Bloomberg

Equity Implied Volatility Surface

April 17, 2017

Quantitative Analytics
Bloomberg L.P.

Abstract

The implied volatility surface is a fundamental object for the pricing and risk management of derivatives. The construction of this surface from listed option prices typically proceeds in two stages. First, since forward prices are not directly quoted in the listed markets, a forward curve has to be implied in a manner consistent with the observed option prices. Second, given the forward, an implied volatility model has to be calibrated to the observed option prices. In this document, we describe how to construct, interpolate and extrapolate such a forward curve and its corresponding dividend curve from options for both equities and equity indices, properly accounting for American exercise and dividend seasonality. Then we describe a lognormal mixture model along with details surrounding calibration, interpolation and extrapolation. We illustrate the performance of the methodology with sample numerical results on selected equity and index underlyings.

Keywords. Implied forwards, dividends, put-call parity, implied volatility, option greeks, dividend seasonality, dividend extrapolation, lognormal mixture model, equity derivatives.

First published June 21, 2011.

Contents

List of Figures

Part I

Implied Forwards and Dividends

1 Summary

1.1 Background

The forward price of an equity is a fundamental quantity impacting option pricing, hedging and implied volatility, yet forward prices are not directly quoted in the listed equity markets. Futures are traded on some equity indices as well as European equities, but their maturities (typically quarterly) may not include all option maturities. Interpolating the futures curve at an intermediate option maturity is a nontrivial task because future dividends payments and payment times, aside from those already announced for the very near term, are generally unknown. Hence forward prices must be implied from market observables, *e.g.* exchange-traded instruments like equity futures, equity options, dividend futures, and OTC instruments such as dividend swaps and total return swaps.

In this article, we focus on inferring a forward curve and a corresponding dividend curve based on exchange-traded vanilla options only, deferring the usage of dividend futures and OTC instruments to future work.

1.2 Difficulties

In principle, forward prices are straightforward to derive via put-call parity of European option prices:

$$F(T) = C(K,T) - P(K,T) + K, \tag{1}$$

where $F(T)$ is the forward at time T, $C(K,T)$ and $P(K,T)$ are respectively the future values of European call and put options struck at K as of the expiration time T. In practice, many difficulties impede the usage of this simple relationship. In some cases, the underlying asset and options on the asset trade during different times on different exchanges. For example, stocks in the Nikkei 225 index trade on the Tokyo Stock Exchange between 0900 – 1130 hrs and 1230 – 1500 hrs Tokyo time. Nikkei index options, however, trade 0900 – 1515 hrs (pit) and 1630 – 0300 hrs (electronic) on the Osaka Stock Exchange, hence the spot price of the underying is unavailable during substantially large periods of option trading as seen in Figure 1. Even when the spot market is open, the index spot price being an average of non-contemporaneous traded prices of the constituent members, can be unreliable, especially within the first few minutes of the opening of a new trading session when not all constituents have started trading yet. The same figure shows the futures price of the prompt contract consistently higher than the spot price (implying a negative dividend yield on the index), therefore indicating an unreliable spot price. Option prices can also be quite erratic, especially close to the opening and closing times of the trading session. At other times too, it is not uncommon to see a complete lack of quoting activity (even for near-the-money strikes), or a one-sided market, or two-sided markets with unreasonably wide bid-ask spreads; see Figure 2.

Even when high quality data are available, option prices at different strikes will typically produce inconsistent forwards in (1) for the same maturity; see Table 1 where the Nikkei 1-month implied forward using mid prices of options ranges from approximately 14496 to 14507 across different strikes. In more pathological, yet realistic, cases with poor option data, this range could span a hundred index points or more. As derived and non-tradable quantities, the mid prices of options do not satisfy a put-call parity relationship, hence do not identify the correct forward. Using this incorrect forward along with mid prices of options will result in unequal put and call implied volatilities at the same strike. Rather, bid/ask prices of options imply corresponding bid/ask ranges on the forward. Any value within the intersection of these ranges across different strikes is an admissible forward price in the sense that it is consistent with option prices at those strikes. If the forward is correctly identified, one should see nearly identical implied volatilities for the put and the call at that strike. In Table 1, this criterion is closely met for the 14500 strike, implying that the correct forward is approximately 14500.

This forward price is the result of the stock price drift composed of three components: the risk-free interest rate, the dividend and the borrow cost. Since the risk-free interest rate is observable, the forward implies a combined dividend and borrow cost. At this time, we do not isolate the borrow cost, and references to dividend in the sequel include the borrow cost as well.

1.3 Methodology

In our approach, we calculate option-implied dividends approximately every 5 minutes, reusing them in between snapshots to recalculate the forward from the current spot price.

The main inputs to this calculation are the spot price, options prices of all strikes and maturities, and the risk-free interest rate curve. For equities paying a discrete cash dividend, we require a dividend schedule consisting of dates and amounts of declared dividends, along with estimated dates of future dividends, and this information is sourced from BDVD, a source of Bloomberg dividend forecasts.

The main output is an implied dividend curve, which:

- matches any declared dividends in the near term;
- is stable across time;
- yields forward prices satisfying put-call parity; and
- yields stable implied volatilities.

In each snapshot, the calculation bootstraps the implied dividend curve, one option maturity at a time, starting with the shortest maturity. For each maturity, put and call bid/ask prices at near-the-money strikes are selected based on certain admissibility criteria. Then the mid prices (de-Americanized, if necessary) are used to derive a raw implied dividend via put-call parity. This implied dividend is denoted by a discrete cash payment for equities and by a continuous yield for indices. For equities that do not pay a dividend, the implied dividend is denoted by a continuous yield, with the tacit understanding that this is actually a borrow cost. The raw implied dividend is noisy, and is smoothed using a Kalman filter, which averages, in a certain sense, the current raw implied dividend with its past values. The weighting given to the current raw implied dividend in this averaging process depends on the number of past observations used to establish the previous dividend level, and whether the previously established dividend level satisfies put-call parity given the current option prices. The smoothed implied dividends produced by the bootstrapping process constitutes the implied dividend curve. The Kalman filter persists across snapshots, producing a smooth implied dividend time series at any listed maturity, while adapting to changing market perceptions of dividend as implied by options prices.

These calculations are detailed in the subsequent sections.

Figure 1: Nikkei 225 spot and futures prices.

2 Data Filtering

As seen in Figure 2, some preprocessing checks are needed to ensure that market data of an acceptable quality are used in our calculations.

The first, and perhaps the most crucial, market data item to check is spot price availability and consistency. To describe the crux of the methodology, let us assume for the moment that spot and options markets are open during the same trading hours, and that consistent and contemporaneous spot and option quotes are available. (Some exceptions are treated in §6.)

Next, option prices are filtered according to the following criteria:

1. For options trading on multiple exchanges, the exchange offering the most liquidity is selected.
2. Stale traded prices older than 3 days are eliminated.
3. Bid/ask quotes that are too small (less than 0.001) are eliminated.
4. Additionally, quotes are also eliminated based on criteria on the bid/ask spread. In particular, if the ask price is greater than 3 times bid price, or if the spread is more than 3 standard

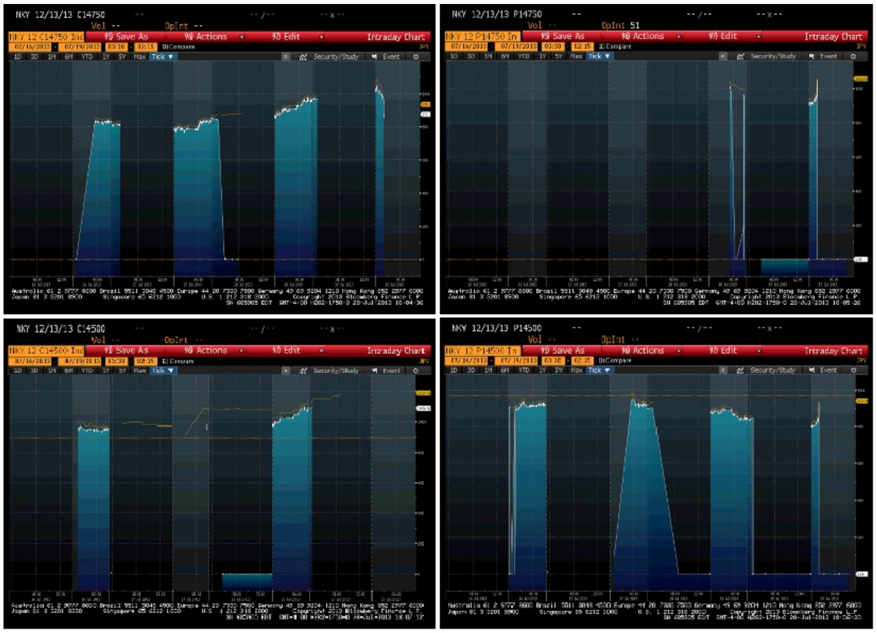


Figure 2: Nikkei 225 near-the-money option prices. The forward prices for the trading session shown were approximately in the 14600 – 14700 range.

Strike	Impl. Fwd	Call Vols (%)			Put Vols (%)		
		bid	ask	mid	bid	ask	mid
12500	14503.68	17.189	41.100	34.770	33.832	34.062	33.947
12750	14502.66	24.334	38.334	32.927	32.355	32.542	32.449
13000	14507.64	28.202	35.297	32.048	30.590	31.324	30.961
13250	14507.61	27.242	33.064	30.315	29.124	29.721	29.424
13500	14502.59	26.413	31.221	28.896	28.405	28.887	28.649
13750	14500.07	24.521	30.239	27.452	27.038	27.852	27.448
14000	14505.05	25.045	28.584	26.828	26.116	26.826	26.472
14250	14500.02	24.613	26.554	25.584	25.260	25.906	25.584
14500	14500.00	24.562	25.184	24.873	24.563	25.185	24.874
14750	14499.98	24.191	24.509	24.350	23.874	24.830	24.352
15000	14507.46	23.845	24.193	24.020	22.445	24.544	23.501
15250	14502.43	23.687	24.090	23.890	20.727	26.452	23.693
15500	14492.41	23.656	24.154	23.906	19.141	29.138	24.651
15750	14496.89	23.586	24.357	23.977	16.583	29.347	24.375

Table 1: Forwards implied from option mid prices using put-call parity. Mid volatlity is the implied volatility corresponding to the mid option price.

deviations away from the average spread for that maturity, the quote is eliminated. In addition, quotes with bid/ask spreads wider than a certain multiple of the dividend yield are eliminated. This multiple is 0.09% for maturities up to 1 year, and 2.1% for longer maturities.

5. At any fixed maturity, call and put prices at all available strikes are restricted to the longest monotonic subsequences, with any offending quotes eliminated.

3 Raw Implied Dividend

European option prices imply a foward and a corresponding dividend via put-call parity. With American options, we need to first calculate the corresponding European option prices prior to using put-call parity. This requires simultaneously calculating, at any given maturity, an implied dividend and an implied volatility smile that matches the American option prices of all strikes at that maturity using the following iterative procedure we call *de-Americanization*.

Input: Begin with an input yield curve and a dividend schedule consisting of declared dividends (dates and amounts) and a forecast for future dividends drawn from BDVD. For indices, dividends are aggregated and represented as continuous yields across option maturities. This gives an estimate of the current forward.

European Prices: Skip this step for European options. For American options at any given maturity, compute the Black implied volatilities from put and call prices and the current forward. (Ideally, this calculation should use a single local volatility model consistent with options of all strikes at this maturity, however a Black implied volatility for each strike is a much faster calculation yielding results of acceptable accuracy.) Finally, use the current forward and the implied volatilities in the Black model to compute the corresponding European option prices.

Implied Forward & Dividend: Choose the two nearest strikes on either side of the current forward that have admissible bid/ask quotes for both puts and calls. Take the new candidate forward to be the average of the forward obtained at each strike by applying put-call parity to the option mid prices, and back out a corresponding implied dividend.

Iterate: Evaluate the termination criteria:

- Does the new candidate forward lie within the two chosen strikes?
- Is the relative change in the forward between the current iteration and the previous one less the 0.1%?

If both criteria are satisfied, then stop here; estimates of the implied forward and dividend have been computed. If not, go back to the "European Prices" step above, taking the new candidate forward as the current guess, and iterate.

For some equities, multiple option maturities may be preceded by a single dividend date, hence multiple implied forwards to impute a corresponding implied dividend. To get a dividend consistent with option prices at all the relevant maturities, a dividend bid/ask range is implied for each maturity, and the mid-point of their intersection is taken as the implied dividend attributed to that dividend date. If the intersection of these ranges is empty, then we take the implied dividend to be the mid-point of the range from the nearest option maturity, with the tacit assumption that discrepancies in the remaining forwards are to be attributed to borrow cost.

