

Article

Asymptotic Expansion of Risk-Neutral Pricing Density

Thomas Mazzoni

Department of Economics and Finance, University of Greifswald, 17489 Greifswald, Germany; thomas.mazzoni@uni-greifswald.de; Tel.: +49-3834-420-2498

Received: 14 December 2017; Accepted: 27 February 2018; Published: 12 March 2018

Abstract: A new method for pricing contingent claims based on an asymptotic expansion of the dynamics of the pricing density is introduced. The expansion is conducted in a preferred coordinate frame, in which the pricing density looks stationary. The resulting asymptotic *Kolmogorov*-backward-equation is approximated by using a complete set of orthogonal *Hermite*-polynomials. The derived model is calibrated and tested on a collection of 1075 European-style ‘Deutscher Aktienindex’ (DAX) index options and is shown to generate very precise option prices and a more accurate implied volatility surface than conventional methods.

Keywords: *Kolmogorov*-backward-equation; asymptotic expansion; *Hermite*-polynomials; implied volatility surface

JEL Classification: C61; G13

1. Introduction

Modern financial markets contain a rich variety of liquidly traded vanilla and exotic contracts, contingent on a large number of underlyings. A key requirement in such dense markets is the consistent valuation of novel and existing derivative contracts to rule out arbitrage opportunities. Because of market incompleteness, there is no unique risk-neutral Martingale measure, and hence no unique risk-neutral probability density for valuing such contracts. The necessary information on risk premia has to be extracted from the observed derivative prices. This is accomplished by calibrating a specific model to the available data. Usually, such models are stochastic path models for the underlying, which are calibrated to fit observed option prices (prominent examples are [Bates 1996](#); [Heston 1993](#); [Kou 2002](#)), or observed *Black–Scholes* implied volatilities (for example the SABR model of [Hagan et al. 2002](#)). A closely related and successful approach is to parameterize the implied volatility smile and skew, for example as suggested in ([Gatheral 2004, 2006](#), p. 37), for different times to maturity, and to smoothly connect the time slices, eliminating calendar spread arbitrage opportunities. Sufficient conditions for elimination of static arbitrage are provided in [Carr and Madan \(2005\)](#), and an efficient method for computing arbitrage-free implied volatility surfaces was introduced by [Fengler \(2009\)](#).

A conceptionally different idea is to estimate the arbitrage-free pricing density directly from available European plain vanilla call prices. This approach is based on the observation of [Breedon and Litzenberger \(1978\)](#) that the pricing density is given by the undiscounted second derivative of the European plain vanilla call price. Some suggested methods of this kind can be found in the works of [Ait-Sahalia and Duarte \(2003\)](#); [Ait-Sahalia and Lo \(1998\)](#); [Bondarenko \(2003\)](#); [Figlewski \(2010\)](#); [Hlavka and Svojik \(2009\)](#); [Huynh et al. \(2002\)](#); [Yatchew and Haerdle \(2006\)](#). A recent approach by [Filipović et al. \(2012\)](#) stipulates a so-called ‘Master Equation’ for the time evolution of a fairly general class of admissible pricing densities, and provides some examples. Even though the idea is vaguely similar, the method suggested in this paper is entirely different.

The key idea in the approach suggested here is to generate modified dynamics of the arbitrage-free pricing density by asymptotic expansion around the classical *Black-Scholes* dynamics of complete markets. Asymptotic analysis has proven a very potent tool in deriving new results over the last fifteen years (see for example [Basu and Ghosh 2009](#); [Hagan et al. 2002](#); [Kim 2002](#); [Mazzoni 2015](#); [Medvedev and Scaillet 2003](#); [Uchida and Yoshida 2004](#); [Whalley and Willmot 1997](#)) whenever certain parts of a problem can be assumed as small. The first step in this approach is to express the complete market dynamics of the risk-neutral pricing density in a new coordinate frame, where it looks stationary. In such a frame, the transition density has to be *Dirac's* delta function. If the dynamics of an incomplete market do not deviate too heavily from those of the complete market, it can be assumed that the singular transition density is only a first-order approximation of a narrow transition kernel of the order $O(\varepsilon\Delta t)$ for short time intervals Δt . Pursuing this avenue leads to an asymptotic version of the *Kolmogorov*-backward-equation for the excess dynamics of the incomplete market. This partial differential equation (PDE) can be solved approximately with the help of a complete set of orthogonal functions and a few additional assumptions, primarily related to the smallness of the asymptotic terms.

The idea of using a complete set of orthogonal functions to represent an unknown probability density function is not new. [Ait-Sahalia \(2002\)](#) advanced the *Gram-Charlier*-series of type A, utilizing *Hermite*-polynomials orthogonal to the weighting function $e^{-x^2/2}$ to represent an unknown probability density. This approach has the advantage that the expansion has a leading *Gaussian* term and the coefficients are proportional to the cumulants of the approximated density. Since then, *Hermite*-polynomials or cumulant expansions, respectively, were also used in derivative pricing (c.f. [Habtemicael and SenGupta 2016a, 2016b](#); [Mazzoni 2010](#); [Xiu 2014](#)). Even though the suggested approach is superficially similar to the work of [Xiu \(2014\)](#), it is based on a completely different idea and has very different implications. [Xiu \(2014\)](#) follows the classical way of representing the unknown pricing density by a *Gram-Charlier*-series and solving the resulting *Feynman-Kac* PDE. There are two major drawbacks involved. Firstly, following the derivation of [Ait-Sahalia \(2002\)](#), there are powers of the infinitesimal generator of the original diffusion involved in the computation of the cumulants. Those terms are exceedingly complicated and have to be evaluated with a computer algebra system, but more importantly, they are model-dependent. Secondly, the *Gram-Charlier*-series is not an asymptotic series in the proper sense (for an excellent discussion of this issue see [Blinnikov and Moessner 1998](#)). Thus, the more expansion terms are involved, the faster the series degenerates, even if the deviation from normality is merely moderate. The approach suggested here is completely model-independent. It derives an asymptotic version of the *Kolmogorov*-backward-PDE from first principles. This equation contains unknown functions to be represented in terms of an orthogonal series expansion using *Hermite*-polynomials, orthogonal with respect to the weighting function e^{-x^2} . This version of the *Hermite*-polynomials is much more robust with respect to deviations from normality and is usually only limited by numerical issues. Of course the coefficients of the orthogonal series expansion have no connection to the cumulants of the unknown density function. They are instead determined explicitly from a system of ordinary differential equations and are related to the empirically observed deviations from normality.

Several aspects of the resulting valuation method are investigated, based on a collection of 1075 European-style index options, contingent on the 'Deutscher Aktienindex' (DAX) index. Because of the decreasing interest rate term structure and some other exceptional market conditions due to the Euro crisis, and the large number of available contracts with bid-offer spreads below 1%, the DAX index is an optimal laboratory to survey the properties of the suggested method. In particular, it is shown that it generates very precise in- and out-of-sample option prices, and that the characteristics of the implied volatility surface are reconstructed quite satisfactorily over wide ranges of moneyness and time to maturity. The remainder of the paper is organized as follows:

Section 2 sets the scene for the asymptotic expansion of the incomplete market transition kernel. Departing from the classical risk-neutral geometrical *Brownian* motion and the corresponding time-dependent probability density function, the stationary coordinate frame transformation is

introduced. Subsequently, the transition density is asymptotically expanded to derive the general equation for the excess dynamics due to market incompleteness. Finally, the unknown functions in this equation are expressed in a suitable way for model fitting.

In Section 3 a complete set of orthogonal functions based on *Hermite*-polynomials is introduced, in order to approximately solve for the incomplete market dynamics. In the process, the partial differential equation is transformed into a solvable linear system of ordinary differential equations. Furthermore, the constituents can be computed recursively. It turns out that there is an intimate connection between the resulting formula for the time-dependent pricing density and option pricing by quadrature methods, which is also elaborated at the end of this section.

In Section 4 the resulting model is calibrated to market data. To this end, a quadratic objective function is defined, which is to be minimized. It is shown that the gradient of this objective function can be computed analytically, which is instrumental for parameter estimation with quasi-*Newton* methods. The results of the calibration procedure are discussed and compared across different model configurations. The order of relative pricing error in the suggested framework is reduced to the order of bid-offer spreads of the contracts in the calibration sample.

Section 5 investigates the quality of the implied volatility surface, generated by the calibrated model. Because many contract types are highly vega-sensitive, implied volatility characteristics are of particular importance. The suggested method is benchmarked against two state-of-the-art approaches: the SABR model of Hagan et al. (2002) and the stochastic volatility inspired (SVI) parametrization of the local volatility surface by Gatheral and Wang (2012), associated with a most likely path approximation. It is shown that both alternatives provide an inferior fit, compared to the conditional density approach suggested here.

In Section 6 a collection of 171 European-style capped call and put options are valued. Those options were not included in the calibration sample and hence form an independent validation sample. It is also shown how to value contracts with arbitrary payoff functions with Monte Carlo simulation. This matter is not trivial, because one is not able to draw directly from the arbitrage-free pricing distribution. Two alternatives—a multinomial approximation and an importance sampling method—are detailed. The results are again in favor of the conditional pricing density approach.

Section 7 concludes the paper with a summary of the results and a discussion of the pros and cons of the suggested method.

2. Asymptotic Expansion of the Pricing Density

Assume a probability space (Ω, \mathcal{F}, P) is fixed, equipped with a natural filtration $\mathcal{F}_0 \subseteq \mathcal{F}_t \subseteq \mathcal{F}$, generated by the P -measurable price processes of the underlying and all derivatives contingent on it, and with all null sets contained in \mathcal{F}_0 . Classical theory (Black and Scholes 1973; Black 1976) entails a unique risk-neutral *Ito*-process

$$dF_t = \sigma F_t dW_t \quad (1)$$

under the T -forward measure Q_T , such that the value of an arbitrary vanilla type contract¹, contingent on its payoff at maturity, is given by

$$V(S_t, t) = B_0(t; T) E^{Q_T} [V(F_T, T) | \mathcal{F}_t]. \quad (2)$$

In Equation (2), $B_0(t; T)$ is a zero-coupon bond with unit face value maturing at time T , and F_t is the forward price of the underlying S . The classical model is rejected with overwhelming empirical proof, partly because of oversimplified assumptions, for example the volatility σ in (1) is assumed constant and known, and partly because it does not properly reflect all sources of (systematic) risk.

¹ In this context, a contract is called vanilla, if it is not path dependent and contains no embedded decisions.

An example for the latter are jump risks. An attempt to overcome this problem is the jump diffusion model of Merton (1976) but in order to preserve market completeness and hence the uniqueness of the pricing measure, jump risks have to be considered purely idiosyncratic, which is barely a realistic assumption (cf. Lewis 2002). Other risks not accounted for are liquidity risks, default risks, and even model risks.

Even though other models like those of Heston (1993), Bates (1996) or Hagan et al. (2002)—which are designed to work properly in incomplete markets after calibration to market data—are extraordinary successful, results of the Black–Scholes model are approximately correct in many situations. Therefore, it seems quite natural to expand around the Black–Scholes solution to obtain a valid result in an incomplete market setup, as long as the deviation from completeness is not too extreme. To set the scene for such an expansion, the time-dependent probability density of the path model (1) is subjected to some basic transformations.

