, in particular, had an open interest of only about 30. However, the market in the 0.5%-strike option was more active, with decent open interest (ranging from 2000 to 8000 in the months of February and March, comparable to the 25 percentile open interest in the put options with strikes  $\sigma^*$  to  $1.5\sigma^*$  away from the current futures rate in Table 2). Therefore, at least the 0.5%-strike option merits a serious look.

Figure 7a shows the pricing error  $P^{model} - P^{ob}$  for the 0.5%-strike option from the B-MN, MLN, and QG3 models, for every business day from February 4 to April 28, 2008. The MLN and QG3 models (and the CIR2 model, not shown) generated much larger errors than the B-MN model in this period, and these errors were systematically negative, i.e., the QG3 and MLN models tended to underprice the option. These errors were particularly large during the first half of March; this is precisely the period where

the B-MN model indicates a substantial weight of the  $\delta$ -function piece ( $w$  in eq. (52)),

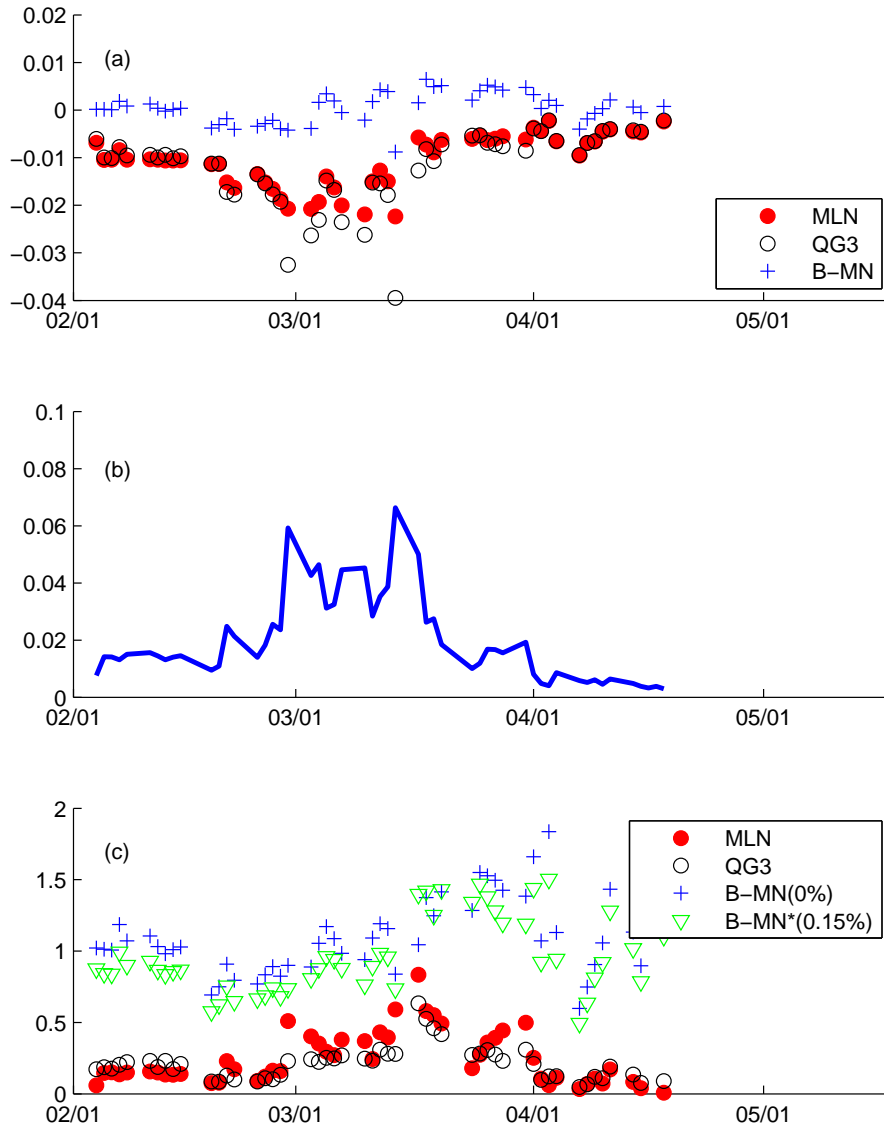


Figure 7: (a) Pricing errors  $P^{model} - P^{ob}$  for  $K = 0.5\%$ . (b) Weight of the Dirac delta function in the B-MN pdf ( $w$  in eq. (52)). (c) The ratio  $P^{model}/P^{ob}$  for  $K = 0.5\%$ . The B-MN\*(0.15%) pdf computations were done without using put options with strikes below 1%.