4 Dividend Smoothing

In implementation, our dividend curve is bootstrapped as a series of spot yields (indices) or discrete cash dividends (equities) across successive option maturities using mid option prices at each market snapshot. That is, given a spot price S, a continuously compounded risk-free rate r and forward F for a fixed option maturity T, the dividend yield d up to T for indices is given by

$$d = \frac{1}{T} \ln \left(\frac{S \exp(rT)}{F} \right),$$

and the discrete cash dividend for equities is given by

$$d = S \exp(rT) - F.$$

In the remainder of this section, we refer to either case as simply "dividend". Since the dividend over any time interval can be sensitive to noise in the option prices, some form of smoothing is necessary to avoid erratic variations in the term structure, or instabilities in the dividend curve from one snapshot to the next.

We use a classical Kalman filter to achieve this smoothing. To fix notation, assume that the snapshots are spaced Δt apart in time, and consider the dividend to a fixed option maturity. At the k^{th} snapshot, the Kalman filter hypothesis is that the true (unobservable) dividend D_k^* (the Kalman state) evolves according to

$$D_k^* = D_{k-1}^* + w_k,$$

where $w_k \sim N(0, \sigma^2 \Delta t_k)$ is white noise. The raw dividend d_k, a noisy measurement of the true dividend D_k^*, is assumed to follow the process

$$d_k = D_k^* + v_k$$

where $v_k \sim N(0, \alpha^2)$ is another source of white noise independent of w_k. The *a posteriori* estimate of the dividend D_k given the measurement d_k is given by

$$D_k = D_{k-1} + G_k(d_k - D_{k-1})$$

where G_k is the Kalman gain. This leads to *a priori* and *a posteriori* error estimates $e_k^- = D_k^* - D_{k-1}$ and $e_k = D_k^* - D_k$ with respective variances, say, P_k^- and P_k. It is straightforward to show that the optimal gain that minimizes the *a posteriori* error variance P_k is then given by

$$G_k = \frac{P_{k-1} + \sigma^2 \Delta t_k}{P_{k-1} + \sigma^2 \Delta t_k + \alpha^2}.$$

The Kalman filtering process then consists of iterating the "Predict" and "Update" steps below:

Initialize

$$
\begin{aligned}
D_0 &= d_0 && \text{(initial estimate of state)} \\
P_0 &= \alpha^2 && \text{(initial estimate of error variance)}
\end{aligned}
$$

Predict

$$
P_k^- = P_{k-1} + \sigma^2 \Delta t_k \qquad (\textit{a priori} \text{ variance prediction})
$$

Update

$$
\begin{aligned}
G_k &= \frac{P_{k-1} + \sigma^2 \Delta t_k}{P_{k-1} + \sigma^2 \Delta t_k + \alpha^2} && \text{(Kalman gain update)} \\
D_k &= D_{k-1} + G_k (d_k - D_{k-1}) && (\textit{a posteriori} \text{ state update}) \\
P_k &= (1 - G_k) P_k^- && (\textit{a posteriori} \text{ variance update})
\end{aligned}
$$

In practice, the implied dividend from the market may not follow any of the dynamics modeled above, and regime changes in the implied dividend may result in put-call parity violations. Such violations are easily detected by testing whether the dividend estimateD_k lies within the admissible bid/ask dividend bounds $(d_k^{\text{bid}}, d_k^{\text{ask}})$ implied by option bid/ask prices. When a parity violation is detected, the Kalman gain is temporarily boosted to catch up with the regime change. Even without sudden regime changes, closely tracking the implied mid dividend requires that every new measurement d_k be given some nontrivial weight in updating the state estimate D_{k-1}. Yet, as the state estimate is refined after many observations, the gain update formula leads to an arbitrarily small gain, hence an arbitrarily small weight assigned to newer observations. This can be alleviated by flooring the Kalman gain to a positive value. Specifically, we use the gain update formula

$$
G_k = \begin{cases} \max\left\{0.03, \dfrac{P_{k-1} + \sigma^2 \Delta t_k}{P_{k-1} + \sigma^2 \Delta t_k + \alpha^2}\right\}, & \text{if } d_k^{\text{bid}} < D_k < d_k^{\text{ask}} \\ 0.5, & \text{otherwise} \end{cases}
$$

Since the state update for D_k does not explicitly guarantee a positive dividend, we floor the dividend at zero in each step:

$$
D_k = \max\left\{0, D_{k-1} + G_k (d_k - D_{k-1})\right\}.
$$

For equities that do not pay a dividend and total return indices, this flooring eliminates small and spurious negative dividends implied from noisy option prices, however potentially at the expense of losing information on borrow cost, which can be negative in some cases.

Normalizing the variances P_k^- and P_k by $\sigma^2 \Delta t$, it can be seen that the filtering process depends on α and σ only through their ratio; we set $\alpha/\sigma = 1.0$. Implied dividend snapshots are spaced $\Delta t =$ 5 minutes apart, commencing at market open and lasting the entire trading session until market close, except as noted earlier for Asian indices. Examples of this dividend smoothing procedure can be seen in Figure 4 – Figure 9. The Kalman filter state is preserved across trading days, with any required ex-dividend adjustments made immediately prior to the start of a new trading session. For indices, these ex-dividend adjustments are aggregated over all index constituents and converted to an equivalent adjustment in yield. To ensure that very stale option data do not contaminate the current implied dividend, the Kalman filter is removed when 7 days have elapsed without a new measurement. The implied dividend at this maturity is then interpolated or extrapolated until new option data become available, at which time the Kalman filtering process resumes. Dividend interpolation and extrapolation require treatment of dividend seasonality, all of which are discussed next.

5 Dividend Seasonality, Interpolation and Extrapolation

5.1 Seasonality

Interpolation or extrapolation of an implied dividend needs to account for the fact that many equities and indices exhibit a strong seasonality in their dividends. US companies tend to pay the same dividend during each quarter of the same fiscal year, but some European and Asian companies may have a smaller interim dividend, then a larger final dividend within the same fiscal year. When aggregated over constituent members, this phenomenon carries forward to indices too. An extreme example of this is the KOSPI 200 index which pays around 95% of its annual dividend within a one week period in December.

Our simple approach to seasonality is equally amenable to discrete cash dividends and continuous dividend yields.

Sampling: The first step is to assemble a working sample of dividend data. This is taken to be the BDVD forecast, which typically extends to about three years into the future. In case the BDVD forecast covers less than one year, we augment it with historical dividends so that we obtain at least one year of dividend data. If the BDVD forecast is entirely missing, we assemble up to 10 years of historical dividend data.

Binning: Then we parition the calendar year into equally spaced bins: 12 monthly bins for equities, and 52 weekly bins for indices.

Weighting: For each calendar year in the sample, we calculate the percentage of the annual cash dividend falling into each bin. The percentage allocated to each bin is averaged over all the calendar years in the sample to yield an average seasonality weight for that bin.

Examples of calculated seasonality weights for some major indices are shown in Figure 3.

5.2 Interpolation

To interpolate an implied dividend between two times, the implied cash dividend accrued between those two times is allocated according to the average seasonality weight into the monthly (equity) or weekly (index) bins. The cash dividends can then be accumulated again up to the interpolated maturity of interest (equity), or restated as an equivalent annualized yield (index). This interpolation applies regardless of whether the interpolated maturity is a listed expiry or not.

For example, if a company XYZ pays dividend \$0.4 on April 22nd and \$0.6 on July 22nd in 2016. BDVD forecast that there will be dividends paid on April 21st and July 21st in 2017. There are options expire in March 2017 and September 2017 only. For illustration purpose let's further assume the interest rate is 0. If the put-call parity implied forwards show that there are \$1.2 dividend paid between March 2017 and September 2017, then we set the dividend value \$0.48 to April 21st 2017 and \$0.72 to July 21st 2017 so that their ratio is 2/3, the same as in 2016.

5.3 Extrapolation

To extrapolate the dividend curve to a maturity (whether listed or not) beyond the last available implied dividend, we first identify a forward yield over a sample period as follows. If implied dividends are available at listed expiries longer than 1 year, then the sample period is the most long-dated pair of listed expiries at least 1 year apart, and we take the annualized forward yield over this period. If implied dividends are not available past 1 year, then we include BDVD forecasts to augment the period to at least 1 year. In this case, the implied cash dividend coupled with the BDVD forecast amounts are combined and restated as an equivalent annualized forward yield. If the implied dividends are not available past 1 year, and no BDVD forecast is available, then we include historical dividends to augment the period to at least 1 year. Again, the implied cash dividend is coupled with the historical dividend payments, and restated as an annualized forward yield.

For indices, this annualized forward yield is then extrapolated flat past the last listed maturity. For equities, the extrapolated yields are converted back to discrete cash amounts for maturities up to 10 years, beyond which they are retained as yields.

6 Extended Trading Hours

Spot and option quotes are not always contemporaneous because they may trade on different exchanges or during different hours. A prominent example is the S&P 500 index: in addition to the regular trading session (RTS) during US trading hours, there is an extended trading session (ETS) for options during European trading hours, when the US spot markets are closed. A reliable spot price is needed during this session to compute an implied dividend, implied volatility and option greeks. A derived spot S during the ETS is obtained by assuming that the ratio of the spot to the discounted futures[1] during the ETS remains the same as the ratio of the spot to the discounted implied forward during the previous RTS.

To elaborate, we identify the reference index underlier with the most liquid futures trading; for S&P 500 index, this is the S&P 500 E-mini index (ESA Index). Denote its spot price (adjusted for dividend payouts from stocks going ex-dividend the next day) at the end of the previous regular trading session by S_{ref}. We then identify the maturity T of the most active futures contract for the reference index (for ESA Index, this is just the front month contract); denote the implied forward and the risk-free rate to T at the end of the previous regular trading session by F_{ref} and

[1] We prefer to use the futures price during the ETS since option prices and the corresponding implied forward may be less reliable.

r_{ref} respectively. We assume that, during the ETS, the spot S, the risk-free rate r and the futures F to maturity T satisfy

$$\frac{S}{F\exp(-rT)} = \frac{S_{\text{ref}}}{F_{\text{ref}}\exp(-r_{\text{ref}}T)},$$

leading to the derived spot

$$S = S_{\text{ref}} \cdot \left[\frac{F\exp(-rT)}{F_{\text{ref}}\exp(-r_{\text{ref}}T)}\right].$$

At present, the discounted futures methodology applies only to the S&P 500 extended trading session. Various Asian indices have similar issues, *e.g.* the Hang Seng index with its mid-day break, and the Nikkei 225 index with the electronic option trading session during New York trading hours. Even when both spot and options markets are open, spot and option prices may not be consistent. Whereas equity spot prices, whether last traded price or bid/ask quotes, are readily available at market open, index spot prices can take some time to stabilize as quoting activity picks up on all the index constituents, and dividends implied from option prices during this period can be erratic, even negative. In the future, we plan to extend the discounted futures approach to more indices to alleviate these problems.

If a liquid index futures market doesn't exist, the discounted futures approach is not applicable. An alternative approach is to adjust the discounted implied forward from the first listed maturity by the BDVD dividend forecasts aggregated over the index constituents to arrive at an option-implied spot. OVDV presently uses this calculation for Asian indices in the first 5 minutes of the trading session when the index spot has not stabilized, although index option prices are reliable.

7 Blended Dividends

Vanilla options determine only the forward price, yet implied volatilities and exotic pricing of equity derivatives rely on the full structure of future dividends. Whereas short-term dividends may be declared or forecast with some accuracy, long term dividend forecasts are subject to much more uncertainty. Therefore it is convenient to model equity cash dividends as being discrete cash payments in the short term, a blend of discrete cash and discrete stock (*aka.*proportional) dividends in the medium term, and a continuous yield in the long term.

To this end, for each discrete dividend time (measured from "today") $0 = t_0, t_1, \ldots, t_n, t_{n+1} = T$ we assume weights as follows:

- for cash dividends: $\omega_1^{\text{d}}, \ldots, \omega_n^{\text{d}}$;
- for stock dividends: $\omega_1^{\text{y}}, \ldots, \omega_n^{\text{y}}$, and
- for continuous yields: $\omega_1^{\text{q}}, \ldots, \omega_n^{\text{q}}$.