Let $x_t = \log F_t$ be the logarithm of the forward price of the underlying. Due to Ito’s lemma, the (risk-neutral) probability density function of x is governed by the Fokker–Planck-equation

$$\frac{\partial}{\partial t} q_X(x, t) = \frac{1}{2} \sigma^2 \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2} \right) q_X(x, t), \quad q_X(x, t_0) = \delta(x - x_0), \tag{3}$$

with δ indicating Dirac’s delta function. The solution to this PDE problem can be obtained with standard methods like Fourier-transform and is known to be

$$q_X(x, t) = \frac{1}{\sqrt{2\pi\sigma^2(t - t_0)}} e^{-\frac{1}{2} \left(\frac{x - x_0 + \frac{1}{2}\sigma^2(t - t_0)}{\sigma\sqrt{t - t_0}} \right)^2}, \quad \text{for } t > t_0. \tag{4}$$

Note that this density is singular at $t = t_0$, which is not a problem because one valid definition of the delta function is in terms of the limit of a sequence of functions like Equation (4), $\delta(x - x_0) = \lim_{t \rightarrow t_0} q_X(x, t) = q_X(x, t_0)$, (cf. Lighthill 1980, chp. 2.2). One merely has to remember that the initial density is not given by Equation (4), but by its limit. This is an important point for the following transformation, which is singular at $t = t_0$, but the limit relation still holds. Define new coordinates $z(x, t)$ and $\tau(x, t)$, with

$$\begin{aligned} z &= \frac{x - x_0 + \frac{1}{2}\sigma^2(t - t_0)}{\sigma\sqrt{t - t_0}}, \\ \tau &= \sqrt{t - t_0}. \end{aligned} \tag{5}$$

After the transformation $q_X(x, t)dx = q_Z(z, \tau)dz$, the risk-neutral probability density is $q_Z(z, \tau) = \phi(z)$, with $\phi(z) = (2\pi)^{-1/2} \exp(-z^2/2)$ indicating the standard normal probability density function. In this new coordinates the probability density is stationary and standardized, making this particular coordinate system appear more fundamental than all others (a proof of the stationarity of the transformation is provided in Appendix A). It serves indeed as a laboratory frame for investigating the deviations from the Black–Scholes solution in that the asymptotic expansion is constructed in this frame. The PDE problem in the (z, τ) coordinates, corresponding to the problem in Equation (3) in the (x, t) -coordinates, is

$$\frac{\partial}{\partial \tau} q_Z(z, \tau) = 0, \quad q_Z(z, 0) = \phi(z). \tag{6}$$

As discussed previously, one has to remember that $q_Z(z, 0)$ is only the limit of $q_Z(z, \tau)$ when $\tau \rightarrow 0$, because the coordinate transformation is singular at $\tau = 0$.

The universal statement implied by Equation (6) is that in the Black-Scholes world, the standardized risk neutral pricing density is Gaussian and remains Gaussian at all times. One would expect the pricing density to deviate from this stationary density in incomplete markets, reflecting the unhedgeable

systematic risk structure of such markets. This deviation is implemented in the next paragraph by asymptotic expansion.

2.1. Asymptotic Deviation from Market Completeness

In order to determine the mechanism for the deviation from the *Black–Scholes* solution, write the pricing density in terms of the law of total probability

$$q_Z(z, \tau + \Delta\tau) = \int_{-\infty}^{\infty} q_Z(z, \tau + \Delta\tau|y, \tau)q_Z(y, \tau)dy. \tag{7}$$

Observe that in the *Black–Scholes* framework the transition density has to be given by $q_Z(z, \tau + \Delta\tau|y, \tau) = \delta(y - z)$ in order for Equation (7) to obey the degenerate dynamics in Equation (6). The key idea of the approach is to assume that in incomplete markets this transition density deviates from the delta function, and that the systematic risk structure, however it may be composed, is encoded in the way the transition kernel deviates.

To make this idea more precise, account for some boundary conditions. First, in the limit $\Delta\tau \rightarrow 0$ the transition density has to be the delta function because Equation (7) becomes an identity. Thus, one can conclude that the space–time interval, occupied by the transition kernel, has to be proportional to $\Delta\tau$ for short times ($\Delta\tau \ll 1$). Second, the *Black–Scholes* framework often generates useful approximative results, thus the spatial expanse of the transition kernel per unit of $\Delta\tau$ should be small, indicated by ε . Putting these arguments together, one concludes that the space–time volume occupied by the transition kernel should be roughly of order $\varepsilon\Delta\tau$. Next, *Taylor*-expand the initial density $q_Z(y, \tau)$ around z to obtain

$$q_Z(y, t) = \sum_{n=0}^{\infty} \frac{(y - z)^n}{n!} \left(\frac{\partial}{\partial z}\right)^n q_Z(z, \tau), \tag{8}$$

and define the auxiliary functions

$$M_n(z, \tau; \Delta\tau) = \int_{-\infty}^{\infty} (y - z)^n q_Z(z, \tau + \Delta\tau|y, \tau)dy. \tag{9}$$

Now Equation (7) can be expressed in terms of a *Taylor*-like series expansion

$$q_Z(z, \tau + \Delta\tau) = \sum_{n=0}^{\infty} \frac{M_n(z, \tau; \Delta\tau)}{n!} \left(\frac{\partial}{\partial z}\right)^n q_Z(z, \tau), \tag{10}$$

which is very similar to the *Kramers–Moyal*-backward-expansion (Riskén 1989, chp. 4.2). However, this similarity is only superficial, because the integration in Equation (9) is with respect to the conditioning variable y , which means that M_n is not a transition moment and Equation (10) is merely a formal series expansion. Because the transition kernel becomes the delta function in the limit $\Delta\tau \rightarrow 0$, it follows immediately that $M_0(z, \tau; 0) = 1$ and $M_n(z, \tau; 0) = 0$ for all $n \geq 1$. If $q_Z(z, \tau + \Delta\tau|y, \tau)$ is sufficiently smooth, which is a very mild requirement, $M_n(z, \tau; \Delta\tau)$ can be expanded itself around $\Delta\tau = 0$ and one obtains

$$M_n(z, \tau; \Delta\tau) = 0 + f_n(z, \tau)\varepsilon^n\Delta\tau + O(\varepsilon^n\Delta\tau^2), \tag{11}$$

for $n \geq 1$, with the yet unknown functions $f_n(z, \tau)$. Remember that the transition kernel occupies a space–time volume of order $\varepsilon\Delta\tau$. Therefore, the n -th order auxiliary function has to be roughly of order $\varepsilon^n\Delta\tau$. Putting all pieces together one obtains

$$\frac{\partial}{\partial \tau} q_Z(z, \tau) = \lim_{\Delta\tau \rightarrow 0} \frac{q_Z(z, \tau + \Delta\tau) - q_Z(z, \tau)}{\Delta\tau} = \sum_{n=1}^{\infty} \frac{f_n(z, \tau)\varepsilon^n}{n!} \left(\frac{\partial}{\partial z}\right)^n q_Z(z, \tau). \tag{12}$$

Clearly one cannot compute the entire sum on the right hand side of Equation (12) and thus usually a decision has to be made with respect to the terms to abandon. Often terms up to $O(\varepsilon^2)$ are considered and all higher orders are neglected. The situation here is different. Because $q_Z(z, \tau)$ is a probability density, which by definition is nonnegative everywhere, the *Pawula* theorem applies (cf. Risken 1989, chp. 4.3). This remarkable theorem proves that considering the first two terms of the expansion in Equation (12) is the best possible approximation available, without considering the last term at infinity. In this case, the contribution from higher-order terms diminishes, because of their order in ε . Hence, if it is assumed that the contribution of terms of infinite powers of ε vanishes, it can be concluded that the approximation including $O(\varepsilon^2)$ terms has to be exact. One therefore obtains instead of Equation (6) an asymptotic version of the *Kolmogorov*-backward-equation

$$\frac{\partial}{\partial \tau} q_Z(z, \tau) = \varepsilon f_1(z, \tau) \frac{\partial}{\partial z} q_Z(z, \tau) + \frac{\varepsilon^2}{2} f_2(z, \tau) \frac{\partial^2}{\partial z^2} q_Z(z, \tau), \quad q_Z(z, 0) = \phi(z), \tag{13}$$

with the yet unknown functions $f_1(z, \tau)$ and $f_2(z, \tau)$, encoding all information about the deviation of the systematic risk structure from the classical *Black–Scholes* world. The next step is to determine these unknown functions.

2.2. Decoding Market Information

Because of the extremely rich structure of systematic market risk, model-guided determination of the functions $f_n(z, \tau)$, apart from a few special cases discussed subsequently, may be generally impossible. Thus, some assumptions have to be made, allowing for the tractable incorporation of observed empirical information. The following discussion is focused on the $O(\varepsilon)$ function $f_1(z, \tau)$ but all arguments carry over to the $O(\varepsilon^2)$ term. The first assumption is that the function is time separable and that it has the form

$$\varepsilon f_1(z, \tau) = e^{-\gamma \tau} a(z), \tag{14}$$

where the small number ε is soaked up in the function $a(z)$. There are several reasons for this particular choice:

- Recall that the (z, τ) coordinates are already dynamically scaled and hence, only the excess dynamics are to be modeled. These dynamics are governed by additional risk structure, unfolding over time. For example, liquidity risk is of minor importance in short-term scenarios but has to be accounted for over longer holding periods. Jump risk contributes to the steepness of the short term implied volatility smile, but does not affect the long-term structure. If all the additional risk structure is fully deployed, the deformation of the pricing density is completed. This behavior is induced in Equation (14).
- This particular choice reproduces some known standard results. For example, in the limit $\gamma \rightarrow \infty$ one obtains the classical *Black-Scholes* solution. As a second example, imagine a completely illiquid market, such that even static hedging is not possible. The choice $\gamma = 0$ and $a(z) = -\mu/\sigma$ yields the solution $q_Z(z, \tau) = \phi(z - \mu\tau/\sigma)$, which after retransformation to (x, t) -coordinates is immediately recognized as the time-honored actuarial pricing density under P (cf. Derman and Taleb 2005).
- The plain exponential model for the time dependence is the most parsimonious parametrization of the problem. By this choice, the subsequent calibration procedure is simplified considerably. Even if this model is oversimplified in that it implies the whole risk structure to unfold in a synchronized way, it seems to be at least a good starting point.

The second assumption, essential for recursive computation of the pricing density as shown in the next section, is that $a(z)$ is sufficiently smooth to be expanded into a power series, $a(z) = \sum_{k=0}^{\infty} a_k z^k$. This is again a relatively mild technical condition. The whole problem now becomes

$$\frac{\partial}{\partial \tau} q_Z(z, \tau) = e^{-\gamma \tau} \left(\sum_{k=0}^{\infty} a_k z^k \frac{\partial}{\partial z} q_Z(z, \tau) + \sum_{k=0}^{\infty} b_k z^k \frac{\partial^2}{\partial z^2} q_Z(z, \tau) \right), \tag{15}$$

again with initial condition $q_Z(z, 0) = \phi(z)$. Obviously, the sums in Equation (15) cannot be calculated either. However, one can expect very few coefficients to contribute to the sums for the following line of reasoning: Departing from the initial standard *Gaussian* density, roughly 32% of the probability mass is located at $|z| > 1$. The outer region of the density would be exposed to violent deformations if for large k the term $a_k z^k$ or $b_k z^k$ contributes, respectively. Because the *Black–Scholes* solution is a good approximation, the deviation from normality has to be moderate, and hence higher order coefficients have to be minute.

3. Computation of the Pricing Density

In order to solve the PDE (15) at least approximately, the pricing density is rewritten in terms of a complete orthogonal system

$$q_Z(z, \tau) = \sum_{n=0}^{\infty} c_n(\tau) \psi_n(z), \quad \text{with} \quad \psi_n(z) = \frac{H_n(z) e^{-\frac{z^2}{2}}}{\sqrt{2^n n!} \sqrt{\pi}}. \tag{16}$$

In Equation (16), $H_n(z)$ represents the n -th *Hermite*-polynomial, orthogonal to the weight function e^{-z^2} defined by $H_n(z) = (-1)^n e^{z^2} (d/dz)^n e^{-z^2}$. Observe that this is neither *Gram–Charlier*-, nor *Edgeworth*-expansion², which are both constructed from orthogonal functions with respect to the weight function $e^{-z^2/2}$, but a generalized *Fourier*-series with correctly normalized orthogonal functions, such that

$$\int_{-\infty}^{\infty} \psi_n(z) \psi_m(z) dz = \delta_{n,m} \tag{17}$$

holds, with the *Kronecker*-delta $\delta_{n,m}$. Observe further that every orthogonal function $\psi_n(z)$, apart from a constant, contains a standard *Gaussian* term. Thus, this orthogonal system is particularly well-suited for the problem at hand. Of course, one has to fix a maximum number of expansion terms to be included, but for small deviations from the normal distribution the series converges well (Blinnikov and Moessner 1998).

