- How to estimate parameters?
- How to calculate standard errors of the estimated parameters?
- How to calculate standard errors of the pricing errors?
- How to test the model?

Chapter12 Regression-based Tests of linear Factor Models

- Time-Series Regressions
- Cross-Sectional Regressions
- Fama-MacBeth Procedure

Data structure

- N Assets， T moments R_t^{ei}

$$\begin{bmatrix} R_1^{e,1} & R_2^{e,1} & \cdots & R_{T-1}^{e,1} & R_T^{e,1} \\ R_1^{e,2} & R_2^{e,2} & \cdots & R_{T-1}^{e,2} & R_T^{e,2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ R_1^{e,N-1} & R_2^{e,N-1} & \cdots & R_{T-1}^{e,N-1} & R_T^{e,N-1} \\ R_1^{e,N} & R_2^{e,N} & \cdots & R_{T-1}^{e,N-1} & R_T^{e,N} \end{bmatrix}$$

12.1 Time-Series Regressions

- Express the asset pricing model as:

$$E\left(R^{ei}\right) = \beta_i \lambda$$

The Betas are defined by regression coefficients:

$$R_t^{ei} = \alpha_i + \beta_i f_t + \varepsilon_t^i \qquad (12.1)$$

- In the model, the factor is an excess return, the test assets are all excess returns

- The model states:

$$E\left(R^{ei}\right)=\beta_i E\left(f\right) \qquad (12.2)$$

since the factor is also an excess return,

$$E\left(f\right)=1\times\lambda$$

- Comparing (12.2)(12.1),

α_i are equal to the pricing errors

$$E\left(R_t^{ei}\right)=\alpha_i+\beta_i E\left(f_t\right)$$

Black,Jensen and Scholes(1972) suggested: Run time-series regression

1. Estimate the factor risk premium,

$$\hat{\lambda} = E_T(f) = \frac{1}{T}\sum_{t=1}^{T} f_t$$

2. Run time-series regression for each test asset

$$R_t^{ei} = \alpha_i + \beta_i f_t + \varepsilon_t^i, i = 1, 2, \cdots, N$$

3. Use standard OLS formulas for a distribution theory of the parameters, t test

4. Jointly test the pricing errors

- Assuming $E\left(\varepsilon_t^i \varepsilon_t^j\right) \neq 0$, no autocorrelation, homoskedastic , $E\left(\varepsilon_t^i \varepsilon_t^i\right) = \sigma_i^2, E\left(\varepsilon_t \varepsilon_{t-j}\right) = 0, j \neq 0$
- χ^2 test:

$$T\left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2\right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \sim \chi_N^2 \qquad (12.3)$$

$\hat{\Sigma}$ is the residual covariance matrix , I.e., the sample estimate of $E\left(\varepsilon_t \varepsilon_t'\right) = \Sigma$

- F test for finite-sample: (ε are normal)

$$\frac{\text{T-N-1}}{\text{N}}\left[1+\left(\frac{\mathrm{E_T}(f)}{\hat{\sigma}(f)}\right)^2\right]^{-1}\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}\sim F_{N,T-N-1} \qquad (12.4)$$

This is Gibbons, Ross and Shanken (1989) or “GRS” test statistic. This distribution is exact in a finite sample

- The test can also be interpreted as a test whether f is ex ante mean-variance efficient after accounting for sampling error.

$$\frac{T-N-1}{N}\frac{\left(\mu_q/\sigma_q\right)^2-\left(E_T\left(f\right)/\hat{\sigma}\left(f\right)\right)^2}{1+\left(E_T\left(f\right)/\hat{\sigma}\left(f\right)\right)^2}$$

- Multi-factors:

The regression equation is :

$$R^{ei} = \alpha_i + \beta_i^{'} f_t + \varepsilon_t^i$$

The asset pricing model is :

$$E\left(R^{ei}\right) = \beta_i^{'} E\left(f\right)$$

Assuming normal I.I.d. errors , the test is :

$$\frac{T-N-K}{N}\left(1+E_T\left(f\right)^{'}\hat{\Omega}^{-1}E_T\left(f\right)\right)^{-1}\hat{\alpha}^{'}\hat{\Sigma}^{-1}\hat{\alpha}\sim F_{N,T-N-K}$$