Those weights determine the percentage of each type that makes up the total dividend at time t_i, hence sum up to unity:

$$\omega_i^{\text{d}} + \omega_i^{\text{y}} + \omega_i^{\text{q}} = 1.$$

For dividend paying equities, we set

- $\omega_i^{\mathrm{d}} = 1$ (purely discrete cash dividend) for t_i up to 1 year;
- $\omega_i^{\mathrm{d}} = \omega_i^{\mathrm{y}} = 0.5$ (50-50 blend of discrete cash and stock dividend) for t_i between 1 year and 2 years;
- $\omega_i^{\mathrm{y}} = \omega_i^{\mathrm{q}} = 0.5$ (50-50 blend of discrete stock dividend and continuous dividend yield) for t_i between 2 years and 10 years; and
- $\omega_i^{\mathrm{q}} = 1$ (purely continuous yield) for t_i greater than 10 years.

For non-dividend paying equities and indices, we set $\omega_i^{\mathrm{q}} = 1$ (pure dividend yield) for all t_i.

We use the following strategy to attribute the total dividend up to time T into each of the three components: cash dividends d_i and stock dividends y_i at payment times t_i, as well as the continuous dividend yield q applicable over the period $(0, T]$. To simplify the discussion, we assume there is no seasonality. Let F_0 and F_T be the known forwards at time t_0 and T.

1. Imply $\hat{d}$ from
$$F_T = F_0 \mathrm{e}^{rT} - \hat{d} \sum_{i=1}^{n} \mathrm{e}^{r(T-t_i)}$$
2. Imply q from
$$F_0 \mathrm{e}^{(r-q)T} = F_0 \mathrm{e}^{rT} - \hat{d} \sum_{i=1}^{n} \omega_i^{\mathrm{q}} \mathrm{e}^{r(T-t_i)}$$
3. Imply d from
$$F_T = F_0 \mathrm{e}^{(r-q)T} - d \sum_{i=1}^{n} \mathrm{e}^{(r-q)(T-t_i)}$$
4. For $i = 1, \ldots, n$ do
   (a) Let
$$d_i = d \frac{\omega_i^{\mathrm{d}}}{\omega_i^{\mathrm{d}} + \omega_i^{\mathrm{y}}}$$
   (b) Let
$$F_i = F_{i-1} \mathrm{e}^{(r-q)(t_i - t_{i-1})} - d$$
   (c) Imply y_i from
$$F_i = F_{i-1} \mathrm{e}^{(r-q)(t_i - t_{i-1})} (1 - y_i) - d_i$$

8 Limitations

Despite the simplicity of the put-call parity relationship, reliable estimates of option-implied dividends are difficult to obtain due to market data illiquidity and idiosyncracies, with added complexity coming from American exercise and dividend seasonality.

Sample results in §9 show that the methodology typically produces implied dividends that are:

- consistent with observed option prices;
- stable within and across trading sessions;
- in close agreement with other sources (*e.g.* analyst dividend estimates, dividend swap quotes).

The implied dividend could be erroneous in cases where options are pricing in other pending corporate actions (*e.g.* spining off a business, M&A transaction, transfer of assets or liabilities *etc*) to which our method is agnostic. Other market circumstances such as a temporary halting of trading or a short-sale ban would also result in dislocated option prices, hence potentially inaccurate implied dividends and forwards.

9 Numerical Results

We show seasonality calculations on some major indices in Figure 3. These plots show the cumulative fraction of the annual dividend as a function of calendar week.

We summarize in Figure 4 – Figure 9, a few examples of our implied dividend methodology on selected equities and indices across North America, Europe and Asia: Ford Motor Company (F US Equity), Eni Spa (ENI IM Equity), BNP-Paribas SA (BNP FP Equity), the Dow Jones index (DJX Index), the EuroStoxx-50 index (SX5E Index) and the Hang Seng China Enterprises index (HSCEI Index). Each plot shows the smoothed implied dividend as well as the admissible bid/ask range ($d^{\text{bid}}, d^{\text{ask}}$) from option quotes at the two strikes containing the forward, based on snapshots spaced 5 minutes apart, and spans trading days from 05-May-2014 through 21-May-2014, except the last plot for BNP-Paribas which spans trading days from 27-May-2014 through 02-Jun-2014.

The plots support several immediate inferences. First, the dividend bid/ask bounds, hence their mid value tracked by the Kalman filter, are clearly noisy, however the calculated implied dividend is much less noisy, demonstrating the smoothing achieved by the Kalman filter. Second, the smoothed dividend closely tracks the mid dividend from the bid/ask bounds, hence the implied dividend is robust to spikes in the bid/ask spread, as long as the mid is relatively unchanged (EuroStoxx-50, T = 1.62y). Third, whenever the implied dividend touches or overshoots either of the bid/ask bounds signifying a parity violation, the filter corrects the estimate very quickly, whether it's due to tight spreads (Ford, T = 0.28y), a temporary dislocation of the options market (HSCEI, T = 0.14y), or a more persistent regime change (Ford, T = 1.70y). Fourth, when option data are missing (HSCEI T = 0.40y, 28-Aug-2014 expiry), the implied dividend is held stable at a reasonable level.

Looking a little more closely, we see that declared dividends for equities are correctly reproduced. For example, Eni Spa declared a €0.55 dividend on 23-Jan-2014 with an ex-dividend date of 19-May-2014, which correctly matches the large implied dividend drop for T = 0.62y (18-Dec-2014 expiry). For indices too (Dow Jones and EuroStox-50), the effect of constituent stocks going ex-dividend is clearly visible, although the impact can be slightly muted by dividend bid/ask spreads that are too wide relative to the dividend itself (for example, Dow Jones index T = 1.62y, 19-Dec-2015 expiry).

Finally, we highlight how the methodology correctly tracks the BNP-Paribas implied dividend drop on 30-May-2014 as soon as rumors and news broke out that the US government was seeking to impose fines in excess of $10 billion; see Bloomberg news story `NSN N6DAZA6S973N<Go>`.

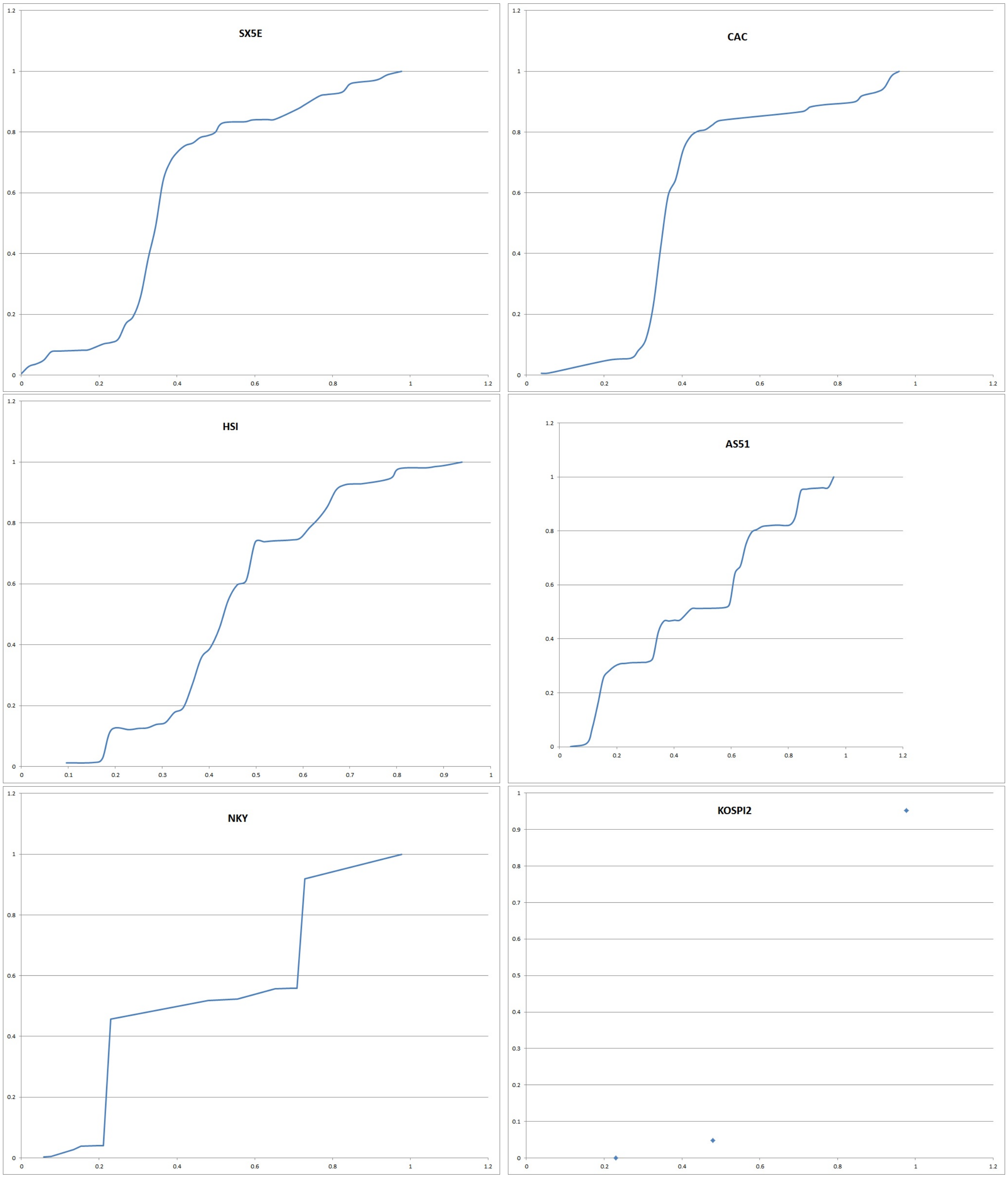


Figure 3: Dividend seasonality for major indices: cumulative fraction of annual dividend as a function of weekly bins within the calendar year. KOSPI2 is shown as a scatter plot with only 3 dividend paying weeks, the last of which accounts for over 95% of the annual dividend.

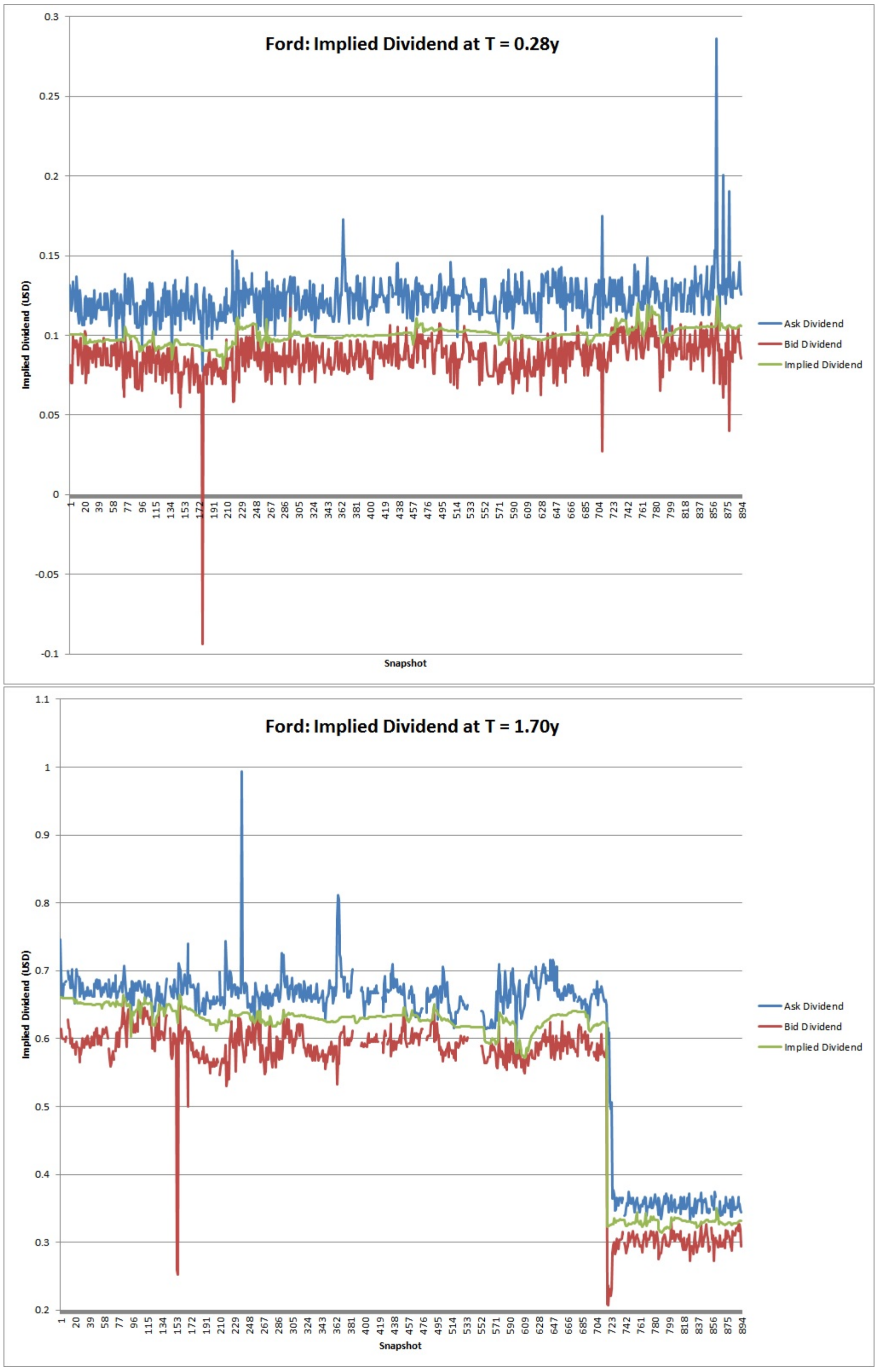


Figure 4: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for Ford Motor Company.