The advantage offered by the *Fourier*-series expansion is the separation of time and spatial dependence. Using this advantage, the time dependent values of the *Fourier*-coefficients in Equation (16) can be computed by solving an ordinary first-order differential equation system.

Proposition 1. *The Fourier-coefficients $c_n(\tau)$ for $n = 0, 1, 2, \dots$ are determined by the solution of the infinite dimensional matrix/vector differential equation*

$$\begin{aligned} \frac{d}{d\tau} c(\tau) &= e^{-\gamma\tau} \left(\sum_{k=0}^{\infty} a_k A^{(k)} + \sum_{k=0}^{\infty} b_k B^{(k)} \right) c(\tau), \quad \text{with} \\ A_{n,m}^{(k)} &= \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{d}{dz} \psi_m(z) \right) dz \quad \text{and} \quad B_{n,m}^{(k)} = \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{d^2}{dz^2} \psi_m(z) \right) dz, \end{aligned} \tag{18}$$

where $c(\tau)$ is the vector containing the coefficients $c_0(\tau), c_1(\tau), c_2(\tau)$, and so forth.

² See Blinnikov and Moessner (1998) for an excellent survey of both expansions and their properties.

Proof. Computing the n th Fourier-coefficient and using Equation (15), one obtains

$$\begin{aligned} \frac{d}{d\tau}c_n(\tau) &= \int_{-\infty}^{\infty} \psi_n(z) \left(\frac{\partial}{\partial\tau}q_Z(z, \tau) \right) dz \\ &= e^{-\gamma\tau} \sum_{k=0}^{\infty} \left(a_k \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{\partial}{\partial z}q_Z(z, \tau) \right) dz \right. \\ &\quad \left. + b_k \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{\partial^2}{\partial z^2}q_Z(z, \tau) \right) dz \right). \end{aligned} \tag{19}$$

Again substituting the complete orthogonal system in Equation (16) for the density function $q_Z(z, \tau)$ yields

$$\begin{aligned} \frac{d}{d\tau}c_n(\tau) &= e^{-\gamma\tau} \sum_{m=0}^{\infty} c_m(\tau) \sum_{k=0}^{\infty} \left(a_k \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{\partial}{\partial z}\psi_m(z) \right) dz \right. \\ &\quad \left. + b_k \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{\partial^2}{\partial z^2}\psi_m(z) \right) dz \right). \end{aligned} \tag{20}$$

Identifying the integrals as elements

$$A_{n,m}^{(k)} = \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{d}{dz}\psi_m(z) \right) dz \quad \text{and} \quad B_{n,m}^{(k)} = \int_{-\infty}^{\infty} z^k \psi_n(z) \left(\frac{d^2}{dz^2}\psi_m(z) \right) dz \tag{21}$$

of the matrices $A^{(k)}$ and $B^{(k)}$, the problem in Equation (20) can be rewritten in matrix/vector form as

$$\frac{d}{d\tau}c(\tau) = e^{-\gamma\tau} \left(\sum_{k=0}^{\infty} a_k A^{(k)} + \sum_{k=0}^{\infty} b_k B^{(k)} \right) c(\tau), \tag{22}$$

which is the desired result. \square

It turns out that the matrices $A^{(k)}$ and $B^{(k)}$ can be computed recursively, exploiting particular properties of the *Hermite*-polynomials contained in the orthogonal functions. This is a very convenient fact, because recursive patterns can be efficiently implemented on a computer. The procedures are detailed in the next paragraph.

3.1. Recursive Computation of the Matrix Entries

The computation scheme for the entries of the matrices $A^{(k)}$ and $B^{(k)}$ is given in the subsequent proposition.

Proposition 2. For $C^{(k)} = A^{(k)}$ or $C^{(k)} = B^{(k)}$, the following recursion holds

$$C_{n,m}^{(k)} = \sqrt{\frac{n}{2}} C_{n-1,m}^{(k-1)} + \sqrt{\frac{n+1}{2}} C_{n+1,m}^{(k-1)}, \tag{23}$$

with initial conditions

$$\begin{aligned} A_{n,m}^{(0)} &= \sqrt{\frac{m}{2}} \delta_{n,m-1} - \sqrt{\frac{m+1}{2}} \delta_{n,m+1} \quad \text{and} \\ B_{n,m}^{(0)} &= \sqrt{\frac{m(m-1)}{2}} \delta_{n,m-2} - \frac{2m+1}{2} \delta_{n,m} + \sqrt{\frac{(m+1)(m+2)}{2}} \delta_{n,m+2}. \end{aligned} \tag{24}$$

Proof. In order to compute the entries of the matrices $A_{n,m}^{(k)}$ and $B_{n,m}^{(k)}$, two essential properties of *Hermite*-polynomials are used. First, the recursive relation between the polynomials

and their derivatives $(d/dz)H_m(z) = 2mH_{m-1}(z)$, and second, the recurrence relation $H_{n+1}(z) = 2zH_n(z) - 2nH_{n-1}(z)$ (cf. Abramowitz and Stegun 1970, p. 782). From property number 1 and the definition of the orthogonal functions (16), one immediately obtains

$$\frac{d}{dz}\psi_m(z) = \sqrt{\frac{m}{2}}\psi_{m-1}(z) - \sqrt{\frac{m+1}{2}}\psi_{m+1}(z), \tag{25}$$

and thus for $k = 0$,

$$\begin{aligned} A_{n,m}^{(0)} &= \sqrt{\frac{m}{2}}\delta_{n,m-1} - \sqrt{\frac{m+1}{2}}\delta_{n,m+1} \quad \text{and} \\ B_{n,m}^{(0)} &= \sqrt{\frac{m(m-1)}{2}}\delta_{n,m-2} - \frac{2m+1}{2}\delta_{n,m} + \sqrt{\frac{(m+1)(m+2)}{2}}\delta_{n,m+2} \end{aligned} \tag{26}$$

follow from the orthonormality in Equation (17). From the recurrence relation, one obtains in terms of the orthogonal functions

$$z\psi_n(z) = \sqrt{\frac{n}{2}}\psi_{n-1}(z) + \sqrt{\frac{n+1}{2}}\psi_{n+1}(z), \tag{27}$$

which yields the recursive relation

$$C_{n,m}^{(k)} = \sqrt{\frac{n}{2}}C_{n-1,m}^{(k-1)} + \sqrt{\frac{n+1}{2}}C_{n+1,m}^{(k-1)} \tag{28}$$

for both matrices $C^{(k)} = A^{(k)}$ and $C^{(k)} = B^{(k)}$, respectively. \square

Note that in the exact case of Equation (18) there are infinitely many matrices of infinite dimensions. Thus, one has to decide how many terms of the orthogonal expansion to include in the computation, and which terms of the power series expansion of the unknown functions $a(z)$ and $b(z)$ to abandon. The former is primarily a technical question of convergence of the Fourier-series, while the latter is a question of approximating the dynamics of the risk structure correctly. Both are strongly related to the degree of deviation from normality, but only the coefficients a_k and b_k immediately affect the model calibration process. Because it is impossible to determine beforehand which terms of both expansions might be neglected, different alternatives are compared in Section 4.

3.2. Fourier-Coefficients and Pricing Density

Once the approximations are fixed, the Fourier-coefficients can be calculated immediately. Knowing all constituents of the matrices $A^{(k)}$ and $B^{(k)}$ from the recursive scheme in Proposition 2, the system of ordinary differential Equation (18) can be solved

$$\begin{aligned} c(\tau) &= \exp \left[\int_0^\tau e^{-\gamma s} \left(\sum_{k=0}^{K_a} a_k A^{(k)} + \sum_{k=0}^{K_b} b_k B^{(k)} \right) ds \right] c(0) \\ &= \exp \left[\frac{1 - e^{-\gamma\tau}}{\gamma} \left(\sum_{k=0}^{K_a} a_k A^{(k)} + \sum_{k=0}^{K_b} b_k B^{(k)} \right) \right] c(0), \end{aligned} \tag{29}$$

with $c_0(0) = (4\pi)^{-1/4}$ and $c_n(0) = 0$ for $n \geq 1$, and $\exp[\dots]$ denoting the matrix exponential. There are several alternative methods for calculating a matrix exponential (cf. Moler and van Loan 2003), such that Equation (29) is quite explicit. Thus, the whole pricing density can be approximated by

$$q_Z(z, \tau) \approx \sqrt[4]{4\pi} \phi(z) \sum_{n=0}^N \frac{c_n(\tau)}{\sqrt{2^n n!}} H_n(z), \tag{30}$$

or after retransformation into (x, t) -coordinates

$$q_X(x, t) \approx \frac{\sqrt[4]{4\pi}}{\sigma\sqrt{t-t_0}} \phi\left(\frac{x-x_0 + \frac{1}{2}\sigma^2(t-t_0)}{\sigma\sqrt{t-t_0}}\right) \times \sum_{n=0}^N \frac{c_n(\sqrt{t-t_0})}{\sqrt{2^n n!}} H_n\left(\frac{x-x_0 + \frac{1}{2}\sigma^2(t-t_0)}{\sigma\sqrt{t-t_0}}\right). \tag{31}$$

The retransformation however is of little practical use, because the payoff function of an arbitrary contingent claim can be easily transformed into (z, τ) -coordinates.

Determination of the power series coefficients a_k and b_k is a matter of calibration to market data and has to be done numerically. Completing this procedure results in a model for the time evolution of the arbitrage-free pricing density, conditional on the information set \mathcal{F}_t , representing the market view of future risks as it stands today. There is no requirement for interpolation or even extrapolation like in case of nonparametric estimates of implied volatility surfaces. However, one big assumption has been made about the time separability of the unknown functions in Equation (13) and the temporal structure of risks in Equation (14). As stated above, this assumption may be oversimplified, leaving a margin for enhancement of the model fit, at the cost of increased complexity.

3.3. Pricing Vanilla Contracts

One major drawback of the pricing density (Equations (30) and (31)) is that analytical valuation of vanilla contracts is much harder than in the *Black–Scholes* case, where all higher polynomial terms H_n vanish. However, there is a very convenient and computationally efficient way of pricing derivatives numerically, based on *Gauss–Hermite*-quadrature. The value of a vanilla contract at time $t_0 = 0$, maturing at time $t = T$ is given by

$$V(S_0, 0) = B_0(0; T) E^{Q_T} \left[V(F_T, T) \middle| \mathcal{F}_0 \right] = B_0(0; T) \int_{-\infty}^{\infty} V\left(F_0 e^{z\sigma\sqrt{T} - \frac{1}{2}\sigma^2 T}, T\right) q_Z(z, \sqrt{T}) dz. \tag{32}$$

The connection between the forward price and the stock price at t_0 is given by $F_0 = S_0/B_0(0; T)$. Because $q_Z(z, \tau)$ has a leading standard normal density function, the integral can be approximated by a weighted sum

$$V(S_0, 0) \approx B_0(0; T) \sum_{j=1}^J w_j V\left(F_0 e^{z^{(j)}\sigma\sqrt{T} - \frac{1}{2}\sigma^2 T}, T\right) \frac{\sqrt[4]{4\pi}}{\sqrt{2^n n!}} \sum_{n=0}^N c_n(\sqrt{T}) H_n(z^{(j)}), \tag{33}$$

with w_j indicating the *Gauss–Hermite*-quadrature weights and $z^{(j)}$ the corresponding quadrature points. All necessary information about weights and points can be extracted from the eigensystem of the matrix,

$$M = \begin{pmatrix} 0 & \sqrt{1} & 0 & \dots & 0 \\ \sqrt{1} & 0 & \sqrt{2} & & \vdots \\ 0 & \sqrt{2} & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & \sqrt{j-1} \\ 0 & \dots & 0 & \sqrt{j-1} & 0 \end{pmatrix}, \tag{34}$$

cf. [Golub \(1973\)](#). The eigenvalues of M are the quadrature points, whereas the corresponding weights are given by the squared first components of the corresponding normalized eigenvectors. The quadrature is exact for polynomials up to a degree of $2j - 1$, indicating an intimate relation between the *Hermite*-polynomials involved and the necessary number of quadrature points. However,

the payoff function of plain vanilla calls and puts is not polynomial, leaving Equation (33) as an approximation.

