Empirical Asset Pricing

Johnew Zhang

April 16, 2018

Contents

1 Introduction

1.1 Risk Free Investment

Let i_t be the nominal interest rate. Then the return is $(1 + i_t) = R^f_{t+1}$. Holding period return on a pure discount bond which pays \$1 at $t + n$, $P^{(n)}_t$. Therefore the return can be calculated as

$$R_{t+1} = \frac{P^{(n-1)}_{t+1}}{P^{(n)}_t}$$

1.2 Stock

Stock return P_t is $\frac{P_{t+1}+D_{t+1}}{P_t} = \frac{\text{Payoff}}{\text{Price}}$. The rates of return are $R_{t+1} - 1 = \exp(r_{t+1})$ where r_{t+1} is the continuous compounded rate of return. $r^{(n)}_{t+1} = \log P^{(n-1)}_{t+1} - \log P^{(n)}_t$.

1.3 Option

If you invested in an option with strike K, the payoff is

$$\max(0, S_{t+1} - K)$$

and the return is

$$\frac{\max(0, S_{t+1} - K)}{C_t}$$

where

$$C_t = \text{cost today for the call option}$$

1.4 Forward

Now let $S_t = \frac{\$}{€}$ and $F_{t,n}$ = Forward Exchange Rate at t for $t+n$. The payoff on purchasing € forward is

$$S_{t+m} - F_{t,m}$$

Invest \$1 in € money market where there is a risk free of $(1 + i^{€}_t)$.

1. Convert \$1 to € to get $\frac{1}{S_t}$
2. Invest $\frac{1}{S_t}(1 + i^{€}_t)$
3. Convert to \$, $\frac{S_{t+1}(1+i^{€}_t)}{S_t}$ in return. It is although exposed to the foreign exchange risk.

We can eliminate the uncertainty by selling the € interest forward. Then

$$\frac{F_{t,1}}{S_t}(1 + i^{€}_t) = \text{known first period return}$$

This should be equal to the risk return and it is called covered interest rate parity

$$1 + i^{\$}_t = \frac{F_{t,1}}{S_t}(1 + i^{€}_t)$$

1.5 Euler Equation

The Euler equation for investor has the intuition that the marginal benefit of return on investment is equal to the marginal cost of the forgone consumption. Secondly, the price consumption level $P_t = \frac{\$}{\text{General Goods}}$ where $\frac{1}{P_t}$ is the goods sacrificed. Then

$$\frac{1}{P_t} MU_t = \text{utility marginal cost}$$

so

$$R_{t+1} = \$ \text{ in the future}$$

$$\frac{R_{t+1}}{P_{t+1}} = \text{goods in the future}$$

$$\frac{R_{t+1}}{P_{t+1}}\beta MU_{t+1} = \text{MU in the future discounted to present}$$

Then

$$\frac{1}{P_t}MU_t = E_t\left[\beta MU_{t+1}\frac{R_{t+1}}{P_{t+1}}\right]$$

is the Euler equation. In time series class, we let the MU be the $C_t^{-\gamma}$. Here, we know the left hand side today so we can divide it into the expectation. Thus we have

$$1 = E_t\left[\beta\frac{MU_{t+1}}{MU_t}\frac{R^\$_{t+1}}{P_{t+1}/P_t}\right] = E[m^\$_{t+1}R^\$_{t+1}]$$

where $m^\$_{t+1}$ is the pricing kernel or stochastic discount factor for all assets.

$$\frac{R^\$_{t+1}}{(1+\pi_{t+1})} = \text{real return}$$

$$\pi_{t+1} = \frac{P_{t+1} - P_t}{P_t} \text{ is the rate of inflation}$$

For real returns, we have the real pricing kernel, that is $\beta\frac{MU_{t+1}}{MU_t} = m_{t+1}$ while the nominal pricing kernel is $m^\$_{t+1} = \frac{m_{t+1}}{(1+\pi_{t+1})}$. Now we can write

$$E[m_{t+1}R_{t+1}] = 1$$

1.6 Notes Regarding the Above Equation

Risk free return can be written as

$$R^f_{t+1} = \frac{1}{E[m_{t+1}]}$$

because we can let the return to be the risk free return and solve for it. However, due to Jensen's inequality,

$$E[R^f_{t+1}] = E\left[\frac{1}{E_t[m_{t+1}]}\right] \neq \frac{1}{E[m_{t+1}]}$$

For other assets,

$$E_t[m_{t+1}R_{t+1}] = 1$$

has the covariance decomposition as the following

$$C_t(m_{t+1}, R_{t+1}) = E_t[m_{t+1}R_{t+1}] - E_t[m_{t+1}]E_t[R_{t+1}]$$

$$E_t[R_{t+1}] = \frac{1}{E_t[m_{t+1}]} - \frac{1}{E_t[m_{t+1}]}C_t(m_{t+1}, R_{t+1}) = R^f_{t+1} - R^f_{t+1}C_t(m_{t+1}, R_{t+1})$$

The expected excess return is $-R^f_{t+1}C_t(m_{t+1}, R_{t+1})$. Here the covariance is negative.

$$E[R_{t+1} - R^f_{t+1}] = -\frac{1}{E[m_{t+1}]}C(m_{t+1}, R_{t+1})$$

$$\frac{E[R_{t+1} - R^f_{t+1}]}{\sigma(R_{t+1})} = -\rho(m_{t+1}, R_{t+1})\frac{\sigma(m_{t+1})}{E[m_{t+1}]}$$

where the left side is the Sharpe Ratio. The maximum Sharpe Ratio is when the asset is perfectly negatively correlated to the stochastic discount factor. In this case the largest is $\frac{\sigma(m_{t+1})}{E[m_{t+1}]}$. If our utility function is CRRA $= \frac{C_t^{1-\gamma}}{1-\gamma}$ and $MU = C_t^{-\gamma}$ and $m_{t+1} = \beta\frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} = \beta\exp(-\gamma\Delta c_{t+1})$ where $c_{t+1} = \log C_t$. The variance of m_{t+1} larger, the greater variance of the consumption on growth or large γ.

1.7 Minimum Second Moment Asset

Consider an asset $R^m_{t+1} = \frac{m_{t+1}}{E[m^2_{t+1}]}$. Then we know that

$$E\left[m_{t+1}\frac{m_{t+1}}{m^2_{t+1}}\right] = 1$$

This is the minimum second moment asset.

Proof. Let $R^z_{t+1} = \frac{Z_{t+1}}{E[m^2_{t+1}]}$ is arbitrary asset and we need to show $E[{R^z_{t+1}}^2] \geq E[{R^m_{t+1}}^2]$.