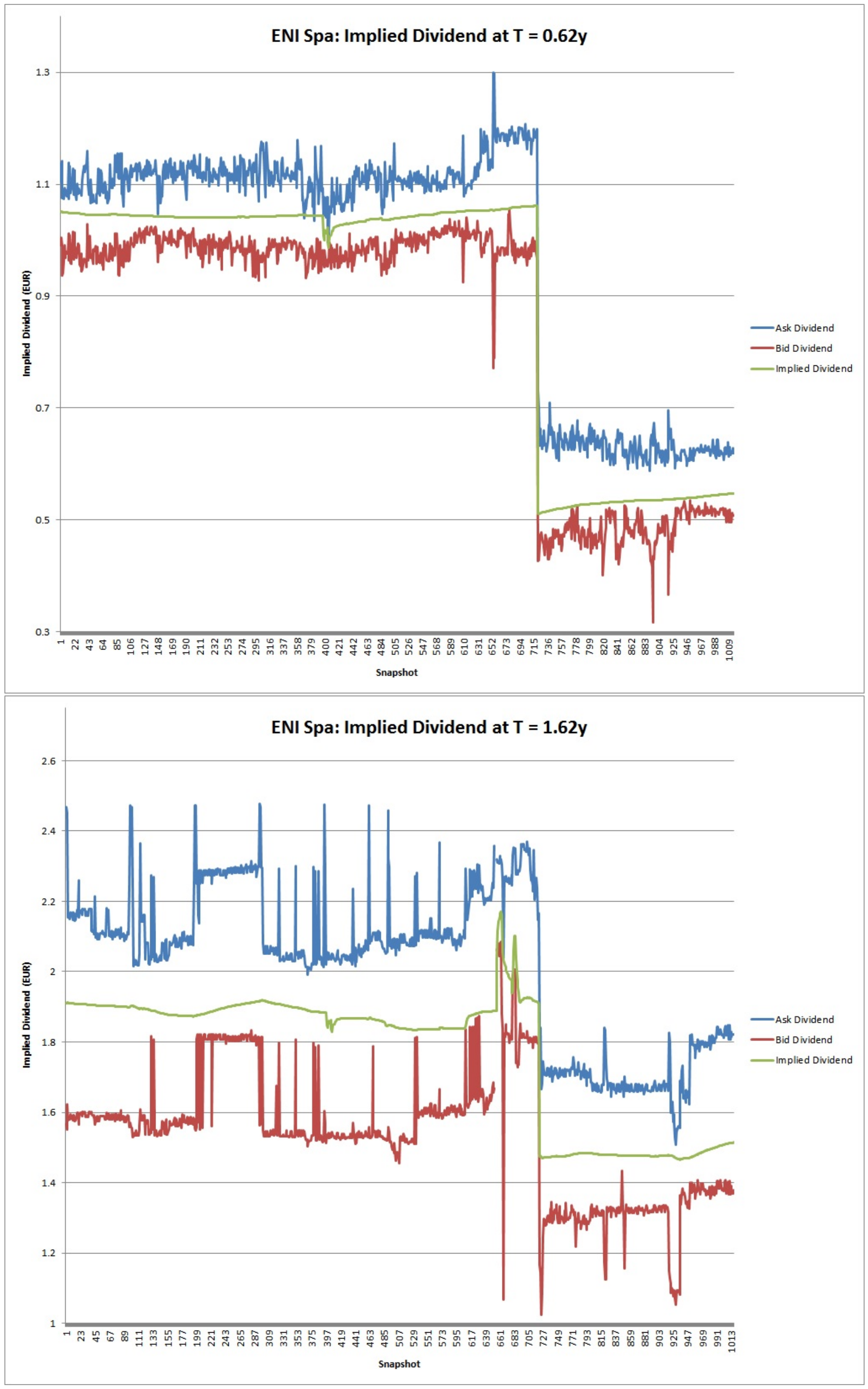


Figure 5: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for Eni Spa.

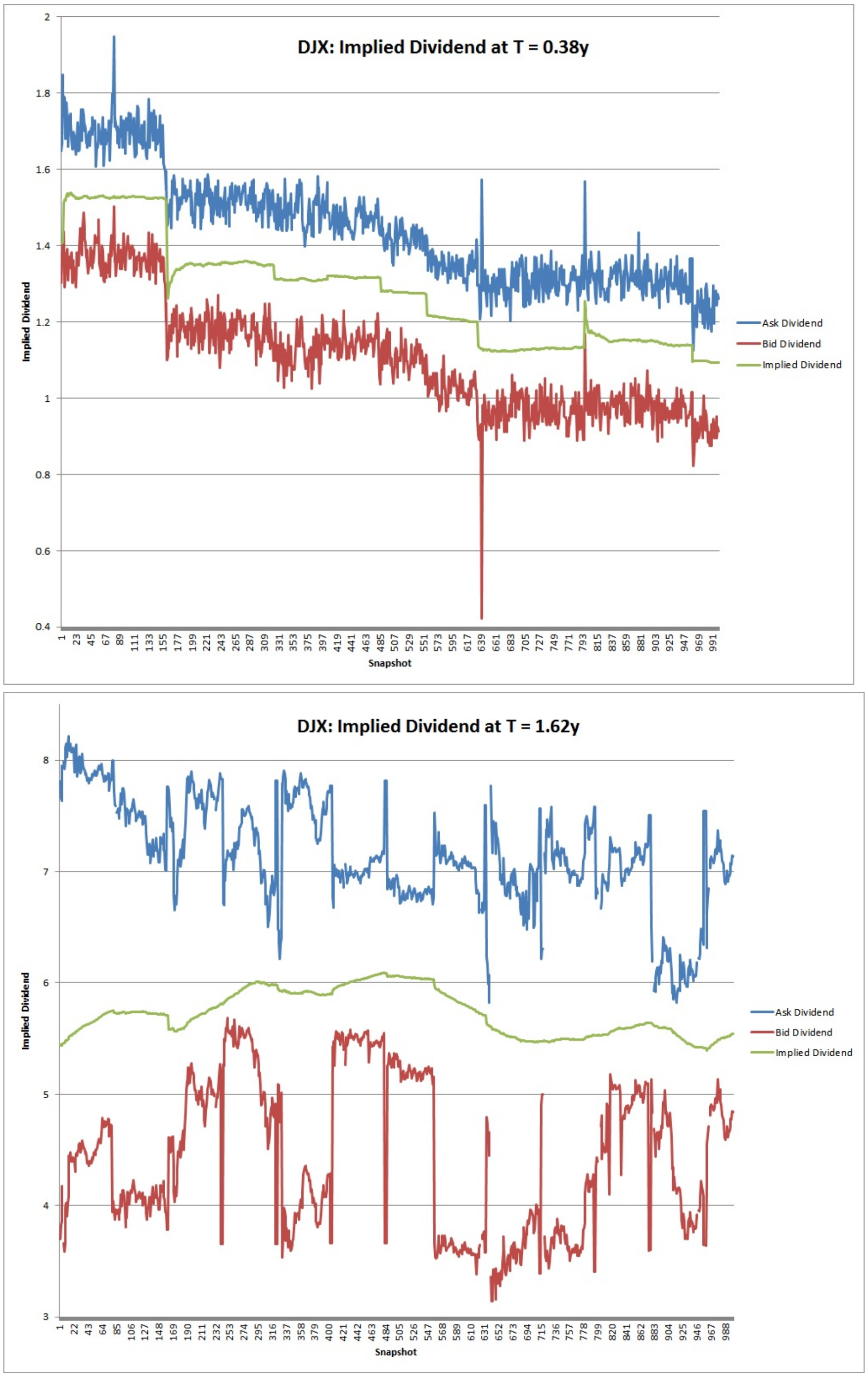


Figure 6: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for the Dow Jones index.

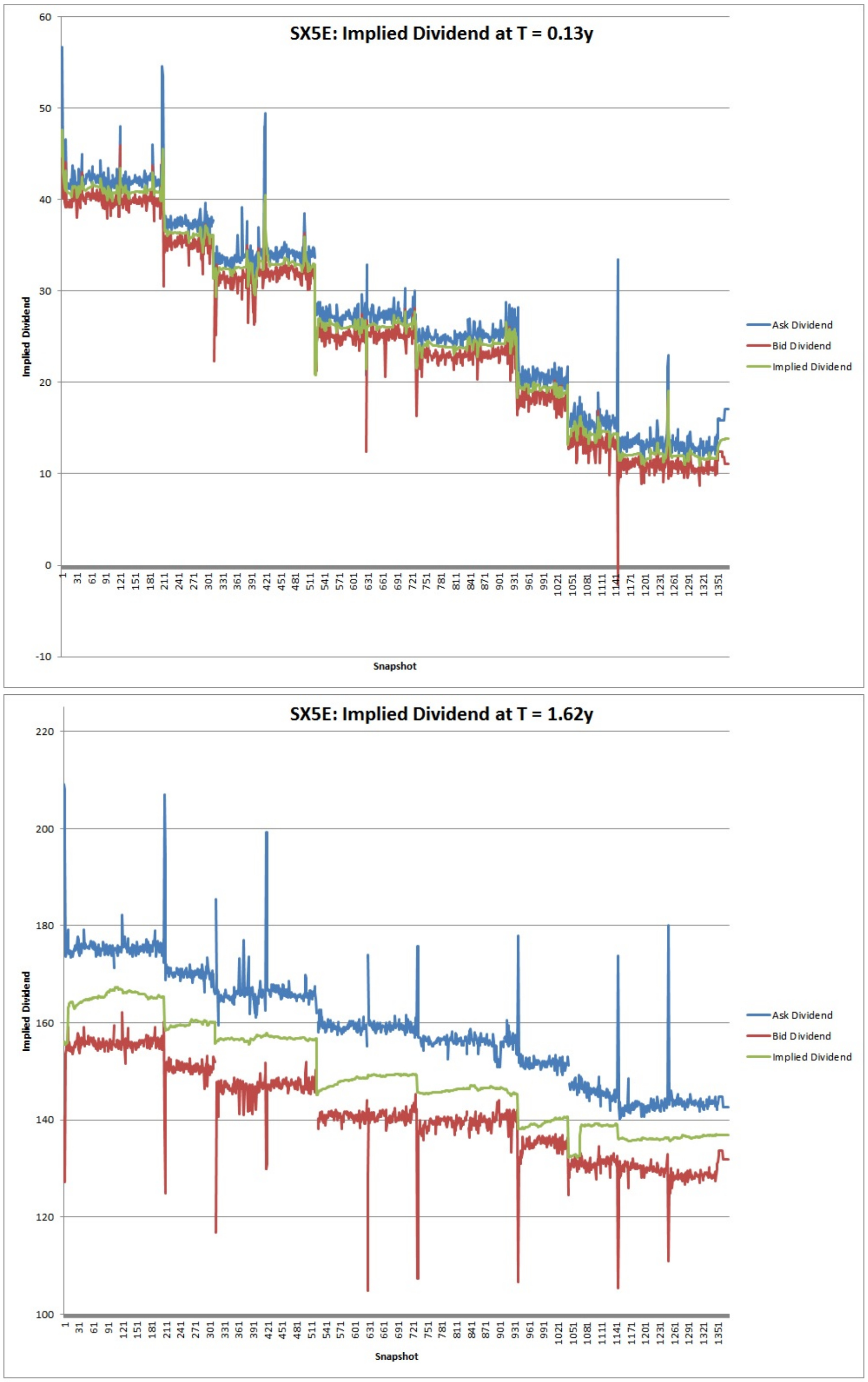


Figure 7: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for the EuroStoxx-50 index.