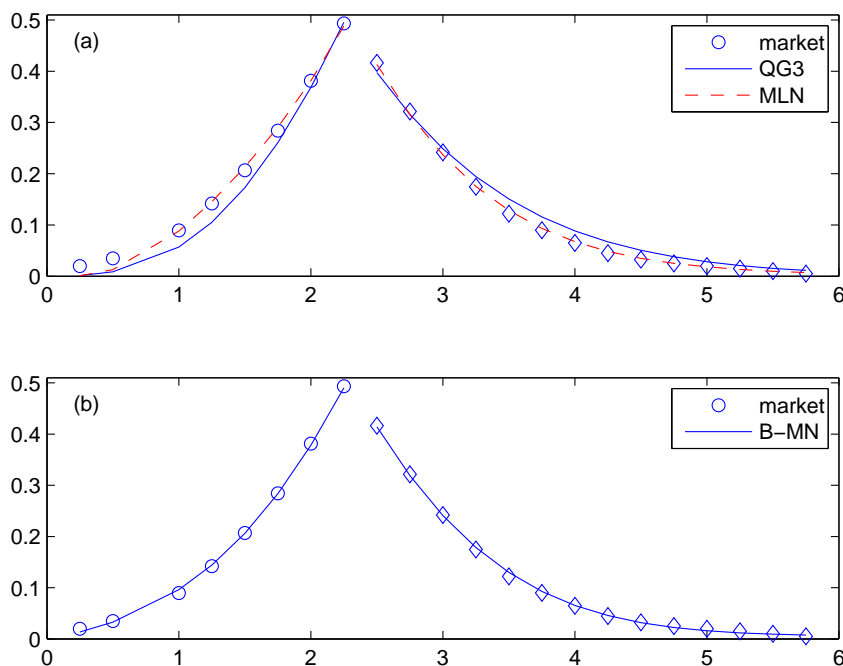


Figure 8: Observed put option prices (circle) and call option prices (diamond) and model predictions on March 10, 2008.

shown in Figure 7b. Although not shown, the pricing errors for  $K = 0.25\%$  tell a similar story. Figure 7c shows the ratio  $P^{model}/P^{ob}$  for the B-MN, MLN, and QG3 models. Note that the ratio is much smaller than 1 (sometimes orders of magnitude smaller) for the MLN and QG3 models, i.e., these models grossly underpriced the  $K = 0.5\%$  option.

To show a representative day from this period, the fit of put and call option prices on March 10, 2008 is shown in Figure 8. Note that the QG3 model fits the put and call option prices rather poorly; the CIR2 model (not shown) gives similar results. The B-MN model prices put and call options fairly well across strikes. The MLN model also prices put and call option price well, except at the strikes of 0.25% and 0.5%.

The model-implied pdfs for this date, shown in Figure 9, help explain this result. The MLN model pdf has a bimodal structure with a pronounced peak in the low interest rate region (centered at about 0.5%) and a broader peak at about 2.5%. This produces a fairly good fit of option prices for most of the strikes, but since the left tails of both peaks have to vanish at 0, it leads to low prices of put options at strikes of 0.5% and 0.25%. The QG3 model also has a pdf that vanishes in the zero limit (recall from Section 4.1 that the pdf of the 3-factor positive-definite QG model has the behavior  $f(r) \propto \sqrt{r}$ ,  $r \rightarrow 0$ ), thus producing much lower prices of low-strike options than the B-MN model