$$E[m_{t+1}R^2_{t+1}] = 1$$

$$E[m_{t+1}Z_{t+1}] = E[m^2_{t+1}]$$

$$E[(R^z_{t+1} - R^m_{t+1})^2] \geq 0$$

$$E[Z^2_{t+1}] - 2E[Z_{t+1}m_{t+1}] + E[m^2_{t+1}] > 0$$

$$E[Z^2_{t+1}] - 2E[m^2_{t+1}] + E[m^2_{t+1}] > 0$$

$$E[Z^2_{t+1}] \geq E[m^2_{t+1}]$$

$$E[{R^m_{t+1}}^2] = \frac{E[m^2_{t+1}]}{E[m^2_{t+1}]^2}$$

$$E[{R^z_{t+1}}^2] = \frac{E[Z^2_{t+1}]}{E[m^2_{t+1}]^2}$$

Thus

$$E[{R^z_{t+1}}^2] \geq E[{R^m_{t+1}}^2]$$

□

1.8 Conditional CAPM

In this section, we are going to get $E_t[R^i_{t+1} - R^f_{t+1}] = \beta_{it}E_t[R^b_{t+1} - R^f_{t+1}]$ where

$$\beta_{it} = \frac{C_t(R^i_{t+1}, R^b_{t+1})}{V_t(R^b_{t+1})}$$

$$R^{min}_{t+1} = \frac{m_{t+1}}{E_t[m_{t+1}]^2}$$

Define a benchmark return

$$R^b_{t+1} = \omega_t R^{min}_{t+1} + (1-\omega_t)R^f_{t+1}$$

where the conditional variance of this benchmark return is

$$V_t(R^b_{t+1}) = \omega_t^2 V_t(R^{min}_{t+1}) = \omega_t^2[E_t[{R^{min}_{t+1}}^2] - E_t[R^{min}_{t+1}]^2]$$

$$E_t[R^i_{t+1} - R^f_{t+1}] = -R^f_{t+1}C_t(m_{t+1}, R^i_{t+1})$$

$$V_t(R^b_{t+1}) = \omega_t^2\left[\frac{E_t[m^2_{t+1}]}{(E_t[m^2_{t+1}])^2} - \frac{(E_t[m_{t+1}])^2}{(E_t(m^2_{t+1}))^2}\right]$$

Then

$$E_t[R^i_{t+1} - R^f_{t+1}] = -R^f_{t+1}\omega_t^2\left[\frac{E_t[m^2_{t+1}]}{(E_t[m^2_{t+1}])^2} - \frac{(E_t[m_{t+1}])^2}{(E_t(m^2_{t+1}))^2}\right]\frac{C_t(R^i_{t+1}, R^b_{t+1})}{V_t(R^b_{t+1})}$$

We know that

$$\frac{\omega_t}{E_t[m^2_{t+1}]}C_t(m_{t+1}, R^i_{t+1}) = C_t\left(\omega_t\frac{m_{t+1}}{E_t[m^2_{t+1}]}, R^i_{t+1}\right) = C(R^b_{t+1}, R^i_{t+1})$$

.

$$E_t[R^i_{t+1} - R^f_{t+1}] = -R^f_{t+1}\omega_t\left[1 - \frac{E_t[m_{t+1}]^2}{E_t[m^2_{t+1}]}\right]\frac{C_t(R^i_{t+1}, R^b_{t+1})}{V_t(R^b_{t+1})}$$

Multiply this by $-R^f_{t+1}\omega_t$ into the bracket and add subtract R^f_{t+1} inside. Thus above equation is just

$$= \left[R^f_{t+1} - R^f_{t+1} - \omega_t R^f_{t+1} + \omega_t R^f_{t+1}\frac{E_t[m_{t+1}]^2}{E_t[m^2_{t+1}]}\right]\beta_{it}$$

$$= \left[(1-\omega_t)R^f_{t+1} + \omega_t E_t[R^{min}_{t+1}] - R^f_{t+1}\right]\beta_{it}$$

$$E_t[R^i_{t+1} - R^f_{t+1}] = \beta_{it}E_t[R^b_{t+1} - R^f_{t+1}]$$

Hence the conditional CAPM holds for benchmark returns.

$$R^b_{t+1} = \omega_t R^{min}_{t+1} + (1-\omega_t)R^f_{t+1}$$

$$R^{min}_{t+1} = \frac{m_{t+1}}{E_t[m^2_{t+1}]}$$

1.9 Systematic vs Idiosyncratic or Unsystematic Risk

Systematic risk implies a covariance with m_{t+1} giving rise to risk premium and unsystematic uncertainty is uncorrelated with m_{t+1} that does not give rise to the risk premium. Rational expectations

$$R^i_{t+1} = E_t[R^i_{t+1}] + \varepsilon^i_{t+1}$$

$$m_{t+1} = E_t[m_{t+1}] + \varepsilon^m_{t+1}$$

$$C_t(m_{t+1}, R^i_{t+1}) = C_t(\varepsilon^m_{t+1}, \varepsilon^i_{t+1}), \text{ source of risk premium}$$

$$\varepsilon^i_{t+1} = \beta^i_t\varepsilon^m_{t+1} + \nu^i_{t+1}$$

where $\nu^i_{t+1} \perp \varepsilon^m_{t+1}$

$$\beta^i_t = \frac{C_t(\varepsilon^m_{t+1}, \varepsilon^i_{t+1})}{V_t(\varepsilon^m_{t+1})}$$

$$E_t[R^i_{t+1} - R^f_{t+1}] = -C_t(m_{t+1}, R^i_{t+1})R^f_{t+1} = -R^f_{t+1}C_t(\varepsilon^m_{t+1}, \varepsilon^i_{t+1}) = -R^f_{t+1}\beta^i_t V_t(\varepsilon^m_{t+1}) = -\beta^i_t\lambda_t$$

where $\lambda_t = R^f_{t+1}V_t(\varepsilon^m_{t+1})$

Only systematic risk is priced and price of risk is λ_t.

1.10 Factor Model

$R^p_t = \omega' R_t$ and $R_t \sim N(\mu, \Sigma)$.

Suppose we maximize portfolio using the mean-variance maximizer

$$\max_\omega\{\omega'\mu - \frac{\gamma}{2}\omega'\Sigma\omega\}$$

Then the FOC condition is

$$\mu - \gamma\Sigma\omega = 0$$

$$\omega = \frac{1}{\gamma}\Sigma^{-1}\mu$$

.

$$E[R^p_t] = \frac{1}{\gamma}\mu'\Sigma^{-1}\mu$$

$$V(R^p_t) = \frac{1}{\gamma}\mu'\Sigma^{-1}\Sigma\Sigma^{-1}\mu\frac{1}{\gamma} = \frac{1}{\gamma^2}\mu'\Sigma^{-1}\mu$$

$$\text{Sharpe Ratio} = \frac{E[R^p_t]}{V(R^p_t)} = \frac{\frac{1}{\gamma}\mu'\Sigma^{-1}\mu}{\frac{1}{\gamma}\sqrt{\mu'\Sigma^{-1}\mu}} = \sqrt{\mu'\Sigma^{-1}\mu}$$

1.10.1 Hansen-Jagannathan Bounds

It relates. the standard deviation of pricing kernel to asset returns.

$$E_t[m_{t+1}(R^i_{t+1} - R^f_{t+1})] = 0$$

Let $R^e_{t+1} =$ excess return. We don't observe m_{t+1} but consider theoretical regression of m_{t+1} onto $1, R^e_{t+1}$ some vector

$$m_{t+1} = \alpha + \beta' R^e_{t+1} + \varepsilon_{t+1}$$

$$\beta = \Sigma^{-1} C(m_{t+1}, R^e_{t+1}) = \Sigma^{-1}(E_t[m_{t+1}R_{t+1}] - E_t[m_{t+1}]E_t[R^e_{t+1}]) = -\Sigma^{-1}E[m_{t+1}]E[R^e_{t+1}]$$

$$V(m_{t+1}) \geq V(\beta' R^e_{t+1}) = \beta'\Sigma\beta = E[m_{t+1}]\mu'\Sigma^{-1}\Sigma\Sigma\Sigma^{-1}\mu$$

$$\frac{\sigma(m_{t+1})}{E[m_{t+1}]} \geq \sqrt{\mu'\Sigma^{-1}\mu}$$

1.11 Risk Neutral Probabilities

$$E_t[m_{t+1}R^i_{t+1}] = 1$$

consider S states of the world with probability $\pi_t(s)$

$$\sum_{s=1}^{S} \pi_t(s) m_{t+1}(s) R^i_{t+1}(s) = 1$$

$$\sum_{s=1}^{S} \pi_t(s) m_{t+1}(s) = \frac{1}{R^f_{t+1}} = E_t[m_{t+1}]$$

Let's define $\pi^*_t(s) = \frac{\pi_t(s)m_{t+1}(s)}{E_t[m_{t+1}]}$. These are all positive and sum to 1. They are like probabilities.

$$\sum_{s=1}^{S} \pi^*_t(s) = 1$$

We have

$$\sum_{s=1}^{S} \frac{\pi_t(s)m_{t+1}(s)R_{t+1}(s)}{E_t[m_{t+1}]} = \frac{1}{E_t(m_{t+1})} = R^f_{t+1}$$

$$\sum_{s=1}^{S} \pi^*_t(s) R^f_{t+1} = R^f_{t+1}$$

$$E^Q_t(R^i_{t+1}) = R^f_{t+1}$$

where $\pi^*_t(s)$ is the risk neutral probability.

