

Pricing derivatives if trading in the underlying is limited: Froot and Stein, and Merton '98

As one of various suggested applications of their model, Froot and Stein describe the pricing of derivatives on assets that are not (cost-effectively) tradable.¹ They derive a partial differential equation for the value a bank would assign to an option on an asset that cannot be traded easily (e.g. due to lack of liquidity or short-selling constraints), but for which observable market prices are available. The derivation is as follows:

The bank assumes that the observable market price S of the underlying nontraded asset (e.g. an illiquid share of a common stock) follows a geometric Brownian motion (lognormal diffusion process) with drift (instantaneous expected return) μ_S and volatility σ_S :

$$dS = \mu_S S dt + \sigma_S S dz$$

As the price of the untraded asset is observable (and consistent with the assumed pricing model for tradable assets), the option price is a continuous and twice differentiable function of time t and the underlying asset price², so Ito's lemma³ can be applied on this process – and subsequent taking of expectations gives an expression for the expected change of the option value F :

¹ page 73 in: Risk Management, Capital Budgeting, and Capital Structure Policy for Financial Institutions: An Integrated Approach, Kenneth A. Froot and Jeremy C. Stein, The Journal of Financial Economics, 1998, no. 47, 55-82.

² That the option value is a twice differentiable function of underlying price and time (and hence Ito's lemma is applicable) is a derived conclusion – the derivation for traded assets is given in Robert C. Merton '77 – On the Pricing of Contingent Claims and the Modigliani Miller Theorem, Journal of Financial Economics, Vol. 5, 1977, p. 241-249. If trading in the underlying is not possible, Merton argues in Applications of Option Pricing Theory: 25 Years Later, American Economic Review, Vol. 88, Issue 3, June 1998, p. 323 – 349, that one can still define a derivative value function as if the underlying was tradable, and a hedging strategy that approximates this function by dynamically allocating funds between a portfolio of tradable risky assets to track the underlying, and the riskless asset. For the hedging strategy that approximates the derivative value function best, the weights of the assets in the tracking portfolio are chosen so that the variance of the instantaneous tracking error, i.e. the difference between the instantaneous return of the tracking portfolio and the untradeable asset, is minimized. This implies that the instantaneous tracking error is uncorrelated with any traded asset's instantaneous return, as Merton has shown in his book Continuous Time Finance – in the revised edition of this book the proof can be found on page 396 within the section on theorem 15.3. An alternative proof is given below, see pages 5-6. It follows that the tracking error is also uncorrelated with the market portfolio (or any priced factor), and therefore, Merton states, on markets where only systematic risk is priced (as under the Froot and Stein pricing model assumption for traded assets used here – see the earlier posts in this blog and references given there for details), the price of the derivative will be the same as if the underlying asset was traded – see Merton '98, page 333. For additional, informal intuition why this is the case, note that with continuous rebalancing, the probability distribution of the underlying return over any instantaneous period is the same as the probability distribution of the tracking portfolio return minus the tracking error. The underlying can hence be replicated with the tracking portfolio minus an additional random payment series corresponding to the tracking error at the end of each instantaneous period – which as the latter constitutes only non-systematic risk, and as it does not require cash outlay at the beginning of the period, has an expected value of zero, and a volatility such that the variance of the tracking portfolio plus the variance of these payments to match the tracking error equal the untradeable underlying's variance. And with a replication strategy for the underlying found in this manner, the conditions for the application of Ito's lemma are met as well as in the case with observable prices.

³ See the appendix A.1 on Ito's Lemma.

$$I. \quad E(\Delta F) = \frac{\partial F}{\partial S} E(\Delta S) + \left(\frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S^2} \sigma^2 S^2 \right) \Delta t$$

In return notation, the Froot and Stein hurdle rate for a new, small exposure is:⁴

$$II. \quad \mu_N^* = r + \gamma \sigma_{NM} + G_R \sigma_{Np}$$

With G_R corresponding to the Rubinstein measure of relative risk aversion for the bank⁵, $\sigma_{Np} = \text{cov}(r_N, r_p)$ being the covariance of the return of the untradeable component of the new investment with the return of the existing portfolio of untradeable assets, $\sigma_{NM} = \text{cov}(r_N, r_M)$ the covariance of the new investment's return with the priced market factor, r the riskless interest rate, γ the market price of risk and μ_N^* the expected return required by the bank for a small investment into the new untradeable asset.⁶

Accordingly, the instantaneous hurdle rate μ_F^* for an option with price F can be written as follows (with r now being expressed as annual rate with continuous compounding, and Δt with unit years):

$$\mu_F^* = r\Delta t + \gamma \text{cov}\left(\frac{\Delta F}{F}, r_M\right) + G_R \text{cov}\left(\frac{\Delta F}{F}, r_p\right) = r\Delta t + \gamma \sigma_{FM} + G_R \sigma_{Fp}$$

The change of the option value and hence the option return is a function of the instantaneous return of the underlying stock r_S :⁷

$$\Delta F = F_S \Delta S = F_S S \frac{\Delta S}{S} = F_S S r_S$$

$$\frac{\Delta F}{F} = F_S \frac{S}{F} r_S$$

And therefore both the covariance of the option with the market and its covariance with the untradeable part of the existing portfolio (σ_{FM} and σ_{Fp}) can be expressed as functions of the respective covariance parameters of the underlying stock (σ_{SM} and σ_{Sp}):

$$\mu_F^* = r\Delta t + F_S \frac{S}{F} (\gamma \sigma_{SM} + G_R \sigma_{Sp})$$

⁴ For notes on the derivation of this hurdle rate see the October '14 post in this blog and the references given there – note that the rate here is before subtracting financing costs (risk free rate).