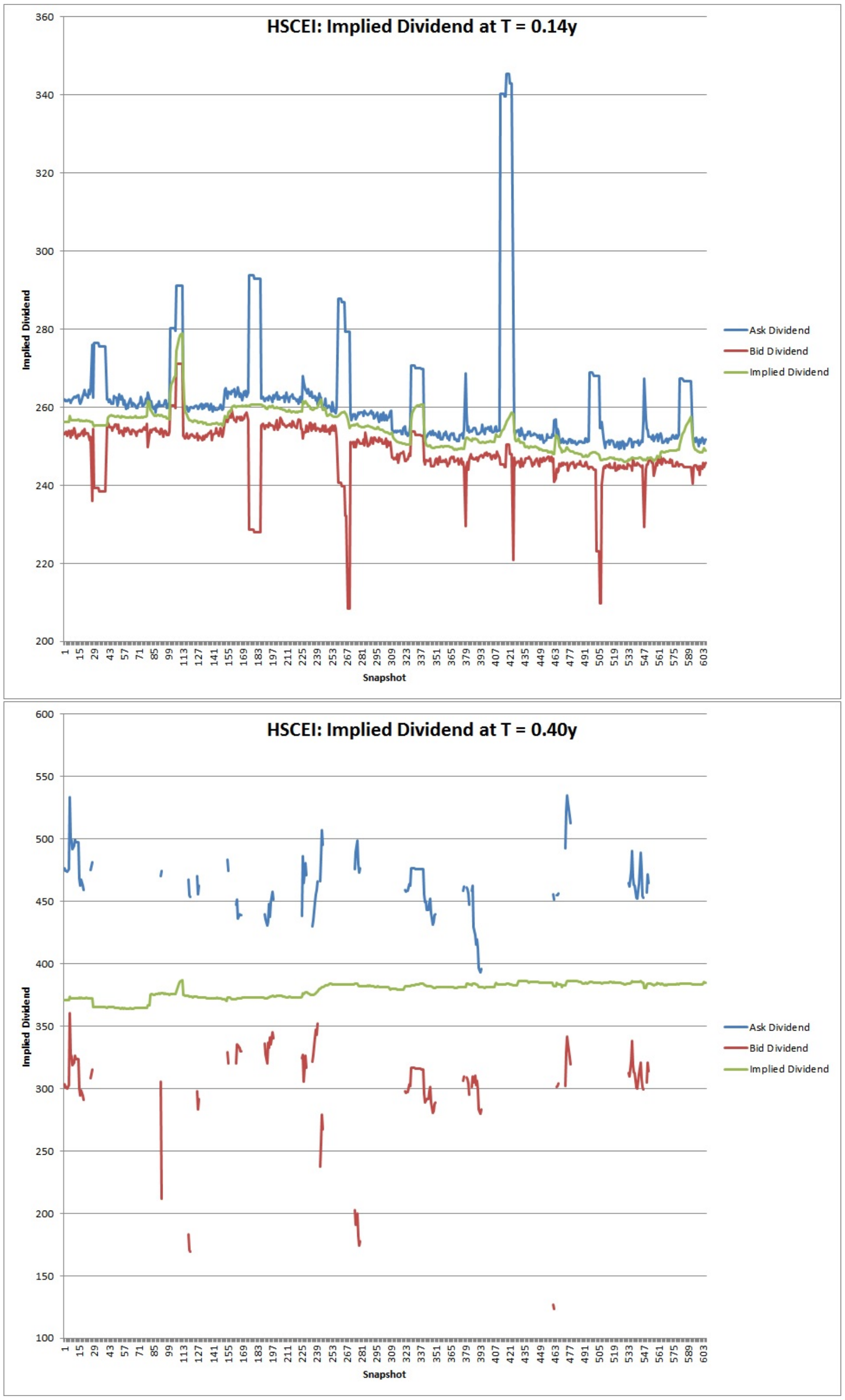


Figure 8: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for the Hang Seng China Enterprises index.

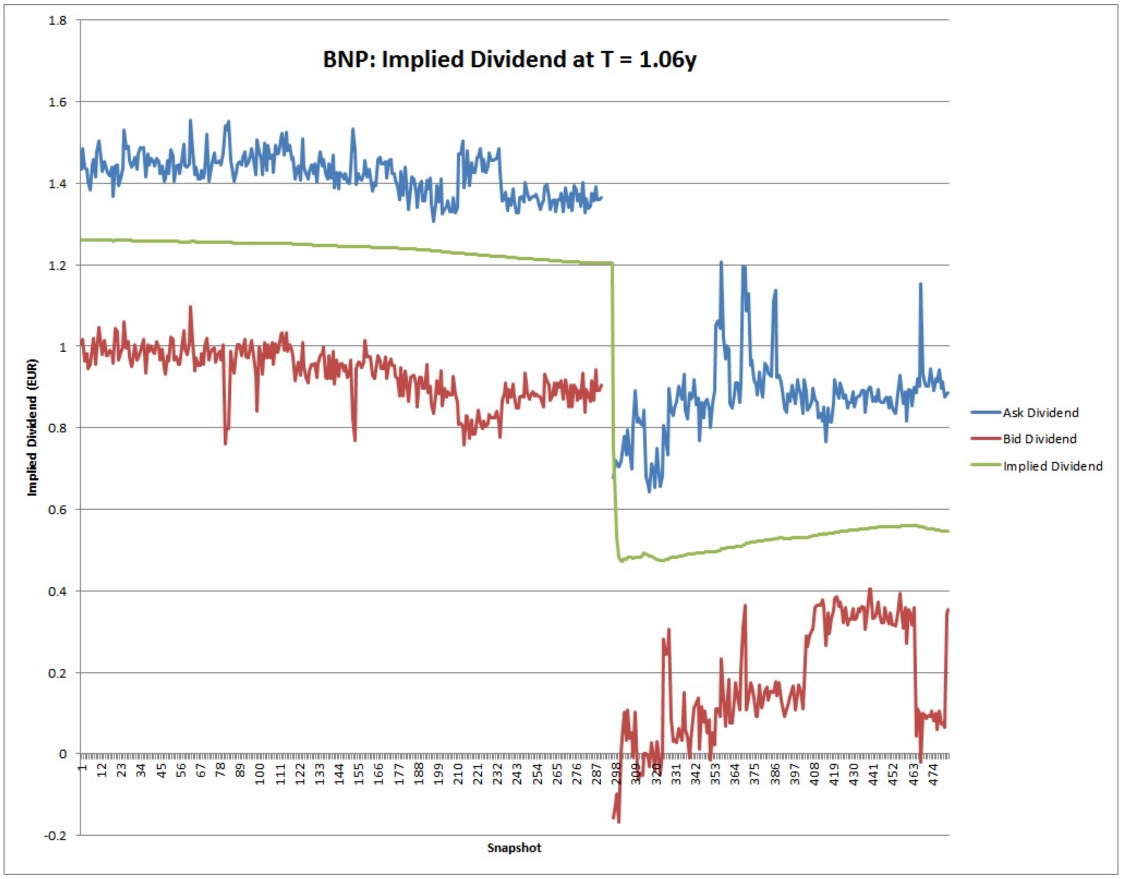


Figure 9: Intraday time series (5-minute snapshots) of smoothed implied dividend, along with option-implied bid/ask bounds, for BNP-Paribas SA.

Part II

Implied Volatility Surface

10 Summary

The implied volatility surface is a fundamental object in the pricing and hedging of derivatives. Our construction of the volatility surface relies on option prices from listed markets, and proceeds by building implied volatility smiles for one listed maturity at a time. Given the implied forward at that maturity, option prices are first filtered to determine which quotes are used in the calibration. A mixture of four lognormal distributions, coupled with a point mass at zero (the "default probability"), is then calibrated to the selected prices. The model offers several advantages:

- the model calibrates well to option prices in the market, with a plausible shape for the implied volatility smile both near the quoted strikes and in the extrapolated regions;
- by construction, the model yields legitimate risk-neutral density, hence an implied volatility smile that is arbitrage-free in strike (although our interpolation scheme in maturity does not guarantee the absence of calendar arbitrage);
- the model also offers fast interpolation of implied volatility with a computational cost comparable to that of the simple Black-Scholes model [BS].

In the sequel, we provide details of the volatility surface calibration, interpolation and extrapolation, and point out some connections between lognormal mixtures and stochastic volatility models.

11 Data Filtering

We determine whether to use bid/ask or settlement data as follows:

1. For each maturity, the bid/ask data at the maturity is deemed to be good if at least 10% of the quotes and at least 3 options have valid bid/ask prices.
2. Use bid/ask data to build the surface if at least 10% of maturities or at least 3 maturities have good bid/ask data; otherwise use exchange settlement data ("last").

The chosen data are then filtered as follows:

1. Filter out maturities that are shorter than 1 day.
2. Eliminate in-the-money options; only out-of-the-money options are used to construct the volatility surface.
3. Eliminate last traded prices older than 3 days.
4. Eliminate "bad" price data with too small bid/ask/last prices or too wide bid/ask spreads.
5. To ensure monotonicity of prices, we take the longest increasing/decreasing subsequence from call/put prices at each maturity, discarding the remaining prices.

6. Prices corresponding to abnormally high or low implied volatilities are eliminated in two stages. First, implied volatilities smaller than one-fifteenth of, or larger than 15 times, the global median implied volatility are eliminated. Second, for each maturity, volatilities smaller than 0.2 times or larger than 5 times the median implied volatility of that maturity are eliminated.

7. We calculate the median of the number of remaining prices at each maturity. Any maturity whose remaining prices number less than 20% of the of the median are entirely eliminated from the surface calibration, and will be interpolated or extrapolated.

12 Reference Volatility Surface

12.1 At-The-Money Treatment

There are cases that no near at-the-money options are available at certain maturities. This is often seen in the last one or two maturities of Hang Seng index and Nikkei 225 index options. If we use only market options in those cases, the resulting smile could have a much higher level then its normal value. To mark the correct smile level, we use a reliable *reference volatility surface* which is known to be good, typically the end-of-day surface of previous business day. For example, assume that there is no option within 90% to 110% forward moneyness at the 2 month maturity and there are at-the-money options at the 1 month maturity. We first get the difference of of at-the-money volatilities at $1M$ between current data and the reference surface. Then we use this difference as the shift amount to adjust the $2M$ near at-the-money volatilities obtained from the reference. At last, we insert those shifted points into the current data to calibrate the model parameters at $2M$. Figure 10 and Figure 11 illustrate the effects using the market opening Hang Seng option data as of 2016-10-05. The reference surface used is the end-of-day 2016-10-04 surface. The 2017-06-29 expiry contains only 5 quoted strikes, all below 70% forward moneyness, with the calibration resulting in an implausible smile shape and ATM term structure. With the reference surface, the smile shape and ATM term structure are much improved.

12.2 Wing Treatment

A calibration is considered successful if the calibrated option prices lie within the bid/ask range at the quoted strikes selected for the calibration. However this criterion is insensitive to wing behavior. In particular, if the valid option prices cover only a narrow range around the forward, a small perturbation of bid/ask prices could lead to a large variation in the wings, leading to large implied volatility jumps in the wings. To minimize this instability, we augment the market quotes with wing samples from the reference volatility surface, adjusting the latter for any observed changes in level or skew from the quoted strike range.

12.3 Coverage

As an auxiliary source of data that is commingled with market quotes in our calibration procedure, the reference surface has to be a trusted surface of high quality. We source the reference volatility

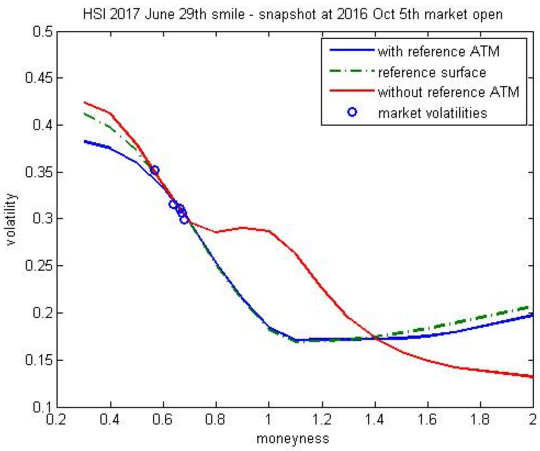


Figure 10: Hang Seng Index 2017-06-29 expiry smile

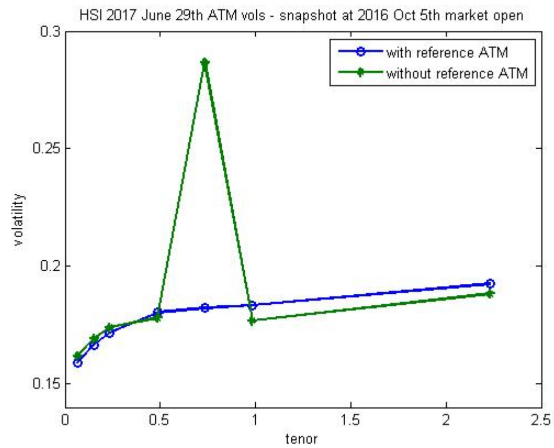


Figure 11: Hang Seng Index at-the-money volatility term structure

surface from broker/dealer quotes of the OTC market. These OTC data presently span the most liquid indices, equities and ETFs, and the coverage is continually expanding as we agument our data contributions.

