XAIA INVESTMENT

Efficient simulation from the risk-neutral density

Jan-Frederik Mai
XAIA Investment GmbH
Sonnenstraße 19, 80331 München, Germany
jan-frederik.mai@xaia.com

Date: May 18, 2016

Abstract The Buchen-Kelly density maximizes entropy among all risk-neutral densities that explain observed option prices for a fixed maturity. Since it is piecewise exponential, the inverse function of the associated distribution function is given in closed form. This makes Monte Carlo simulations from the Buchen-Kelly density extremely efficient via the classical inversion method. Unfortunately, however, the Buchen-Kelly density depends critically on the input option data and tends to be spiky, which is often not desired. We demonstrate how a deliberate choice of the input option data can resolve this issue and yield a smooth Buchen-Kelly density.

1 Introduction

The motivation for the present note stems from the following application. We seek to find a risk-neutral probability distribution for a bivariate random vector (S_T, R_T) describing the value of two underlyings S and R at a fixed future time point $T > 0$. While the risk-neutral distributions of both underylings S_T and R_T can be retrieved from vanilla option data, there is a priori no information available about the dependence structure between S and R. However, in the applications we have in mind this dependence is very strong, e.g. S may be the ITRXX Main spread and R the ITRXX Sen Financial spread. A reasonable dependence model in such a case is hence a copula that results from a small disturbance (in a justifiable sense) of co-monotonicity[1]. When having constructed such a bivariate model, we seek to carry out (extensive) simulation studies in order to estimate the risk of potential positions involving both underlyings. In such a simulation study a sample of (S_T, R_T) is generated by first simulating a sample (U, V) from the underlying copula, and then transforming the uniformly distributed random variables U, V to the desired quantities S_T, R_T via the classical inversion method, i.e. $S_T = F_S^{-1}(U)$ and $R_T = F_R^{-1}(V)$ with F_S, F_R denoting the univariate distribution functions of S_T and R_T, respectively.

To this end, it is convenient if the univariate distributions of the underlyings are given in such a way that the inverses of the distribution functions (the inverse cdfs) of S_T and R_T can be evaluated efficiently. Even though there are numerous methods to infer univariate risk-neutral distributions from option data, unfortunately not many of these approaches provide the inverse cdfs in the required simple form. For some stochastic price evolution models, a simulation of the terminal value requires simulation of a whole path, which is too computation-costly and cannot

[1] We elaborate more on this technique in a future XAIA article.

be combined with the aforementioned copula-method. For other models the distribution function of the terminal value is known in closed form, but the inversion needs to be done numerically vie Newton's or bisection methods, which still results in a costly simulation engine.

The so-called Buchen-Kelly density of Buchen, Kelly (1996); Neri, Schneider (2012, 2013) is a promising candidate for the described task. Besides its nice interpretation in terms of entropy maximization, it is piecewise exponential, hence its inverse cdf is given in closed and simple form, so that millions of simulations can be generated within fractions of a second, as desired. Unfortunately, the Buchen-Kelly density is typically not very smooth. Even though it is continuous by definition, in practice these densities are often quite spiky. In particular, the algorithm described in Neri, Schneider (2013) gets as input arbitrage-free call (mid) prices and spits out the unique Buchen-Kelly density associated with these. Small changes in the input call prices can result in dramatic shape changes of the resulting Buchen-Kelly density. The present note provides practical advice on how to apply this algorithm in order to obtain smooth densities. The key technique here is to choose the input call (mid) prices deliberately. If these prices are computed from some model with a smooth density f, the Buchen-Kelly algorithm is basically a tool which approximates f by a continuous, piecewise exponential function under the side constraint that option prices are matched perfectly. This approximation can further be improved by enhancing the input data by additional option prices outside the observed moneyness range, that are computed from f.

The remaining article is organized as follows. Section 2 introduces notations and definitions. Section 3 describes how the input data for the Buchen-Kelly algorithm should be chosen in order to obtain a smooth density.

