Volatility forecasting for the purpose of DtD estimation

Susan Thomas
Finance Research Group, IGIDR, Bombay
Presentation for StatFin 2018

18 December, 2018

Outline

Part I

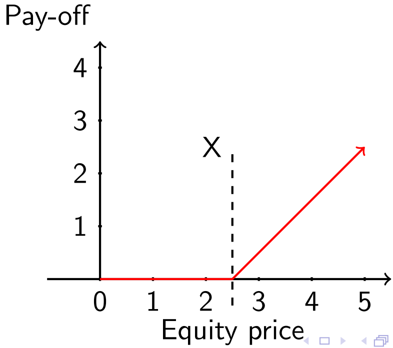
Foundations

Estimating the prob(default) using call option pricing

- Merton (1974) presents that there is a call option payoff *in every firm*.
- If this is true, then we can use the well-established formula to price a call option (Black-Scholes options pricing model (1973), Merton (1974)) to estimate the probability that a firm will default on payments.
- How?

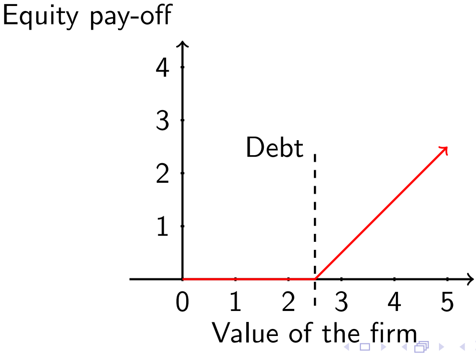
What is a call option?

- A call option gives the investor the right (not obligation) to buy a share at a pre-determined price (X) within a pre-determined period of time (T).
  For this right, the investor pays C.
- If the price of the share rises much higher than X during T, the investor gets to buy the share at X.
- If the price drops below X, the investor has the choice to not buy.



What is a firm

- The firm is constructed out of a combination of Equity + Debt
- Equity shareholders have control over the firm. Debt holders have the promise that they will be paid (a fixed amount).
- When shareholders cannot pay the debt, they lose control to the debt holders. This is called Default.
- As long as shareholders can pay the debt, they have the upside. The upside is like a call option on the value of the firm.

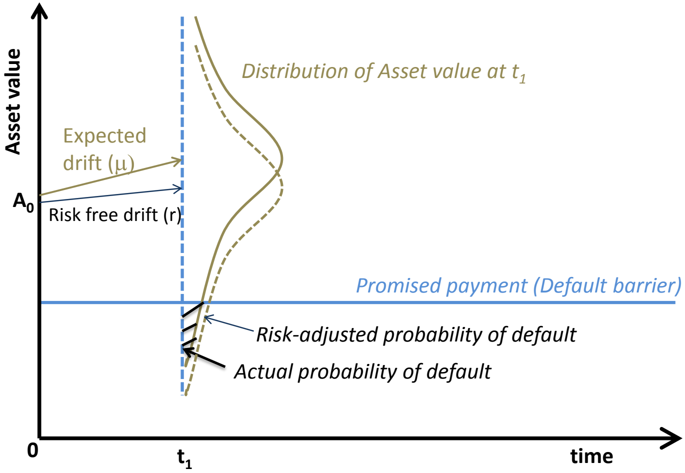


Call option pricing at the heart of estimating prob(default)

- Basic idea:
  1. Equity is a call option on firm's asset value with debt as strike price.
  2. Firm's shareholders hold a call option on the firm's asset.

- Seminal papers:
  1. Merton, R. C., 1974. On the pricing of corporate debt: the risk structure of interest rate. *Journal of Finance* 29, 449-470.
  2. Merton, R. C., 1977. An analytical derivation of the cost of deposit insurance and loan guarantee. An application of modern option pricing theory. *Journal of Banking and Finance* 1, 3-11.

Call option pricing at the heart of estimating prob(default)



Part II

Operationalising Merton, 1974

Operationalising estimation of default risk

- What we know:
  1. We see E, the equity market capitalisation of the firm
  2. We see σ_E the equity volatility of the firm
  3. We know the level of Debt, as the book value recorded in the balance sheet.

- What we do not know:
  1. We do not see V, the total value of the firm
  2. We do not know σ_V, the volatility of the total assets of the firm

We can setup two equations in two unknowns – what the KMV model did.