⁵ This is based on interpreting the payoff function of the bank's future, non-stochastic investment opportunity after deducting financing cost as utility function – for details see the October and November 14 posts on the Froot-Stein model in this blog, and the references given there (note that in the previous posts G_R was labelled G_K).

⁶ As the market price of the stock is consistent with the pricing model for traded assets, $\mu_S = r + \gamma \sigma_{SM}$ and the bank would not pay the observable "market price" for the not perfectly tradable stock (unless $\sigma_{Np} = 0$), but would require a discount (or be willing to pay a premium, depending on the sign of σ_{Np}) – see Froot and Stein page 73 (note that the variance of the untradeable component does not appear here, as for a small investment, the impact on portfolio variance due to the asset's stand-alone variance is considered to be negligible – different to the potential covariance impact).

⁷ see again fn. 2

And with equation II., one can write this hurdle rate as a function of the (minimum) expected stock return μ_S^* the bank would require if it had to hold the stock directly:

$$\mu_F^* = r\Delta t + F_S \frac{S}{F} (\mu_S^* - r\Delta t)$$

μ_F^* is the fair (required) instantaneous expected return of the option from the bank's perspective. Multiplying with the option value gives the required instantaneous expected change of the option value:

$$\text{III.} \quad F\mu_F^* = F \left[r\Delta t + F_S \frac{S}{F} (\mu_S^* - r\Delta t) \right]$$

And as the expected change of the value of the option is also given by I., if the option is priced fairly from the bank's perspective equation I. and III. are equal:⁸

$$\text{IV.} \quad Fr\Delta t + F_S S (\mu_S^* - r\Delta t) = F_S S \mu + \left(F_t + \frac{1}{2} F_{SS} \sigma^2 S^2 \right) \Delta t$$

Slightly rearrange IV.:

$$Fr\Delta t + F_S S (\mu_S^* - r\Delta t - \mu) = \left(F_t + \frac{1}{2} F_{SS} \sigma^2 S^2 \right) \Delta t$$

and with $\pi_S^* = \mu_S^* - \mu_S$:

$$Fr\Delta t + F_S S (\pi_S^* - r\Delta t) = \left(F_t + \frac{1}{2} F_{SS} \sigma^2 S^2 \right) \Delta t$$

Further rearranging gives equation 17 in Froot and Stein:

$$(r\Delta t - \pi_S^*) S F_S + \left(F_t + \frac{1}{2} \sigma^2 S^2 F_{SS} - rF \right) \Delta t = 0$$

Rearranging again:

$$\frac{1}{2} \sigma^2 S^2 F_{SS} \Delta t + (rS\Delta t - \pi_S^* S) F_S + F_t \Delta t - rF \Delta t = 0$$

And with π_S^* now also expressed as annualized figure (and hence multiplied with Δt):

$$\frac{1}{2} \sigma^2 S^2 F_{SS} \Delta t + (rS\Delta t - \pi_S^* S \Delta t) F_S + F_t \Delta t - rF \Delta t = 0$$

⁸ The same principle as for the alternative derivation of their option value differential equation by Black and Scholes (see the appendix A.2 and the earlier post in this blog on that derivation) is applied here – the expected change of the option value derived with Ito's lemma is compared to the expected change of the value required by the investor – while in Black-Scholes the required expected change is the market's required expected change determined with the CAPM, here an additional component is included by the bank to compensate for the impact of the bank's portfolio's untraded risk.

Now divide the whole equation by Δt :

$$\frac{1}{2}\sigma^2 S^2 F_{SS} + (rS - \pi_S^* S)F_S + F_t - rF = 0$$

This is essentially the differential equation given by Merton for European options on stocks paying a constant dividend yield π_S^* .⁹ Hence the corresponding option pricing formulas can be applied to calculate a fair price from the bank's perspective, by setting the constant continuous dividend yield equal to π_S^* . These formulas for stocks paying a constant dividend yield lead to the same result as discounting the stock price with the dividend yield, and then applying the corresponding Black-Scholes-Merton formula for a European option on a non-dividend paying stock, with this discounted stock price (and a slightly modified volatility measure), i.e. by setting the underlying spot price for the option valuation to:¹⁰

$$S_0 e^{-\pi_S^* T}$$

An intuitive reason for this equivalence is that after discounting the price of the dividend paying stock with the certain dividend yield (and adjusting the volatility), the probability distribution of that stock price at T will be the same as the probability distribution for the price at T of an (otherwise equal) non-dividend-paying stock with current price equal to that discounted price.¹¹

Pricing derivatives on untraded assets if prices are not continuously observable

Merton's reasoning for pricing an option on an untradeable asset, when its price is only observable at the time the option is written and its expiry date, is based on the approach with observable prices described above in footnote 2. As in the case with observable prices, if the underlying is not traded, a portfolio of traded assets can be used to track the underlying's return as closely as possible. The return of the underlying in an instantaneous interval is the difference of this tracking portfolio's return and the

⁹ See Robert C. Merton, Theory of Rational Option Pricing, The Bell Journal of Economics and Management Science, Vol. 4, No. 1 (Spring, 1973), pp. 141-183, equation 44 and Footnote 62, p.171. The option pricing formula given there implies that the dividend rate is a riskless instantaneous rate, paid in every infinitesimal small time interval during the life of the option. An explicit derivation for this special case can be found in John C. Hull, Options, Futures & Other Derivatives, Ninth edition (in the following referred to as "Hull", p. 372). Hull further shows, that the put call parity also holds for options on stocks paying a constant dividend yield and derives the valuation formula for a put on a dividend paying stock (Hull p.372/373 and technical note 6, <http://www-2.rotman.utoronto.ca/~hull/TechnicalNotes/TechnicalNote6.pdf>). Note that instead of the partial derivative of the option price with respect to the current time variable t , Merton uses the partial derivative with respect to time till expiration – which leads to the different sign of that partial derivative in Merton '73.