1.12 International Implications

$$E_t[m^{\$}_{t+1}R^{\$}_{t+1}] = 1$$

where

$$m^{\$}_{t+1} \text{ is USD SDF}$$

$$E_t[m^{€}_{t+1}R^{€}_{t+1}] = 1$$

where

$$m^{\$}_{t+1} \text{ is EURO SDF}$$

In the Euler equation theory,

$$m^{\$}_{t+1} = \frac{\beta\mu'(C^{\$}_{t+1})(1/p^{\$}_{t+1})}{u'(c^{\$}_t)(1/p^{\$}_t)}$$

and

$$m^{€}_{t+1} = \frac{\beta\mu'(C^{€}_{t+1})(1/p^{€}_{t+1})}{u'(c^{€}_{t})(1/p^{€}_{t})}$$

Return in \$ to \$1 invested in European assets

$$R^{\$}_{t+1} = \frac{1}{S_t} R^{€}_{t+1} S_{t+1}$$

where $S_t = \$/€$

$$E_t[m^{\$}_{t+1} R^{€}_{t+1} \frac{S_{t+1}}{S_t}] = 1$$

if markets are complete, then

$$m^{€}_{t+1} = m^{\$}_{t+1} \frac{S_{t+1}}{S_t}$$

take log

$$s_{t+1} - s_t = \log m^{€}_{t+1} - \log m^{\$}_{t+1}$$

where the left hand side is the continuously compounded rate of appreciation of euro vs dollar.

$$V(s_{t+1} - s_t) = V(\log m^{€}_{t+1}) + V(\log m^{\$}_{t+1}) - 2\rho(\log m^{\$}_{t+1}, \log m^{€}_{t+1})\sigma_{\log m^{€}_{t+1}}\sigma_{\log m^{\$}_{t+1}}$$

Solve for

$$-\rho = \frac{V(s_{t+1} - s_t) - V(\log m^{€}_{t+1}) + V(\log m^{\$}_{t+1})}{2\sigma_{\log m^{€}_{t+1}}\sigma_{\log m^{\$}_{t+1}}}$$

$$\rho = 1 - \frac{0.1^2}{2\sigma_{\log m^{€}_{t+1}}\sigma_{\log m^{\$}_{t+1}}}$$

if $\sigma_{\log m^{€}_{t+1}} = \sigma_{\log m^{\$}_{t+1}} = 0.5$ and suppose $V(s_{t+1} - s_t) = 0.1^2$.

Then $\rho = 0.98$.

2 Factor Models

$$E_t[M_{t+1} R^e_{i,t+1}] = 0$$

$$E_t[M_{t+1}] = \frac{1}{R^f_t}$$

$$M_{t+1} = a - b f_{t+1}$$

where $a = 1$.

Therefore,

$$E[M_{t+1}, R^e_{i,t+1}] = E[M_{t+1}]E[R^e_{i,t+1}] + Cov(M_{t+1}, R^e_{t+1})$$

$$E[R^e_{i,t+1}] = -\frac{Cov(M_{t+1}, R^e_{i,t+1})}{E[M_{t+1}]} = b\frac{Cov(f_{t+1}, R^e_{i,t+1})}{E[M_{t+1}]}$$

If f_{t+1} is the expected excess returns, then

$$E[f_{t+1}] = b\frac{V(f_{t+1})}{E[M_{t+1}]} \text{ or } \frac{b}{E[M_{t+1}]} = \frac{E[f_{t+1}]}{V(f_{t+1})}$$

$$E[R^e_{i,t+1}] = \frac{Cov(f_{t+1}, R^e_{i,t+1})}{V(f_{t+1})} E[f_{t+1}]$$

Now what is $\frac{Cov(f_{t+1}, R^e_{i,t+1})}{V(f_{t+1})}$? The beta for this factor.

$$R^e_{i,t+1} = \alpha_i + \beta_i f_t + \varepsilon_{i,t+1}$$

$$\hat{\alpha}_i = E[R^e_{i,t+1}] - \hat{\beta}_i E[f_t] = 0$$

Here, ε_{it} is homoskedastic and normally distribution and

$$\frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)} = \text{t-distribution}$$

$$R^e_{i,t+1} = \alpha_1 + \beta_1 f_t + \varepsilon_{1,t+1}$$

$$\vdots$$

$$R^e_{N,t+1} = \alpha_N + \beta_N f_t + \varepsilon_{N,t+1}$$

The OLS GMM orthogonality condition is

$$E\left[\varepsilon_{1,t+1}\begin{pmatrix}1\\ f_t\end{pmatrix}\right] = 0$$

$$\vdots$$

$$E\left[\varepsilon_{N,t+1}\begin{pmatrix}1\\ f_t\end{pmatrix}\right] = 0$$

Hence we will write

$$g_T(b) = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\varepsilon_t\\ \varepsilon_t f_t\end{pmatrix} = E_T\left[\begin{pmatrix}\varepsilon_t\\ \varepsilon_t f_t\end{pmatrix}\right]$$

and $b = (\alpha, \beta)'$.

$$\hat{\alpha} = E_T[R^e_t] = \hat{\beta} E_T[f_t]$$

$$\hat{\beta} = \frac{E_T[(R^e_t - E_T(R^e_t))f_t]}{E_T((f_t - E_T(f_t))f_t]} = \frac{Cov_T(R^e_t, f_t)}{Var_T(f_t)}$$

Then by GMM,

$$\sqrt{T}(\hat{b} - b) \to N(0, (D'_T S_T^{-1} D_T)^{-1})$$

where $D_T = \bar{V}_b g_T(b)$ and

$$S_T = \begin{pmatrix} E_T(\varepsilon_t\varepsilon'_t) & E_T(\varepsilon_t\varepsilon'_t f_t)\\ E_T(\varepsilon_t\varepsilon'_t f_t) & E_T(\varepsilon_t\varepsilon'_t f_t^2)\end{pmatrix}$$

$$D_T = \nabla g_T(b) = \nabla\begin{pmatrix}\frac{1}{T}\sum_{t=1}^{T}(R^e_t - \alpha - \beta f_t)\\ \frac{1}{T}\sum_{t=1}^{T}(R^e_t f_t - \alpha f_t - \beta f_t^2)\end{pmatrix} = \begin{pmatrix}-I_N & -I_N E_T(f_t)\\ -I_N E_T(f_t) & -I_N E_T(f_t^2)\end{pmatrix} = -\begin{pmatrix}-1 & -E_T(f_t)\\ -E_T(f_t) & -E_T(f_t^2)\end{pmatrix}\otimes I_N$$