2 Notation

We consider European call and put options on a stock $S = \{S_t\}_{t\geq 0}$ with fixed maturity $T > 0$. Their respective market prices in dependence on the exercise strike price are denoted by $C(K)$, resp. $P(K)$, in the sequel. We furthermore denote by $DF(t)$ today's price of a risk-free zero coupon bond with maturity $t \geq 0$. The fair strike price of an equity forward on the stock S with maturity T, entered into at $t \in [0, T]$, is denoted by $F(t, T)$. The equity forward strike price equals the market's expectation about the stock price at maturity, i.e. $F(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T \,|\, \mathcal{F}_t]$, which implies that the stochastic process $\{F(t, T)\}_{t\in[0,T]}$ is a $\mathbb{Q}$-martingale, cf. Bernhart, Mai (2015). Troughout, the quantities $F(t, T)$ and S_t are assumed to be related via

$$S_t = F(t, T)\, e^{\delta\,(T-t)}\, \frac{DF(T)}{DF(t)}, \tag{1}$$

with the parameter $\delta \geq 0$ accounting for potential differences between $F(0, T)$ and $S_0/DF(T)$, which are both observable quantities. It can be interpreted either as a continuous repo margin or a continuous dividend yield – i.e. a continuous spread on top of the risk-free rate that is earned by stock owners[2]. The option market prices in general can be represented as

$$C(K) = DF(T)\,\mathbb{E}^{\mathbb{Q}}[(S_T - K)_+], \quad P(K) = DF(T)\,\mathbb{E}^{\mathbb{Q}}[(K - S_T)_+].$$

In the classical Black–Scholes model the martingale $\{F(t,T)\}_{t\geq 0}$ is a geometric Brownian motion with volatility parameter $\sigma > 0$, i.e.

$$F(t,T) = F(0,T)\,e^{-\frac{\sigma^2}{2}\,t+\sigma\,W_t} = \frac{S_0\,e^{-\delta\,T}}{DF(T)}\,e^{-\frac{\sigma^2}{2}\,t+\sigma\,W_t}, \quad t \in [0,T],$$

with a standard Brownian motion $W = \{W_t\}_{t\geq 0}$. Introducing the notation of *log-moneyness* $k := \log(K/F(0,T))$, the option prices under the Black–Scholes model assumption are given by

$$\begin{aligned} C_{BS}\big(\sigma,k,S_0,\delta\big) &:= e^{-\delta\,T}\,S_0\Big(\Phi\big(d_+(\sigma,k)\big) - e^k\,\Phi\big(d_-(\sigma,k)\big)\Big), \\ P_{BS}\big(\sigma,k,S_0,\delta\big) &:= e^{-\delta\,T}\,S_0\Big(e^k\,\Phi\big(-d_-(\sigma,k)\big) - \Phi\big(-d_+(\sigma,k)\big)\Big), \\ d_\pm(\sigma,k) &:= -\frac{k}{\sigma\,\sqrt{T}} \pm \frac{\sigma\,\sqrt{T}}{2}, \end{aligned}$$

with Φ denoting the cdf of a standard normally distributed random variable. For the sake of notational convenience we follow the seminal reference Lee (2004) in expressing the Black–Scholes formula in terms of k instead of K. Due to the assumed relationship (1), any three of the four variables $S_0, F(0,T), \delta, DF(T)$ implies the fourth. Hence, in general the option formula needs to depend on three of the four variables. The switch from K to k eliminates one variable (namely $F(0,T)$) so that the remaining formulas only depend explicitly on S_0 and δ. However, implicitly the dependence on $F(0,T)$ is still there, it is only hidden in k and (1).
For given market put and call prices $K \mapsto P(K)$ and $K \mapsto C(K)$, we choose $F(0,T)$ as the unique root of the function $K \mapsto C(K) - P(K)$ and define the associated *implied volatility smile* $K \mapsto \sigma_{BS}(K)$ implicitly as the unique root of the equation

$$\begin{aligned} C(K) &\stackrel{!}{=} C_{BS}\Big(\sigma_{BS}(K), \log\Big(\frac{K}{F(0,T)}\Big), S_0, \delta\Big), \text{ if } K > F(0,T), \\ P(K) &\stackrel{!}{=} P_{BS}\Big(\sigma_{BS}(K), \log\Big(\frac{K}{F(0,T)}\Big), S_0\Big), \text{ if } K \leq F(0,T). \end{aligned}$$

Section A in the Appendix explains this definition.

3 Buchen-Kelly density

For a given battery of call option prices $C(K_1), \ldots, C(K_n)$ and a forward value $F(0,T)$, (Neri, Schneider, 2013, Corollary 5.3) states that there exists a unique function g with the following properties:

(a) $DF(T)\,\int_{K_i}^{\infty} g(x)\,(x - K_i)\,\mathrm{d}x = C(K_i)$ for $i = 1, \ldots, n$.