¹⁰ Hull, p. 372. the volatility measure has to be adjusted, as with a lower stock price a given price change is relatively larger (Hull p.343 fn. 12 – note that the adjustment is labelled there as "approximation" as it is introduced in the more general context with dividends not necessarily being paid with a constant yield – however for the special case with a constant instantaneous dividend yield like here, this approximation is exact).

¹¹ Ibid. "having the same probability distribution" can here be interpreted as "having the same value in every state of the world". Further, recall that the additional return component required by the bank in Froot and Stein can be interpreted as a discount (or potential premium, in case of a negative covariance of the position with the existing portfolio of untradeable assets) on the stock price. see again fn. 6 and the reference given there.

tracking error (with the tracking error defined as the difference between instantaneous tracking portfolio return and instantaneous underlying return). With prices being only observable at issue and expiry date of the option, during the life of the option the value of the tracking portfolio constitutes the best estimate of the value of the underlying. The instantaneous tracking errors are unobservable, and cumulate to a total tracking error that becomes only visible at option expiry. At this moment the true value of the underlying gets revealed and the best estimate jumps from the value of the tracking portfolio to the true value.¹² This means that the underlying asset price can be modelled as a special case of a mixed-jump diffusion process – more details further below.

To derive the properties of the tracking portfolio mentioned in footnote 2, note that the instantaneous return r_s of a tracking portfolio consisting of positions in all tradable assets i, \dots, n with portfolio weights w_i is given by:

$$r_s = \sum_{i=1}^n w_i r_i$$

With the instantaneous return of the untradeable asset labelled r_v , the instantaneous tracking error x is defined as follows:

$$x = r_s - r_v = \sum_{i=1}^n w_i r_i - r_v$$

The expected instantaneous tracking error is therefore:

$$\mu_x = E(r_x) = E(r_s - r_v) = E\left(\sum_{i=1}^n w_i r_i - r_v\right) = \sum_{i=1}^n w_i \mu_i - \mu_v$$

and the variance of the instantaneous tracking error:

$$\sigma_x^2 = \sigma_s^2 - 2\sigma_{sv} + \sigma_v^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} + \sigma_v^2 - 2 \sum_{i=1}^n w_i \sigma_{iv}$$

The tracking portfolio is supposed to track the underlying return as closely as possible, i.e. the variance of the tracking error is minimized, which gives the following set of conditions:

$$\frac{\partial \sigma_x^2}{\partial w_i} = 2 \sum_{j=1}^n w_j \sigma_{ij} - 2\sigma_{iv} = 0$$

¹² Merton '98, p. 335

$$\sum_{j=1}^n w_j \sigma_{ij} = \sigma_{iv}$$

$$\sigma_{is} = \sigma_{iv}$$

I.e., when the tracking variance is minimized, the covariance of the tracking portfolio with any tradable asset equals the covariance of the untradeable underlying with that tradable asset.

It follows for the covariance of the tracking error of the optimal tracking portfolio with any tradable asset:

$$\sigma_{ix} = \sum_{j=1}^n w_j \sigma_{ij} - \sigma_{iv} = 0$$

Therefore, the tracking error is uncorrelated with the tracking portfolio, and hence:

$$\sigma_s^2 + \sigma_x^2 = \sigma_v^2$$

Or with annualized instantaneous variances dt :

$$\sigma_s^2 dt + \sigma_x^2 dt = \sigma_v^2 dt$$

Merton assumes that the tracking error is normally distributed.¹³ Further, the replication portfolio is rebalanced continuously, and all assets' volatilities and expected returns are constant, so the volatility and expected value of the tracking error are constant.¹⁴ The instantaneous tracking error can therefore be represented by a stochastic differential equation for a variable X (with an arbitrary starting value X_0):

$$dX = \mu_x X dt + \sigma_x X dz_x$$

such that $x = dX/X$ is the instantaneous tracking error and with z_x being a Wiener process. Merton assumes a starting value of $X_0=1$. X_t then can be interpreted as the value per notional dollar of a portfolio which in $t=0$ consists of a long position in the tracking portfolio "financed" with a short position with equal notional size in the untradeable asset. Merton refers to X_t as cumulative proportional tracking error.

As assets are jointly normal¹⁵, and if no dividends are assumed, the process assumed by Merton for the tracking portfolio of traded assets can be written as follows:

¹³ See the process for the tracking error in Merton '98, p. 330, where the stochastic component contains a Wiener process.

¹⁴ This follows from Merton '98, page 330 where it says that at each point in time, the weights in the replicating portfolio are chosen as to minimize the tracking error variance, and the assumption that the instantaneous covariance matrix including the asset to be hedged and all traded assets stays constant over time.

