Data Informed VC

The objective of this research is to try to give subjective private company valuations an objective basis. The techniques below can be used to create an estimate for the performance of private companies in a specific remit, by assimilating private company valuations with public company valuations.

Sector drift and volatility (μ_M and σ_M) can be approximated and used under the assumption that at any given time t, a company can be valuated at V_t over continuous time. This can be used to probabilistically forecast company valuation **in between funding rounds**.

The model proposed below is entirely designed with the hope to minimise the number of companies required to make an informed approximation of sector drift and volatility.

The data required for these estimations is simply a big list of companies in a remit M, a post money valuation after a fundraising round, and a pre-money valuation before the next funding round, as well as the time in between the events.

Model Core Assumptions

1. σ_M and μ_M are time independent
2. Valuation is a continuous geometric brownian

The statistical approach below is aimed at building probability distributions for σ_M and μ_M.

Stochastic model core

At the core of our model, we have the subjective-ish value of a business V_t in time-slices excluding the occasional dramatic fundraising/loss events (V_t may represent business value in stable operating periods).

$$dV_t \approx V_t(\mu_M dt + \sigma_M dW_t)$$

$$V_t = V_0 \exp\left[\left(\mu_M - \frac{1}{2}\sigma_M^2\right)t + \sigma_M W_t\right]$$

This gives us the time dependent distribution of V_t:

$$V_t \sim \textbf{LogNormal}\left(\log V_0 + \left(\mu_M - \frac{1}{2}\sigma_M^2\right)t, \ \sigma_M t\right)$$

So the probability density function is given by

$$p_V(v_0, v_t, t) = \frac{1}{v_t \sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log x - \log v_0 - \frac{1}{2}t(\mu_M - \frac{1}{2}\sigma_M^2)^2}{\sigma_M t}\right)\right]$$

So our goal is to estimate parameters μ_M, σ_M from market data. Using Scaled Beta distributions, we create prior probability distributions for volatility and drift with fixed hyperparameters.

$$\hat{\mu} \sim \mathbf{Beta}_{I_\mu}(\alpha_\mu, \beta_\mu)$$

$$\hat{\sigma} \sim \mathbf{Beta}_{I_\sigma}(\alpha_\sigma, \beta_\sigma)$$

$$\hat{\theta} = (\hat{\mu}, \hat{\sigma})$$

The support of each beta distribution given in the subscript (I_μ and I_σ) which are intervals on $\mathbf{R}$, which should be scaled relative to the time scale.

Bayesian inference

These distribution choices are almost arbitrary, with comparably low kurtosis and obeying $\hat{\sigma} \in I_\sigma$ and $\hat{\mu} \in I_\mu$

Lets say we make observe some data $\mathcal{D} = (v_0, v_\tau, \tau)$, comprising of pre-money valuation v_τ after time τ from the previous post money valuation v_0. By Bayes' theorem, we may update our belief upon observing data $\mathcal{D}$:

$$\mathbf{P}(\theta \mid \mathcal{D}) = \frac{\mathbf{P}(\mathcal{D}, \theta)}{\mathbf{P}(\mathcal{D})} = \frac{\mathbf{P}(\mathcal{D} \mid \theta)\, \mathbf{P}(\theta)}{\mathbf{P}(\mathcal{D})}$$

Where $p(\mathcal{D})$ is given by

$$\mathbf{P}(\mathcal{D}) = \iint_{I_\sigma \times I_\mu} \mathbf{P}(\mathcal{D} \mid \mu, \sigma)\, p_{\hat{\mu}}(\mu)\, p_{\hat{\sigma}}(\sigma)\, d\mu\, d\sigma$$

$$= \iint_{I_\sigma \times I_\mu} p_V(v_0, v_\tau, \tau)\, p_{\hat{\mu}}(\mu)\, p_{\hat{\sigma}}(\sigma)\, d\mu\, d\sigma$$

Which can be a difficult computation.

So for a series of observations of various companies we collect data $\mathcal{D} = (\mathcal{D}_i)_{i=1}^n$ we can iteratively update our probability model.

$$\mathbf{P}(\theta \mid \mathcal{D}_1, \ldots, \mathcal{D}_n) = \mathbf{P}(\theta) \prod_{i=1}^n \frac{\mathbf{P}(\mathcal{D}_i \mid \theta, \mathcal{D}_1, \ldots, \mathcal{D}_{i-1})}{\mathbf{P}(\mathcal{D}_i \mid \mathcal{D}_1, \ldots, D_{i-1})}$$

So for each company, simply rinse and repeat the process: posterior becomes the prior and so on. By the end, we have a probability distribution for the drift and volatility in the valuations of private companies, which can be used to gauge how promising/popular a specific remit of companies is.

Limitations and Issues

Of course this model requires good data. This involves supplying it with a representative sample of companies, including ones that go bust before being able to do another funding round.