Hence

$$(D'_T S_T^{-1} D_T)^{-1} = D_T^{-1} S_T D_T^{-1}$$

$$\begin{aligned} Var(\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}) &= \frac{1}{T}\left\{\left(\begin{pmatrix}1 & E_T(f_t)\\ E_T(f_t) & E_T(f_t^2)\end{pmatrix}\otimes I_N\right)^{-1}\begin{pmatrix}E_T(\varepsilon_t\varepsilon'_t) & E_T(\varepsilon_t\varepsilon'_t f_t)\\ E_T(\varepsilon_t\varepsilon'_t f_t) & E_T(\varepsilon_t\varepsilon'_t f_t^2)\end{pmatrix}\left(\begin{pmatrix}1 & E_T(f_t)\\ E_T(f_t) & E_T(f_t^2)\end{pmatrix}\otimes I_N\right)^{-1}\right\}\\ &= \hat{\Omega}\end{aligned}$$

$$\hat{\alpha}'\hat{\Omega}^{-1}_{\alpha\alpha}\hat{\alpha} \sim \chi^2(N)$$

If ε_t is serially uncorrelated and conditionally homoskedastic, then

$$S_T = \begin{pmatrix}1 & E_T(f_t)\\ E_T(f_t) & E_T(f_t^2)\end{pmatrix}\otimes\Sigma$$

where $\Sigma = E_T[\varepsilon_t\varepsilon'_t]$ Let

$$A = \begin{pmatrix}1 & E_T(f_t)\\ E_T(f_t) & E_T(f_t^2)\end{pmatrix}$$

$$(D_T'S_T^{-1}D_T)^{-1} = [-A \otimes I_N]^{-1}[A \otimes \Sigma][-A \otimes I_N]^{-1} = A^{-1}AA^{-1} \otimes \Sigma = A^{-1} \otimes \Sigma$$

where

$$A^{-1} = \frac{1}{E_T(f_t^2) - E_T(f_t)^2} \begin{pmatrix} 1 & E_T(f_t) \\ E_T(f_t) & E_T(f_t^2) \end{pmatrix} = \frac{1}{Var_T(f_t)} \begin{pmatrix} 1 & E_T(f_t) \\ E_T(f_t) & E_T(f_t^2) \end{pmatrix}$$

$$Var(\hat{\alpha}) = \frac{1}{T}\frac{E_T(f_t^2)}{Var(f_t)}\Sigma = \frac{1}{T}\frac{[Var(f_t) + E(f_t^2)]}{Var(f_t)}\Sigma = \frac{1}{T}\left[1 + \frac{E_T(f_t)^2}{Var(f_t)}\right]\Sigma$$

$$\hat{\alpha}'(Var(\hat{\alpha}))^{-1}\hat{\alpha} = T\left(1 + \frac{E_T(f_t)^2}{Var(f_t)}\right)^{-1}\alpha'\Sigma^{-1}\alpha$$

Thus the GRS test for a small sample is

$$\frac{T}{T-2}\frac{T-N-1}{N}\left(1 + \frac{E_T(f_t)^2}{Var_T(f_t)}\right)^{-1}\hat{\alpha}'(Var(\hat{\alpha}))^{-1}\hat{\alpha} \sim F_{N,T-N-1}$$

2.1 Non-traded Factor

What if the factor is not traded, then

$$E[m_t R_t^e]$$

where $m_t = 1 - bf_t$

$$E[R_t^e] = bCov(R_{t+1}^e, f_{t+1}) = bVar(f_{t+1})Var^{-1}(f_{t+1})Cov(R_{t+1}^e, f_{t+1})$$

where $\lambda = bVar(f_{t+1})$ and $\beta = Var^{-1}(f_{t+1})Cov(R_{t+1}^e, f_{t+1})$.

Then $E[R_{1,t}^e] = \beta_i\lambda$ but $\lambda \neq E[f_t]$. This leads to $\alpha_i's$ are not zero. The goal is to test

$$E[R_{it}^e] = \lambda\beta_i$$

that is a cross sectional testing where time series regression is ran and if the betas have a linear relationship with the return.

$$g_T(b) = \begin{pmatrix} E_T[R_t^e - a - \beta f_t] \\ E_T[(R_t^e - a - \beta f_t)f_t] \\ E_T[R_t^e - \beta\lambda] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

GMM

$$g_T(b)Wg_T(b)$$

$$\frac{\partial g_T(b)'}{\partial b}Wg_T(b) = D_T'Wg_T(b)$$

where $b = (\alpha, \beta, \lambda)$. This is equivalent to solve

$$ag_T(b) = 0, a = D_T'W$$

(1) Efficient GMM

$$S_T = E_T\begin{pmatrix} \varepsilon_t\varepsilon_t' & \varepsilon_t\varepsilon_t'f_t & \varepsilon_t(R_t^e - \beta\lambda)' \\ \varepsilon_t\varepsilon_t'f_t & \varepsilon_t\varepsilon_t'f_t^2 & \varepsilon_t f_t(R_t^e - \beta\lambda)' \\ (R_t^e - \beta\lambda)\varepsilon_t' & (R_t^e - \beta\lambda\beta_t\varepsilon_t' & (R_t^e - \beta\lambda)(R_t^e - \beta\lambda)' \end{pmatrix}$$

$$D_T = \frac{\partial g_T}{\partial b} = E_T\begin{pmatrix} -I_N & -f_t \otimes I_N & 0 \\ -f_t \otimes I_N & -f_t \otimes I_N & 0 \\ 0 \otimes I_N & -\lambda \otimes I_N & \beta \end{pmatrix} = \left[-\begin{pmatrix} 1 & E_T(f_t) \\ E_T(f_t) & E_T(f_t^2) \\ 0 & \lambda \end{pmatrix} \otimes I_N : \begin{pmatrix} 0 \\ 0 \\ \beta \end{pmatrix}\right]$$

then

$$\sqrt{T}\left[\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\lambda} \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \\ \lambda \end{pmatrix}\right] \to N(0, (D_T'S_T^{-1}D_T)^{-1})$$

$$Tg_T(\hat{\beta})'S_T^{-1}g_T(\hat{\beta}) \sim \chi^2(3N - (2N-1))$$

(2)

$$a = \begin{pmatrix} I_{2N} & 0 \\ 0 & \beta' \end{pmatrix}$$

$$\hat{\lambda} = (\hat{\beta}'\hat{\beta})^{-1}\hat{\beta}' E_T(R_t^e)$$

$$a g_T(b) = 0 \implies \hat{\beta}, \hat{a} \text{ are OLS estimates}$$

$$\sqrt{T}(\hat{b} - b) \to N(0, \hat{\Omega})$$

where

$$\hat{\Omega} = (D_T' W D_T)^{-1} D_T' W S_T W D_T (D_T' W D_T)^{-1} = (aD)^{-1} a S_T a' (aD)'^{-1}$$

(3)

$$a = \begin{pmatrix} I_{2N} & 0 \\ 0 & \beta'\Sigma^{-1} \end{pmatrix}$$

(GLS)

$$\hat{\lambda} = (\beta'\Sigma^{-1}\beta)^{-1}.\beta'\Sigma^{-1} E_T[R_t^e]$$

(4) LM approach

$$R_{it.}^e = \beta_i \lambda + \beta_i(\hat{f}_t - \mu_p) - \varepsilon_{it}$$

$$E\left[\begin{pmatrix} \varepsilon_t \\ \varepsilon_t f_t \end{pmatrix}\right] = 0$$

$$E[f_t] - \mu_f = 0$$

Use efficient GMM

2.2 Fama MacBeth

$$\begin{pmatrix} R_{1t} \\ \vdots \\ R_{NT} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix} \hat{\lambda}_t + \begin{pmatrix} \hat{\alpha}_{1t} \\ \vdots \\ \hat{\alpha}_{Nt} \end{pmatrix}$$

(1) Estimate β_i's from time series regression.