(b) g is the density of a probability distribution with mean $F(0,T)$.

[2] The classical Black–Scholes model relies on the simplifying assumption $\delta = 0$, which is too restrictive in practice.

(c) g is piecewise exponential, i.e. on each interval[3] $[K_{i-1}, K_i]$ has the form $g(x) = \alpha_i \, e^{\beta_i \, x}$ for certain $\alpha_i > 0$ and $\beta_i \in \mathbb{R}$, $i = 1, \ldots, n+1$.

(d) g is continuous.

The function g is called Buchen-Kelly density, named after Buchen, Kelly (1996), and is known to maximize entropy among all functions satisfying (a) and (b). It can be computed along the iterative algorithm described in Neri, Schneider (2012, 2013).
One weakness of the Buchen-Kelly density is that it depends critically on the input prices $C(K_1), \ldots, C(K_n)$. Small changes in these prices may result in dramatic shape changes of the Buchen-Kelly density, and very unsmooth densities are commonly obtained. For instance, the top plot in Figure 1 visualizes a Buchen-Kelly density for the DAX (German stock index). The green lines in the background indicate the strike levels $K_1, \ldots, K_n$ for which option prices are given. The input prices $C(K_1), \ldots, C(K_n)$ for the algorithm are the quoted mid prices. One observes that the density is continuous, so it is the unique Buchen-Kelly density, i.e. the spiky shape is not the result of bad numerical behavior of the algorithm but of the poor choice of mid prices. The middle plot in Figure 1 visualizes another Buchen-Kelly density. The sole difference to the top plot is that the input mid prices $C(K_1), \ldots, C(K_n)$ have been chosen differently according to the method described below in paragraph 3.1. The density is obviously much smoother than before, but still has two spikes near the boundary strike prices K_1 and K_n. This issue is resolved in the bottom plot of Figure 1. Here, five additional call prices with strikes outside the observed moneyness range have been included. These additional strike prices are indicated by the red lines in the background. The inclusion of additional call prices deviates from the concept of entropy maximization because model assumptions are necessary for their computation. It is a trade-off for the sake of smoothness.

3.1 How to choose the input prices?

If the input call prices $C(K_1), \ldots, C(K_n)$ are computed from some density f, then the resulting Buchen-Kelly density g is intuitively expected to approximate f. In other words, smoothness of f is expected to carry over to smoothness of g. Consequently, in a first step we fit a smooth density f to the observed option data. To this end, we use the three-parametric SVI model mentioned at the end of Section B in the Appendix, which provides an excellent fit in the illustrated example, see Figure 2. The obtained SVI-density f is also illustrated in the middle and bottom plots of Figure 1; it is computed from the formula in paragraph B.2. Apparently, the difference between the SVI-density and the Buchen-Kelly density is minimal in the bottom plot of Figure 1. The Buchen-Kelly density, however, has the advantage that its associated inverse cdf is given in closed form. This enables very efficient Monte Carlo simulation. In this example, the exact simulation of one million independent samples from the Buchen-Kelly density only took 0.18 seconds on a standard PC in MATLAB via the classical inversion method. This is a striking improvement

[3] Here, $K_0 := 0$ and $K_{n+1} := \infty$.

compared to the effort it would take to invert the SVI-distribution function (2) numerically.

4 Conclusion

It was outlined how the Buchen-Kelly algorithm of Neri, Schneider (2012, 2013) may be enhanced by a deliberate choice of input (mid) call prices in order to retrieve a smooth Buchen-Kelly density. Furthermore, it was highlighted that the piecewise exponential form of the Buchen-Kelly density, implying exact and efficient evaluation of the associated inverse cdf, is a desirable feature in applications involving extensive Monte Carlo simulations from the risk-neutral probability distribution.