13 Lognormal Mixture Calibration

In the lognormal mixture model [BM], the risk-neutral density of the spot price at a fixed maturity is modeled as a weighted sum of lognormal densities with different means and variances. Specifically, the risk-neutral probability density function of the stock price at any future time $T > 0$ is assumed to be in the following form

$$\mathsf{pdf}(T, S) = \sum_{l=1}^{N} p_l(T) \cdot \mathsf{lognormalpdf}(S; \xi_l(T)F(T), \Sigma_l(T))$$

where

- N is the number of lognormals,
- $F(T)$ is the forward price,
- $0 \leq p_l(T) \leq 1$ is the time-dependent weight of the l-th lognormal,
- $\xi_l(T) > 0$ is the time-dependent multiplicative means of the l-th lognormal, and
- $\Sigma_l(T) > 0$ is the time-dependent standard deviation of the l-th lognormal.

The multiplicative means $\xi_l(T)$ must reprice the forward, hence

$$\sum_{l} p_l(T)\xi_l(T) = 1.$$

Additionally, the weights $p_l(T)$ must result in a valid probability density, therefore

$$\sum_{i=1}^{N} p_l(T) = 1.$$

In §15 we give an example of the risk-neutral density and implied volatility smile produced by such a lognormal mixture model.

We modify this model by augmenting the state space with a "default" state where the stock price drops to zero. The rationale behind this modification is to facilitate calibration to steep, short-term equity put skews far out of the money by loading a point mass at zero with finite probability. To accommodate this probability of default $Q(T)$, the last condition is modified as

$$Q(T) + \sum_{i=1}^{N} p_l(T) = 1.$$

The number N of lognormal densities in the mixture is a delicate choice. If N is too small, the volatility smile would not have enough flexibility to fit all the relevant market data. If N is too large, the model will overfit the data, resulting in an unstable calibration and large fluctuations in interpolated volatilities. In practice, we choose $N = 4$ to obtain enough flexibility to calibrate tightly to market data, while restricting the parameters by appropriate constraints to obtain a stable calibration. In particular, to mitigate the possibility of calendar arbitrage arising from completely independent calibrations across different market maturities $T_1, \ldots, T_n$, we constrain the term structure of the lognormal parameters: for each $l = 1, \cdots, N$,

$$\begin{aligned} 0 < \Sigma_l(T_1) \quad &< \cdots < \Sigma_l(T_n) \\ 1 \geq \sum_{l=1}^{4} p_l(T_1) &\geq \cdots \geq \sum_{l=1}^{4} p_l(T_n). \end{aligned}$$

The calibration is then formulated as a constrained, nonlinear least squares problem, where the optimal values of the model parameters are determined to match, in the least squares sense, the chosen European option prices at a given maturity. American option prices are de-Americanized as described earlier, and the calibration is performed one maturity at a time.

14 Volatility Interpolation, Extrapolation and Spreads

14.1 Strike Interpolation and Extrapolation

Once the optimal parameter values are determined, the price of an European call/put option at a calibrated maturity T and any strike K is given by

$$C(T, K) = \sum_{l=1}^{N} p_l(T) \cdot \mathsf{blsprice}(\xi_l(T) S_0, K, r, T, \Sigma_l(T)/\sqrt{T}) \tag{2}$$

where blsprice(S, K, r, T, σ) is the Black-Scholes price of a call/put option with spot S, strike K, risk-free rate r, time to expiration T and volatility σ. Thus pricing a European option or interpolating an implied volatility at any strike have a computational cost comparable to that of the Black-Scholes model. Further, the model yields a legitimate probability density, hence the volatilities interpolated or extrapolated in strike are devoid of call-spread and butterfly arbitrage.

Extreme wing strikes When the strike is extremely small or large, the numerical error in the implied volatility calculation is big because the vega is too small. To resolve this numerical difficulty, we extrapolate the implied volatility linearly with respect to the logarithm of forward moneyness $\log(k/F)$ according to Roger Lee's formula [Lee]. For a given maturity, let V_{est} be 3 times of the implied volatility at the forward moneyness. Let d_1 and d_2 be defined as those in the Black-Schoels formula,

$$d_{1,2}(k) = \frac{\log(F/k)}{V_{est}\sqrt{T}} \pm \frac{1}{2} V_{est}\sqrt{T}$$

We then define a minimum strike $k_{\min}$ with $d_2(k_{\min}) = 4.0$ and a maximum strike $k_{\max}$ with $d_1(k_{\max}) = -5.0$. For strikes within this range, we calculate the implied volatilities from the out-of-the-money European option price. For strikes outside this range, we use Lee's formula.

14.2 Maturity Interpolation and Extrapolation

For notational ease, let us define

$$\alpha(t) = \frac{T_{i+1} - t}{T_{i+1} - T_i}$$
$$\eta_l(t) = \log\left(\frac{\xi_{l+1}(t)}{\xi_l(t)}\right).$$

Together with $\sum_l p_l(t)\xi_l(t) = 1$, the $\eta_l(t)$ uniquely determine the $\xi_l(t)$.

Hazard rate damping Assuming a Poisson default process, we can imply the default intensity or hazard rate $\lambda(t)$ consistent with the survival probability $P(t) = 1 - Q(t)$ as

$$P(t) = 1 - Q(t) = \sum_l p_l(t) = \mathrm{e}^{-\lambda(t)\,t} \tag{3}$$

Whereas a positive hazard rate can help match steep put skews at short-dated market maturities, it may force a very steep skew for short-term extrapolation, so some damping of this hazard rate is needed for extremely low strikes.

1. **At a market maturity,** $t = T_i$**:** Define $k_{\min}$ to be the minimum of all market strikes and 90% moneyness, and let $x_{\mathrm{m}} = \log(k_{\min}/F(T_i))$. For all strikes $k < k_{\min}$, let $x = \log(k/F(T_i))$ and the damped hazard rate is defined as

$$\lambda_{\mathrm{new}} = \lambda\, \mathrm{e}^{\frac{x_{\mathrm{m}}^2 - x^2}{2T_i}} \tag{4}$$

For strikes above the threshold ($k \geq k_{\min}$), no hazard rate damping is applied.

2. **Prior to the first market maturity,** $t < T_1$**:** For $k < 0.9\,F(t)$, we first get a λ_{new} in strike dimension as described above for T_1, but with $F(T_1)$ replaced by $F(t)$. Then the damped hazard rate is defined as

$$\hat{\lambda}_{\text{new}} = \lambda_{new}\, \mathrm{e}^{\frac{x^2}{2}\left(\frac{1}{T_0} - \frac{1}{t}\right)}$$

For $k \geq 0.9\,F(t)$, we keep the default probability the same. It effectively sets

$$\hat{\lambda}_{\text{new}} = \lambda_1 \frac{T_1}{t},$$

where λ_1 is the hazard rate at the first market maturity T_1.

3. **Between market maturities,** $T_i < t < T_{i+1}$**:** We first compute a base λ that

$$\lambda = \frac{\lambda_i T_i \alpha + \lambda_{i+1} T_{i+1} (1 - \alpha)}{t}$$

Then we compute k^i_{min} and k^{i+1}_{min} respectively for market maturities T_i and T_{i+1} as described above. Then define $k_{\text{min}} = \alpha(t) k^i_{\text{min}} + (1 - \alpha(t)) k^{i+1}_{\text{min}}$. For all strikes $k < k_{min}$, we use the same formula (4) to get the new hazard rate.

4. For times bigger than the last market maturity $t > T_n$, no hazard rate damping is needed.

Now we can describe interpolation and extrapolation in maturity.

Short-term extrapolation For $t < T_1$, we damp the hazard rate as described above, and calculate the weights $p_l(t)$ by keeping the ratio p_{l+1}/p_l the same as that at T_1. The variance of each lognormal is extrapolated in such a way that the volatility is kept the same, thus $\Sigma_l(t) = \Sigma_l(T_1)\, t/T_1$. The means of each lognormal is extrapolated as

$$\eta_l(t) = \eta_l(T_1) \sqrt{\frac{t}{T_1}}$$

Interpolation between market maturities For $T_i < t < T_{i+1}$, we again damp the hazard rate as described above, and find the corresponding survival probability $P(t)$. Then we set the weights $p_l(t)$ as weighted average of those on the 2 bracketing maturities

$$p_l(t) = \left(\frac{p_l(T_i + 1)}{P(T_{i+1})} \frac{\sqrt{t} - \sqrt{T_i}}{\sqrt{T_{i+1}} - \sqrt{T_i}} + \frac{p_l(T_i)}{P(T_i)} \frac{\sqrt{T_{i+1}} - \sqrt{t}}{\sqrt{T_{i+1}} - \sqrt{T_i}} \right) P(t)$$

The variance is interpolated linearly in time as

$$\Sigma^2_l(t) = (1 - \alpha(t))\, \Sigma^2_l(T_{i+1}) + \alpha(t) \Sigma^2_l(T_i).$$

The means of each lognormal is interpolated by the following formula

$$\eta^2_l(t) = (1 - \alpha(t))\, \eta^2_l(T_{i+1}) + \alpha(t) \eta^2_l(T_i).$$

Long-term extrapolation The sum of the weight $\sum_l p_l(T)$ can be thought as the survival probability. For $t > T_n$, the last market maturity, we bound the survival probability at infinity by half of its value at T_n,

$$P(T) = \frac{1}{2}P(T_n) + \frac{1}{2}\mathrm{e}^{-\lambda_n t} = \frac{1}{2}\mathrm{e}^{-\lambda_n T_n} + \frac{1}{2}\mathrm{e}^{-\lambda_n t}$$

The weights are extrapolated as

$$p_l(t) = p_l\left(T_n\right)\frac{P(t)}{P(T_n)}$$

The volatility of each lognormal process is extrapolated flat for $t > T_n$, so

$$\Sigma_l^2(t) = \Sigma_l^2(T_n)\,\frac{t}{T_n}$$

The lognormal means are extrapolated as

$$\eta_l(t) = \eta_l(T_n)\sqrt{\frac{t}{T_n}}.$$

When there is no historical references on the long term ATM volatility level can be used, after obtaining all these parameters, we perform a final adjustment based on the Heston model. In the Heston stochastic volatility model

$$\begin{aligned} dS(t) &= \mu S(t)\,dt + \sqrt{V(t)}S(t)\,dW_S(t) \\ V(t) &= \kappa(\theta - V(t))\,dt + \xi\sqrt{V(t)}\,dW_V(t) \\ \langle dW_S(t), dW_V(t)\rangle &= \rho\,dt, \end{aligned}$$

the at-the-money (ATM) implied volatility $\sigma_{\mathrm{H}}(t)$ can be approximated [Gat] as

$$\sigma_{\mathrm{H}}^2(t) \approx \theta + (V_0 - \theta)\,\frac{1 - \mathrm{e}^{-\kappa t}}{\kappa t} \tag{5}$$

in terms of the mean reversion speed κ, initial variance V_0 and long term variance θ.

We calibrate κ, V_0 and θ to the ATM implied volatilities at the market maturities. This establishes via (5) an approximate reference level for the ATM volatility. For Heston based ATM vol extrapolation, for $t > T_n$ where the ATM volatilities obtained using $\xi_l(t), p_l(t), \Sigma_l(t)$ are lower than the corresponding $\sigma_{\mathrm{H}}(t)$, we find a new parameter $\beta(t)$ such that the ATM implied volatility from the Black-Scholes call price

$$\hat{C}(t, F(t)) = \sum_{l=1}^{N} p_l(t)\cdot \mathsf{blsprice}(\xi_l(t)S_0, F(t), r, t, \beta(t)\Sigma_l(t)/\sqrt{t}) \tag{6}$$

equals to $\sigma_{\mathrm{H}}(T)$. The extrapolated model parameters are then given by $\xi_l(t), p_l(t)$, and $\beta(t)\Sigma_l(t)$.