(2) Run cross-sectional regression

(3)

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T} \hat{\lambda}_t$$

$$\hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T} \hat{\alpha}_{it}$$

$$\sigma^2(\hat{\lambda}) = \frac{1}{T}\sum_{J=-K}^{K} w_j E_T[(\hat{\lambda}_t - \hat{\lambda})(\hat{\lambda}_t - \hat{\lambda})']$$

When $K = 0$,

$$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_N \end{pmatrix} \to Cov(\hat{\alpha}) = \frac{1}{T}\sum_{t=1}^{T}(\hat{\alpha}_t - \hat{\alpha})(\hat{\alpha}_t - \hat{\alpha})'$$

Then

$$\hat{\alpha}' Cov(\hat{\alpha})^{-1}\hat{\alpha} \sim \chi^2(N-1)$$

$$R^e_{i,t\cdot} = \beta_i\lambda + \varepsilon_{it}, i = 1, \cdots, N; t = 1, \cdots T$$
$$R^e_t = \beta\lambda + \varepsilon_t$$
$$R_t = \begin{pmatrix} R^e_1 \\ \vdots \\ R^e_T \end{pmatrix} = \begin{pmatrix} \beta \\ \vdots \\ \beta \end{pmatrix} \lambda + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

or

$$R = B\lambda + \varepsilon$$

Then

$$\hat{\lambda}_{OLS} = (B'B)^{-1}BR$$
$$B'B = T\beta'\beta$$
$$B'R = \beta'\sum_{t=1}^{T} R_t$$
$$(B'B)^{-1} = \frac{1}{T}(\beta'\beta)^{-1}$$
$$E[\varepsilon\varepsilon'] = \Omega$$
$$Cov(\hat{\lambda}_{OLS}) = (B'B)^{-1}B'\hat{\Omega}B(B'B)^{-1}$$

where

$$B'\Omega B = T\beta'\Sigma\beta = \frac{1}{T}(\beta'\beta)^{\cdot -1}.T\beta'\Sigma\beta\frac{1}{T}(\beta'\beta)^{-1} = \frac{1}{T}(\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1}$$

estimate Σ with

$$E_T[\hat{\varepsilon}_t\hat{\varepsilon}'_t]$$

and

$$\hat{\varepsilon}_t = R_t - \beta\hat{\lambda}_{OLS}$$

$$E_T(R^e_t) \text{ on } \beta$$
$$E_T(R^e_t) = \beta\lambda + E_T(\varepsilon_t)$$
$$\lambda_{XS} = (\beta''\beta)^{-1}\beta' E_T(R^e_t)$$
$$\sigma^2(\hat{\lambda}_{XS} = (\beta'\beta)^{-1}\beta' Cov(E_T(\varepsilon_t))\beta(\beta'\beta)^{-1}$$
$$Cov(E_T(\varepsilon_t)) = \frac{1}{T}\Sigma$$
$$\sigma^2(\hat{\lambda}_{XS}) = \frac{1}{T}((\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1}$$

In particular for FM.

$$\lambda_t = (\beta'\beta)^{-1}\beta' R^e_t$$
$$\hat{\lambda}_{FM} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_t = (\beta'\beta)^{-1}\beta'\frac{1}{T}\sum_{t=1}^{T} R^e_t = (\beta'\beta)^{-1}\beta' E_T(R^e_t)$$
$$Cov(\hat{\lambda}_{FM}) = \frac{1}{T}Cov(\hat{\lambda}_t) = \frac{1}{T}(\beta'\beta)^{-1}\beta'\Sigma((\beta'\beta)^{-1}\beta')'$$
$$\hat{\lambda}_t = (\beta'\beta)^{-1}\beta' R^e_t = (\beta'\beta)^{-1}\beta'(\beta\lambda + \varepsilon_t) = \lambda + (\beta'\beta)^{-1}\beta'\varepsilon_t$$
$$\frac{1}{T}\hat{\lambda}_t = \hat{\lambda}_{FM} = \lambda + \frac{1}{T}\sum_{t=1}^{T}(\beta'\beta)^{-1}\beta'\varepsilon_t$$
$$Var(\hat{\lambda}_{FM} - \hat{\lambda}) \to N(0, \Omega)$$

$$\Omega = \frac{1}{T}\sum_{t=1}^{T}(\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1}$$

To sum up,

$$m_t = a + bf_t$$

$$E[(a + b'\tilde{f}_t)R_{i,t}] = 1$$

$$aE[R_{it}] = 1 - b'E[\tilde{f}_t R_{it}]$$

$$E[R_{it}] = \frac{1}{a} - \frac{b'}{a}E[\tilde{f}_t R_{it}] = \frac{1}{a} - E[R_{it}f_t']E[\tilde{f}_t\tilde{f}_t']^{-1}E[\tilde{f}_t\tilde{f}_t']\frac{b}{a} = \frac{1}{a} + \beta_i'\lambda$$

where $\beta_i = E[\tilde{f}_t\tilde{f}_t']^{-1}E[f_t R_{it}]$. $\tilde{f}_t = f_t - E[f_t]$ and $\lambda = -E[\tilde{f}_t f']\frac{b}{a}$

That is

$$E[R_{it}] = \gamma + \beta_i'\lambda \iff m_{,t} = a + b'\tilde{f}_t$$

2.3 Horse Race - Multi-factor Factor Model

Given f_{1t} factor, do you need f_{2t} factor. SDF $m_t = a + b_1' f_{1t} + b_2' f_{2t}$ use a set of test asset and GMM to estimate $\hat{b}_1$ and $\hat{b}_2$.

$$\hat{b}_2' var(\hat{b}_2)^{-1}\hat{b}_2 \sim \chi^2(\#b_2)$$

This is Wald test - estimate the alternative and test the zero restrictions. You can also use likelihood ratio test using GMM

- Estimate with unrestricted model J^{UR}.
- Estimate the restricted model with S_{UR}^{-1} as weighting matrix

  $$m_t = a + b_1' f_{1t}$$

  to get J^{RES}

  $$TJ_T^{Res} - TJ_T^{UR} \sim \chi^2(\text{\# of restriction})$$

2.3.1 Testing the CAPM - Sharpe-Lintner

Miller and Scholes (1972) say β's are measured with error. The estimation bias will shift down the security market line. They propose to form portfolios on $\hat{\beta}_i$ and re-estimate the β's of portfolios.

$$R_{it}^e = \beta_i R_{mt}^e. + \varepsilon_{it}, \sigma(R_{mt}^e) \approx 15\%(pa), \sigma(\varepsilon_{it}) \approx 20 - 40\%(pa)$$

$$\frac{\bar{R}_{it}^e}{\sigma_t/\sqrt{T}} \text{ cannot reject all average returns are equal and equal to zero}$$

and CAPM worked with the β sorted portfolios for 13 years.

Size Rolf Banz found the small firm effect sorted by market equity (decile)

$$E[R_{small,t}^e] = \alpha + \beta_{small}E[R_{mt}^e], \alpha > 0$$

David Booth, Founder of Dimensional Fund Advisor (DFA), capitalizes on this result. SMB is the small minus big portfolio.

Book to Market is book equity divided by market equity. High book-to-market firms have high returns - value stocks (Ben Graham, "Value Investing") and low book-to-market are growth (glamor) stock with low $E[R]$. HML factor is high-low.

Multi-factor model Merton (1973) suggests

$$E_t[R^e_{i,t+1}] = \beta_{it} E_t[R^e_{m,t+1}] + \sum_{j=1}^{N} \delta_{i,j,t} E_t[R^e_{j,t+1}]$$

where $\delta_{ij,t} =$ exposure of asset i to risk factor j. N factors describes the changes in investment opportunities. Need to find factors that spans the relevant multi-factor efficient set.

$$R_{it} = \sum_{j=1}^{N} \beta_{ij} R_{jt} + \varepsilon_{it}, i = 1, \cdots, N = 3 - 5$$

(The premise is does asset manager require stock picking ability?)