Observe that not all terms in (33) rely on the full information of an individual contract. The *Fourier*-coefficients $c_n(\sqrt{T})$ for example only depend on the time to maturity of the contract, whereas the *Hermite*-polynomials $H_n(z^{(j)})$ do not depend on the contract at all. This suggests an efficient way of pricing individual contracts. Defining the vector $V^{(m)}$ with components $V_j^{(m)} = w_j V_m(F_0 e^{z^{(j)} \sigma \sqrt{T} - \frac{1}{2} \sigma^2 T}, T)$ and the matrix H , with $H_{j,m} = \sqrt[4]{4\pi} / \sqrt{2^n n!} H_n(z^{(j)})$, the fair value of the m -th contract is given by

$$V_m(S_0, 0) \approx B_0(0; T) (V^{(m)})' H c(\sqrt{T}), \tag{35}$$

where H has to be computed only once, and $c(\sqrt{T})$ only once for each expiry of the whole set of contracts.

Equipped with both a model for the time evolution of the arbitrage-free pricing density and a method for pricing plain vanilla options in this framework, the next step is calibrating the model to market data.

4. Calibration to Market Data

The model calibration process primarily contains three more or less interdependent tasks:

- Determination of a sufficient number of *Fourier*-terms to be included in the approximation.
- Determination of the optimal model order K_a and K_b .
- Estimation of the model parameters γ, a_k, b_k based on the available empirical data.

The determination of the optimal number of *Fourier*-coefficients is only weakly related to the model order. It is primarily affected by the degree of deviation from normal. The more extensive the deviation is, the more terms are required for the orthogonal series expansion to converge. Theoretically, an arbitrary probability density can be approximated with sufficient precision by simply including enough expansion terms. Practically, numerical problems have to be considered if the desired density function deviates extensively from the standard normal. Because of finite numerical precision, at some point including additional terms is no longer beneficial because of rounding errors, effectively limiting the manageable degree of deviation from normal. Beyond this limit, artifacts like local negative densities may occur, which cannot be removed or may even be amplified by involving more expansion terms.

Both remaining determinations are strongly interdependent in that a sufficient model order can only be identified by judging the fit accomplished by different models. To this end, all potentially qualifying candidates have to be estimated. This is done numerically by a *Newton–Raphson* type scheme, associated with a prespecified objective function. Usually, a weighted sum of squared errors is to be minimized. In this case

$$Q = \sum_{m=1}^M e^{-\beta_m} \left(\frac{V_m - V_m^{Obs.}}{V_m^{Obs.}} \right)^2 \tag{36}$$

is used, with $V_m^{Obs.}$ indicating the observed mid-price of the m -th contingent claim. The weight factor β_m may be chosen to reflect uncertainty induced by the magnitude of the bid-offer spread. In the study at hand, a large number of vanilla derivatives with spreads below 1% was available, and thus the individual weight factors were set to $\beta_m = 0$ for $m = 1, \dots, M$.

4.1. Data Description

In this analysis, a total of 433 European plain vanilla call options and 471 put options on the ‘Deutscher Aktien’ (DAX) index, quoted as per closing prices on 23 of July 2012, were available. In this case, 501 of these 904 contracts exhibited a bid-offer spread smaller or equal to 1% and thus were used

for model calibration. Additionally, 95 capped calls and 76 capped puts of the same date were used as a validation sample, although they could have been as well used for calibration³.

Figure 1 shows the relative pricing error under the *Black–Scholes* model for the 501 low spread contracts. The DAX index itself was quoted at 6419.33 points and the annualized at-the-money (ATM) implied volatility was about 22.5%.

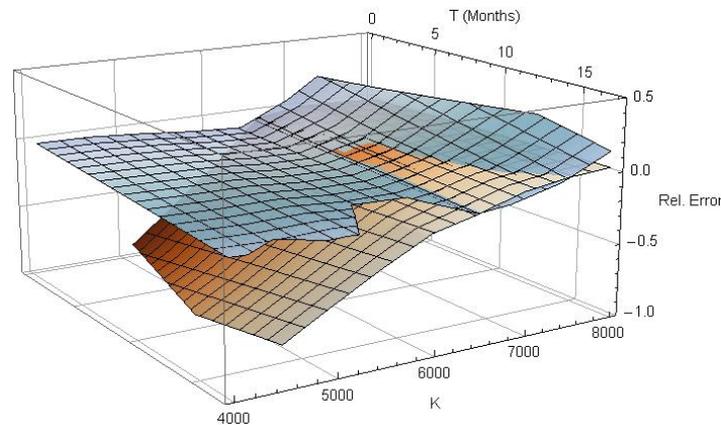


Figure 1. Relative pricing error of European plain vanilla calls (blue) and puts (red) under *Black–Scholes*.

The interpolated call (blue) and put (red) surfaces in Figure 1 indicate that the relative error under *Black–Scholes* is moderate for in-the-money contracts, but grows formidable for out-of-the-money contracts. By using the right pricing density, both surfaces should be flattened out in time-to-maturity, as well as in moneyness direction.

The term structure and hence the prices of zero-coupon bonds of different maturities were extracted from the calibration sample as well by using put-call parity. After simple algebraic manipulations the bond price is explicitly given by

$$B_0(0; T) = \frac{S_0 - C(S_0; K, T) + P(S_0; K, T)}{K}. \tag{37}$$

The yield curve extracted from the zero-coupon bond prices is inverted, falling from a 2.25% return for 4 weeks time to maturity to approximately 0.42% for an 18-month bond. This shape of the yield curve is due to the sovereign debt crisis, affecting the Euro area since 2010.

Based on the empirical data, the next step is calibrating the model and identifying a suitable model order. To this end, the whole set of parameters has to be estimated for different model alternatives.

4.2. Gradient of the Objective Function

Define the whole parameter vector $\theta = (\gamma, a_0, \dots, a_{K_a}, b_0, \dots, b_{K_b})'$, for a given model. A suitable estimate for θ can be obtained recursively, departing from a given initial configuration, by an iterative *Newton–Raphson* type scheme

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \alpha_i \mathcal{H}(\hat{\theta}^{(i)})^{-1} \nabla Q(\hat{\theta}^{(i)}). \tag{38}$$

In Equation (38) α_i indicates an individual step size factor, determined by step halving or trust region methods⁴, \mathcal{H} is a model *Hessian*, for example the identity matrix, resulting in a steepest descent

³ All data was provided by a service of ‘SIX Financial Information’ (<http://www.six-financial-information.com>) and ‘Smarthouse Media GmbH’ (<http://www.smarthouse.de>).

⁴ For an excellent treatment of numerical optimization techniques see [Dennis and Schnabel \(1983\)](#).

algorithm, and ∇Q is the gradient of the objective function, defined componentwise by $\nabla Q_j = \partial Q / \partial \theta_j$. In the present analysis, the BFGS-method of [Broyden \(1970\)](#); [Fletcher \(1970\)](#); [Goldfarb \(1970\)](#); [Shanno \(1970\)](#) has been used, because it converges rapidly and no second derivatives are involved in the computation of the *Hessian* model. Thus, an analytical expression for the gradient of the objective function eliminates the need for finite difference approximations entirely. It turns out that such an expression can be derived, at least approximately. First, note that the partial derivative of the m -th term of the objective function in Equation(36), with respect to the j -th parameter is given by

$$\frac{\partial Q_m}{\partial \theta_j} = 2 \frac{Q_m}{V_m^{Obs.}} \frac{\partial V_m}{\partial \theta_j} \approx 2 \frac{Q_m}{V_m^{Obs.}} B_0(0; T) (V^{(m)})' H \frac{\partial c(\sqrt{T})}{\partial \theta_j}, \tag{39}$$

where the approximate value of the contract (Equation (35)) was plugged in on the right-hand side of Equation (39). Obviously, the partial derivatives of Q are linear functions of the partial derivatives of the *Fourier*-coefficients given by Equation (29). They are given here as a proposition, the proof of which can be found in Appendix B.

Proposition 3. *The partial derivatives of the Fourier coefficient vector $c(\tau)$ with respect to γ , a_k and b_k are given by*

$$\frac{\partial c(\tau)}{\partial \gamma} = \frac{(1 + \gamma\tau)e^{-\gamma\tau} - 1}{\gamma^2} Xc(\tau) \tag{40}$$

$$\frac{\partial c(\tau)}{\partial a_k} = \begin{bmatrix} I & 0 \end{bmatrix} \exp \left[\frac{1 - e^{-\gamma\tau}}{\gamma} \tilde{A}^{(k)} \right] \begin{bmatrix} 0 \\ c(0) \end{bmatrix} \tag{41}$$

$$\frac{\partial c(\tau)}{\partial b_k} = \begin{bmatrix} I & 0 \end{bmatrix} \exp \left[\frac{1 - e^{-\gamma\tau}}{\gamma} \tilde{B}^{(k)} \right] \begin{bmatrix} 0 \\ c(0) \end{bmatrix}, \text{ with}$$

$$X = \sum_{k=0}^{K_a} a_k A^{(k)} + \sum_{k=0}^{K_b} b_k B^{(k)} \quad \text{and} \quad \tilde{A}^{(k)} / \tilde{B}^{(k)} = \begin{bmatrix} X & A^{(k)} / B^{(k)} \\ 0 & X \end{bmatrix},$$

with $[\dots]$ indicating a block matrix and I the $(n \times n)$ identity matrix.

Now, one is able to estimate different models and to compare their fit with respect to their model order and the residual square error.

4.3. Results of Model Calibration

Table 1 shows the results of the calibration process involving $n = 45$ *Fourier*-coefficients, which turned out to be sufficient over all model orders. Each cell shows the residual root-mean-square error (RMSE) and the estimated standardized pricing density $q_Z(z, 1)$ for contracts with time to maturity $T = 1$ year. This is quite close to the stationary density for the most models. Obviously, the pricing density exhibits significant skewness and a pronounced left tail. The deviation from normal is excessive, resulting in invalid density estimates for some model candidates, indicated in gray in Table 1. The degeneration of the density estimates is due to the numerical limits of the orthogonal series expansion and could not be remedied in the present analysis by involving more *Fourier*-terms. Nevertheless, there are some valid and parsimonious candidates with small RMSE, like the (2, 2)-model.

Table 1. Model calibration for $n = 45$ Fourier-coefficients. RMSE: root-mean-square error.

