Pricing Variance Swaps with Cash Dividends

Timothy Klassen

Abstract

We derive a simple formula for the price of a variance swap when the underlying has cash dividends.

1 Introduction

The last years have seen renewed interest in the modeling of dividends for equity derivatives. This is partially due to the general increase in sophistication and maturity of the market, partially due to the increased dividend yields in the wake of the bursting of the dotcom bubble. More recently, the deflation of the credit/housing bubble led to an increased awareness of dividend risk.

There are two main questions relevant for dividend modeling: One is the proper mix of “sticky” (cash-like) and proportional (yield-like) dividends one should assume, the other is what exactly a suitable process for the stock price would be that incorporates the dividend assumption.

Note that without answering both of these questions one can not even begin to talk about inferring some “implied volatility” from a market price for a vanilla option. Note also that this question is just as relevant for indices as for single stocks – just because the dividends of an index come in frequent small amounts does not mean that they can be treated as yield-like.

As far as vanilla option prices are concerned, one could argue that dividend (and volatility) modeling is effectively just an interpolation/extrapolation algorithm between liquid and less liquid parts of the market. To some extent this is true, but even then, improper dividend modeling can lead to problems when extrapolating from the listed to the OTC market (which might also mean a switch from American to European style exercise). Furthermore, the value of greeks depend on the dividend modeling assumption, even when different models are perfectly calibrated to the same prices. Perhaps most importantly, exotic structures can have very significant sensitivity to the dividend model (see eg [1]). This is even true for “mild” exotics like dividend or variance swaps.

The purpose of this note is to discuss this issue for variance swaps. The standard model used in the market follows Derman et al [2] and relies on the now well-known “option strip formula” that (essentially) statically replicates the variance swap in terms of vanillas. Its derivation, which will be reviewed as part of the generalization below, assumes that all dividends are yield-like.

In practice, even though the option strip formula is the starting point, one finds that variance swaps always seem to trade below the option strip prices. Various recipes are used by different dealers to cut off the option strip to match the market. Practitioners have been debating for some time whether this is purely a supply-demand issue, or whether there are other, more fundamental, contributing factors.

This motivated us to explore the effect of cash dividends on variance swap pricing. We present a simple formula for the fair value of a variance swap in terms of market prices of vanillas and
dividends. The only other published research in this direction we are aware of is [3], where a rough approximate formula is given for the case of one cash dividend and assuming rates, skews and volatility term-structure are weak. We consider completely general dividends, rates and volatility skews and derive an exact formula to first order in the (effective) dividend yield. We start by laying out our notation and assumptions.

2 Dividends, Forwards, and SDEs

Relative to today, \( t_0 = 0 \), let us denote by \( t_i > 0 \) the ex-dividend dates on which the underlier \( S_t \) drops by some amount \( D_i \). For a given time horizon \( T \) we will index the dividends as \( i = 1, \ldots, N_T \). The dividend can be a mix of a cash amount, \( d_i \), and a proportional amount \( \delta_i S_{t^{-}} \), viz \( D_i = d_i + \delta_i S_{t^{-}} \). Here \( t^{-} \) denotes an infinitesimal instant before time \( t \) (at \( t = t_i \) the dividend jump has already occurred). We also allow an arbitrary continuous dividend yield via the drift \( \mu_t \) of the underlier, which one can think of as \( \mu_t = r_t - q_t \), where \( r_t \) is the discount rate and \( q_t \) the dividend yield (any funding spread can be absorbed in \( q_t \)).

Concerning the first question of the introduction, the simplest reasonable assumption about the mix of cash and proportional dividends is to assume a smooth (or stepwise) transition from pure cash dividends, \( d_i > 0, \delta_i = 0 \), to purely proportional dividends \( d_i = 0, \delta_i > 0 \). This simply expresses the intuition that dividends tend to be sticky in the short term, but in the long run it is more reasonable to assume that they go up or down with the stock price.

This assumption seems sensible and is simple enough to allow for an analytic expression for forward prices \( F_t = E[S_t] \). Namely, noting that between dividend dates the forward grows with rate \( \mu_t \), and at a dividend date drops by a factor \( 1 - \delta_i \) and a cash amount \( d_i \) we can write:

\[
F_{t_i} = \exp\left(\int_{t_{i-1}}^{t_i} \mu_t \, dt\right) F_{t_{i-1}} (1 - \delta_i) - d_i
\]

This recursion has the solution

\[
F_T = f_p(T) \left( S_0 - \sum_{i : t_i \leq T} \frac{d_i}{f_p(t_i)} \right)
\]

where we introduced the proportional forward factor

\[
f_p(t) := \exp\left(\int_0^t \mu_t \, dt\right) \prod_{i : t_i \leq t} (1 - \delta_i)
\]

which is the proper discount (growth) factor to use for the cash-dividends (stock). Note that it is also the delta of the forward.