Another Anomaly Novy-Marx (2013) operating profitability

$$OP_t = \frac{Revenue_t - CGS_t - Interest_t - SGA_t}{\text{End of Period Book Value}_t}$$

Within size quintile, increasing OP implies higher average return. (RMW, robust versus weak)

Investment

$$\frac{\text{Growth in Asset (fiscal } t-1)}{\text{Total Assets } (t-2)}$$

Firms with high investment and have low return (CMA, conservative minus aggressive)

Momentum is a big anomaly sorted on $R_{t-12,t-1}$ invested at t. Winners outperform losers. This gives UMD portfolio (ups minus down)

11 Anomalies Stambaugh and Yuan, "Mispricing Factors", RFS 2017 developed a 4 factor model with MKT, SMB (different from FF), 2 factors capture 11 anomalies related to FF3.

1. Net stock issuance, Ritter (1991). Equity issuers underperform non-issuers. Annual log change split adjusted shares outstanding.
2. Composite equity issue, Daniel, Titman (2006). Growth in total market equity minus rate of return per share over 12 months.
3. Accruals, Sloan (1996). High accrual firms worse than low accrual.
4. Net operating assets. Hirschleifer et. al (2004). Operating assets minus operating liability over the total assets. Low predict low return.
5. Asset growth, Cooper et. al (2008).

$$\frac{AT_t - AT_{t-1}}{AT_{t-1}}$$

   with four month lag. High bad.
6. Investment to assets, Titman, Wei, Xing (2004), Xing (2008).

$$\frac{\text{Gross PPE}_t + \Delta Inventory}{AT_{t-1}}$$

   with 4 month lag. High bad.
7. Distress, Campbell, Hilscher and Szilagyi (2008). Model failure probability. High probability of failure implies low return.
8. O-Score, Ohlson (1980). Static model of bankruptcy probability. High probability of bankruptcy implies lower return.
9. Momentum, Jegadesesh and Titman (1993). Carhart (1997) added UMD to FF model.
10. Gross profitability, Novy-Marx (2013). High profit implies high return.

11. Return on Assets, FF (2006).

$$\frac{\text{Income before extraordinary items}}{ATQ}$$

They sorted them into 2 groups and ranked on anomaly, take 10% most over-valued and 10% most under-valued.

Method 1 Run $R^e_{it} = \alpha_i + b_i MKT_t + c_i SMB_t + u_{it}$. Compute 11×11 covariance matrix of u_{it}. "Clustering Method", Ward (1963) is used. Two groups came out, MGMT (cluster of management group) and PERF (cluster of performance).

Method 2 Generate a z-score on the anomaly ranking

$$z_j = \frac{s_j - \bar{s}}{\sigma_s}, s_j = \text{ raw rank}$$

Then do the same regression as method 1.

MGMT are net stock issuance, composite equity issuance, accruals, net operating assets, asset growth, investment to assets. PERF are distress, O-Score, momentum, gross profit, ROA. They equal-weight within the two clusters, P_1 and P_2 for each firm. Mispricing factors are 2×3 sorts.

1. Size median NYSE
2. Independently sort on P_1 and P_2, 20% and 80% combined NYSE, AMEX, NASDAQ.

SMB is small 60% unused minus big 60% unused.

3 Options

A call option is a contract that gives the buyer the right but not the obligation to buy asset at predetermined strike price X at maturity European (at maturity) and American (prior at t). If S_t is the asset price at time t, the payoff at $T = \max(S_T - X, 0)$. c_t = option price and $m_{t,T}$ = SDF between t and T. Then

$$c_t = E_t[m_{t,T} \max(0, S_T - X)]$$

Returns is

$$\frac{\max(0, S_T - X)}{c_t}$$

A put option is a contract that buyers buy the right. to sell asset at. X at T to the seller of the option (European option). The payoff is $\max(0, X - S_T)$.

Straddle Purchase of a call option and put option at the same strike price. The payoff is $\max(0, S_T - X) + \max(0, X - S_T)$. This is a bet on volatility.

To write an option is to sell it. Writing out-of-money options generate cash flow (put: $X < S_t$, call: $X > S_t$). At the money is $X = S_t$ or forward price. In the money call $X < S_t$ and in the money put $X > S_t$.

3.1 Put-Call Parity

$$c_t = E_t[m_{t,T} \max(0, S_T - X)], p_t = E_t[m_{t,T} \max(0, X - S_T)]$$

Buy a call and sell a put,

$$c_t - p_t = E_t[m_{t,T}(S_T - X)]$$

For a non-dividend paying stock,

$$E_t[m_{t,T} S_T] = S_t, E_t[m_{t,T} X] = \frac{X}{1 + i_t}$$

Therefore, the put-call parity for non-dividend paying stock is

$$c_t - p_t = S_t - \frac{X}{1 + i_t}$$

3.2 No Arbitrage Binomial Pricing (One-Period)

Find a portfolio of stock and borrowing that replicates the payoff on call. Suppose

$$s_t = \$150, \text{interest rate} = 0.5\%$$

What is the price of the call option with strike \$152? In addition, we know that $s_{t+1} = 145$ or 155.

Let's buy z shares of stock and borrow y dollars. Then the cost of our portfolio is $150z - y$. In the bad state, we will have $145z - 1.005y = 0$ and in the good state, we have $152z - 1.0005y = 3$. Then we can solve the linear equations to get $y = \$43.28, z = 0.3$. Then by no arbitrage, we have $c = \$150 \times 0.3 - \$43.28 = \$1.72$.

You can solve the option prices recursively through a binomial tree. Note that we did not know the probabilities of up and down.

- Just with the magnitudes.
- The replicating portfolio involves leverage. Expected return on the call > $E[R]$ (10-30 times)

3.3 Introduction to Continuous Time, Stochastic Processes

Discrete time random walk:

$$z_t - z_{t-1} = \varepsilon_t$$
$$Var(z_{t+2} - z_t) = 2Var(z_{t+1} - z_t)$$

The variance scales with time directly. We can define $z_{t+\delta} - z_t \sim N(0, \Delta)$ for a small δ. increments in $z(t)$ are independent of $z(t)$.

$$dz_t = z_{t+\delta} - z_t, \text{for arbitrarily small } \delta$$

The stochastic integral defines the level of z_t relative to z_0.

$$z_t - z_0 = \int_{\delta=0}^{t} dz_\delta$$

Because the variance scale with time, the standard deviation scales with the square root of time. The standard deviation describes a typical size change of a normally distributed random variable so $z_{t+\delta} - z_t$ has typical size $\sqrt{\Delta}$. Therefore, $\frac{z_\delta - z_t}{\Delta}$ has typical size $\frac{1}{\sqrt{\Delta}}$. Thus, sample path of z_t are continuous but not differentiable. Now, $E_t[dz_t] = 0$ since dz_t is the forward increment and variance of dz_t is $E_t[dz_t^2] = dt$ where dt is the limit as Δ gets small. Here, dz_t is the brownian motion process and is the building block of all diffusion models.

3.4 Processes

$$dx_t = \mu(\cdot)dt + \sigma(\cdot)dz_t$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are function of t information set (all conditional on time t)

Random walk with a drift is

$$dx_t = \mu dt + \sigma dz_t$$

Take the integral both sides

$$x_t - x_0 = \mu(t-0) + \sigma(z_t - z_0) \implies x_t = x_0 + \mu t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2 t)$$

AR(1): $x_t = (1-\rho)\mu + \rho x_{t-1} + \varepsilon_t$ where μ is the long run mean. Subtract x_{t-1} from both sides

$$x_t - x_{t-1} = -(1-\rho)(x_{t-1} - \mu) + \varepsilon_t = -\phi(x_t - \mu) + \varepsilon_t$$
$$dx_t = -\phi(x_t - \mu)dt + \sigma dz_t$$

This is called Ornstein-Uhlenbeck process. Square root process

$$dx_t = -\phi(x_t - \mu)dt + \sigma\sqrt{x_t}dz_t$$
$$E_t[\sigma\sqrt{x_t}dz_t)^2] = \sigma^2 x_t dt$$

volatility varies with x_t and as x_t goes to 0, the drift pulls x_t toward μ. If $\mu > 0$ and $\phi > 0$, then $2\phi\mu > \sigma^2$ guarantees x_t always positive. "Feller condition".