	$K_b = \emptyset$	$K_b = 0$	$K_b = 2$	$K_b = 4$
$K_a = \emptyset$	RMSE: 22.44% 	RMSE: 20.02% 	RMSE: 11.91% 	RMSE: 11.42%
$K_a = 0$	RMSE: 21.56% 	RMSE: 15.52% 	RMSE: 4.97% 	RMSE: 1.66%
$K_a = 1$	RMSE: 17.00% 	RMSE: 15.51% 	RMSE: 3.07% 	RMSE: 1.43%
$K_a = 2$	RMSE: 2.43% 	RMSE: 2.42% 	RMSE: 1.54% 	RMSE: 1.47%
$K_a = 3$	RMSE: 1.99% 	RMSE: 1.73% 	RMSE: 1.50% 	RMSE: 1.16%
$K_a = 4$	RMSE: 1.48% 	RMSE: 1.67% 	RMSE: 1.03% 	RMSE: 1.09%

Note that only models with even orders of K_b are reported. This is due to the definition of the auxiliary functions in Equation (9). Because M_2 has a quadratic kernel, the function $f_2(z, \tau)$ in Equation (11) should always be positive and thus a power series approximation of this function should be given by a polynomial of an even degree.

Figure 2 shows the relative pricing error for the calibration sample of 501 European plain vanilla call- and put-options. Obviously, the observed prices are reconstructed very precisely by the (2,2)-model across the whole spectrum of moneyness and time to maturity. The exact parameter estimates for this model are $\hat{\gamma} = 2.2636$, $\hat{a} = (-1.4561, 0.1086, 0.4159)'$ and $\hat{b} = (-0.1694, 0.1490, 0.0496)'$. The root-mean-square error is 1.54%, which is roughly the order of the bid-offer spreads of the valued contracts, suggesting that a sufficient fit has been accomplished.

Therefore, the (2,2)-model will be used in all subsequent benchmarks and numerical computations. All models were estimated with initial parameter setting $\gamma = 1$ and $a_k = b_k = 0$.

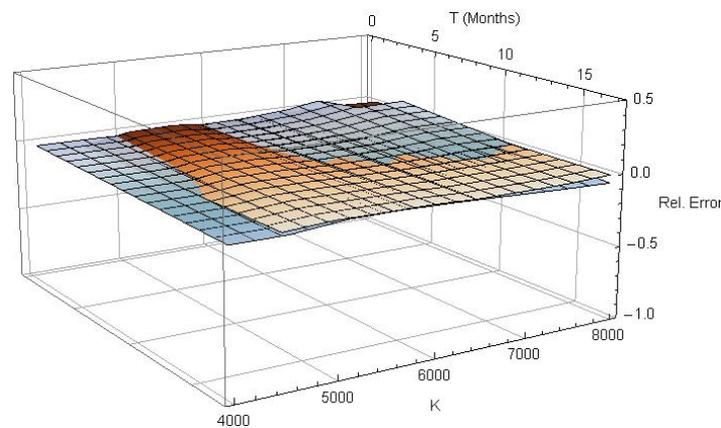


Figure 2. Relative pricing error of European plain vanilla calls (blue) and puts (red) for $K_a = 2, K_b = 2$ and $n = 45$.

5. Implied Volatility Surface

In this section, the implied volatility surface, induced by the preferred (2,2) model of Section 4, is analyzed and compared with other methods. The benefit of this investigation is twofold: on the one hand, volatility surfaces are a widely used tool for calibration of option pricing models to market data (for an excellent survey on this subject see Gatheral 2006). Their strengths and weaknesses are well known, and hence they are a convenient instrument for assessing the quality of the suggested method.

On the other hand, conditional pricing density estimation has one important conceptual drawback: since the whole density is globally conditioned on the information set \mathcal{F}_0 , there is no way to get access to the transition density between times s and t for $s > t_0$. This implies that valuation of path-dependent contracts by Monte Carlo simulation is not possible directly. However, this can be remedied by a kind of reverse engineering. One can use the Dupire-equation (Dupire 1994) to express the local volatility in terms of implied volatility (for details see for example Van der Kamp 2009, sct. 2.3). Simulation can then be performed using a geometrical Brownian motion under local volatility as model for the underlying.

The implied volatility surface is compared with the resulting surfaces of two standard approaches: the SABR model of Hagan et al. (2002) and a local volatility surfaces parametrization suggested by Gatheral and Wang (2012).

5.1. The SABR Model

Hagan et al. (2002) suggested a parametrization of implied volatility based on an asymptotic analysis of a parsimonious stochastic volatility model with singular perturbation methods. Their model is widely used because it is extraordinary easy to fit and generates correct implied volatility dynamics. Their general asymptotic formula is

$$\sigma_{imp}(K, T) = \frac{\alpha}{(F_0 K)^{\frac{1-\beta}{2}} \left(1 + \frac{(1-\beta)^2}{24} \log^2[F_0/K] + \frac{(1-\beta)^4}{1920} \log^4[F_0/K] + \dots \right)} \cdot \frac{z}{\chi(z)} \cdot \left(1 + \left(\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(F_0 K)^{1-\beta}} + \frac{1}{4} \frac{\rho \beta v \alpha}{(F_0 K)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24} v^2 \right) T + \dots \right), \tag{42}$$

with

$$z = \frac{v}{\alpha} (F_0 K)^{\frac{1-\beta}{2}} \log[F_0/K] \quad \text{and} \quad \chi(z) = \log \left[\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho} \right]. \tag{43}$$

Define the (inverse) log-moneyness $k = \log[K/F_0]$ and observe that the backbone of the implied volatility surface in Figure 3 (top left) does not drift vertically in time. Thus, one can set $\beta = 1$ (for details see Hagan et al. (2002)). With these new parameters one obtains

$$\sigma_{imp}(k, T) \approx -\frac{vk}{\chi(z)} \left(1 + \left(\frac{\rho v \alpha}{4} + \frac{2 - 3\rho^2}{24} v^2 \right) T \right), \tag{44}$$

with $z = -vk/\alpha$ and $\chi(z)$ as in Equation (43). This can be easily fitted to the calibration sample, and the resulting implied volatility surface is shown in Figure 3 (bottom right). It is however not entirely fair to calibrate the SABR model to the entire volatility surface, because it does not provide temporal dynamics by construction.

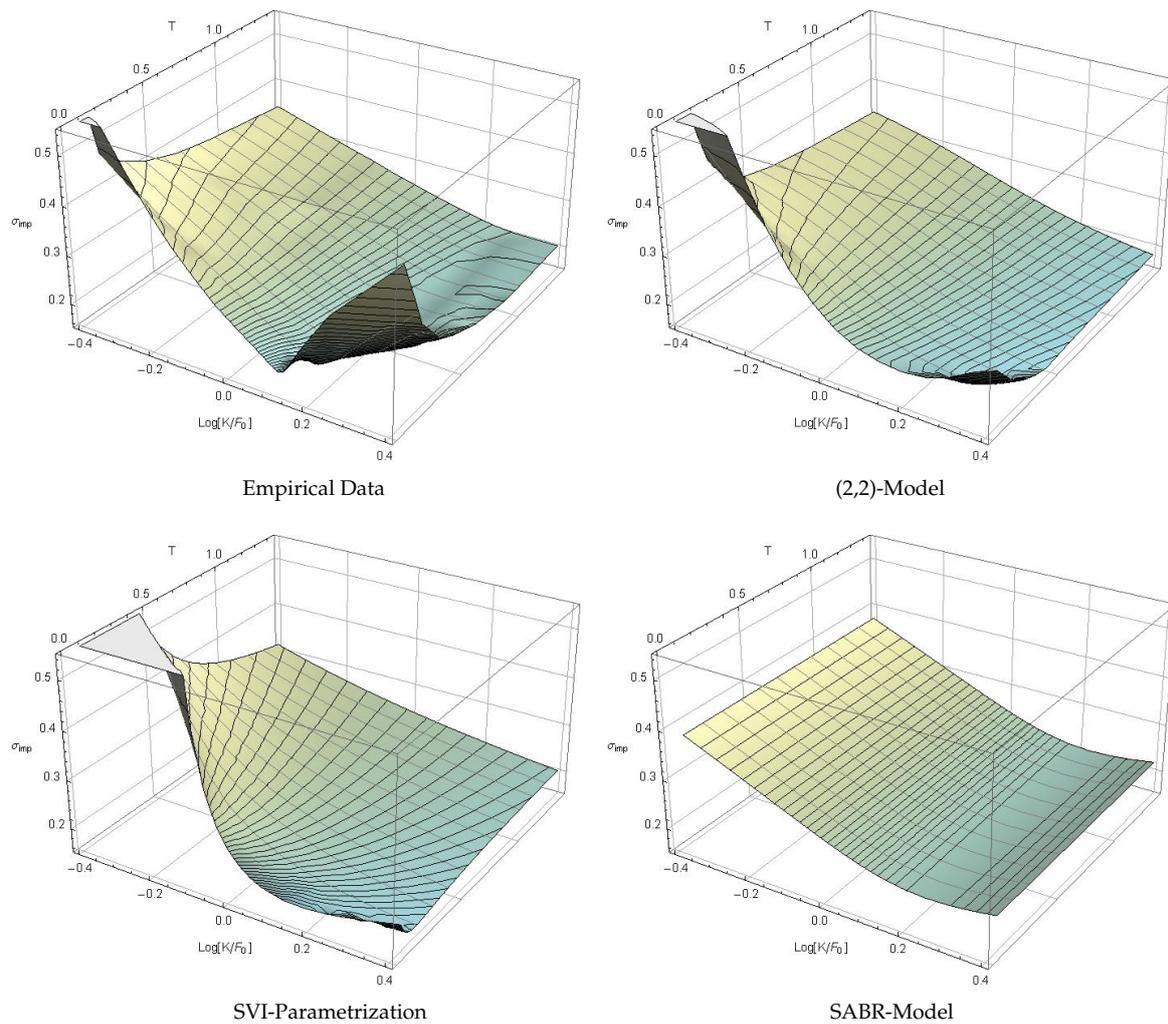


Figure 3. Implied volatility surfaces—**top left:** linear interpolated data, **top right:** estimated (2,2)-model, **bottom left:** stochastic volatility inspired (SVI) parametrization, **bottom right:** SABR model.

5.2. The SVI Parametrization of the Local Volatility Surface

In their paper, Gatheral and Wang (2012) suggest a parametrization of the local volatility surface, motivated by the structure of stochastic volatility models (‘stochastic volatility inspired’, SVI)

$$\sigma_{loc}^2(k, T) = a + b \left(\rho \left(\frac{k}{\sqrt{T}} - m \right) + \sqrt{\left(\frac{k}{\sqrt{T}} - m \right)^2 + \delta^2 T} \right), \tag{45}$$

which is effectively a hyperbola in the log-strike k . There is an intimate connection between local and implied volatility. Implied variance is approximately the average over local variance

$$\sigma_{imp}^2(k, T) \approx \frac{1}{T} \int_0^T \sigma_{loc}^2(\tilde{x}(t), t) dt \tag{46}$$

along the most likely path $\tilde{x}(t)$ from $\tilde{x}(0) = 0$ to $\tilde{x}(T) = k$ (cf. Gatheral 2006, chp. 3). Usually it is very difficult to compute this path and several approaches have been suggested (an incomplete list covers the work of Berestycki et al. 2002; Gatheral and Wang 2012; Gatheral et al. 2012; Guyon and Henry-Labordere 2011; Reghai 2006). However, it turns out that the straight line in the log-strike space is a reasonable first guess. Under this assumption, the line integral in Equation (46) can be expressed as

$$\sigma_{imp}^2(k, T) \approx \frac{1}{T} \int_0^1 \sigma_{loc}^2(\alpha k, \alpha T) d\alpha. \tag{47}$$

The integral in Equation (47) with respect to the SVI parametrization (45) can be computed analytically. Neglecting the constant of integration, one obtains

$$\begin{aligned} \int \sigma_{loc}^2(\alpha k, \alpha T) d\alpha &= \alpha(a - b\rho) + \frac{2bk\rho\alpha^{3/2}}{3\sqrt{T}} \\ &+ \frac{bg \left(2h^2\alpha - hkm\sqrt{\alpha T} + m^2T(2\delta^2T^2 - k^2) \right)}{3h^2} \\ &+ \frac{bk\delta^2m^3T^3 \log \left[h\sqrt{\alpha} + \sqrt{T}(g\sqrt{h} - km) \right]}{h^{5/2}}, \end{aligned} \tag{48}$$

with

$$h = k^2 + \delta^2T^2 \quad \text{and} \quad g = \sqrt{\left(\sqrt{\frac{\alpha}{T}}k - m \right)^2 + \alpha\delta^2T}. \tag{49}$$

Differentiating (48) with respect to α shows that the integral is indeed correct.