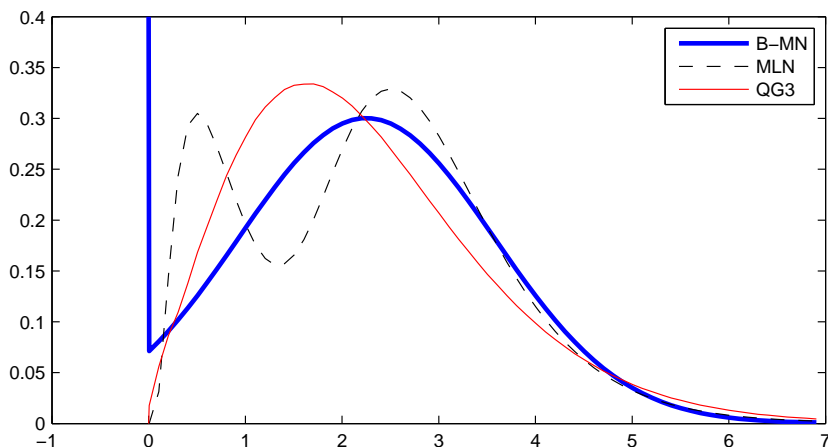


Figure 9: Pdfs on March 10, 2008.

which has a delta function singularity with weight of about 0.04. It is also interesting to note that the modal implications of the pdfs are quite different, e.g., the QG3 pdf indicates that the most likely value is less than 2%, while the B-MN pdf indicates that the most likely value is larger than 2%.

Several alternative computations of the pdf were done to check the robustness of the results. For example, put options with strikes below 1% were dropped in the computation (10), but this led to very similar results as the original computations. In other words, the good fit of the low-strike put option in the B-MN model is not due to their inclusion in the computation.

Another robustness check is with respect to the choice of the lower-boundary; as discussed in Section 5.2, there are reasons to consider a nonzero  $\underline{r}$ . The computation of the B-MN model pdf with  $\underline{r} = 0.15\%$  again matched the price of the 0.5%-strike put option reasonably well. Figure 7c shows, with inverted triangle symbols, the  $P^{model}/P^{ob}$  ratio for this case (this computation also dropped options with strikes less than 1%); it can be seen that the ratio is still fairly close to 1, and much larger than the values predicted by the QG3 and MLN models.<sup>34</sup>

Lastly, note that 2002-3 is another interesting period from which the near-zero behavior of the risk-neutral pdf can be studied, but this paper forgoes a treatment comparable to that given to the 2008 episode, for space considerations.

<sup>34</sup>Two remarks: (1) Note that the  $P^{model}/P^{ob}$  ratios for the QG3 and MLN models in Figure 7c were computed from the pdfs that assumed  $\underline{r} = 0$ . If these models also used  $\underline{r} = 0.15\%$  (via the functional form  $f(r - \underline{r})$ ), they would have predicted *even smaller* ratios. (2) In the case of the 0.25%-strike put option (not shown in figures), the  $\underline{r} = 0.15\%$  lower bound for the B-MN model led to a  $P^{model}/P^{ob}$  ratio that is about half of the  $\underline{r} = 0$  case, but this still represents much larger values than the corresponding numbers from the QG3, CIR2, and MLN models.

## 6. CONCLUDING REMARKS

The main conclusions of this paper can be summarized as follows. (1) Black’s boundary model can capture the low-strike put option prices in February and March of 2008, which would have seemed puzzlingly high from the viewpoint of the popular zero-cdf term structure models. While some might alternatively argue that the “high” prices of these options are due to low liquidity or irrational pricing, we can say, at least, that the Black’s boundary behavior provides one possible rational explanation for them. (2) Well-known non-negative interest rate models like the multifactor (positive-definite) CIR/QG models not only have difficulty capturing the low-strike option prices in early 2008 but also generate large and biased option pricing errors at other times as well. We have seen this specifically in the context of 2-factor and 3-factor CIR/QG models, but the theoretical arguments presented in Sec 4.1 suggest that having more factors would not change this conclusion. We have also seen that options data contradict the prediction that the pdf is always positively skewed. These constitute strong evidence against well-known term structure models that have zero cdf at the zero boundary.

This paper also has broader implications:

(a) The encouraging results from the flexible parametric pdf model with Black’s boundary behavior developed here suggests that such a model may be useful for practical applications to other settings, for example, the extraction of risk-neutral pdfs from the Japanese interest rate options market (euroyen futures options).