In cases with limited quoted maturities, say less than 6 months, the Heston calibration can be sensitive to the short-term ATM volatilities, and this in turn can induce substantial swings in the long-term ATM volatility. Whereas short-term volatilities may fluctuate, long-term volatilities are expected to be stable. For such cases, we stabilize the long-term extrapolation using a *reference surface*, which is typically a reliable surface covering long maturities. For any fixed extrapolated maturity T, we guide the ATM extrapolation $\sigma_{\text{ATM}}(T)$ using the ATM volatilities of the reference surface $\sigma_{\text{refATM}(T)}$ as follows:

$$\sigma_{\text{ATM}}(T) = \sigma_{\text{refATM}}(T) + \Delta\sigma_{\text{ATM}}(T_{\text{last}}) \times \sqrt{\frac{T_{\text{last}}}{T}} \tag{7}$$

where the T_{last} is the last market expiry to which a volatility smile was calibrated, T is any extrapolated tenor after the last market option expiry, $\Delta\sigma_{\text{ATM}}(T_{\text{last}})$ is the difference between the market ATM volatility and the reference ATM volatility at T_{last}.

Once $\sigma_{\text{ATM}}(T)$ has been determined, we need to adjust the extrapolated model parameters $\xi_l(t), p_l(t)$, and $\beta(t)\Sigma_l(t)$ obtained above need to be adjusted to match this ATM level. This is first attempted by adjusting $\beta(T)$, which scales the volatilities of each lognormal process. This produces mainly a shift in the smile level, and a relatively smaller change in the smile shape. If this procedure succeeds, it generally produces a good smile extrapolation resembling the market smile from T_{last}.

If this procedure fails, we determine the ATM volatility $\sigma(T)$ corresponding to the parameters $\xi_l(t), p_l(t)$, and $\Sigma_l(t)$, and scale each $\Sigma_l(T)$ by the ratio $\sigma_{\text{ATM}}(T)/\sigma(T)$. Then we adjust the hazard rate $\lambda(T)$ until the ATM volatility matches the target value $\sigma_{\text{ATM}}(T)$. If this procedure succeeds, we will get a smile effect in the extrapolation, although not necessarily similar to the one at T_{last}.

If this procedure fails, we adjust the relative distance between the lognormal means by scaling $\eta_l(T)$ until the target value $\sigma_{\text{ATM}}(T)$ is matched. As the $\eta_l(T)$ approach zero, the lognormal mixture degenerates to a pure lognormal density, whose volatility can be adjusted to match the target value $\sigma_{\text{ATM}}(T)$, albeit at the cost of flattening the extrapolated smile.

14.3 Volatility Spreads

Once the mid implied volatility is calculated as described above, bid/ask volatility spreads are interpolated linearly and extrapolated flat in strike. In maturity, we use Hermite interpolation and flat extrapolation of the bid/ask volatility spread.

15 Numerical Results

We illustrate in Figure 12 the key differences between a lognormal mixture (blue) of two components with different means and variance and an equivalent lognormal density (red) with the same mean and variance as the lognormal mixture. The lognormal mixture has a higher mode and heavier tails, resulting in a significantly different implied volatility smile.

The remaining plots show examples of volatility surface calibrations of the lognormal mixture model with 4 lognormal components augmented by a default state for an equity (IBM US Equity) and an index (SPX Index). In Figure 13 and Figure 14, we show calibration errors via implied volatility smiles at listed market maturities ranging from around 1 month to 18 months. Figure 15 shows the ATM implied volatility extrapolation based on the Heston reference level. Finally, Figure 16 shows the entire volatility surface.

References

[BS] F. Black and M. Scholes. The pricing of options and corporate liabilities, *Journal of Political Economy* 81: 637-659.

[BM] D. Brigo and F. Mercurio (2002), Lognormal-Mixture Dynamics and Calibration to Market Volatility Smiles, *International Journal of Theoretical & Applied Finance* 5(4), 427-446.

[Gat] J. Gatheral. *The Volatility Surface.* John Wiley & Sons, Inc., 2006

[Lee] R. Lee (2004), The Moment Formula for Implied Volatility at Extreme Strikes, *Mathematical Finance* 14(3): 469-480

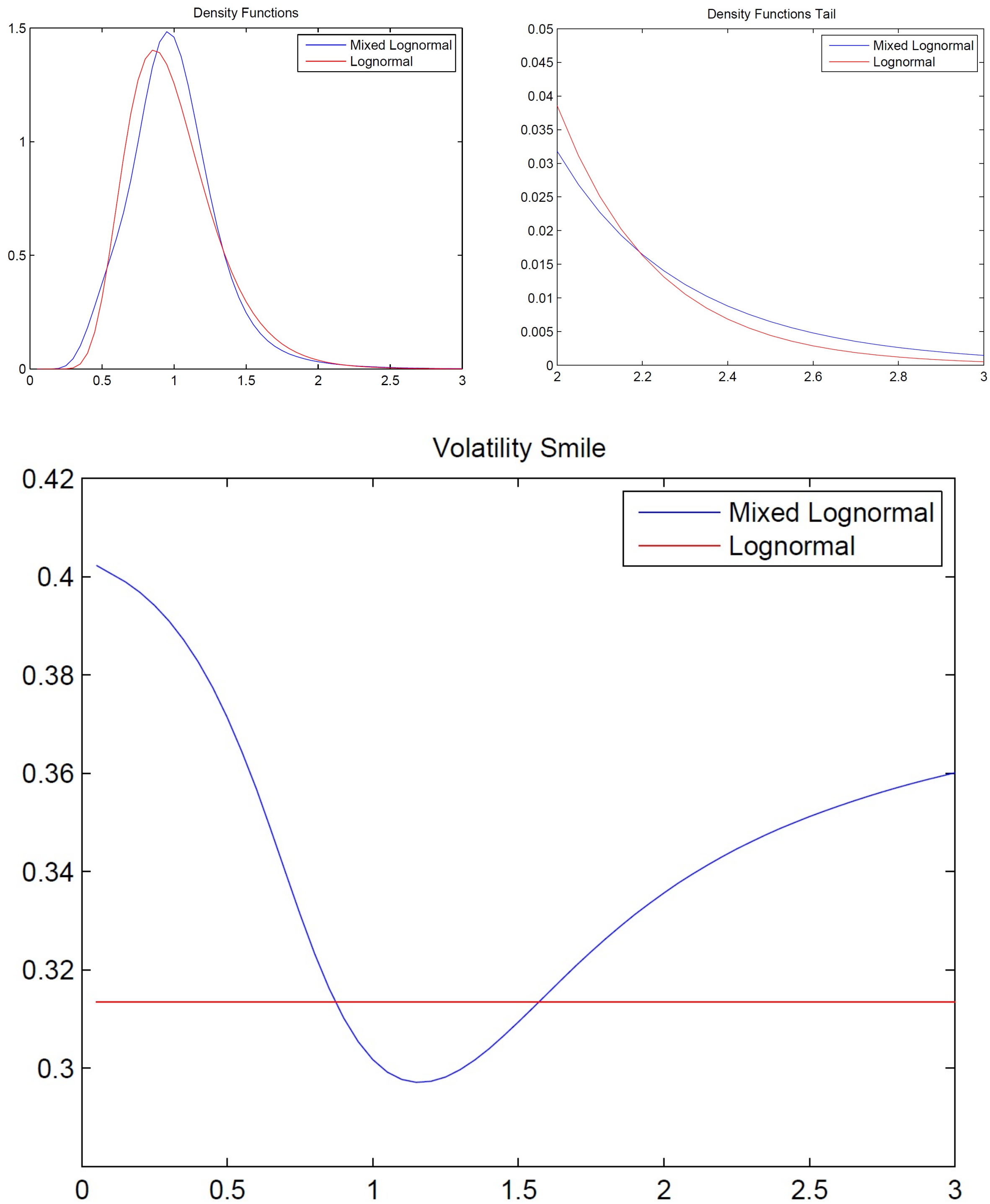


Figure 12: Top panel: Density functions corresponding to a mixture of two lognormals (blue) and one equivalent lognormal (red). $N = 2$, $\xi_1 = 0.95$, $\xi_2 = 1.05$, $\sigma_1 = 0.40$, $\sigma_2 = 0.20$, $p_1 = p_2 = 0.5$, $T = 1$, $F = 1$. The mean and standard deviation of the single lognormal are chosen to match those of the lognormal mixture. Bottom panel: Implied volatility smiles produced by the lognormal mixture (blue) and the equivalent lognormal (red).

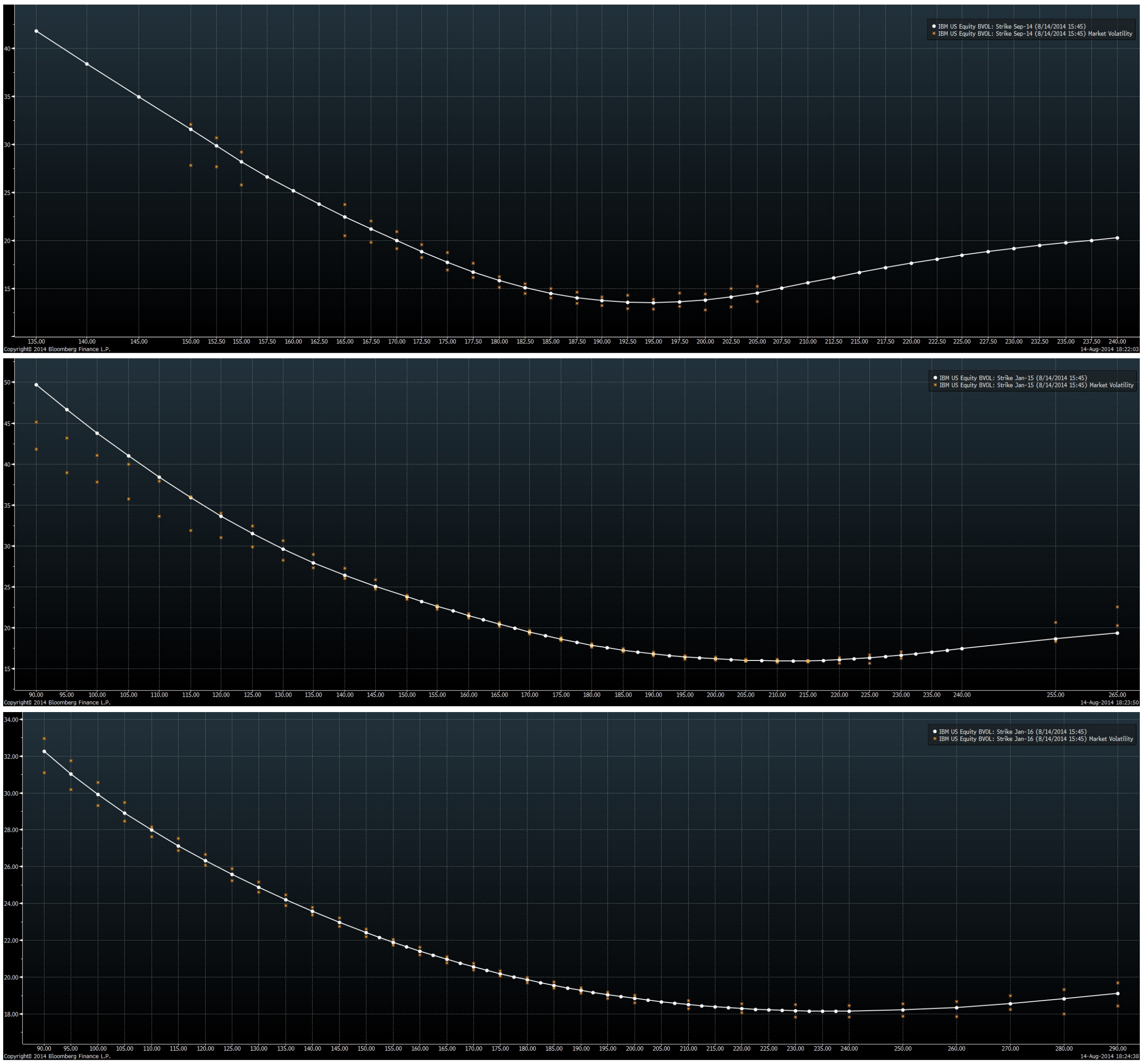


Figure 13: Volatility smile and calibration errors for IBM for the Sep-2014, Jan-2015 and Jan-2016 expiries using the lognormal mixture model with 4 components augmented by a default state. The model implied volatility (solid while) is compared with the market bid/ask implied volatilities (yellow stars).