A The choice of $F(0,T)$

Market (mid) quotes for puts $K \mapsto P(K)$ and calls $K \mapsto C(K)$ are arbitrage-free if and only if there exists a non-negative random variable S_T such that $P(K) = DF(T)\,\mathbb{E}^{\mathbb{Q}}[(K - S_T)_+]$ and $C(K) = DF(T)\,\mathbb{E}^{\mathbb{Q}}[(S_T - K)_+]$ for all K. The probability law of S_T under $\mathbb{Q}$ is uniquely determined by either the function $K \mapsto C(K)$ or by the function $K \mapsto P(K)$. In particular, the mean

$$F(0,T) := \mathbb{E}^{\mathbb{Q}}[S_T] = \frac{C(0)}{DF(T)} = \frac{\lim\limits_{K\to\infty} \{P^{'}(K)\,K - P(K)\}}{DF(T)},$$

which is the forward equity strike, is uniquely defined. In practice, however, a typical situation is that the prices $K \mapsto C(K)$ are arbitrage-free, and also the prices $K \mapsto P(K)$ are arbitrage-free, but not put and call prices jointly. This is most obviously seen by the following check: $(S_T - K)_+ + K = (K - S_T)_+ + S_T$ implies the put-call parity $C(K) - P(K) + DF(T)\,K = DF(T)\,F(0,T)$. The left-hand side of the last equation in theory must be invariant with respect to K (because the right-hand side is), but in practice is not. In the sequel, we are going to show how this problem is resolved in practice.

In case that $C(K) - P(K) + DF(T)\,K$ is non-constant, it is impossible to find a single random variable S_T explaining both call and put prices jointly. Consequently, we must simplify our target by deciding which market prices we want to explain and which we do not need to match perfectly with our model. Out-of-the-money options are traded much more liquidly than in-the-money options. Consequently, it is (a) more important to replicate those appropriately, and (b) the information from out-of-the-money options is more reliable. As a consequence, we simplify our task by seeking a choice of $F(0,T) = \mathbb{E}^{\mathbb{Q}}[S_T]$ such that our resulting model for S_T perfectly explains the function $K \mapsto C(K)$ only on $[F(0,T),\infty)$ (the *call wing*) and $K \mapsto P(K)$ only on $[0, F(0,T)]$ (the *put wing*). A solution $F(0,T)$ to this problem must necessarily be chosen as the unique (and existing) root of the function $K \mapsto C(K) - P(K)$, as the following lemma shows.

Lemma A.1 (Necessary condition on $F(0,T)$)
Let $F(0,T)$ be a number in $(0,\infty)$. There exists a random variable S_T with mean $\mathbb{E}^{\mathbb{Q}}[S_T] = F(0,T)$ satisfying

$$\begin{aligned} C(K) &= DF(T)\,\mathbb{E}^{\mathbb{Q}}[(S_T - K)_+], \quad K \in [F(0,T),\infty), \\ P(K) &= DF(T)\,\mathbb{E}^{\mathbb{Q}}[(K - S_T)_+], \quad K \in [0, F(0,T)], \end{aligned}$$

only if $C\big(F(0,T)\big) = P\big(F(0,T)\big)$.

Proof
Assuming the existence of a random variable S_T as claimed, it is readily observed that

$$C\big(F(0,T)\big) - P\big(F(0,T)\big) = DF(T)\,\mathbb{E}^{\mathbb{Q}}[S_T - F(0,T)] = 0.\square$$

While Lemma A.1 shows that the choice of $F(0,T)$ is unambiguous, it still does not guarantee that we can actually find a random variable explaining call and put wing perfectly. A sufficient condition guaranteeing this is that the function

$$\tilde{C}(K) := \begin{cases} P(K) + DF(T)\,\big(F(0,T) - K\big) & \text{, if } K \leq F(0,T), \\ C(K) & \text{, else} \end{cases},$$

for $K \in [0, \infty)$, is convex, which can be seen from the put-call-parity and (Hirsch, Roynette, 2012, Proposition 2.1). For this, it is sufficient to postulate that $C^{'}\big(F(0,T)\big) \geq P^{'}\big(F(0,T)\big) - DF(T)$, which in probabilistic terms means that the call-implied risk-neutral distribution attributes at least as much probability mass to the interval $[0, F(0,T)]$ as the put-implied risk-neutral distribution. Interpreting Lemma A.1 in terms of the implied volatility smile, choosing $F(0,T)$ as the root of $K \mapsto C(K) - P(K)$ is the only choice that makes the implied volatility smile $\sigma_{BS}(K)$ which is derived from $\tilde{C}(K)$ a continuous function, which it must be (cf. Rogers, Tehranchi (2008)).