(b) The evidence against zero-cdf models (and evidence for Black’s boundary models) documented here indicates that, in modeling the postwar US economy there is little motivation to search for “special” macroeconomic mechanisms that would generate behaviors in which the short rate always stays above zero or “bounces off” from zero.<sup>35</sup>

(c) While the tractability of the affine and QG models may provide some incentive to consider the positive-definite versions of these models as *approximate* models (even if the true boundary behavior were Black’s boundary behavior), the results in this paper cautions that they would be *poor* approximations, at least in certain contexts.

(d) This paper provides some justification for favoring models that can have negative nominal interest rates. It may appear counter-intuitive to deliberately ignore a clearly reasonable constraint about nominal interests (positive-definiteness). Indeed, Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003) considered it a virtue that the QG model’s parameters can be restricted so as to guarantee the positivity of nominal interest rates, and to my knowledge, all published papers that studied the QG model empirically have imposed such restriction. In so doing, however, they rule out specifications such as

$$(62) \quad r_t = c + x_{1t}^2 - x_{2t}^2.$$

---

<sup>35</sup>Note that macroeconomic models that optimize the standard cost function  $E_t[\sum_i \beta^i (\alpha x_{t+i}^2 + (1 - \alpha)\pi_{t+i}^2)]$  (the GDP gap  $x_t$  and inflation  $\pi_t$  being unbounded random variables), with the condition  $r_t \geq 0$  introduced via a *Lagrange multiplier*, would lead to a Black-like boundary behavior, as opposed to a CIR-like behavior.

Though it allows negative nominal rates, a specification like this accommodates negatively skewed pdfs (as observed in Sec. 5.4). The key point is that such specification is very similar to Black's model specification (which guarantees non-negative interest)

$$(63) \quad r_t = \max[0, c + x_{1t}^2 - x_{2t}^2],$$

except when rates are close enough to zero for the zero bound to have substantial effects;<sup>36</sup> hence most of the times it corresponds to Black's boundary model. In situations in which the zero bound matters, eq. (62) would no longer be a good approximation of (63), but in this case the positive-definite versions of the QG model would not be a good approximation of (63) either.

The empirical design of this paper had the following features: (i) the focus on the risk-neutral measure, and (ii) day-by-day implementations using only options data (and the current futures rate) for that day. Though this design was sufficient for detecting the problems with zero-cdf models (such as the multifactor positive-definite CIR/QG models) and yielded useful techniques for modeling pdfs for low interest rate scenarios, in the future it would be useful to extend the design to also examine the physical-measure processes and distributions, and utilize the time-series as well as the cross-sectional aspects of both options data and yield curve data, to achieve a more detailed understanding of the near-zero boundary processes.

#### APPENDIX A. THE $r \rightarrow 0$ BEHAVIOR OF THE PDF FOR MULTIFACTOR CIR/QG MODELS

This appendix derives the small- $r$  asymptotic expansion of the probability density function  $f(r)$  for the distribution  $r \sim b_1 \chi_{\nu_1, \lambda_1}^2 + \dots + b_n \chi_{\nu_n, \lambda_n}^2$ .

For this, we will use the Laplace transform of the probability density function

$$(64) \quad \bar{f}(\omega) = \int_0^\infty e^{-r\omega} f(r) dr.$$

The inverse Laplace transform

$$(65) \quad f(r) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{r\omega} \bar{f}(\omega) d\omega$$

gives back the pdf from  $\bar{f}(\omega)$ .

The Laplace transform of the pdf, i.e.  $E(e^{-\omega r})$ , is simply related to the characteristic function  $E(e^{i\omega r})$  as  $\bar{f}(\omega) = \hat{f}(i\omega)$ . Therefore, from eq. (45), we have

$$(66) \quad \bar{f}(\omega) = \prod_{j=1}^n \frac{e^{-\lambda_j b_j \omega / (1+2b_j \omega)}}{(1 + 2b_j \omega)^{\nu_j/2}}.$$

The basic idea is to expand this in powers of  $1/\omega$ , which will then lead to an inverse transform in powers of  $r$ .

---

<sup>36</sup>How close is "close enough to zero" may depend on the context of the problem; e.g., out-of-the-money put option pricing would be more sensitive to this issue than at-the-money option pricing.