¹⁵ Merton does not explicitly state this assumption, however, tracking error and underlying are normally distributed and uncorrelated, and the sum of two uncorrelated normal variables is also normal – so to ensure that this holds here, the returns of the traded assets and the untraded asset must be jointly normal.

$$dS = \mu_S S dt + \sigma_S S dz_S$$

With z_S being a Wiener process (this is essentially equation 4 in Merton '98, here adjusted for zero dividends). The instantaneous return of the tracking portfolio is hence:

$$r_S = \frac{dS}{S} = \mu_S dt + \sigma_S dz_S$$

The untraded asset follows a similar process – see assumption 2 in Merton '98, again here adjusted for the case of zero dividends:

$$dV = \mu_V V dt + \sigma_V V dz_V$$

with z_V being a Wiener process. Hence, the instantaneous return of the untraded asset (which is unobservable) equals:

$$r_V = \frac{dV}{V} = \mu_V dt + \sigma_V dz_V$$

With the process for S shown above, the value of S at the option's expiry date, S_T , can be (somewhat informally) written as follows:¹⁶

$$S_T = S_0 e^{\left(\mu_S - \frac{\sigma_S^2}{2} \right) T + \sigma_S \sum_{t=1}^N dz_{S,t}}$$

Where the period T is divided into infinitesimal small time intervals with length dt , numbered from 1 till N .

And correspondingly V_T and X_T can be expressed as:

$$V_T = V_0 e^{\left(\mu_V - \frac{\sigma_V^2}{2} \right) T + \sigma_V \sum_{t=1}^N dz_{V,t}}$$

and:

$$X_T = X_0 e^{\left(\mu_X - \frac{\sigma_X^2}{2} \right) T + \sigma_X \sum_{t=1}^N dz_{X,t}}$$

The instantaneous tracking error is the difference of the instantaneous returns of the replication portfolio and the untradeable asset:

$$\frac{dX}{X} = \mu_X dt + \sigma_X dz_X = \frac{dS}{S} - \frac{dV}{V} = \mu_S dt + \sigma_S dz_S - \mu_V dt - \sigma_V dz_V$$

¹⁶ See the section on the stock price after a period longer than dt in the appendix A.1 on Ito's Lemma.

As Merton argues, if the CAPM (or another equilibrium model that implies that the market prices only systematic risk) holds and as the tracking error is uncorrelated with any tradable asset and hence the market portfolio (or priced factor), i.e. the tracking error is purely nonsystematic, then the instantaneous expected returns of untradeable asset and replication portfolio are equal and the expected instantaneous tracking error is zero.

Substituting $\mu_X = 0$ in the equation above for X_T :

$$X_T = X_0 e^{-\frac{\sigma_X^2}{2}T + \sum_{t=1}^N \sigma_X dz_{X,t}}$$

As $X_0 = 1$:

$$X_T = e^{-\frac{\sigma_X^2}{2}T + \sum_{t=1}^N \sigma_X dz_{X,t}}$$

And hence:

$$S_T X_T =$$

$$S_0 e^{\left(\mu_S - \frac{\sigma_S^2}{2}\right)T + \sigma_S \sum_{t=1}^N dz_{S,t} - \frac{\sigma_X^2}{2}T + \sum_{t=1}^N \sigma_X dz_{X,t}} =$$

$$S_0 e^{\left[\mu_S - \left(\frac{\sigma_S^2 + \sigma_X^2}{2}\right)\right]T + \sigma_S \sum_{t=1}^N dz_{S,t} + \sum_{t=1}^N \sigma_X dz_{X,t}}$$

And with

$$\sigma_S^2 + \sigma_X^2 = \sigma_V^2, \quad \mu_S dt = \mu_V dt \quad \text{and} \quad \sigma_V dz_V = \sigma_S dz_S + \sigma_X dz_X$$

This simplifies to:

$$S_T X_T = S_0 e^{\left(\mu_V - \frac{\sigma_V^2}{2}\right)T + \sigma_V \sum_{t=1}^N dz_{V,t}}$$

Because $V_0 = S_0$:

$$S_T X_T = V_0 e^{\left(\mu_V - \frac{\sigma_V^2}{2}\right)T + \sigma_V \sum_{t=1}^N dz_{V,t}} = V_T$$

Or:

$$V. \quad V_T = S_0 e^{\left[\mu_S - \left(\frac{\sigma_S^2 + \sigma_X^2}{2} \right) \right] T + \sigma_S \sum_{t=1}^N dz_{S,t} + \sum_{t=1}^N \sigma_X dz_{X,t}}$$

Where $\sum_{t=1}^N \sigma_X dz_{X,t}$ is the realized cumulative tracking error.

This corresponds to a special case of Merton's Mixed Jump-Diffusion model¹⁷, with the underlying jumping exactly one time, the jump component X_T being lognormally distributed, the expected jump size being zero, and the jump occurring at T . As Merton shows, under these conditions the value of an option in $t=0$ is the standard Black Scholes value, except that the per unit time variance to be used is:¹⁸

$$VI. \quad \sigma_V^2 = \sigma_S^2 + \frac{\delta^2}{T}$$

Where δ^2 is the variance of $LN(X_T)$, which is in $t=0$ equal to $Ndt\sigma_X^2 = T\sigma_X^2$, as can be seen from:

$$X_T = e^{-\frac{\sigma_X^2}{2}T + \sum_{t=1}^N \sigma_X dz_{X,t}}$$

So that:

$$\frac{\delta^2}{T} = \sigma_X^2$$

And hence the variance measure for V to be used for pricing at $t=0$ is

$$\sigma_V^2 = \sigma_S^2 + \sigma_X^2$$

At later points in time ($t>0$), one cannot use σ_X^2 , because the variance δ^2 of the logarithm of X_T does not change during the life of the option, as the tracking error is not observable before expiry. To calculate an annualized per unit time variance, one must instead divide the constant δ^2 by the remaining time to expiry (labelled τ by Merton), i.e. the variance per unit time for V to be used is then $\sigma_V^2 = \sigma_S^2 + \frac{\delta^2}{\tau}$, such that with τ approaching zero, the variance of the continuously compounded return of V till maturity, $\tau\sigma_V^2$, does not approach zero as would be the case without the jump, but approaches δ^2 .¹⁹