The Kealhofer-Merton-Vasicek (KMV) model

- Debt: homogenous with maturity at time, T
- Capital structure of the firm: $V_A(t) = D(t) + E(t)$
- Perfect markets: no coupons, no dividends, no frictions on trading
- Asset dynamics: assets are traded, and prices follow GBM.

$$dV_A = \mu_A V_A dt + \sigma_A V_A dW$$

where V_A is the value of the asset, σ_A is its volatility, μ_A is the drift and dW is a Wiener process.

KMV methodology

- Using the analogy of the Black-Scholes model,
  1. Equity, E = Call option, C
  2. Debt, D = Strike, K
  3. Value of the firm, V_A = Equity price, S
- Then if the call option pricing formula is:

$$C(t) = S(t)\Phi(d_1) - e^{-r(T-t)}K\Phi(d_2)$$

- Equity, E can be priced as:

$$E(t) = V_A(t)\Phi(d_1) - e^{-r(T-t)}D\Phi(d_2)$$

- Ito's formula is used to show that:

$$\sigma_E = \left(\frac{V_A}{V_E}\right)\left(\frac{\partial V_E}{\partial V_A}\right)\sigma_A$$

KMV model

- To find V_A, σ_A, solve the non-linear system of equations:

$$
\begin{aligned}
f_1(V_E, \sigma_E) : V_A(t)\Phi(d_1) - e^{-r(T-t)}K\Phi(d_2) - V_E(t) &= 0 \\
f_2(V_E, \sigma_E) : \frac{V_A}{V_E}\Phi(d_1)\sigma_A - \sigma_E &= 0
\end{aligned}
$$

- The solution is unique since $\partial f_1 / \partial V_A = \Phi(d_1)$ (analogous to δ in the original B-S).
- f_1 is increasing in $V_A \rightarrow f_1(V_A)$ has a unique solution.
- Similarly, we can see that $f_2(\sigma_E)$ has a unique solution also.

The ultimate prize - Distance from default (DtD)

- Default is the instance when firm value falls below debt or $V_A \leq D$.
- DtD(t) is the distance between the expected firm value and the default point:

$$\mathrm{DtD}(t) = \frac{\log\left(\frac{V_A(t)}{D}\right) + (r - 0.5\sigma_A^2)(T - t)}{\sigma_A\sqrt{T - t}}$$

- Probability of default: substitution into a normal CDF gives:

$$Pr(def)(t) = P[V_A \leq D] = \Phi(-\mathrm{DtD})$$

- When DtD is measured as:

$$\mathrm{DtD} = \frac{V_A - D}{\sigma_A V_A}$$

It is interpreted as the number of standard deviations the firm value is away from the default trigger.

Part III

The importance of volatility forecasting

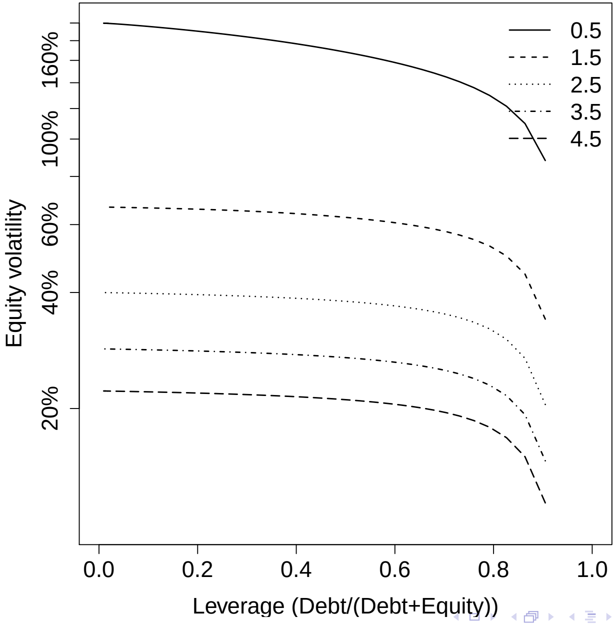
A weak link in operationalising KMV: σ_E forecasts

- What the model requires is a forecast of σ_E over the coming one year.
- The typical model of volatility forecasts use daily data as input.

  These forecast daily volatility.
- A thoughtful user of KMV scales this daily forecast by $\sqrt{252}$.
- However, this is not necessarily optimal.
- Some alternatives include:
  1. Use the forecasting model to forecast daily values for the next 252 days, and then take an average before scaling by $\sqrt{(}252)$.
  2. If volatility is persistent, then a weighted average may be more optimal.
  3. Misspecification problems of using higher frequency data models for lower frequency requirements (Drost and Nijman, 1993).