The time scale(s) of the cash-to-proportional dividend transition can be estimated either from historical data about stock prices and dividends, or by observing how forwards of different maturities move when the spot \( S_0 \) moves.

The next question is how exactly the stock evolves. There seem to be three main models in use among practitioners. We can present them in a somewhat unified manner as follows. Write the stock price as

\[
S_t = (F_t - \Delta_t) X_t + \Delta_t,
\]
in terms of a positive martingale $X_t$ with expected value $E[X_t] = 1$. Here $\Delta_t$ is a dividend-dependent shift, and its precise form is what distinguishes the three models. Note that by construction the above will always lead to the correct forward, $E[S_t] = F_t$.

- Partial Hybrid Model [4]: The idea here is to consider the stock as consisting of a riskless piece $\Delta_t$, the dividends up to maturity $T$, and the fluctuating remainder:

$$\Delta_t = \Delta_{t,T} := f_p(t) \sum_{i:t < t_i < T} \frac{d_i}{f_p(t_i)}$$

Note that $\Delta_T = 0$, whereas $\Delta_0$ is the full discounted dividend sum up to maturity, which is why this model is sometimes referred to as the spot-adjustment model.

- Full Hybrid Model [5]: In the partial hybrid model the shift $\Delta_t$ is actually maturity-dependent, which means that (4) does not really describe one consistent process; instead it describes a separate stock process for each maturity. A consistent model, as described in [5], is obtained by letting the shift contain all dividends from time $t$ to infinity:

$$\Delta_t = \Delta_{t,\infty} = f_p(t) \sum_{i:t_i > t} \frac{d_i}{f_p(t_i)}$$

In this model one can really think of the stock $S_t$ as a hybrid of a cash component, the discounted value of all cash dividends from $t$ to infinity (the cash dividend stream therefore has to be cut off at some point in the future), and the fluctuating remainder.

- Spot Model [1, 6]: This model is defined by the SDE

$$dS_t = \mu_t S_t dt - \sum_i (d_i + \delta_i S_{t_i^-}) \delta(t - t_i) + \sigma_t S_t dW_t$$

(5)

where $\delta(t)$ is Dirac’s delta function. As pointed out in [6], an accurate approximate solution of the European vanilla pricing equation for this SDE can be written in an intuitive manner in terms of a spot and strike adjustment to the standard Black-Scholes formula (viz, subtracting the “near-term dividends” from the spot, and adding the “far-term dividends” to the strike). Equivalently, in terms of eqn (4), this approximate solution corresponds to

$$\Delta_t \equiv -\delta K_t, \quad \delta K_t = f_p(t) \sum_{i:t_i \leq t} \frac{t_i d_i}{t f_p(t_i)}$$

where $\delta K_t$ is the aforementioned strike adjustment (aka the “far-term dividends”) for time $t$.

All three models satisfy the desired property that the stock drops precisely by the cash dividend amount across a dividend date, $S_{t_i^-} - S_{t_i} = \Delta_{t_i} - \Delta_{t_i^-} = F_{t_i} - F_{t_i^-} = d_i$.

Otherwise, each of these models has some theoretical or practical quirk, which perhaps explains why none of them is universally preferred. The partial hybrid model seems to be popular among, for example, options market makers (or at least used to be in the past), but since it does not correspond to one consistent stochastic process for all maturities, it is difficult to use in an integrated vanilla and exotic business where such consistency is important.

The other two models do not have this problem. The full hybrid model is a consistent and sensible approach. It does require one to maintain a finite time scale for the transition from cash
to proportional dividends, otherwise the $\Delta_t$ can get very large, leading to large implied volatilities (since they only apply to the piece $S_t - \Delta_t$). However, even with reasonable dividend assumptions can implied volatilities (and deltas) be significantly different from the more Black-Scholes-like values of the other models, which might make some people uncomfortable. One might also be concerned that in this model implied volatilities for a given maturity depend on what one assumes about the cash-dividend stream after maturity. This is unlike the spot model, but fully consistent with the logic of the hybrid model.

The spot model as defined above has the issue that there is a non-zero probability for the stock to go negative at some dividend date. In practice this is not much of a problem [6], except for very large maturities, if the cash dividend stream is not suitably truncated. One could modify the SDE to floor the stock at zero, which one will effectively do anyhow if one uses a numerical method (finite difference or tree) to solve the pricing PDE rather than use the spot-strike-adjustment approximation. In this modified version, or in the spot-strike adjustment approximation, the spot model is perhaps the most widely used approach.