¹⁷ Option Pricing When Underlying Stock Returns Are Discontinuous, Robert C. Merton, Journal of Financial Economics, Vol. 3, 1976, p. 124-144. See page 128 of this source, where the general description of the jump event is introduced in a notation analogous to the notation here ($V=SX$), and further Merton's explanation to equation 18 on page 135, and note that here a special case with the conditions given above applies.

¹⁸ See footnote 13., page 136 in Merton '76 and page 335 in Merton '98.

¹⁹ Merton '98 page 335.

Combining Froot and Stein and Merton '98

In summary, as it is certain that exactly one jump will occur during the life of the option, and as X_T is lognormal with an expected value of one, the distribution of the final stock price will exactly equal the distribution the final stock price would have if there was no jump but if the stock's (annualized) variance was the sum of the variance of the logarithm of the cumulative proportional tracking error divided by time to expiry, and the tracking portfolio's (annualized) variance²⁰ – which can be seen from equations V and VI.²¹ Further, it is known that the jump will happen at the end of the life of the option. For these reasons, the final stock price from the perspective at any point in time before expiry has the same probability distribution as if the stock was traded and would follow a lognormal diffusion process with the mentioned non-proportional-in-time variance.²² It follows that at any point in time, the price of an option on this stock, the probability distribution of the instantaneous change of this option price and hence the instantaneous return distribution of the option are identical to those of an option on a traded but otherwise equal stock with a variance equal to the non-proportional-in-time variance.

Hence the bank should derive the hurdle rate for the expected return of an option on an illiquid asset with unobservable prices analogously to the case with observable prices. As mentioned in the first section, the risk premium required by the bank above the market's required expected return can be treated as a dividend paid at a constant rate – hence amending the pricing formula to adjust for the additional premium can straightforwardly be done by discounting the underlying's initial spot price with this rate (and adjusting total volatility and cumulative tracking error to reflect the lowered price)²³.

Recall that the bank will charge a premium exceeding the market's required return only for the untradeable part of the option's risk. The component of the instantaneous risk premium to be added to the market requirement is therefore a function of the covariance of the untradeable part – the instantaneous tracking error – with the bank's existing portfolio of untradeable assets. As mentioned above, the variance of the logarithm of the cumulative proportional tracking error does not decrease with time, as during the life of the option no information about the cumulative tracking error gets revealed, and hence from the perspective of the bank it has the same distribution at option expiry no matter the time of valuation. For the same reason also the adjustment to the underlying spot price to compensate for the contribution of the tracking error to the bank's aggregate portfolio of untradeable risks stays constant over time when prices are unobservable – which requires discounting the stock price for valuation at later times also with the instantaneous “dividend” rate (the additional premium charged by the bank to compensate for the impact on untradeable portfolio risk) multiplied with the initial time to expiry T , as in the first instant.

²⁰ This sum will, following Merton, in the remainder of this text be referred to as “non-proportional-in-time variance” – see again Merton '76, p. 136, fn. 13. It is only in $t=0$ equal to the stock's instantaneous (annualized) variance, as noted above.

²¹ This is analogous to Hull's reasoning for the treatment of dividends – see page 4 and the references given there.

²² At a time later than $t=0$, with the current price of the underlying set equal to the price of the tracking portfolio – as the underlying's price is unobservable – the impact of all tracking error realizations gets revealed only at option expiry, without distinction of realizations that may be attributable to a time before the current point in time.

²³ see fn. 10 and the reference given there for the volatility adjustment

Appendix:

A.1: Ito's Lemma

Preliminary:

Taylor polynomial:

$$T_n f(x; a) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

Taylor series (the Taylor polynomial when n approaches infinity):²⁴

$$f(x) \approx \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k = f(a) + f'(a)(x-a) + \frac{f''(a)}{2} (x-a)^2 + \dots$$

Taylor series in several variables:²⁵

For a function of two variables, $G(S, t)$, the Taylor series to second order about the point (S_0, t_0) is:

$$G(S, t) = G(S_0, t_0) + \frac{\partial G}{\partial S} (S - S_0) + (t - t_0) \frac{\partial G}{\partial t} + \frac{1}{2} (S - S_0)^2 \frac{\partial^2 G}{\partial S^2} + (S - S_0)(t - t_0) \frac{\partial^2 G}{\partial S \partial t} + (t - t_0)^2 \frac{\partial^2 G}{\partial t^2} \dots$$

(with the **cross partial derivative** $\frac{\partial G}{\partial S \partial t} = \frac{\partial G / \partial S}{\partial t} = \frac{\partial G_S}{\partial t}$ being “the first derivative with respect to t of the first derivative of G with respect to S ”)²⁶. It gives an approximate value of G if both S and t change by a small amount.