3.4.1 Pricing Processes

$$dp_t = p_0 \mu d - t + p_t \sigma dz_t$$

Return

$$\frac{dp_t}{p_t} = \mu dt + \sigma dz_t$$

Generally

$$\frac{dp_t}{p_t} = \mu(\cdot)(\cdot)dt + \sigma dz_t$$

Local mean is $\mu(\cdot)dt$ and local variance

$$E_t[(dp_t/p_t - \mu(\cdot)dt)^2] = \sigma(\cdot)^2 dt$$

3.4.2 It$\hat{o}$'s Lemma

If $y_t = f(x_t)$ and x_t follows a diffusion process what is y_t?

Take a second-order Taylor series,

$$dy_t = \frac{\partial f}{\partial x} dx_t + \frac{1}{2}\frac{\partial^2 f}{\partial x_t^2} dx_t^2$$

$$dx_t = \mu_x dt + \sigma_x dz_t$$

$$(dx_t)^2 = (\mu_x^2 (dt)^2 + 2\mu_x \sigma_x dt dz_t + \sigma_x^2 dz_t^2)$$

Set $(dt)^2 = 0$, $dtdz_t = 0$.They go to 0 faster than dt

$$dy_t = \frac{\partial f}{\partial x} dx_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_x^2 dt = (\frac{\partial f}{\partial x}\mu_x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_x^2)dt + \frac{\partial f}{\partial x}\sigma_x dz_t$$

(this is like the Jensen's inequality)

Now let's apply this to our call option.

$$c_t = C(S_t, t)$$

$$dc_t = c - td_t + c_s dS_t + \frac{1}{2} c_{ss} dS_t^2$$

We need a continuous time discount factor:

$$p_t \Lambda_t = E_t[\Lambda_{t+1} p_{t+1}]$$

$$\frac{\partial \Lambda_t}{\Lambda_t} = -rdt - \frac{\mu - r}{\sigma} dz_t - \sigma_w dw_t$$

$$dz_t = \text{brownian motion driving stocks}$$

$$dw_t = \text{orthogonal to } dz_t$$

$$\frac{dS_t}{S_t} = \mu dt + \sigma dz_t$$

$$E_t \frac{dS_t}{S_t} - rdt = -E_t \frac{d\Lambda_t}{\Lambda_t}\frac{dS_t}{S_t}$$

$$(\mu - r)dt = -E_t(-rdt - \frac{\mu - r}{\sigma} dz_t - dw_t)(\mu dt + \sigma dz_t) = E[\frac{\mu - r}{\sigma}\sigma dz_t^2] = (\mu - r)dt$$

$$c_0 = \text{price of a call option at time } 0$$

Then

$$
\begin{aligned}
c_0 &= E_0 \frac{\Lambda_T}{\Lambda_0} \max(0, S_T - X) \\
&= \int \frac{\Lambda_T}{\Lambda_0} \max(0, S_T - X) df(\Lambda_T, S_T)
\end{aligned}
$$

$$
c_0 = S_0 N(d_1) - X e^{-rT} N(d_2)
$$

where $N(k) = \int_{-\infty}^{k} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$ and

$$
d_1 = \frac{\log S_0/X + (r + \sigma^2/2)T}{\sigma(\sqrt{T})}
$$

$$
d_2 = \sigma\sqrt{T} - d_1
$$

3.5 Coval and Shumway

$$
\begin{aligned}
E_t[m_T R_T] = 1 &\implies E_t[R_T - R_T^f] = -\frac{Cov_t(m_T, R_T)}{E_t[m_T]} \\
R_T^c = \text{ return on a call option } &= \frac{\max(0, S_T - X)}{c_t} \\
E_t[R_T^c - R_{TT}^f] &= -Cov(m_T, \frac{\max(0, S_T - X)}{c_t})
\end{aligned}
$$

On right hand side, move c_t out and multiply by S_t/S_t and move $\frac{1}{S_t}$ in, then

$$
\begin{aligned}
&= -Cov_t\left(\frac{m_T}{E_t(m_T)}, \frac{\max(0, S_T - X)}{S_t}\right)\frac{S_t}{c_t} \\
&= -Cov_t\left(\frac{m_T}{E_t[m_T]}, \frac{S_T}{S_t}\right)\frac{S_t}{c_t} + Cov_t\left(\frac{m_T}{E_t[m_T]}, \frac{\max(0, X - S_T)}{S_t}\right)\frac{S_t}{c_t} \\
(R_T^c - R_T^f) &= E_t[R_T - R_T^f]\frac{S_t}{C_t} + Cov_t\left(\frac{m_T}{E_t[m_T]}, \frac{\max(0, X - S_T)}{S_t}\right)\frac{S_t}{c_t}
\end{aligned}
$$

4 Term Structure of Interest Rate

$$
\begin{aligned}
P_t^{(N)} &= \text{Price of a zero coupon bond paying \$1 at } t + N \\
\ln(P_t^{(N)}) &= p_t^{(N)} \\
y_t^{(N)} &= \text{Continuously compounded yield to maturity} \\
P_t^{(N)} &= \exp(-N y_t^{(N)}) \implies p_t^{(N)} = -N y_t^{(N)}
\end{aligned}
$$

Zero coupon yields are the basis of discounting. Risk-free bond pays C and \$1 at maturity, then

$$
P_t = \sum_{j=1}^{N} \frac{CF_{t+j}}{\exp(j y_t^{(j)})}, CF_{t+j} = C, j = 1, \cdots, N-1, CF_{t+N} = 1 + C
$$

Holding period return on N period bond

$$
HPR_{t+1}^{(N)} = \frac{P_{t+1}^{(N-1)} - 1}{P_t^{(N)}}
$$

$$hpr_{t+1}^{(N)} = p_{t+1}^{(N-1)} - p_t^{(N)} = -(N-1)y_{t+1}^{(N-1)} + Ny_t^{(N)}$$

Forward rates are implicit in the term structure at what rate can you contract today to borrow or lend starting at N period in the future for 1 period.

$$F_t^{N\to N+1} = \frac{P_t^{(N)}}{P_t^{(N+1)}} = \frac{\exp(-Ny_t^{(N)})}{\exp(-(N+1)y_t^{(N+1)})}$$

$$f_t^{N\to N+1} = p_t^{(N)} - p_t^{(N+1)} = (N+1)y_t^{(N+1)} - Ny_t^{(N)} = y_t^{(N+1)} + N(y_t^{(N+1)} - y_t^{(N)})$$

Forward rate above yield when yields are upward sloping.

$$f_t^{N\to N+1} + N(y_t^{(N)}) = (N+1)y_t^{(N+1)} = \text{ return on \$1 invested for } N+1 \text{ period}$$

so

$$\begin{aligned}
p_t^{(N)} &= (p_t^{(N)} - p_t^{(N-1)}) + (p_t^{(N-1)} - p_t^{(N-2)}) + \cdots + (p_t^{(2)} - p_t^{(1)}) + p_t^{(1)} \\
&= -f_t^{N-1\to N} - f_t^{N-2\to N-1} - \cdots - f_t^{1\to 2} - y_t^{(1)} \\
p_t^{(N)} &= -\sum_{j=0}^{N-1} f_t^{j\to j+1} \\
P_t^{(N)} &= \exp\left(-\sum_{j=0}^{N-1} f_t^{j\to j+1}\right)
\end{aligned}$$

Price of N-period zero coupon bond is discounted value of \$1 when discount rates are the forward rates. There are three ideas about how yields are determined

1. N-period yield is the average of expected future 1 period yields plus risk premium:

$$y_t^{(N)} = \frac{1}{N}E_t[y_t^{(1)} + y_{t+1}^{(1)} + \cdots + y_{t+N-1}^{(1)}] + rpy_t^{(N)}$$

   where rp has Jensen's inequality as well as risk.