Note that the exponent in the exponential in eq. (66) can be written

$$(67) \quad -\frac{\lambda_j b_j \omega}{1 + 2b_j \omega} = -\frac{\lambda_j}{2} \left( 1 - \frac{1}{2b_j \omega} + O(1/\omega^2) \right)$$

and that

$$(68) \quad (1 + 2b_j \omega)^{\nu_j/2} = (2b_j \omega)^{\nu_j/2} \left( 1 + \frac{\nu_j}{4b_j} \frac{1}{\omega} + O(1/\omega^2) \right).$$

Therefore,

$$(69) \quad \begin{aligned} \bar{f}(\omega) &= \frac{\exp(-\sum_j \lambda_j/2 + (\sum_j \lambda_j/4b_j)1/\omega + O(1/\omega^2))}{(\prod_j (2b_j)^{\nu_j/2}) \omega^{\sum_j \nu_j/2} (1 + (\sum_j \nu_j/4b_j)1/\omega + O(1/\omega^2))} \\ &= \frac{\exp(-\sum_j \lambda_j/2)}{(\prod_j (2b_j)^{\nu_j/2})} \left( \omega^{-\sum_j \nu_j/2} + \left( \sum_j \frac{\lambda_j - \nu_j}{4b_j} \right) \omega^{-\sum_j \nu_j/2-1} + O(\omega^{-\sum_j \nu_j/2-2}) \right). \end{aligned}$$

Because the inverse Laplace transform of  $\omega^{-q}$  is simply  $r^{q-1}/\Gamma(q)$ , we obtain immediately eq. (48).

#### APPENDIX B. OPTION PRICES FOR RISK-NEUTRAL DISTRIBUTION DESCRIBED BY POSITIVE LINEAR COMBINATION OF INDEPENDENT NONCENTRAL $\chi^2$ RANDOM VARIABLES

This appendix derives an approximate (but very accurate) formula for option prices for the distribution of short rate implied by CIR/QG models.

We want to evaluate the expressions

$$(70) \quad C(K) = e^{-rf\tau} E((r - K) \mathbf{I}(r - K))$$

$$(71) \quad P(K) = e^{-rf\tau} E((K - r) \mathbf{I}(K - r))$$

where  $E$  is an expectation with respect to the distribution  $r \sim b_1 \chi_{\nu_1, \lambda_1}^2 + \dots + b_n \chi_{\nu_n, \lambda_n}^2$ . Since  $E(\mathbf{I}(K - r))$  is the cumulative distribution function  $F(r)$ , we have

$$(72) \quad \begin{aligned} C(K) &= e^{-rf\tau} (E(r \mathbf{I}(r - K)) - K(1 - F(K))) \\ P(K) &= e^{-rf\tau} (KF(K) - E(r \mathbf{I}(K - r))). \end{aligned}$$

Thus we need to evaluate  $E(r \mathbf{I}(\cdot))$ . For this, it is useful to define a new probability measure, with Radon-Nikodym derivative  $r/E(r)$ . (This can be done because  $r$  is a positive random variable.) In terms of the new measure (denoted with superscript  $\dagger$ ), we have

$$(73) \quad E(r \mathbf{I}(\cdot)) = E(r) E^\dagger(\mathbf{I}(\cdot)),$$

therefore

$$(74) \quad \begin{aligned} C(K) &= e^{-rf\tau} (E(r)(1 - F^\dagger(K)) - K(1 - F(K))) \\ P(K) &= e^{-rf\tau} (KF(K) - E(r)F^\dagger(K)), \end{aligned}$$

where  $F^\dagger(K)$  denotes the cdf in the new measure. From eq. (50),  $E(r)$  is simply  $E(r) = \sum_i b_i (\nu_i + \lambda_i)$ .