Note that $\frac{\partial G}{\partial t}$ as partial derivative of G with respect to t implies holding S constant – an alternative notation to emphasize this is: $\left(\frac{\partial G}{\partial t} \right)_S$.²⁷

Ito process:

$$dx = a(x, t)dt + b(x, t)dz \quad \text{i.e. } a \text{ and } b \text{ are functions of } x \text{ and } t.$$

with z following a Wiener process, so that:

$$\Delta z = \varepsilon \sqrt{\Delta t}$$

²⁴See for example http://en.wikipedia.org/wiki/Taylor_series

²⁵ See e.g. http://en.wikipedia.org/wiki/Taylor_series#Taylor_series_in_several_variables

²⁶See e.g. http://en.wikipedia.org/wiki/Partial_derivative#Higher_order_partial_derivatives

²⁷ http://en.wikipedia.org/wiki/Partial_derivative#Notation

(ε is a standard normal variable and dz the limit for Δt approaching zero)

Discrete time version: $\Delta x = a(x, t)\Delta t + b(x, t)\Delta z$ or $\Delta x = a(x, t)\Delta t + b(x, t)\varepsilon\sqrt{\Delta t}$

Derivation of Ito's Lemma:²⁸

Step 1: apply a Taylor series in several variables to a function G of time and a variable S that follows an Ito process:

with $\Delta S = S - S_0$ and $\Delta t = t - t_0$:

$$G(S, t) = G(S_0, t_0) + \frac{\partial G}{\partial S} \Delta S + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} (\Delta S)^2 + \frac{\partial G}{\partial S \partial t} \Delta S \Delta t + \frac{\partial^2 G}{\partial t^2} (\Delta t)^2 \dots$$

Step 2: subtract G at the beginning of the time period considered (Δt), to get the price change:²⁹

$$\Delta G(S, t) = G(S, t) - G(S_0, t_0) = \frac{\partial G}{\partial S} \Delta S + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} (\Delta S)^2 + \frac{\partial G}{\partial S \partial t} \Delta S \Delta t + \frac{\partial^2 G}{\partial t^2} (\Delta t)^2 + \dots$$

$$\text{A.I.} \quad \Delta G = \frac{\partial G}{\partial S} \Delta S + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} (\Delta S)^2 + \frac{\partial G}{\partial S \partial t} \Delta S \Delta t + \frac{\partial^2 G}{\partial t^2} (\Delta t)^2 + \dots$$

In “ordinary” calculus, typically terms of higher order (products containing variables with exponents larger than one) would be ignored, because when ΔS and Δt approach zero, these terms would approach zero faster. However, here this is not completely the case for $\frac{1}{2} \frac{\partial^2 G}{\partial S^2} (\Delta S)^2$.

To see this, look at the discrete time version of the Ito process:

$$\text{A.II.} \quad \Delta S = a(S, t)\Delta t + b(S, t)\varepsilon\sqrt{\Delta t}$$

And calculate $(\Delta S)^2$:

$$\begin{aligned} (\Delta S)^2 &= (a_t \Delta t + b_t \varepsilon \sqrt{\Delta t})^2 = (a_t \Delta t)^2 + 2(a_t \Delta t)(b_t \varepsilon \sqrt{\Delta t}) + (b_t \varepsilon \sqrt{\Delta t})^2 \\ &= a_t^2 (\Delta t)^2 + 2(a_t \Delta t)(b_t \varepsilon \sqrt{\Delta t}) + b_t^2 \varepsilon^2 \Delta t \\ &= a_t^2 (\Delta t)^2 + 2a_t b_t \varepsilon (\Delta t)^{1.5} + b_t^2 \varepsilon^2 \Delta t \end{aligned}$$

²⁸ The following follows Hull, Options, Futures & Other Derivatives, 9th edition (referred to as “Hull” in the following), p. 319-320 with some additional explanations for which references are given.

²⁹ That's the starting equation 14.A.3 of the derivation of Ito's Lemma in Hull, it can also be found in the slide pack to chapter 14, p. 25. – available at <http://www-2.rotman.utoronto.ca/~hull/ofodslides/index.html>

This contains a term which is a multiple of Δt .

Step 3: determine the approximate limit of $(\Delta S)^2$ as Δt approaches zero:³⁰

When Δt approaches zero, $a^2(\Delta t)^2 + 2ab\varepsilon(\Delta t)^{1.5} + b^2\varepsilon^2\Delta t$ approaches $b^2\varepsilon^2\Delta t$, because $(\Delta t)^2$ and $(\Delta t)^{1.5}$ approach zero faster than Δt .

The expected value of $b^2\varepsilon^2\Delta t$ is $b^2\Delta t$, because $E(\varepsilon)$ is zero and the variance of ε^2 is 1 and hence:

$$\sigma_\varepsilon^2 = 1 = E(\varepsilon^2) - [E(\varepsilon)]^2$$

$$E(\varepsilon) = 0$$

$$[E(\varepsilon)]^2 = 0$$

$$\sigma_\varepsilon^2 = 1 = E(\varepsilon^2)$$

so that $E(b^2\varepsilon^2\Delta t) = b^2\Delta t E(\varepsilon^2) = b^2\Delta t$

Subtract this expected value from $b^2\varepsilon^2\Delta t$ to get the stochastic component of $b^2\varepsilon^2\Delta t$:

$$b^2\varepsilon^2\Delta t - b^2\Delta t = (b\varepsilon\sqrt{\Delta t})^2 - b^2\Delta t$$