2. Forward rate is the expected spot rates plus risk premium

$$f_t^{N\to N+1} = E_t[y_{t+N}^{(1)}] + rpf_t^{(N)}$$

3. The expected holding period return is the risk free rate plus the risk premium

$$E_t[hpr_{t+1}^{(N)}] = y_t^{(1)} + rpr_t^{(N)}$$

Ignore the risk term to start:

$$y_t^{(1)} = 3\%, y_t^{(2)} = 6\%$$

1. $6\% = \frac{1}{2}(3\% + E_t[y_{t+1}^{(1)}])$ and $E_t[y_{t+1}^{(1)}] = 12\% - 3\% = 9\%$
2. $f_t^{1\to 2} = 2y_t^{(2)} - y_t^{(1)} = 12\% - 3\% = 9\%$
3. $E_t[p_{t+1}^{(1)} - p_t^{(2)}] = 3\%$ and $E_t[-y_{t+1}^{(1)} + 2\times 6\%] = 3\%$ so $E_t[y_{t+1}^{(1)}] = 9\%$

We must take the risk into account.

$$E_t[M_{t+1}HPR_{t+1}^{(N)}] = E_t[M_{t+1}\frac{P_{t+1}^{(N-1)}}{P_t^{(N)}}] = 1$$

$$P_t^{(N)} = E_t[M_{t+1}P_{t+1}^{(N-1)}]$$

$$P_{t+1}^{(N-1)} = E_t[M_{t+2}P_{t+2}^{(N-2)}]$$
$$P_t^{(N)} = E_t[M_{t+1}M_{t+2}P_{t+2}^{(N-2)}]$$
$$\vdots$$
$$P_t^{(N)} = E_t[\prod_{j=1}^{N} M_{t+j}], \text{ the term structure provides lots of information about distribution of } M_{t+j}$$

If M_{t+j} is lognormal, then $P_t^{(N)}$ is lognormal

$$P_t^{(N)} = E_t[M_{t+1}P_{t+1}^{(N-1)}] = \exp\left\{E_t(m_{t+1}) + \frac{1}{2}V_t(m_{t+1}) + E_t[p_{t+1}^{(N-1)}] + \frac{1}{2}V_t(p_{t+1}^{(N-1)}) + C_t(m_{t+1}, p_{t+1}^{(N-1)})\right\}$$

$$E_t[M_{t+1}] = \exp(-y_t^{(1)})$$
$$\exp(E_t[m_{t+1}] + \frac{1}{2}V_t(m_{t+1})) = \exp(-y_t^{(1)})$$
$$E_t[m_{t+1}] + \frac{1}{2}V_t(m_{t+1}) = -y_t^{(1)}$$

Hence

$$p_t^{(N)} = -y_t^{(1)} + E_t[p_{t+1}^{(N-1)}] + \frac{1}{2}V_t(p_{t+1}^{(N-1)}) + C_t(m_{t+1}, p_{t+1}^{(N-1)})$$
$$E_t[p_{t+1}^{(N-1)} - p_t^{(N)}] - y_t^{(1)} = -\frac{1}{2}V_t(p_{t+1}^{(N-1)}) - C_t(m_{t+1}, p_{t+1}^{(N-1)})$$

where the above says the expected excess holding period rate of return is equal to the Jensen's inequality term and the risk premium term. We call the right hand side as $rpr_t^{(N)}$

4.1 Canonical Affine Model

$$m_{t+1} = -y_t^{(1)} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}$$

where $\varepsilon_{t+1} \sim N(0, I_k)$ is k-dimensional vector of risks necessary to price bonds: level, slope and curvature as driving processes and λ_t is the prices of risk. Let X_ts are state variables and $X_{t+1} = \mu + \Phi X_t + \Sigma\varepsilon_{t+1}$ where Σ is the square root of X_{t+1}

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$
$$y_t^{(1)} = \delta_0 + \delta_1' X_t$$
$$P_t^{(N)} = \exp(A_N + \beta_N' X_t)$$

where A_N is a constant and B_N is the constant parameters.

Use the method of undermined coefficient to solve for recursions A_N and B_N as functions of $\mu, \Phi, \lambda_0, \lambda_1, \delta_0, \delta_1, \Sigma$.

$$p_t^{(N)} = -y_t^{(1)} + E_t[p_{t+1}^{(N+1)}] + \frac{1}{2}V_t(p_{t+1}^{(N+1)}) + C_t(m_{t+1}, p_{t+1}^{(N+1)})$$

$$\begin{aligned}
A_N + B_N'X_t &= -\delta_0 - \delta_1'X_t + E_t[A_{N-1} + B_{N-1}'(\mu + \Phi X_t)] + \frac{1}{2}B_{N-1}'\Sigma\Sigma'B_{N-1} + C_t(-\lambda_t'\varepsilon_{t+1}, B_{N-1}'\Sigma\varepsilon_{t+1}) \\
&= -\delta_0 - \delta_1'X_t + E_t[A_{N-1} + B_{N-1}'(\mu + \Phi X_t)] + \frac{1}{2}B_{N-1}'\Sigma\Sigma'B_{N-1} - B_{N-1}'\Sigma(\lambda_0 + \lambda_1 X_t) \\
A_N &= -\delta_0 + A_{N-1} + B_{N-1}'\mu + \frac{1}{2}B_{N-1}'\Sigma\Sigma'B_{N-1} - B_{N-1}'\Sigma\lambda_0 \\
B_N' &= -\delta_1' + B_{N-1}'\Phi - B_{N-1}'\Sigma\lambda_1
\end{aligned}$$

We can define $A_1 = -\delta_0$ and $B_1 = -\delta_1$. Then

$$A_N - A_{N-1} = A_1 + B'_{N-1}(\mu - \Sigma\lambda_0) + \frac{1}{2}B'_{N-1}\Sigma\Sigma' B_{N-1}$$

$$B'_N = B'_1 + B'_{N-1}(\Phi - \Sigma\lambda_1)$$

$$p_t^{(N)} = -Ny_t^{(N)}, \text{we have term structures in terms of } A_N, B_N$$

Note, $\mu - \Sigma\lambda_0, \Phi - \Sigma\lambda_1$ are risk adjustments to X processes.

4.2 Campbell-Schiller: Yield Spreads and Interest Rate Movements: A Bird's Eye View

This paper strongly demonstrates the need for time variant risk premium.

$$E_t[p_{t+1}^{(N-1)} - p_t^{(N)}] - y_t^{(1)} = constant$$

a type of expectation hypothesis

1. Start with excess rate of return on n period bond held $m < n$ periods.

$$E_t[-(n-m)y_{t+m}^{(n-m)} + ny_t^{(n)}] - my_t^{(m)} = C$$

   Add and subtract $my_t^{(m)}$

$$(n-m)y_t^{(m)} - (n-m)E_t[y_{t+m}^{(n-m)}] + m(y_t^{(n)} - y_t^{(m)}) = C$$

$$E_t[y_{t+m}^{(n-m)} - y_t^{(n)}] = C + m(y_t^{(n)} - y_t^{(m)})$$

2.