It thus remains to derive expressions for cdfs  $F(K)$  and  $F^\dagger(K)$ . This can be done using the saddle-point formula of Lugannani and Rice (1980) for cdfs,<sup>37</sup> which uses the cumulant generating function  $\mathcal{C}(\omega) = \log E(e^{\omega r})$ :

$$(75) \quad \begin{aligned} F(K) &= \Phi(z) - \phi(z) \left( \frac{1}{y} - \frac{1}{z} \right), \\ z &= \sqrt{2(\omega^* K - \mathcal{C}(\omega^*))} \operatorname{sgn}(\omega^*) \\ y &= \omega^* \sqrt{\mathcal{C}''(\omega^*)}, \end{aligned}$$

where  $\omega^*$  is the solution of the equation

$$\mathcal{C}'(\omega) = K.$$

Note that  $\mathcal{C}(\omega)$  is related to the characteristic function as  $\mathcal{C}(\omega) = \log \hat{f}(-i\omega)$ . Thus, from eq. (45), we have

$$(76) \quad \mathcal{C}(\omega) = \log \left( \prod_{i=1}^n \frac{e^{\lambda_i b_i \omega / (1 - 2b_i \omega)}}{(1 - 2b_i \omega)^{\nu_i / 2}} \right) = \sum_{i=1}^n \left( \frac{\lambda_i b_i \omega}{1 - 2b_i \omega} - \frac{\nu_i}{2} \log(1 - 2b_i \omega) \right).$$

The cumulative distribution function in the new measure  $F^\dagger(K)$  is given by the same expression as that of  $F(K)$ , except that the cumulant generating function  $\mathcal{C}$  in eq. (75) is replaced by the cumulant generating function for the new measure,  $\mathcal{C}^\dagger$ , which is given by

$$(77) \quad \mathcal{C}^\dagger(\omega) = \log E^\dagger(e^{\omega r}) = \log(E(e^{\omega r r})/E(r)).$$

Using the identity

$$(78) \quad E(e^{\omega r r}) = \frac{\partial}{\partial \omega} E(e^{\omega r}) = \frac{\partial}{\partial \omega} \hat{f}(-i\omega),$$

it is straightforward to show

$$(79) \quad \mathcal{C}^\dagger(\omega) = \mathcal{C}(\omega) + \log \left[ \sum_{i=1}^n \left( \frac{\lambda_i b_i}{(1 - 2b_i \omega)^2} + \frac{\nu_i b_i}{1 - 2b_i \omega} \right) \right] - \log \left( \sum_{i=1}^n b_i (\lambda_i + \nu_i) \right).$$

I have checked the Lugannani-Rice formula for the cumulant generating functions  $\mathcal{C}, \mathcal{C}^\dagger$  against the Fourier-inversion formula for cdf (eq. (47)) at several sets of parameter values, and found that in all cases the agreements were excellent across the entire range of  $K$ , i.e., both in the tail and center regions of the distribution. One caveat is that the formula can be numerically unstable when  $K$  happens to be the mean of the distribution (i.e.,  $K = \mathcal{C}'(\omega = 0)$ ,  $K = \mathcal{C}^\dagger(\omega = 0)$ ). However, such situation occurs rarely in our optimization problem (eq. (10)); if it does occur, one can find a nearby parameter vector that does not have the problem but has a very similar pdf content, or use the conventional (slower) formula (47) with characteristic functions corresponding to  $\mathcal{C}, \mathcal{C}^\dagger$ .

---

<sup>37</sup>Lugannani and Rice (1980) derived their formula via a judicious choice of the path of integration for the inverse Fourier (Laplace) transform representation of the cdf. This formula found much use in the statistics literature (e.g., Daniels (1987)); its application to stock option pricing (lognormal-type distributions) was explored by Rogers and Zane (1999).

APPENDIX C. OPTIONS AND FUTURES PRICES FOR FLEXIBLE PDFS WITH BLACK'S  
BOUNDARY BEHAVIOR

This appendix provides formulae that are useful for the implementation of flexible pdf models with Black's boundary behavior.

It is straightforward to show that the option prices for the normal-mixture shadow rate distribution  $f^S$  (eq. (53)) and lower boundary  $\underline{r}$  are given by

$$(80) \quad P(K) = e^{-rf\tau} [w(K - \underline{r}) + \sum_{i=1}^2 \beta_i \sigma_i (k_i (\Phi(k_i) - \Phi(\underline{\rho}_i)) + \phi(k_i) - \phi(\underline{\rho}_i))]$$

$$C(K) = e^{-rf\tau} \sum_{i=1}^2 \beta_i \sigma_i (\phi(k_i) - k_i (1 - \Phi(k_i))),$$

where

$$(81) \quad k_i \equiv (K - \mu_i) / \sigma_i, \quad \underline{\rho}_i \equiv (\underline{r} - \mu_i) / \sigma_i,$$

$$(82) \quad w = \beta_1 \Phi(\underline{\rho}_1) + \beta_2 \Phi(\underline{\rho}_2).$$