When Δt approaches zero, this stochastic component will approach zero considerably faster than the stochastic component in ΔS which is $b\varepsilon\sqrt{\Delta t}$ (see equation A.II), and is hence considered negligible compared to that term and ignored. Therefore, $(\Delta S)^2$ can be replaced with $b^2\Delta t$ in equation A.I:

$$\Delta G = \frac{\partial G}{\partial S} \Delta S + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} b^2 \Delta t + \frac{\partial G}{\partial S \partial t} \Delta S \Delta t + \frac{\partial^2 G}{\partial t^2} (\Delta t)^2 + \dots$$

Step 4: substitute for ΔS in $\frac{\partial G}{\partial S \partial t} \Delta S \Delta t$:

$$\Delta S = a\Delta t + b\Delta z$$

$$\frac{\partial G}{\partial S \partial t} \Delta S \Delta t = \frac{\partial G}{\partial S \partial t} (a\Delta t + b\Delta z) \Delta t$$

Substitute for $\Delta z = \varepsilon\sqrt{\Delta t}$:

$$\frac{\partial G}{\partial S \partial t} \Delta S \Delta t = \frac{\partial G}{\partial S \partial t} (a\Delta t + b\varepsilon\sqrt{\Delta t}) \Delta t = \frac{\partial G}{\partial S \partial t} [a(\Delta t)^2 + b\varepsilon(\Delta t)^{1.5}] = \frac{\partial G}{\partial S \partial t} a(\Delta t)^2 + \frac{\partial G}{\partial S \partial t} b\varepsilon(\Delta t)^{1.5}$$

³⁰ <http://www.sjsu.edu/faculty/watkins/ito.htm>

As both summands are coefficients of Δt raised to higher powers than 1, they get both ignored (like other such terms in the Taylor-series approximation), so that ΔG simplifies to:

$$\Delta G = \frac{\partial G}{\partial S} \Delta S + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} b^2 \Delta t$$

Step 5: substitute for delta ΔS :

$$\Delta S = a\Delta t + b\Delta z$$

$$\Delta G = \frac{\partial G}{\partial S} (a\Delta t + b\Delta z) + \left(\frac{\partial G}{\partial t} \right)_S \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} b^2 \Delta t$$

Step 6: Rearranging:

$$\Delta G = \left[\frac{\partial G}{\partial S} a + \left(\frac{\partial G}{\partial t} \right)_S + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} b^2 \right] \Delta t + \frac{\partial G}{\partial S} b \Delta z$$

Step 7: Taking limits (Δt approaching zero) gives Ito's Lemma:

$$dG = \left[\frac{\partial G}{\partial S} a + \left(\frac{\partial G}{\partial t} \right)_S + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} b^2 \right] dt + \frac{\partial G}{\partial S} b dz$$

A process for stock prices:³¹

$$a = \mu S$$

$$b = \sigma S$$

$$dS = \mu S dt + \sigma S dz$$

or in discrete time:

$$\text{A.III:} \quad \Delta S = \mu S \Delta t + \sigma S \Delta z = \mu S \Delta t + \sigma S \varepsilon \sqrt{\Delta t}$$

Where return statistics (expected value and volatility) are given as p.a. figures and the unit of t is years. This process is known as geometric Brownian motion. If the stock price follows this process, the stock's instantaneous return (discrete return in an (infinitesimal) small interval) is: $\frac{dS}{S} = \mu dt + \sigma dz$.

³¹ See e.g. Hull, p. 308-309

Apply Ito's Lemma to a function G of S and t :

$$\text{A.IV.: } dG = \left[\frac{\partial G}{\partial S} \mu S + \left(\frac{\partial G}{\partial t} \right)_S + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right] dt + \frac{\partial G}{\partial S} \sigma S dz$$

Apply Ito's lemma to $LN(S)$:³²

$$G = LN(S)$$

$$\frac{\partial G}{\partial S} = \frac{1}{S} = S^{-1}$$

$$\frac{\partial^2 G}{\partial S^2} = -S^{-2} = -\frac{1}{S^2}$$

$$\frac{\partial G}{\partial t} = 0$$

So in general, if S follows an Ito process, the process for $G=LN(S)$ is:

$$dG = \left(\frac{1}{S} a - \frac{1}{2S^2} b^2 \right) dt + \frac{1}{S} b dz$$

And more specifically, if S follows a geometric Brownian motion, so that $dS = \mu S dt + \sigma S dz$, i.e. $a = \mu S$ and $b = \sigma S$, the process for dG is:

$$dG = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dz$$

Hence:

$$dG \sim N \left[\left(\mu - \frac{\sigma^2}{2} \right) dt, \sigma^2 dt \right]$$

(i.e. e^{dG} is "log-normally distributed" meaning that its logarithm is normally distributed).

and hence the following holds for the stock price after a short period of time:

$$S_{t+dt} = S_t e^{dG_t} = S_t e^{\left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dz}$$

i.e. with the instantaneous return dS/S being defined as above, dG is the return of the stock expressed as rate of return with continuous compounding during time interval dt ³³.