One can now fit the implied volatility surface to the calibration sample. The result is shown in Figure 3 (bottom left).

5.3. Results of the Benchmark

In order to compute the implied volatility surface, *Black–Scholes* implied volatilities were calculated for all out-of-the-money plain vanilla calls and puts, because they contain the most information about the volatility structure. This leaves 210 observations of the original low spread sample of 501 options, used for model calibration. This sample is also used for estimation of the SVI and SABR parameters in order to fit all models with identical information. The full sample of 904 contracts provides 613 observations of implied volatility. The particular model fits are also benchmarked regarding the full sample.

Figure 3 shows all estimated implied volatility surfaces. In particular, a first-order spline interpolation of the observation data is given in the upper left quadrant of Figure 3. The upper right surface is generated by the calibrated (2, 2)-model for the pricing density of Section 4. The lower left and lower right surfaces are generated by the SVI parametrization of the local volatility surface and by the SABR model, respectively. The meshing on the surfaces indicates slices of identical time to maturity (gray) and identical implied volatility (black), to emphasize the different features of the particular surfaces.

Obviously, none of the suggested models seems to manage the extremely sharp smile in the ultra short-term region, but this conclusion should be drawn with caution. Short-term out-of-the-money options are usually traded rarely and hence, quoted prices are not unconditionally reliable. In the data sample used in this analysis, information about the trading frequency was not provided. The surface

based on the conditional pricing density (top right) nevertheless seems to cover the features of the empirical surface quit well, at least for $k < 0.25$. On the right edge it slightly underestimates the smile. The SVI parametrization (bottom left) generates an adequate long-term skew but an excessive smile for $k < 0$. It also misses the flattening of the surface for $k > 0.2$ and $T > 0.5$. The SABR model surface in the bottom right quadrant seems to cover this particular feature but completely misses the short term structure of the volatility smile.

The difference between the observed implied volatility surface and the values generated by the three competitive approaches is shown in Figure 4, focusing on the central moneyness region. The surface meshing again indicates slices of identical time to maturity (gray) and identical implied volatility (black). The conditional density (2,2)-model (top left of Figure 4) fits the observed implied volatility extremely accurately, whereas the SVI parametrization (top right) underestimates the mid- and long-term skew, and the SABR model (bottom center) does not generate the correct smile. It is evident from Figure 4 that the conditional density model generates the best implied volatility fit of all candidates.

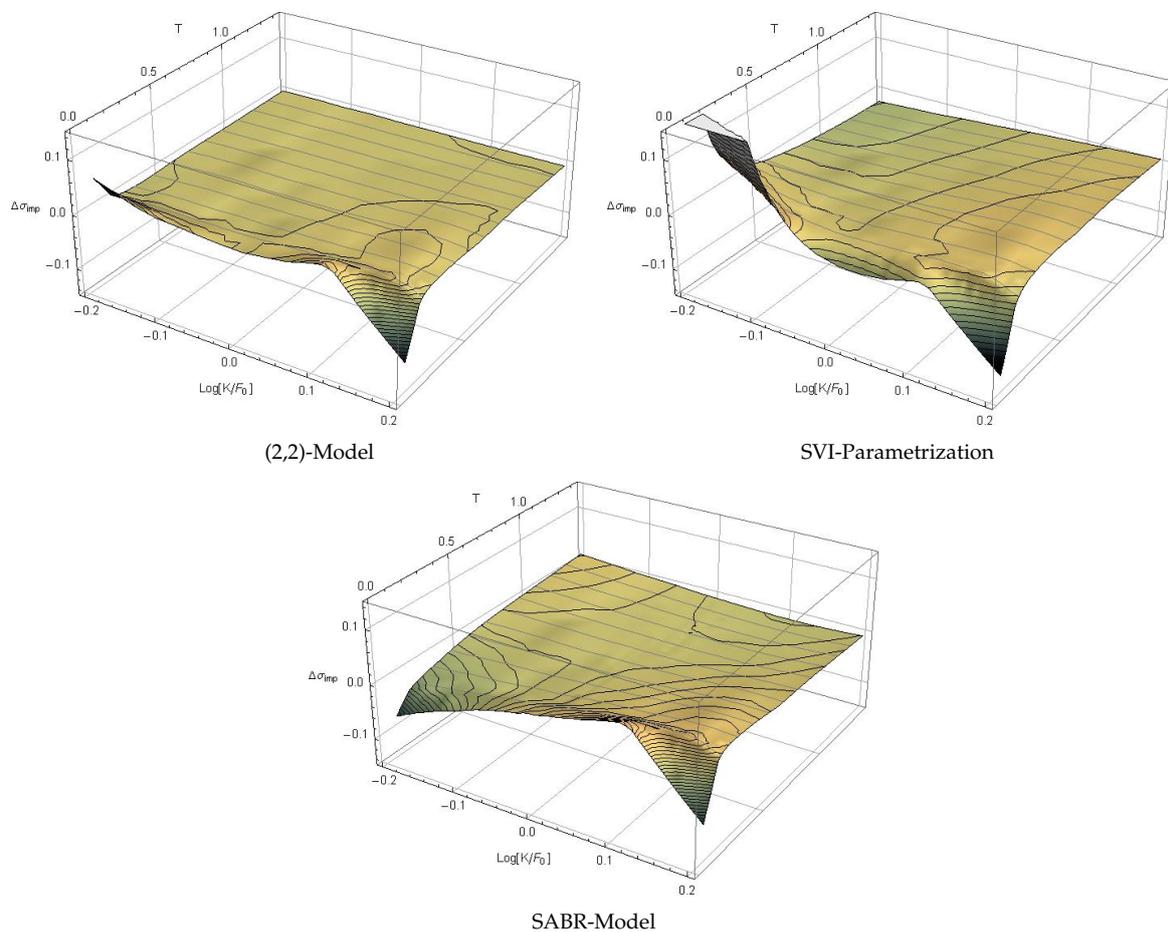


Figure 4. Difference between observed implied volatility and conditional density (2,2)-model (top left), SVI parametrization (top right) and SABR model (bottom center).

Table 2 summarizes all models and compares the root-mean-square errors in both the calibration sample (CS) and the full sample (FS).

Table 2. RMSE of estimated implied volatility surfaces in the calibration sample (CS) and the full sample (FS).

Model	Parameters	RMSE CS ($M = 210$)	RMSE FS ($M = 613$)
(2, 2)	$\gamma = 2.2636$	0.38%	6.69%
	$a_0 = -1.456$		
	$a_1 = 0.1086$		
	$a_2 = 0.4159$		
	$b_0 = -0.169$		
	$b_1 = 0.1490$		
	$b_2 = 0.0496$		
SVI	$a = -0.004$	1.10%	68.68%
	$b = 0.7373$		
	$\delta = 0.4259$		
	$\rho = -0.995$		
	$m = -0.466$		
SABR	$\alpha = 0.2478$	0.82%	7.77%
	$\beta = 1.0000$		
	$\rho = -0.609$		
	$\nu = 0.9550$		

Again, the conditional pricing density model clearly provides the best fit, in particular in the calibration sample, where its root mean square error is smaller than half the RMSE of the SABR model.

6. Valuation under the Conditional Pricing Density Model

In this section an additional validation sample of 95 European vanilla capped calls and 76 puts of the same style is priced and analyzed. This is again accomplished numerically by Gauss–Hermite-quadrature methods like in Section 3.3. An alternative method for valuation is Monte Carlo simulation. Two different simulation approaches are introduced, one immediately related to quadrature methods, and the other based on importance sampling.

6.1. Capped Options Valuation

A European vanilla capped option is a unification of a long and a short position in the same plain vanilla type option, with identical time to maturity but different exercise prices, known as vertical spread. For example, the payoff of a European vanilla capped call option, with strike K and cap $C > K$ is

$$V(F_T, T) = \max[\min[F_T, C] - K, 0]. \tag{50}$$

With this payoff function, the valuation Equations (33) and (35) respectively and immediately apply.

Figure 5 shows the relative misspricing under the classical Black–Scholes model (left) and the estimated (2, 2)-model of Section 4 (right).

Obviously, the pricing error is reduced dramatically. The root-mean-square error under the original Black–Scholes model is 10.72%, whereas the remaining RMSE after conditional pricing density model fitting is 3.19%. The spread of the analyzed capped options varies between 0.2% and 12.5%. Thus, the prices predicted by the (2, 2)-model match the observed mid-prices very closely, apart from a few short-term out-of-the-money contracts.

Nevertheless, the capped option valuation reveals a potential problem of the quadrature based numerical valuation procedure. The payoff function in Equation (50) clips a narrow interval out of the entire pricing density, which possibly contains only a small number of quadrature points. Therefore, numerical results may be inaccurate. There are two possible ways to improve the situation. First, one could simply increase the number of quadrature points involved in the numerical integration procedure. This idea breeds two new problems: On the one hand, only a fraction of the additional points is located in the relevant interval of the payoff function. On the other hand, there are a large

number of quadrature points, with associated weights very close to zero, which means that the effect of a considerable amount of computed quadrature points on the valuation result is negligible. The latter problem at least can be resolved by pruning (cf. Jaeckel 2005).

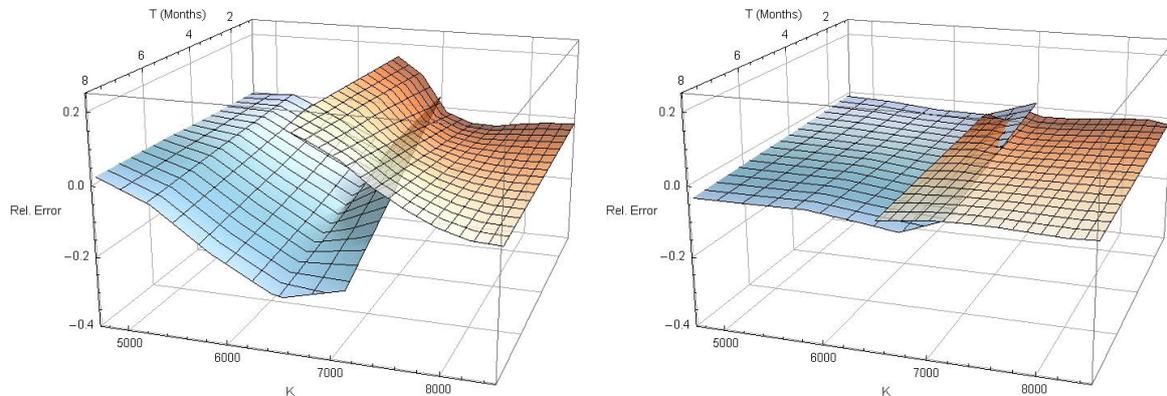


Figure 5. Valuation of European vanilla capped calls (blue) and puts (red) with the *Black–Scholes* model (left) and (2,2)-model of conditional pricing density (right).

Another alternative is Monte Carlo simulation. This is not a trivial task, because one is not able to draw from the arbitrage-free pricing distribution directly. Nevertheless, two indirect sampling methods are detailed in the next paragraph.