     Volatility at lower frequency is lower than that at higher frequency, and forecasts out of daily volatility models can be biased.
- In this paper, we seek to do this optimally.

How sensitive is the DtD estimate to changes in σ_E?



Part IV

Some simple choices

Four plausible measure of volatility

1. Historical volatility (using daily returns)
2. Range(using intra-day price variation)
3. Implied volatility (using options prices data)
4. Exponentially weighted moving average (using daily returns)

Each are computed for a sample of liquid stocks.
Since the DtD is sensitive to both volatility and leverage, the sample is separated into three sub-categories: low, medium and high leverage by the debt-equity ratio.

A simple performance framework

- How do we measure that the forecast is 'working'?
- Performance evaluation:
  1. How well does the measure forecast the next day's volatility measured by returns squared?
     This is the performance of the one-step ahead forecast.
  2. How well does the measure forecast the one year out returns squared?
     This is the performance of the one-year ahead forecast.
- Report the RMSE for each measure, separately across the three different categories of stocks by leverage.
  In this presentation, results are for three stocks: Hindustan Unilever Ltd. (low leverage), Mahindra & Mahindra Ltd. (medium leverage), and Larsen & Toubro Ltd. (high leverage).

Measure 1: Historical volatility (HV)

- The standard deviation of returns is computed over the previous 22 days.
- This is used as the forecast for the next day's variance in the one-step ahead forecast performance measurement.
- This is scaled by 252 as the forecast for the one-year ahead forecast.
  This is also what is used for the prob(def) estimation.

Measure 2: Range volatility

- Range is computed using the daily high and low prices for the stock (Parkinson 1980).
- The formula to compute range is as follows:

$$\text{Range} = \sqrt{\frac{1}{4\log 2}(logH - logL)^2}$$

  Here, H and L refer to the high and low price of the day, respectively.
- Range for $(t-1)$ is used as the forecast for t.
  It is scaled by 252 for the one-year ahead forecast, and used for the DtD estimation.

Measure 3: Implied volatility (IV)

- This is the implied volatility calculated using the options traded on the stock.
- It is analogous to the old CBOE VIX calculation and is the average of 8 near-the-money option IVs, for call and put and for near and next month maturities is taken (Whaley 1993).
- The methodology used in the calculation follows what is used in Grover, Thomas (2012).
- IV for $(t-1)$ is used as the forecast for t in the one-step ahead forecast.
- It is scaled by 252 for the one-year ahead forecast and the DtD estimation.

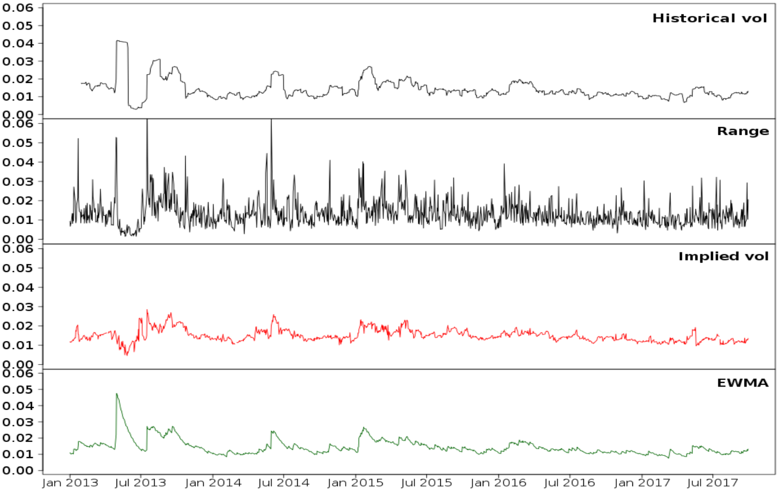
Measure 4: Exponentially weighted moving average, EWMA

- $E(\sigma_{t+1})^2 = \lambda r_t^2 + \lambda^2 r_{t-1}^2 + \lambda^3 r_{t-2}^2 + \ldots + \lambda^{k+1} r_{t-k}^2 + \ldots =$
  $\implies E(\sigma_{t+1})^2 = \lambda r_t^2 + (1-\lambda) E(\sigma_t)^2$
- A higher weight is assigned to more recent returns.
- The smoothing parameter λ is taken to be 0.94 to generate a long time series of exponentially declining weights.