For the purpose of pricing variance swaps with cash dividend the spot model is the most convenient. Compared to the hybrid models, one can easily obtain analytical results, as we now describe.

3 Pricing Variance Swaps

The price of a (newly issued) variance swap is simply related to the expected value of the total variance up to some maturity $T$, defined by

$$w_T := \sum_j \ln^2 \left( \frac{S_{j+1}}{S_j} \right)$$

Here the sum is over all sampling dates (we reserve the index ‘$i$’ for sums over dividend dates). In the usual continuum limit of frequent sampling (for a fixed physical time $T$), we have

$$w_T \approx \int_0^T \sigma_t^2 \, dt + \sum_{i=1}^{N_T} \ln^2 \left( \frac{S_{t_i}}{S_{t_i^-}} \right)$$

in terms of the (possibly stock-dependent and/or otherwise stochastic) local volatility $\sigma_t$ of $S_t$.

The above holds up to terms down by factor of $O(\Delta T)$, in terms of the sampling time interval $\Delta t$, relative to the ones shown. Note that the second term would be absent if the log-return in (6) were dividend-corrected, ie the numerator replaced by $S_{j+1} + D_{j+1}$. In practice this is in fact usually done for single stock variance swaps, but not for indices. In any case, this term is quadratic in the effective dividend yield, and we will only work to first order in this yield, so we neglect it here (its precise contribution is easy to work out along the lines below; it is numerically small in realistic cases).

From Ito’s lemma for processes with jumps we have

$$d \ln S_t = \frac{dS_t^{(c)}}{S_t^-} - \frac{1}{2} \left( \frac{dS_t^{(c)}}{S_t^-} \right)^2 + \sum_i \delta(t - t_i) \ln \left( \frac{S_t}{S_{t_i^-}} \right) \, dt$$

1Then, however, the expression for the forward, eqn (2), does not hold exactly anymore for large maturities. With suitably (realistically) truncated cash dividends it will still be quite accurate, though.
where $S^c_i$ is the continuous part of the stock process. In the spot model, generalized to allow $\sigma_i$ in (5) to be spot-dependent and/or stochastic, this implies

$$E[\ln(S_T/S_0)] = \int_0^T \mu_t \, dt - \frac{1}{2} \int_0^T \sigma_t^2 \, dt + \sum_{i=1}^{N_T} E[\ln(1 - D_i/S_{t_i^-})]$$

or, to the order we are working,

$$\frac{1}{2} E[w_T] = \frac{1}{2} E[\int_0^T \sigma_t^2 \, dt] = \int_0^T \mu_t \, dt - E[\ln(S_T/S_0)] + \sum_{i=1}^{N_T} E[\ln(1 - D_i/S_{t_i^-})] \quad (8)$$

The first term on the right-hand side (rhs) can be calculated from eqn (1) as

$$\int_0^T \mu_t \, dt = \sum_{i=1}^{N_T} \ln \frac{F(t_i) + d_i}{F(t_{i-1}) (1 - \delta_i)} + \ln \frac{F_T}{F_{t_N}} \quad (9)$$

Noting that $1 - D_i/S_{t_i^-} = (1 - \delta_i)S_{t_i}/(S_{t_i} + d_i)$, we see that the $1 - \delta_i$ factors in the first and last term on the rhs of (8) will cancel, and we are left to calculate $E[\ln(S_T/S_0)]$ and a sum of terms of the form $E[\ln(S_{t_i}/(S_{t_i} + d_i))]$.

To this end recall a general result for the expected value of a (generalized) function of the stock price,

$$E[f(S_T)] = f(S_*) + (F_T - S_*)f'(S_*) + \int_0^\infty f''(K) \hat{V}_{\text{OTM}}(T, K) \, dK \quad (10)$$

Here $S_*$ is a reference spot that can, in principle be chosen to take any value. $f'$, $f''$ denote the first, respectively, second derivative of $f$, and $\hat{V}_{\text{OTM}}(T, K)$ denotes the un-discounted value of the out-of-the-money vanilla option (with respect to $S_*$) of maturity $T$ and strike $K$. In other words,

$$\hat{V}_{\text{OTM}}(T, K) = \begin{cases} \hat{P}(T, K) & \text{if } K < S_* \\ \hat{C}(T, K) & \text{if } K > S_* \end{cases}$$

in terms of un-discounted call and put prices, $\hat{C}$ respectively $\hat{P}$.

The above is easily proved after expressing the expectation value in terms of the implied density. It follows after a couple of partial integrations, that can be performed under mild assumptions about the decay of the density in the wings and the function $f$ (chiefly that $f$ is continuous at $S_T = S_*$, but the above formula can be generalized to the case where even this does not hold).