6.2. Monte Carlo Valuation Methods

A key requirement for Monte Carlo simulation is the ability to draw random numbers from the relevant probability distribution. Remember that the conditional pricing distribution for any time to maturity is given by its density function (Equation (30)) in (z, τ) -coordinates. This can be written as

$$\begin{aligned}
 q_Z(z, \tau) &\approx \sqrt[4]{4\pi} \int_{-\infty}^{\infty} \phi(y) \sum_{n=0}^N \frac{c_n(\tau)}{\sqrt{2^n n!}} H_n(y) \delta(y - z) dy \\
 &\approx \sqrt[4]{4\pi} \sum_{j=1}^J \sum_{n=0}^N w_j \frac{c_n(\tau)}{\sqrt{2^n n!}} H_n(y) \delta(y - z^{(j)}),
 \end{aligned}
 \tag{51}$$

with w_j and $z^{(j)}$ again indicating *Gauss–Hermite*-quadrature weights and points, respectively. However, the second line of (51) is just an abusive way of writing a multinomial distribution function with values $z^{(j)}$, occurring with probability

$$q_j = \sqrt[4]{4\pi} w_j \sum_{n=0}^N \frac{c_n(\tau)}{\sqrt{2^n n!}} H_n(z^{(j)}),
 \tag{52}$$

for $j = 1, \dots, J$. It is easy to draw from this multinomial distribution.

Figure 6 (left) shows the pricing density, generated with the (2,2)-model (red), and the distribution of one million draws from the multinomial approximation as histogram (gray). A total of $J = 1000$ quadrature points were used and again $N = 45$ *Fourier*-terms were included. Both densities coincide perfectly. Unfortunately, the multinomial approximation method does not resolve the problem discussed in the previous paragraph.

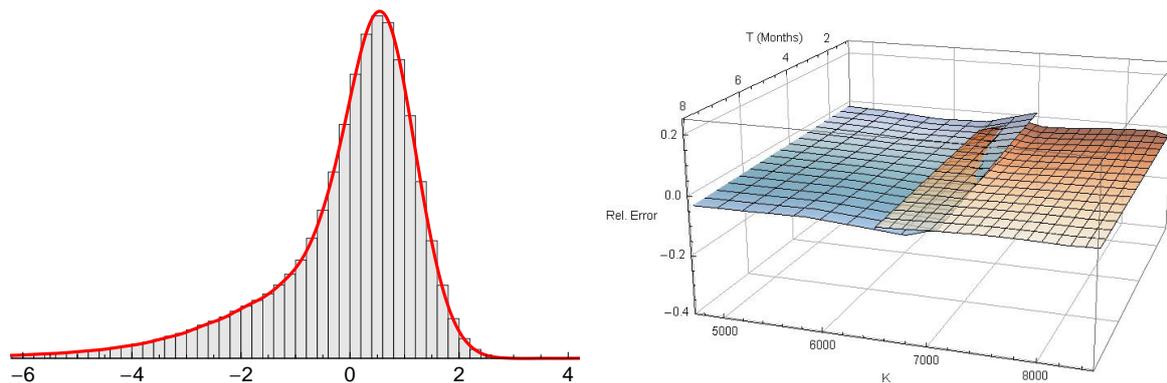


Figure 6. One million draws from the multinomial density approximation for $T = 1$ (left)—capped option valuation with 100,000 draws from $N(0, 2)$ importance distribution (right).

An alternative approach is based on the idea of choosing a suitable importance density that covers the z -support of $q_Z(z, \tau)$ for a desired value of τ , and writes valuation Equation (32) as

$$V(S_0, 0) = B_0(0; \tau^2) \int_{-\infty}^{\infty} \tilde{V}(z, \tau) \frac{q_Z(z, \tau)}{f_Z(z, \tau)} f_Z(z, \tau) dz, \tag{53}$$

with $\tilde{V}(z, \tau) = V(F_0 e^{z\sigma\tau - \frac{1}{2}\sigma^2\tau^2}, \tau^2)$ and the importance density $f_Z(z, \tau)$. Now, an arbitrary sample of J realizations may be drawn from the importance distribution $F_Z(z, \tau)$. An unbiased estimator of (53) is then given by

$$V(S_0, 0) \approx B_0(0; \tau^2) \frac{1}{J} \sum_{j=1}^J \tilde{V}(z^{(j)}, \tau) \frac{q_Z(z^{(j)}, \tau)}{f_Z(z^{(j)}, \tau)}, \tag{54}$$

where the last term on the right hand side of Equation (54) is called the importance weight or likelihood ratio. It is even possible to reduce the variance of this estimator below the initial variance, induced by drawing from the target distribution for a comprehensive treatment of this subject see (Glasserman 2010, sct. 4.6). If the pricing density $q_Z(z, 1)$ is estimated itself with the normal importance distribution $N(0, 2)$, and J and N set as in the previous example, the result is indistinguishable from Figure 6 (left).

The valuation procedure for the whole validation sample was repeated with Monte Carlo simulation based on the $N(0, 2)$ -importance distribution. A total of $J = 100,000$ points were drawn for each contract. The resulting relative pricing errors, with all environmental conditions unchanged, are shown in Figure 6 (right). This is indeed very close to Figure 5 (right), but not identical. The root-mean-square error is slightly reduced, at 3.14%.

7. Conclusions

A new method for estimating the time evolution of the arbitrage-free pricing density, conditioned on the observable market information, was suggested. The key idea of the approach is to model the excess dynamics beyond the classical *Black–Scholes* dynamics. To this end, a coordinate transformation was introduced, under which the pricing density looks stationary. In this ‘laboratory frame’, the excess dynamics are extracted by an asymptotic series expansion, resulting in a *Kolmogorov*-backward-equation with $O(\varepsilon)$ drift and $O(\varepsilon^2)$ diffusion terms. This equation is approximately solved by making a time separable ansatz and using a complete set of orthogonal *Hermite*-polynomials.

The resulting model frame was calibrated to market data of the ‘Deutscher Aktienindex’ (DAX) index and one particular model was singled out and benchmarked against other approaches. It was shown that the pricing error was reduced to the order of the bid-offer spread and that the implied

volatility surface, generated by the new method, is closer to the observed one than those generated by other popular approaches. Finally, a validation sample of 171 capped options was valued. The pricing error was again reduced dramatically, emphasizing the quality of the model fit.

The suggested approach has a number of appealing properties, but also some drawbacks and limitations, which should be summarized to present a balanced view:

- Access to the time evolution of the arbitrage free pricing density is very convenient, because any vanilla contract can be priced immediately and consistently. There is no need for semi-parametric or non-parametric interpolation, or extrapolation of a volatility surface. Furthermore, one is able to draw random samples directly from the correct conditional pricing density for any given time to maturity.
- The estimated pricing density is always conditioned on the present market information \mathcal{F}_0 . One has no access to transition probabilities, because no (pathwise) model, in terms of a stochastic process, is formulated. This is a major drawback, because valuation of path dependent options with Monte Carlo simulation methods is not possible directly. However, those contracts can be valued indirectly by extracting the *Black–Scholes* implied volatility surface and computing local volatilities to be used in a simulation of the corresponding geometrical *Brownian* motion.
- Using a complete set of orthogonal *Hermite*-polynomials is a convenient way of translating the differential operators in the *Kolmogorov*-backward-equation into infinite dimensional matrices. One can confidently expect that a finite number of *Fourier*-terms is sufficient to approximate the density function to the desired level of accuracy. This means there is a finite dimensional, and thus computable, approximation to the problem. Unfortunately, numerical issues impose a limit on the manageable deviation from the normal density. This limit is reached and exceeded in some models listed in Table 1. The only possible remedy is the use of a better-suited complete orthogonal system.
- The assumptions regarding time separability and the functional form of time dependence are somewhat artificial. The functional form is chosen to reproduce some known solutions as special cases and to ensure tractability of the model. Even though the implications of these assumptions are by no means implausible, there is a margin for improving the model fit by imposing a richer time structure. This may possibly also resolve the problem of the short-term implied volatility fit, which is not satisfactory as observed in Figure 3.

Considering all advantages and drawbacks, the suggested method is very promising and well-suited for option pricing, even in difficult markets with exceptional conditions. Furthermore, calibration to market data is easy, because the gradient of the quadratic objective function to be minimized is available analytically. The results obtained are conclusive and the approach was able to produce a better implied volatility fit than conventional models.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Stationary Coordinate Frame Transformation

Departing from the original PDE

$$\frac{\partial}{\partial t} q_X = \frac{1}{2} \sigma^2 \left(\frac{\partial}{\partial x} + \frac{\partial^2}{\partial x^2} \right) q_X \quad (\text{A1})$$

in terms of x and t , where the arguments were suppressed for notational simplicity, the transformations

$$\begin{aligned} z &= \frac{x + \frac{1}{2}\sigma^2 t}{\sigma\sqrt{t}}, \\ \tau &= \sqrt{t} \end{aligned} \quad (\text{A2})$$

were suggested with x_0 and t_0 set to zero, because they merely shift the starting point in the spatial and time directions. The differentials change under this coordinate transformation as

$$\begin{aligned} \frac{\partial q_X}{\partial t} &= \frac{\partial q_X}{\partial \tau} \frac{d\tau}{dt} + \frac{\partial q_X}{\partial z} \frac{\partial z}{\partial t} = \frac{\partial q_X}{\partial \tau} \frac{1}{2\tau} + \frac{\partial q_X}{\partial z} \left(\frac{\sigma}{2\tau} - \frac{z}{2\tau^2} \right) \\ \frac{\partial q_X}{\partial x} &= \frac{\partial q_X}{\partial z} \frac{\partial z}{\partial x} = \frac{\partial q_X}{\partial z} \frac{1}{\sigma\tau} \\ \frac{\partial^2 q_X}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{\partial q_X}{\partial z} \frac{1}{\sigma\tau} \right) = \frac{\partial^2 q_X}{\partial z^2} \frac{1}{\sigma^2\tau^2}. \end{aligned} \tag{A3}$$

Thus, Equation (A1) now becomes

$$\frac{\partial q_X}{\partial \tau} \frac{1}{2\tau} - \frac{\partial q_X}{\partial z} \frac{z}{2\tau^2} = \frac{\partial^2 q_X}{\partial z^2} \frac{1}{2\tau^2}, \tag{A4}$$

or expressed in a more familiar way

$$\frac{\partial q_X}{\partial \tau} = \frac{\partial q_X}{\partial z} \frac{z}{\tau} + \frac{\partial^2 q_X}{\partial z^2} \frac{1}{\tau}. \tag{A5}$$

Using now the identity $q_X dx = q_Z dz$ yields $q_X = q_Z \frac{dz}{dx} = q_Z \frac{1}{\sigma\tau}$, and therefore the derivative of q_X with respect to τ becomes

$$\frac{\partial q_X}{\partial \tau} = \frac{\partial q_Z}{\partial \tau} \frac{1}{\sigma\tau} - q_Z \frac{1}{\sigma\tau^2}. \tag{A6}$$

The derivatives of q_X with respect to z remain intact, which means that they are only multiplied by a factor of $(\sigma\tau)^{-1}$. Again collecting terms, one obtains

$$\frac{\partial q_Z}{\partial \tau} = \frac{1}{\tau} \left(\frac{\partial q_Z}{\partial z} z + \frac{\partial^2 q_Z}{\partial z^2} + q_Z \right). \tag{A7}$$

Because under the coordinate change in Equation(A2), the density q_Z becomes the standard normal density $\phi(z)$, and one has $\frac{\partial q_Z}{\partial z} = -zq_Z$ and $\frac{\partial^2 q_Z}{\partial z^2} = -q_Z + z^2q_Z$. Thus, Equation (A7) yields

$$\frac{\partial q_Z}{\partial \tau} = \frac{1}{\tau} \left(-z^2q_Z - q_Z + z^2q_Z + q_Z \right) = 0, \tag{A8}$$

which proves the stationarity of the new coordinate frame.