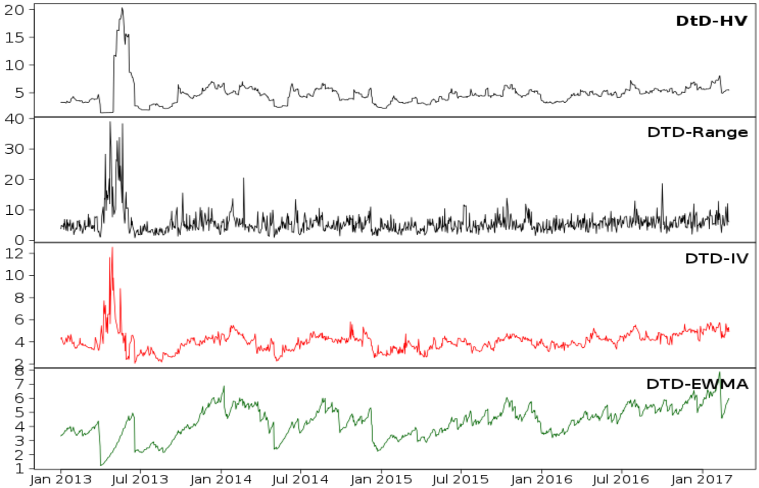
$$\mathrm{EWMA}(t) = \lambda r_t^2 + (1-\lambda)\sigma_{t-1}^2$$

- EWMA(t) is used as the forecast for t in the one-step ahead forecast.
- It is scaled by 252 for the one-year ahead forecast and the DtD estimation.

An example of volatility time series for a low leverage firm (HUL)



DtD for HUL using different volatility measures



Performance evaluation: one-step ahead forecast RMSE

We compare the performance of each of the volatility measures by comparing the estimate at t-1 with actual returns in period t.

$$\mathrm{RMSE} = \sqrt{\sum |\sigma^2_{t-1} - r^2_t|}$$

	RMSE			
	HV	IV	Range	EWMA
HUL	11.8	12.3	10.5	12.1
M&M	13.2	14.8	12.4	13.2
L&T	13.7	15.2	11.3	14.0

Range performs ‘best’ in this setting but the Range-based DtD is also the most volatile.

Part V

This is a non-standard volatility forecasting problem

There is a lot of machinery available

- A great deal of more knowledge about volatility forecasting is in place:
  - ARCH family, deterministic forecasts
  - Implied volatility: forward looking
  - New measures – Realised Volatility, Range
- The most common use case for volatility forecast is in options trading and for calculating the Value at Risk for capital requirements.
  1. For options trading, the most active options are 1-2 months from maturity
  2. For VaR calculations, the typical horizons are from 1 day out to a fortnight out.
- Hence the prime instinct in this domain is to forecast volatility over horizons going out to 1–2 months.
- For the DtD problem, we require 12 month forecasts.
- Another area in finance which requires longer horizon volatility forecasts is portfolio management, where the forecasting horizon can go from one quarter out to five years out.

How does this requirement change our thinking?

- Perhaps vol does behave like an ARMA model.
  It can get shocked, but the average over the next one year will be quite stable and it will be the long-run average.
- In this case, we have no problem!
  We just estimate σ_E using the maximum history and ignore local effects.
- On the other hand, we know that shocks to volatility are extremely persistent!
  While the IGARCH model is not correct, it performs surprising well. (Sarma, Shah and Thomas, 2003)
- Perhaps the local effects persist out into the bulk of the coming year. Then when working out the average for one year, the local effects matter.

A possible middle ground?

Let us say:

H_t	The model's view of variance as of today
$\bar{H}$	The volatility computed over all available data
η_t	$\alpha H_t + (1-\alpha)\bar{H}$

- H_t is our best estimate for today, which varies a lot through time, and
- $\bar{H}$ is the time-invariant maximal-span estimate, and then we have $\eta_t(\alpha)$ which is a linear combination.

Can we do some empirical experiments to discover what is a good α?

Design a test rig

- We see N stocks, daily returns data
- Let's pick K timepoints in the life of each stock
- At each timepoint, we know the true volatility from t to $t + 365$
- We use these true values to compare how alternative forecasters are faring.

Susan Thomas
susant@igidr.ac.in
http://ifrogs.org