The current futures rate  $F = \int_{-\infty}^{\infty} r f(r) dr$  is given by

$$(83) \quad F = w\underline{r} + \sum_{i=1}^2 \beta_i (\sigma_i \phi(\underline{\rho}_i) + \mu_i (1 - \Phi(\underline{\rho}_i))).$$

The option prices for the Gram-Charlier/Edgeworth shadow rate distribution (eq. (54)) and lower boundary  $\underline{r}$  are given by

$$(84) \quad P(K) = e^{-rf\tau} [w(K - \underline{r}) + \sigma (H_0^p(k; \rho) + \frac{\gamma_3}{6} H_3^p(k; \rho) + \frac{\gamma_4}{24} H_4^p(k; \rho) \dots)],$$

$$(85) \quad C(K) = e^{-rf\tau} \sigma (H_0^c(k) + \frac{\gamma_3}{6} H_3^c(k) + \frac{\gamma_4}{24} H_4^c(k) \dots),$$

where

$$(86) \quad H_j^p(k; \rho) \equiv \int_{\rho}^k (k - x) H_j(x) \phi(x) dx$$

$$(87) \quad H_j^c(k) \equiv \int_k^{\infty} (x - k) H_j(x) \phi(x) dx$$

$$(88) \quad k \equiv (K - \mu) / \sigma, \quad \rho \equiv (\underline{r} - \mu) / \sigma.$$

Using the property of Hermite polynomials that  $H_j(x) \phi(x) = (-1)^j \frac{d^j}{dx^j} \phi(x)$ , it is straightforward to show

$$(89) \quad w = \Phi(\rho) - \frac{\gamma_3}{6} H_2(\rho) \phi(\rho) - \frac{\gamma_4}{24} H_3(\rho) \phi(\rho) - \dots.$$

Similarly, and using integration by parts, we have

$$(90) \quad H_j^c(k) = H_{j-2}(k) \phi(k)$$

$$(91) \quad H_j^p(k; \rho) = (k - \rho) H_{j-1}(\rho) \phi(\rho) + H_{j-2}(k) \phi(k) - H_{j-2}(\rho) \phi(\rho)$$

for  $j \geq 2$ . The integrals  $H_0^c(k)$  and  $H_0^p(k)$  are given by

$$(92) \quad H_0^c(k) = \phi(k) - k(1 - \Phi(k)),$$

$$(93) \quad H_0^p(k; \rho) = k(\Phi(k) - \Phi(\rho)) + (\phi(k) - \phi(\rho)).$$

The current futures rate is given by

$$(94) \quad \begin{aligned} F &= w \underline{r} + \mu(1 - \Phi(\rho)) + \sigma\phi(\rho) \\ &+ \frac{\gamma_3}{6}\phi(\rho)(\underline{r}H_2(\rho) + \sigma H_1(\rho)) + \frac{\gamma_4}{24}\phi(\rho)(\underline{r}H_3(\rho) + \sigma H_2(\rho)) + \dots \end{aligned}$$