B Implied volatility

Knowledge about the function $K \mapsto C(K)$ (or the function $K \mapsto P(K)$) is equivalent to knowledge about the risk-neutral probability distribution of S_T under $\mathbb{Q}$. Moreover, this knowledge is equivalent to knowledge about the function $K \mapsto \sigma_{BS}(K)$. From this perspective, the implied volatility smile is just a peculiar way to characterize the risk-neutral probability law of S_T analytically. Because of this it is instructive to collect some facts about the implied volatility that help to understand its link to the law of S_T.

B.1 Risk-neutral cdf

The risk-neutral distribution function of S_T can be written as a function of $\sigma_{BS}(.)$ as

$$\begin{aligned} \mathbb{Q}(S_T \leq x) &= 1 - \Phi\Big(d_-\big(\sigma_{BS}(x), k(x)\big)\Big) \\ &\quad + F(0,T)\,\sqrt{T}\,\sigma_{BS}^{'}(x)\,\varphi\Big(d_+\big(\sigma_{BS}(x), k(x)\big)\Big), \quad x > 0, \end{aligned} \tag{2}$$

where $k(x) = \log(x/F(0,T))$ denotes the log-moneyness. Regarding an interpretation of the last formula, the first summand $1 - \Phi\Big(d_-\big(\sigma_{BS}(x), k(x)\big)\Big)$ equals the probability that a lognormal random variable with volatility parameter $\sigma_{BS}(x)$ is smaller or equal to x, while the second summand represents a correction term.

B.2 Risk-neutral pdf

Assume that S_T is an absolutely continuous random variable on $(0, \infty)$. Introducing the function

$$w(y) := \sigma_{BS}^2\Big(F(0,T)\,e^y\Big)\,T, \quad y \in \mathbb{R}, \tag{3}$$

it is explained in Gatheral, Jacquier (2014) that the risk-neutral density, i.e. the density f_{S_T} of S_T under $\mathbb{Q}$, is given by

$$f_{S_T}(x) = \frac{\partial}{\partial x}\mathbb{Q}(S_T \leq x) = \frac{p\big(\log\big(\frac{x}{F(0,T)}\big) + \delta\,T\big)}{x}, \quad x > 0,$$

where the function[4] $p : \mathbb{R} \to [0,\infty)$ is given by

$$p(y) := \frac{g(y)}{w(y)\,\sqrt{2\,\pi}}\,\exp\Big(-\Big(\frac{y}{\sqrt{w(y)}} + \frac{\sqrt{w(y)}}{2}\Big)^2\Big),$$
$$g(y) := \Big(1 - \frac{y\,w^{'}(y)}{2\,w(y)}\Big)^2 - \frac{w^{'}(y)^2}{4}\,\Big(\frac{1}{w(y)} + \frac{1}{4}\Big) + \frac{w^{''}(y)}{2}.$$

This provides an expression for the risk-neutral density f in terms of the implied volatility smile $K \mapsto \sigma_{BS}(K)$.

B.3 Wing properties

The paper Lee (2004) derives the asymptotic behavior of $\sigma_{BS}(K)$ as $K \to \infty$ and as $K \searrow 0$. This behavior is shown to depend on the number of finite moments of S_T (for the case $K \to \infty$) and of $1/S_T$ (for the case $K \searrow 0$) under $\mathbb{Q}$. As an immediate corollary from the formulas derived in that reference, it follows that $\sigma^{'}_{BS}(0) = -\infty$ unless all positive moments of $1/S_T$ exist. In particular, positive mass at zero, i.e. $\mathbb{Q}(S_T = 0) > 0$, implies $\sigma^{'}_{BS}(0) = -\infty$. Moreover, it follows that if the density of S_T under $\mathbb{Q}$ has a light tail (i.e. one of the form $\exp(-\beta\,x)$ for some $\beta > 0$ as $x \to \infty$), then all moments of S_T exist, and

$$\limsup_{K\to\infty} \frac{\sigma_{BS}(K)}{\sqrt{\log\big(\frac{K}{F(0,T)}\big)\,T}} = 0,$$

i.e. $\sigma_{BS}(K)$ grows significantly slower than $\sqrt{\log(K)}$ with increasing K. For example, this is the case for the entropy-maximizing Buchen-Kelly density.