³² Hull, p. 314-315

The expected value of a log-normally distributed variable w , whose logarithm is normally distributed with mean m and standard deviation σ is:³⁴

$$E(w) = e^{m + \frac{\sigma^2}{2}}$$

So that:

$$E(S) = S_0 E(e^{dG}) = S_0 e^{\mu dt}$$

An informal expression for the stock price after a period longer than dt :

With $S_{t+dt} = S_t e^{dG_t} = S_t e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz}$, and as under the assumption of S following a geometric Brownian motion, μ and σ are constant, one can derive an “informal”³⁵ expression for the random variable S_t after a longer period than dt , for example a period containing two (infinitesimal small) periods, i.e. two periods with the length dt of each approaching zero – with the index t in S_t indicating the number of such periods of length dt that will have passed:

$$\begin{aligned} S_1 &= S_0 e^{dG_1} = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_1} \\ S_2 &= S_1 e^{dG_2} = S_1 e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_2} = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_1} e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_2} = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_1 + \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dz_2} \\ &= S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)dt + \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma(dz_1 + dz_2)} \end{aligned}$$

Correspondingly, with T being the point in time after N such periods with a length dt that approaches zero, the stock price at T is:

$$S_T = S_0 e^{N\left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma \sum_{t=1}^N dz_t}$$

Hence, the expected return $E[LN(S_T) - LN(S_0)]$ with continuous compounding over the total period is:

³³ For the definition of the rate of return with continuous compounding and its relationship to a discrete return under certainty see e.g. Hull, Options Futures & Other Derivatives, 9th edition, page 81. For a discussion of the relationship between an uncertain instantaneous return and the corresponding random return with continuous compounding and stock price, see Hull, section 15.3 (page 325).

³⁴ See Hull technical note 2 to Options, Futures & Other Derivatives, (9th ed).

³⁵ a formal expression would use a stochastic integral over t .

$N\left(\mu - \frac{\sigma^2}{2}\right)dt$, and the variance of the return with continuous compounding is: $N\sigma^2 dt$

And as the starting point was set as $t=0$, T is also the length of the total period measured in years, i.e. $T = Ndt$ - so that the expected return with continuous compounding can be written as:

$$E[LN(S_T) - LN(S_0)] = TN\left(\mu - \frac{\sigma^2}{2}\right)dt = T\left(\mu - \frac{\sigma^2}{2}\right)$$

and the variance of $LN(S_T) - LN(S_0)$ is:

$$TN\sigma^2 dt = T\sigma^2$$

And with the returns in each sub-period being normally distributed (as S follows a geometric Brownian motion), the return with continuous compounding over the total period is a sum of independent normal variables and hence one gets for the annualized return x over the total period:

$$x = \frac{1}{T}[LN(S_T) - LN(S_0)]$$

$$x \sim N\left(\mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{T}\right)$$

A.2: CAPM-based derivation of the Black-Scholes differential equation:³⁶

A stock's prices follow the process specified above (equation A.III.):

$$\Delta S = \mu S \Delta t + \sigma S \Delta z$$

For the change of the price of a derivative the price of which is a function of G and t in a small interval, one gets with Ito's lemma as above (equation A.IV.):

$$\Delta G = \left[\frac{\partial G}{\partial S} \mu S + \left(\frac{\partial G}{\partial t} \right)_S + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right] \Delta t + \frac{\partial G}{\partial S} \sigma S \Delta z$$

Or, (somewhat closer to BS original notation):³⁷

$$\text{A.V.} \quad \Delta G = \frac{\partial G}{\partial S} \Delta S + \left[\frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right] \Delta t$$

³⁶ Fisher Black and Myron Scholes, "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, Vol. 81, Issue 3, 1973, p. 645 - 646

³⁷ Black/Scholes divide by Δt to get an annualized instantaneous return

Black/Scholes assume the CAPM holds for instantaneous returns.³⁸

Derive the option's beta: Write the instantaneous option return as function of the stock return $\frac{\Delta S}{S}$:

$$\frac{\Delta G}{G} = \frac{\partial G}{\partial S} \frac{\Delta S}{S} \frac{S}{G} + \left(\frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) \frac{\Delta t}{G}$$

Covariance of option return with market return:

$$\text{cov}\left(\frac{\Delta G}{G}, r_M\right) = \frac{\partial G}{\partial S} \frac{S}{G} \text{cov}\left(\frac{\Delta S \Delta t}{S}, r_M\right)$$

Hence, the option beta as a function of the underlying stock's beta β_S is:

$$\beta_S \frac{\partial G}{\partial S} \frac{S}{G}$$

If CAPM holds, the instantaneous expected return of the option must equal:

$$E\left(\frac{\Delta G}{G}\right) = r\Delta t + E(r_M - r)\Delta t \beta_S \frac{\partial G}{\partial S} \frac{S}{G}$$

$$E\left(\frac{\Delta G}{G}\right) - r\Delta t = E(r_M - r)\Delta t \beta_S \frac{\partial G}{\partial S} \frac{S}{G}$$

$$E\left(\frac{\Delta G}{G}\right) - r\Delta t = E(r_S - r)\Delta t \frac{\partial G}{\partial S} \frac{S}{G}$$

Where r_S is the instantaneous return of the stock. The expected change in the option price is therefore:

$$\text{A.VI.:} \quad E(\Delta G) = \frac{\partial G}{\partial S} E(\Delta S) + Gr\Delta t - rS\Delta t \frac{\partial G}{\partial S}$$

The expected absolute change results also from taking the expected value of both sides of A.IV.:

$$\text{A.VII.:} \quad E(\Delta G) = \frac{\partial G}{\partial S} E(\Delta S) + \left(\frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) \Delta t$$

Setting A.VI. and A.VII. equal, subtracting $\frac{\partial G}{\partial S} E(\Delta S)$ on both sides, dividing by Δt and simplifying gives the Black-Scholes differential equation.³⁹

$$\frac{\partial G}{\partial t} = rG - rS \frac{\partial G}{\partial S} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 G}{\partial S^2}$$

³⁸ Black/Scholes page 646 and footnote 2 on page 639

³⁹ Black/Scholes , equation 7, page 643