## REFERENCES

- Ahn, D. -H., R. F. Dittmar, and A. R. Gallant (2002), "Quadratic term structure models: theory and evidence", *Review of Financial Studies*, 15, pp243-88.
- Ahn, D. -H., and B. Gao (1999), "A parametric nonlinear model of term structure dynamics," *Review of Financial Studies*, 12, pp721-62.
- Ait-Sahalia, Y. (1996), "Testing continuous-time models of the spot interest rate," *Review of Financial Studies*, 9, pp385-426.
- Bank for International Settlements (1999), *Estimating and Interpreting Probability Density Functions* (Proceedings of the workshop held at the BIS on 14 June 1999).
- Baroni-Adesi, G. and R. Whaley (1987), "Efficient analytic approximation of American option values," *Journal of Finance*, 42, pp301-20.
- Beaglehole, D. and M. Tenney (1992), "Corrections and additions to 'A nonlinear equilibrium model of the term structure of interest rates' " *Journal of Financial Economics*, 32, pp345-53.
- Black, F. (1976), "The pricing of commodity contracts", *Journal of Financial Economics*, 3, pp167-79.
- Black, F. (1995), "Interest rates as options", *Journal of Finance*, 50, pp1371-76.
- Bomfim, A. (2003), "Interest rate as options: Assessing the markets' view of the liquidity trap", FEDS working paper No 2003-45.
- Brace, A., D. Gatarek, and M. Musiela (1997), "The market model of interest rate dynamics", *Mathematical Finance* 7, 127-55.
- Breeden, D. T. and R. H. Litzenberger (1978), "Prices of state-contingent claims implicit in option prices", *Journal of Business*, 51, 621-51.
- Cecchetti, S. G. (1988), "The case of the negative nominal interest rates: new estimates of the term structure of interest rates during the Great Depression", *Journal of Political Economy*, 96, pp1111-41.
- Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, "An empirical comparison of alternative models of the short-term interest rate", *Journal of Finance*, 47, pp1209-27.
- Cox, J.C., J. Ingersoll, and S. A. Ross (1985), "A theory of the term structure of interest rates", *Econometrica*, 53, pp385-407.

- Daniels, H. E. (1987), "Tail probability approximations", *International Statistical Review*, 55, pp37-48.
- Dirac, P. A. M. (1958), *Principles of Quantum Mechanics*, (4th ed.) Oxford Univ. Press.
- Duffie, D. and K. Singleton (1997), "An econometric model of the term structure of interest rate swap yields", *Journal of Finance*, 52, pp1287-1321.
- Feldhütter, P., and D. Lando (2007), "Decomposing the swap spread", working paper.
- Feller, W. (1952), "The parabolic differential equations and the associated semi-groups of transformations", *Annals of Mathematics*, 55, pp468-519.
- Goldstein, R. S. and W. P. Kierstead (1997), "On the term structure of interest rates in the presence of reflecting and absorbing boundaries", working paper.
- Gorovoi, V. and V. Linetsky (2004), "Black's model of interest rates as options, eigenfunction expansions, and Japanese interest rates", *Mathematical Finance*, 14, pp49-78.
- Hördahl, P. (2000), "Estimating the implied distribution of the future short-term interest rate using the Longstaff-Schwartz model", working paper.
- Hördahl, P. and D. Vestin (2005), "Interpreting implied risk-neutral densities: the role of risk premia", *Review of Finance*, 9, pp97-137.
- Jackwerth, J. C. (2004), *Option-Implied Risk-Neutral Distributions and Risk Aversion*, The Research Foundation of AIMR.
- Johnson, N., S. Kotz, and N. Balakrishnan (1994), *Continuous Univariate Distributions*, Volume 2, Wiley.
- Karlin, S. and H. M. Taylor (1981), *A Second Course in Stochastic Processes*, Academic Press.
- Leippold, M. and L. Wu (2003), "Design and estimation of quadratic term structure models", *European Finance Review*, 7, pp47-73.
- Longstaff, F. A. (1992), "Multiple equilibria and term structure models", *Journal of Financial Economics*, 32, pp333-44.
- Longstaff, F. A. and E. S. Schwartz (1992), "Interest rate volatility and term structure: a two-factor general equilibrium model", *Journal of Finance*, 47, pp1259-82.
- Lugannani, R. and S. Rice (1980), "Saddlepoint approximations for the distribution of the sum of independent random variables", *Advances in Applied Probability*, 12, pp475-90.
- McManus, D. J. (1999), "The information content of interest rate futures options", working paper.
- Rogers, L. C. G. and O. Zane (1999), "Saddlepoint approximations to option prices", *Annals of Applied Probability*, 9, pp493-503.
- Shimko, D. (1993), "Bounds of probability", *RISK*, 6(April), 33-37.
- Stuart, A. and K. Ord (1994), *Kendall's Advanced Theory of Statistics*, Volume 1, Halsted Press.

Ueno, Y., N. Baba, and Y. Sakurai (2006), "The use of the Black model of interest rates as options for monitoring the JGB market expectations", Bank of Japan Working Paper Series.