B.4 The SVI parameterization

The most popular parametric model for the implied volatility smile is *Gatheral's stochastic volatility inspired (SVI) model*. It defines the total implied variance (3) as a function of the log-strike $k \in \mathbb{R}$ in terms of the five-parametric function

$$w(k) := a + b\,\Big(\rho\,(k-m) + \sqrt{(k-m)^2 + \sigma^2}\Big),$$

for parameters $a, m \in \mathbb{R}$, $b \geq 0$, $\sigma > 0$, and $\rho \in (-1,1)$. Empirically, this parametric ansatz is well-known to provide an excellent match to most observed implied volatility smiles in the marketplace. Moreover, Zeliade Systems (2009) describe an efficient method to calibrate the five involved parameters to given market quotes.

Unfortunately, sufficient and necessary conditions on the parameters in order for the SVI model to imply a proper risk-neutral density seem to be unknown[5], i.e. the SVI-implied function (2)

[4] The function p equals the density of $\log(S_T/F(0,T)) + \delta\,T$.

[5] However, several necessary conditions on the parameters can be derived from results in Lee (2004); Rogers, Tehranchi (2008).

in general is not a proper distribution function. However, the following sufficient (and almost necessary) conditions for a three-parametric submodel are derived in (Gatheral, Jacquier, 2014, Theorem 4.2): for the three-parametric specification

$$(a, b, \rho, m, \sigma) = \Big(\frac{\theta}{2}\,(1-\rho^2), \frac{\theta\,\varphi}{2}, \rho, -\frac{\rho}{\varphi}, \frac{\sqrt{1-\rho^2}}{\varphi}\Big),$$

with $\rho \in (-1, 1), \theta, \varphi > 0$, the function (2) derived from the associated implied volatility smile is a proper distribution function if

$$-\frac{2}{\sqrt{\theta\,(1+|\rho|)}} \leq \varphi \leq \frac{2}{\sqrt{\theta\,(1+|\rho|)}}, \text{ and } \varphi < \frac{4}{\theta\,(1+|\rho|)}.$$

This three-parametric specification is typically sufficient to match observed market data quite well.

References

G. Bernhart, J.-F. Mai, Consistent modeling of discrete cash dividends, *Journal of Derivatives* **22** (2015) pp. 9–19.

P.W. Buchen, M. Kelly, The maximum entropy distribution of an asset inferred from option prices, *Journal of Financial and Quantitative Analysis* **31** (1996), pp. 143–159.

J. Gatheral, A. Jacquier, Arbitrage-free SVI volatility surfaces, *Quantitative Finance* **14** (2014) pp. 59–71.

F. Hirsch, B. Roynette, A new proof of Kellerer's theorem, *ESAIM: Probability and Statistics* **16** (2012) pp. 48–60.

R.W. Lee, The moment formula for implied volatility at extreme strikes, *Mathematical Finance* **14** (2004) pp. 469–480.

L.C.G. Rogers, M.R. Tehranchi, Can the implied volatility surface move by parallel shifts?, *Finance and Stochastics* **14** (2008) pp. 235–248.

C. Neri, L. Schneider, Maximum entropy distributions inferred from option portfolios on an asset, *Finance and Stochastics* **16** (2012), pp. 293–318.

C. Neri, L. Schneider, A family of maximum entropy densities matching call option prices, *Applied Mathematical Finance* **20** (2013), pp. 548–577.

Zeliade Systems, Quasi-explicit calibration of Gatheral's SVI model, *Zeliade white paper* (2009).