Appendix B. Proof of Proposition 3

The derivative of $c(\tau)$ with respect to γ is an immediate consequence of the *Hadamards* lemma. By this lemma, the following relation holds for a smooth matrix function $X(\gamma)$

$$\begin{aligned} \left(\frac{d}{d\gamma} e^{X(\gamma)} \right) e^{-X(\gamma)} &= \frac{d}{d\gamma} X(\gamma) + \frac{1}{2!} \left[X(\gamma), \frac{d}{d\gamma} X(\gamma) \right] \\ &+ \frac{1}{3!} \left[X(\gamma), \left[X(\gamma), \frac{d}{d\gamma} X(\gamma) \right] \right] + \dots, \end{aligned} \tag{A9}$$

with the commutator $[X, Y] = XY - YX$ of two arbitrary square matrices X and Y of the same dimension. In the linear problem (Equation (29)), $X(\gamma)$ has the particular form

$$\begin{aligned} X(\gamma) &= f(\gamma)X, \quad \text{with } f(\gamma) = \frac{1 - e^{-\gamma\tau}}{\gamma} \\ \text{and } X &= \sum_{k=0}^{K_a} a_k A^{(k)} + \sum_{k=0}^{K_b} b_k B^{(k)}, \end{aligned} \tag{A10}$$

and thus, $(d/d\gamma)X(\gamma) = (df(\gamma)/d\gamma)X$. The scalar derivative can be pulled out of the commutators in (A9), and hence all of them vanish because a square matrix always commutes with itself. One finally obtains

$$\frac{d}{d\gamma} \exp [f(\gamma)X] = \frac{df(\gamma)}{d\gamma} X \exp [f(\gamma)X], \quad (\text{A11})$$

from which the first part of Proposition 3 follows immediately.

For the second part, write the differential equation system in Equation (18) using X as defined in Equation (A10)

$$\frac{d}{d\tau} c(\tau) = e^{-\gamma\tau} X c(\tau). \quad (\text{A12})$$

Now, following an idea of Fung (2004), differentiate both sides of Equation (A12) with respect to a_k

$$\frac{d}{d\tau} \left(\frac{\partial}{\partial a_k} c(\tau) \right) = e^{-\gamma\tau} A^{(k)} c(\tau) + e^{-\gamma\tau} X \frac{\partial}{\partial a_k} c(\tau). \quad (\text{A13})$$

By defining the extended Fourier-coefficient vector $\tilde{c}(\tau) = [(\partial/\partial a_k)c(\tau), c(\tau)]'$, one again obtains a system of linear differential equations

$$\frac{d}{d\tau} \tilde{c}(\tau) = e^{-\gamma\tau} \tilde{X} \tilde{c}(\tau), \quad \text{with} \quad \tilde{X} = \begin{bmatrix} X & A^{(k)} \\ 0 & X \end{bmatrix}. \quad (\text{A14})$$

This system obviously has the solution $\tilde{c}(\tau) = \exp[f(\tau)\tilde{X}]\tilde{c}(0)$, with $\tilde{c}(0) = [0, c(0)]'$. The second part of Proposition 3 follows immediately by extracting the first part of the extended coefficient vector.

Derivatives with respect to b_k are computed analogously by replacing $A^{(k)}$ with $B^{(k)}$ in Equation (A14).

References

- Abramowitz, Milton, and Irene A. Stegun. 1970. *Handbook of Mathematical Functions*. New York: Dover Publications.
- Ait-Sahalia, Yacine, and Jefferson Duarte. 2003. Nonparametric Option Pricing under Shape Restrictions. *Journal of Econometrics* 116: 9–47.
- Ait-Sahalia, Yacine, and Andrew W. Lo. 1998. Nonparametric Estimation of State-Price Densities Implicite in Financial Asset Prices. *Journal of Finance* 53: 499–547.
- Ait-Sahalia, Yacine. 2002. Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach. *Econometrica* 70: 223–62.
- Basu, Arnab, and Mrinal K. Ghosh. 2009. Asymptotic Analysis of Option Pricing in a Markov Modulated Market. *Operations Research Letters* 37: 415–19.
- Bates, David S. 1996. Jumps and Stochastic Volatility: The Exchange Rate Processes Implicit in Deutschemark Options. *Review of Financial Studies* 9: 69–107.
- Berestycki, H., J. Busca, and I. Florent. 2002. Asymptotics and Calibration of Local Volatility Models. *Quantitative Finance* 2: 61–69.
- Black, Fischer, and Myron Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81: 637–54.
- Black, Fischer. 1976. The Pricing of Commodity Contracts. *Journal of Financial Economics* 3: 167–79.
- Blinnikov, Sergei, and Richhild Moessner. 1998. Expansions for nearly Gaussian Distributions. *Astronomy & Astrophysics Supplement Series* 130: 193–205.
- Bondarenko, Oleg. 2003. Estimation of Risk-Neutral Densities Using Positive Convolution Approximation. *Journal of Econometrics* 116: 85–112.
- Breeden, Douglas T., and Robert H. Litzenberger. 1978. Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business* 51: 621–51.
- Broyden, Charles George. 1970. The Convergence of a Class of Double-Rank Minimization Algorithms. *Journal of the Institute of Mathematics and Its Applications* 6: 76–90.

- Carr, Peter, and Dilip B. Madan. 2005. A Note on Sufficient Conditions for No Arbitrage. *Finance Research Letters* 2: 125–30.
- Dennis, John E., and Robert B. Schnabel. 1983. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Upper Saddle River: Prentice-Hall.
- Derman, Emanuel, and Nassim Nicholas Taleb. 2005. The Illusions of Dynamic Replication. *Quantitative Finance* 5: 323–26.
- Dupire, Bruno. 1994. Pricing with a Smile. *Risk* 7: 18–20.
- Fengler, Matthias R. 2009. Arbitrage-Free Smoothing of the Implied Volatility Surface. *Quantitative Finance* 9: 417–28.
- Figlewski, Stephen. 2010. Estimating the Implied Risk-Neutral Density for the US Market Portfolio. In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*. Edited by Tim Bollerslev, Jeffrey R. Russell and Mark W. Watson. Oxford: Oxford University Press, chp. 15, pp. 323–53.
- Filipović, Damir, Lane P. Hughston, and Andrea Macrina. 2012. Conditional Density Models for Asset Pricing. *International Journal of Theoretical and Applied Finance* 15: 1250002-1–24.
- Fletcher, Roger. 1970. A New Approach to Variable Metric Algorithms. *Computer Journal* 13: 317–22.
- Fung, T. C. 2004. Computation of the Matrix Exponential and its Derivatives by Scaling and Squaring. *International Journal of Numerical Methods in Engineering* 59: 1273–86.
- Gatheral, Jim, and Tai-Ho Wang. 2012. The Heat-Kernel Most-Likely-Path Approximation. *International Journal of Theoretical and Applied Finance* 15: 1250001-1–18.
- Gatheral, Jim, Elton P. Hsu, Peter Laurence, Cheng Ouyang, and Tai-Ho Wang. 2012. Asymptotics of Implied Volatility in Local Volatility Models. *Mathematical Finance* 22: 591–620.
- Gatheral, Jim. 2004. A Parsimonious Arbitrage-Free Implied Volatility Parameterization with Application to the Valuation of Volatility Derivatives. Paper presented at Talk at the Global Derivatives & Risk Management Conference, Madrid, Spain, May 26.
- Gatheral, Jim. 2006. *The Volatility Surface—A Practitioner’s Guide*. Hoboken: John Wiley & Sons.
- Glasserman, Paul. 2010. *Monte Carlo Methods in Financial Engineering*. Berlin/Heidelberg and New York: Springer.
- Goldfarb, Donald. 1970. A Family of Variable Metric Updates Derived by Variational Means. *Mathematics of Computation* 24: 23–26.
- Golub, Gene H. 1973. Some Modified Matrix Eigenvalue Problems. *SIAM Review* 15: 318–34.
- Guyon, Julien, and Pierre Henry-Labordere. 2011. From Spot Volatilities to Implied Volatilities. *Asia-Risk*, 59–64. doi:10.2139/ssrn.1663878.
- Habtemicael, Semere, and Indranil SenGupta. 2016a. Pricing Variance and Volatility Swaps for Barndorff-Nielsen and Shephard Process Driven Financial Markets. *International Journal of Financial Engineering* 3: 1650027.
- Habtemicael, Semere, and Indranil Sengupta. 2016b. Pricing Coariance Swaps for Barndorff-Nielsen and Shephard Process Driven Financial Markets. *Annals of Financial Economics* 11: 1650012.
- Hagan, Patrick S., Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward. 2002. Managing Smile Risk. *Wilmott Magazine*, September, pp. 84–108.
- Heston, Steven L. 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options. *The Review of Financial Studies* 6: 327–43.
- Hlavka, Zdenek, and Marek Svojik. 2009. Application of Extended Kalman Filter to SPD Estimation. In *Applied Quantitative Finance*, 2nd ed. Edited by Wolfgang Haerdle, Nikolaus Hautsch and Ludger Overbeck. Berlin, Heidelberg and New York: Springer, pp. 233–47.
- Huynh, Kim, Pierre Kervella, and Jun Zheng. 2002. Estimating State-Price Densities with Nonparametric Regression. In *Applied Quantitative Finance*. Edited by Wolfgang Haerdle, Torsten Kleinow and Gerhard Stahl. Berlin, Heidelberg and New York: Springer, pp. 171–96.
- Jaeckel, Peter. 2005. A Note on Multivariate Gauss-Hermite Quadrature. Paper Published on the World Wide Web. Available online: <http://www.jaeckel.org> (accessed on 8 March 2018).
- Kim, Yong-Jin. 2002. An Asymptotic Valuation for the Option under a General Stochastic Volatility. *Journal of the Operations Research Society of Japan* 45: 404–25.
- Kou, Steven G. 2002. A Jump-Diffusion Model for Option Pricing. *Management Science* 48: 1086–101.
- Lewis, Alan. 2002. Fear of Jumps. *Wilmott Magazine*, December, pp. 60–67.
- Lighthill, Michael J. 1980. *Introduction to Fourier Analysis and Generalised Functions*. Cambridge, London and New York: Cambridge University Press.

- Mazzoni, Thomas. 2010. Fast Analytic Option Valuation with GARCH. *Journal of Derivatives* 18: 18–38.
- Mazzoni, Thomas. 2015. A GARCH Parametrization of the Volatility Surface. *Journal of Derivatives* 23: 9–24.
- Medvedev, Alexey, and Olivier Scaillet. 2003. *A Simple Calibration Procedure of Stochastic Volatility Models with Jumps by Short Term Asymptotics*. Technical Report 93. International Center for Financial Asset Management and Engineering. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=477441 (accessed on 3 March 2018)
- Merton, Robert C. 1976. Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* 3: 125–44.
- Moler, Cleve, and Charles van Loan. 2003. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. *SIAM Review* 45: 1–46.
- Reghai, A. 2006. The Hybrid Most Likely Path. *Risk* 19: 34–35.
- Risken, Hannes. 1989. *The Fokker-Planck Equation. Methods of Solution and Applications*, 2nd ed. Berlin, Heidelberg and New York: Springer.
- Shanno, David F. 1970. Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation* 24: 647–56.
- Uchida, Masayuki, and Nakahiro Yoshida. 2004. Asymptotic Expansion for Small Diffusions Applied to Option Pricing. *Statistical Inference for Stochastic Processes* 7: 189–223.
- Van der Kamp, Roel. 2009. Local Volatility Modelling. Master's thesis, University of Twente, Enschede, The Netherlands.
- Whalley, A. Elizabeth, and Paul Wilmott. 1997. An Asymptotic Analysis of an Optimal Hedging Model for Option Pricing with Transaction Costs. *Mathematical Finance* 7: 307–24.
- Xiu, Dacheng. 2014. Hermite Polynomial Based Expansion of European Option Prices. *Journal of Econometrics* 179: 158–77.
- Yatchew, Adonis, and Wolfgang Haerdle. 2006. Nonparametric State-Price Density Estimation Using Constrained Least Squares and the Bootstrap. *Journal of Econometrics* 133: 579–99.



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).