The most convenient choice here is to take $S_*$ to be the forward of the appropriate maturity, so that the second term on the rhs of (10) vanishes. Putting everything together, using that for $f(S) = \ln(S/(S + d))$ we have $f''(S) = -1/S^2 + 1/(S + d)^2 \approx -2d/S^3$ as a legitimate approximation to the first order in $d/S$ we are working, as well as noting the cancellation of the terms involving the forwards on the rhs of (8), we finally get simply

$$\frac{1}{2} E[w_T] = \int_0^\infty \hat{V}_{\text{OTM}}(T, K) \frac{dK}{K^2} - 2 \sum_{i=1}^{N_T} \int_0^\infty \hat{V}_{\text{OTM}}(t_i, K) \frac{d_i \, dK}{K^2} + O((d/S_0)^2) \quad (11)$$

This generalizes the results of [2] where the second term is absent, while still expressing everything in terms of observable vanillas prices and forwards (up to strike-maturity interpolation/extrapolation issues), and the cash-dividend stream. In addition to the standard option strip at maturity, it now
Table 1: Numerical results for fair volatilities $\sigma_f$ for various maturities $T$, assuming a flat implied volatility of 20%, a $0.02$ dividend paid every 0.01 years (with the first one at 0.005 years) and forwards of $100$ at all dividend dates. The third column should be $-q_{\text{eff}}$ when the approximation $\sigma_f \approx \sigma_{f,0} \left(1 - \frac{1}{2} q_{\text{eff}} T \right)$ is accurate. The last column is the number of standard deviations (in terms of the ATM vol; chosen symmetrically on the call and put side) one has to use in the no-dividend formula to match the cash-dividend fair volatility.

![Table 1](image.png)

Involves a strip for each cash-dividend date. Note that the continuous and discrete proportional dividends do not explicitly appear in the final result (to the order we are working).

Market quotes for variance swap prices are usually expressed in terms of a “fair volatility” $\sigma_f$ (aka “fair strike”), defined by $E[w_T] = \sigma_f^2 T$. In contrast to the convention-independent total variance, it requires a choice of time-convention. The market convention is “business time”; we assume ‘$T$’ has been chosen appropriately.

To gain some intuition, note the following. Each of the integrands in (11) is peaked around the forward of the relevant maturity (roughly, at least for small maturities). The integrand is not perfectly symmetric around this peak, but at least approximately so if the skew is weak. If there is just one cash dividend $d$ just before maturity, we have $\sigma_f \approx \sigma_{f,0} \left(1 - \frac{d}{F_T} \right)$, where $\sigma_{f,0}$ denotes the fair volatility with no cash dividends. If there is just one dividend at time $t$ before maturity, and rates, skews and term-structure are weak, then $\sigma_f \approx \sigma_{f,0} \left(1 - \frac{t}{T} \frac{d}{F_t} \right)$ (cf. [3]).

As a final example, if there is a large number of cash-dividends and they are roughly equally distributed in time, then under the same assumptions we have $\sigma_f \approx \sigma_{f,0} \left(1 - \frac{1}{2} q_{\text{eff}} T \right)$, in terms of the effective dividend yield of the cash-dividends, $q_{\text{eff}}$. As seen in table 1, this approximation is pretty accurate in the flat case for equally-spaced dividends.

For any real-life application one should of course use the exact (to first-order in $q_{\text{eff}}$) eqn (11).\footnote{Perhaps with some “bundling” of the cash dividends for indices (while maintaining the forwards at the relevant maturities), if one wants to avoid doing too many one-dimensional integrals.} Note that a larger skew will increase the magnitude of the correction term, especially for larger maturities. All the usual remarks about how to value variance swaps on the run, greeks, etc can be applied to (11).

Finally, a remark relevant for the market practice of “cutting off the option strip”. In table 1 we also show, for $q_{\text{eff}} = 2\%$, roughly the yield of the SPX, the implied number of standard deviations $n_0$ one has to use in the no-dividend formula to match (11). These numbers are not far from typical market-implied values for the SPX. They are quite insensitive to the overall level of volatility, but with nonzero skew and rates they change somewhat, of course. A detailed study is beyond the scope of this note (since it requires a discussion of how to extrapolate volatilities in the wings, precise dividend streams, how to transition from cash to proportional dividends, effect of stochastic interest rates for very large maturities, etc), and is left for future work.
4 Conclusion

We presented a simple formula, eqn (11), for the value of a variance swap when the underlier has cash dividends. It is exact to first order in the cash-dividend yield. The effect of cash dividends is sizable and goes at least some way towards explaining (and perhaps obviating) the market practice of maintaining a term-structure of implied standard deviations at which to cut off the option strip in the standard zero-dividend formula.

Acknowledgement: I would like to thank Hans Buehler and Jim Gatheral for discussions.

References