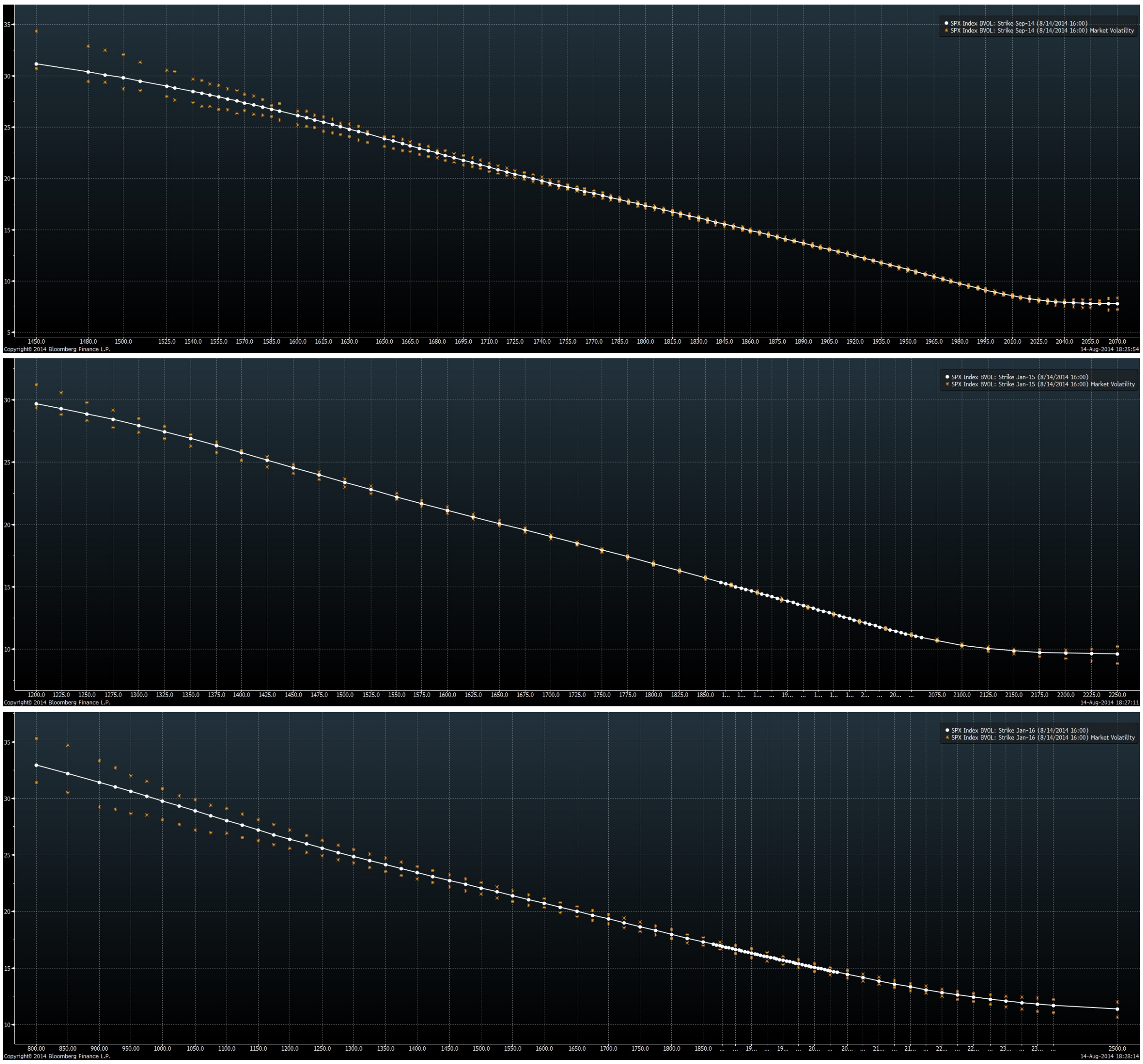
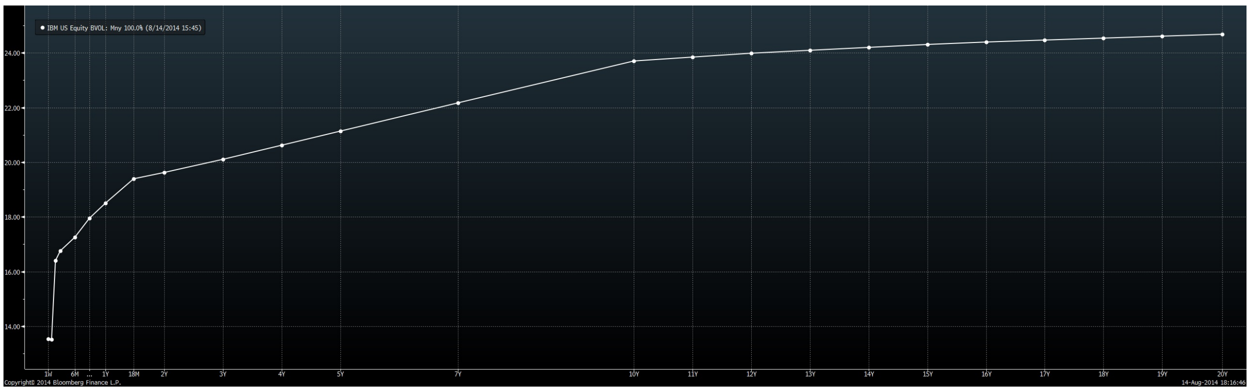


Figure 14: Volatility smile and calibration errors for the S&P 500 index for the Sep-2014, Jan-2015 and Jan-2016 expiries using the lognormal mixture model with 4 components augmented by a default state. The model implied volatility (solid while) is compared with the market bid/ask implied volatilities (yellow stars).



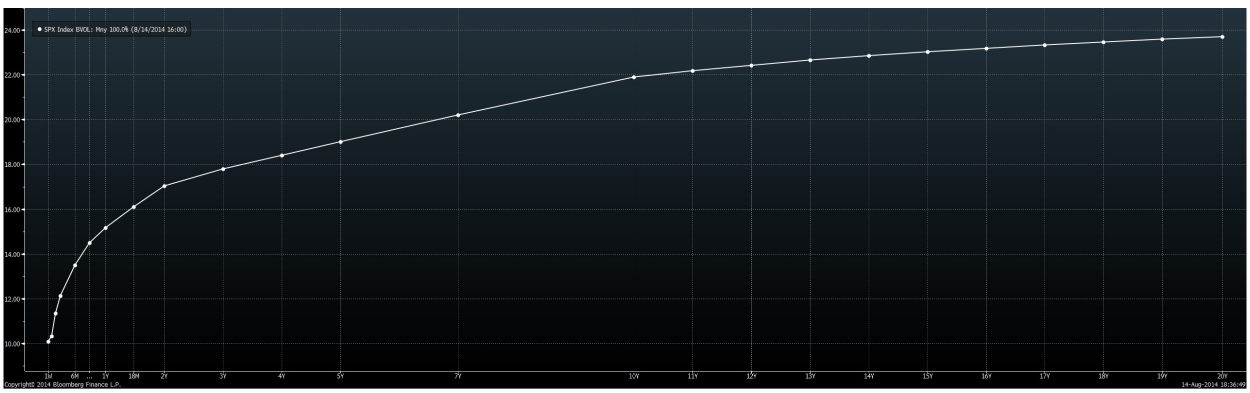
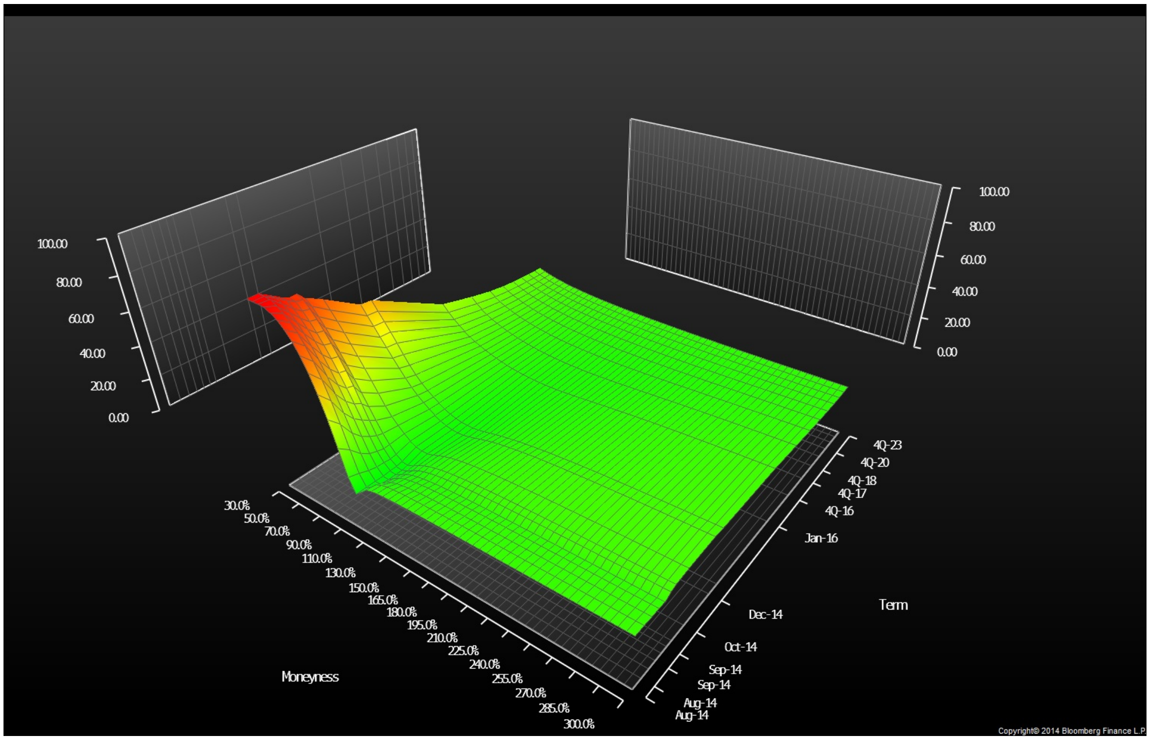


Figure 15: ATM volatility extrapolation using the Heston reference level for IBM (top) and S&P 500 (bottom).



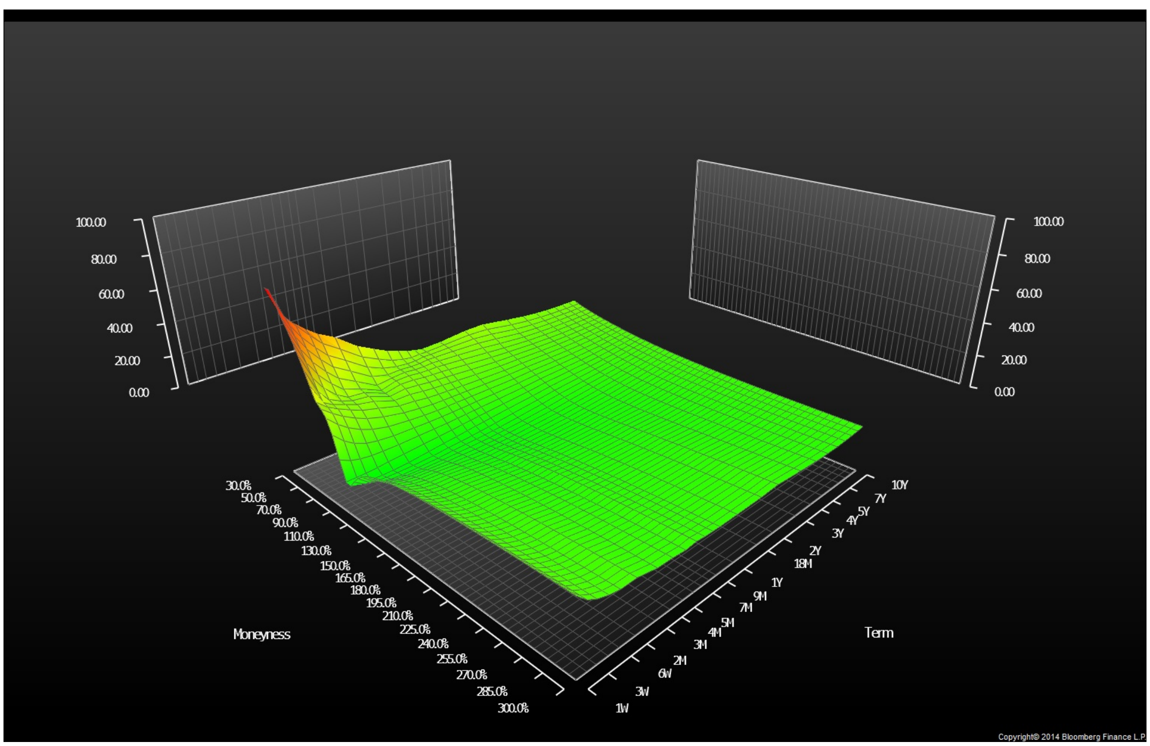


Figure 16: Volatility surface for IBM (top) and S&P 500 (bottom).