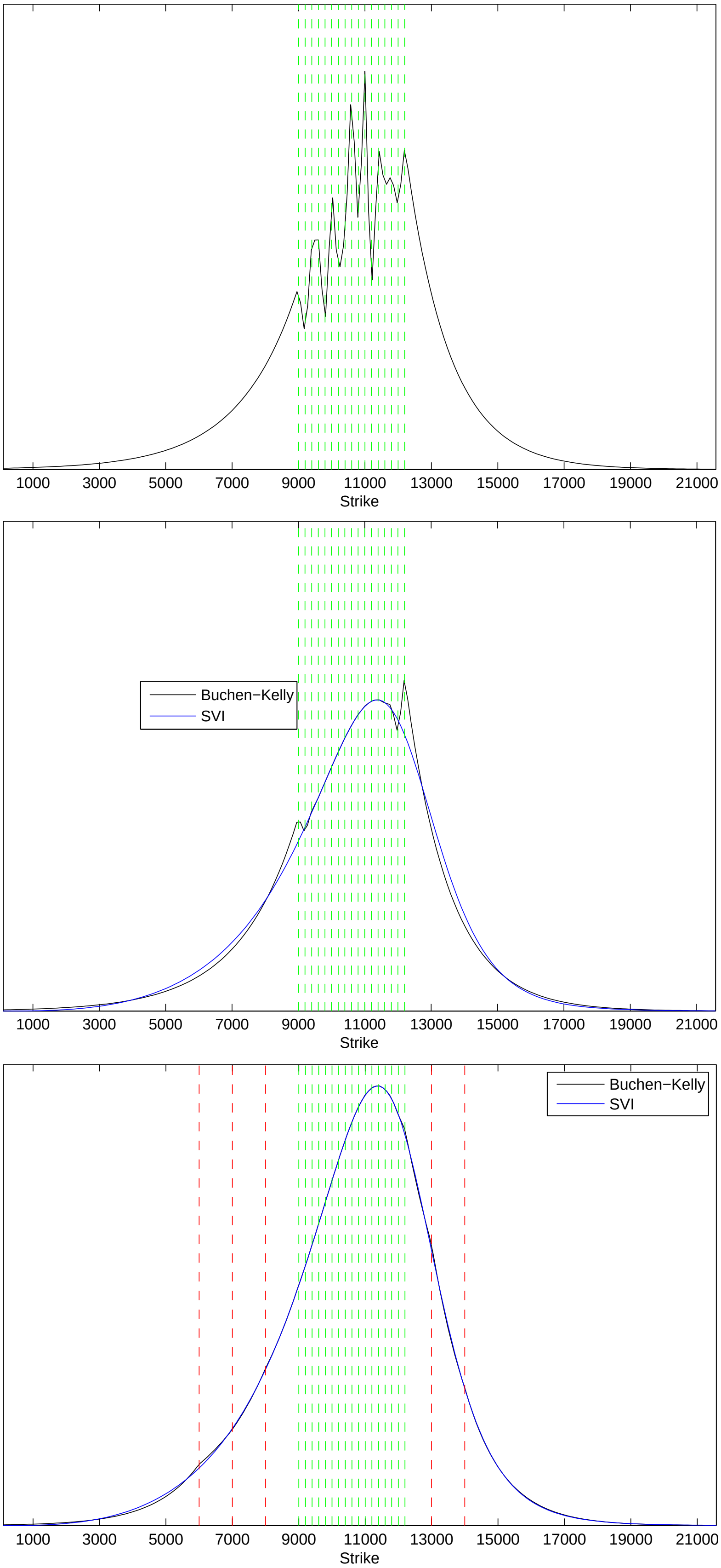

Fig. 1: Buchen-Kelly densities for the value of the DAX at 16 December 2016 retrieved from option data on 4 December 2015. Top: Input option prices are a subset of the quoted mid prices. Middle: Input option prices are SVI-implied (i.e. slightly smoothed) mid prices. Bottom: Input option prices are SVI-implied mid prices enhanced by five additional prices outside the observed moneyness range.

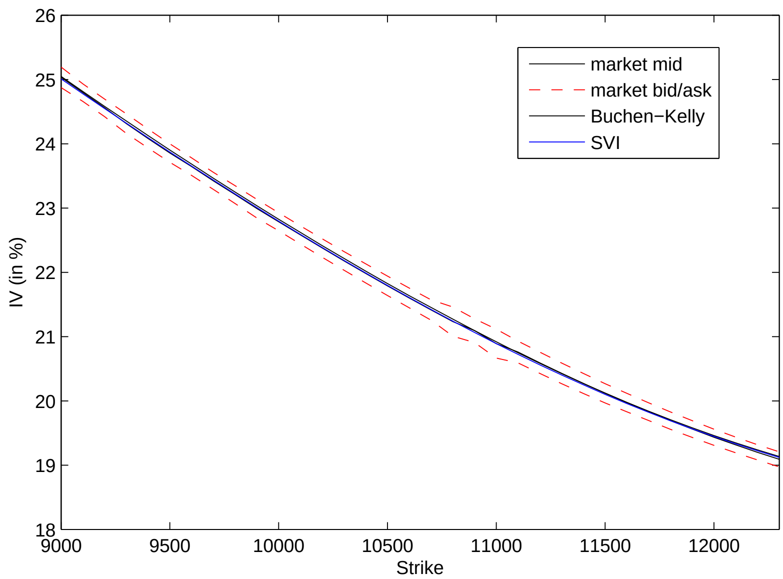

Fig. 2: Implied volatility smile as observed in the market and as computed from the fitted densities.