Derivation of The Chi Statistic and Distributions with General Errors

- Derive (12.3) as an instant of GMM
- Write the equations for all N assets together in vector from:

$$R_t^e = \alpha + \beta f_t + \varepsilon_t$$

- Use the usual OLS moments:

$$g_T(b) = \begin{bmatrix} E_T\left(R_t^e - \alpha - \beta f_t\right) \\ E_T[\left(R_t^e - \alpha - \beta f_t\right) f_t] \end{bmatrix} = E_T\left(\begin{bmatrix} \varepsilon_t \\ f_t \varepsilon_t \end{bmatrix}\right) = 0$$

- Exactly identify, so the a matrix in $a \cdot g_T\left(\hat{b}\right)=0$ is identity matrix, $a=I$
- GMM estimate is : (OLS estimate)

$$\hat{\alpha} = E_T\left(R_t^e\right) - \hat{\beta}\, E_T\left(f_t\right),$$

$$\hat{\beta} = \frac{E_T\left[\left(R_t^e - E_T\left(R_t^e\right)\right) f_t\right]}{E_T\left[\left(f_t - E_T\left(f_t\right)\right) f_t\right]} = \frac{\operatorname{cov}_T\left(R_t^e, f_t\right)}{\operatorname{var}_T\left(f_t\right)}$$

- The d matrix is :

$$d = \frac{\partial g_T(b)}{\partial b'} = -\begin{bmatrix} I_N & I_N E(f_t) \\ I_N E(f_t) & I_N E(f_t^2) \end{bmatrix} = -\begin{bmatrix} 1 & E(f_t) \\ E(f_t) & E(f_t^2) \end{bmatrix} \otimes I_N$$

- The S matrix is :

$$S = \sum_{j=-\infty}^{\infty} \begin{bmatrix} E(\varepsilon_t \varepsilon_{t-j}^{'}) & E(\varepsilon_t \varepsilon_{t-j}^{'} f_{t-j}) \\ E(f_t \varepsilon_t \varepsilon_{t-j}^{'}) & E(f_t \varepsilon_t \varepsilon_{t-j}^{'} f_{t-j}) \end{bmatrix}$$

- Using the GMM variance formula with a=I:

$$\operatorname{var}\left(\hat{b}\right)=\frac{1}{T}(ad)^{-1}aSa'(ad)^{-1}{}' \qquad (11.4)$$

So:

- $$\operatorname{var}\left(\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\end{bmatrix}\right)=\frac{1}{T}d^{-1}Sd^{-1}{}' \qquad (12.7)$$

- Get the standard formulas by assuming:

1. The errors are uncorrelated over time and homoskedastic
2. The factor and error are independent as well as orthogonal (multi-factors)

$$E\left[f\varepsilon\varepsilon'\right] = E\left[f\right]E\left[\varepsilon\varepsilon'\right]$$

$$E\left[f^2\varepsilon\varepsilon'\right] = E\left[f^2\right]E\left[\varepsilon\varepsilon'\right]$$

- Then the S matrix simplifies to :

$$S=\begin{bmatrix} E(\varepsilon_t\varepsilon_t') & E(\varepsilon_t\varepsilon_t')E(f_t) \\ E(f_t)E(\varepsilon_t\varepsilon_t') & E(\varepsilon_t\varepsilon_t')E(f_t^2) \end{bmatrix}=\begin{bmatrix} 1 & E(f_t) \\ E(f_t) & E(f_t^2) \end{bmatrix}\otimes\Sigma$$

- Plug into (12.7), we obtain:

$$\operatorname{var}\left(\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}\right)=\frac{1}{T}\left(\begin{bmatrix} 1 & E(f_t) \\ E(f_t) & E(f_t^2) \end{bmatrix}^{-1}\otimes\Sigma\right)$$

$$\operatorname{var}\left(\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix}\right)=\frac{1}{T}\frac{1}{\operatorname{var}(f)}\left(\begin{bmatrix} E(f_t^2) & -E(f_t) \\ -E(f_t) & 1 \end{bmatrix}\otimes\Sigma\right)$$

- The variance of $\hat{\alpha}$ is :

$$\operatorname{var}\left(\hat{\alpha}\right)=\frac{1}{T}\left(1+\frac{E(f)^{2}}{\operatorname{var}(f)}\right)\Sigma$$

- This is the tradition formula (12.3)
- By simply calculating (12.7), we can easily construct standard errors and test statistic that do not requires these assumptions.

12.2 Cross-Sectional Regressions

- Start with the K factor model,

$$E\left(R^{ei}\right) = \beta_i^{'}\lambda \quad ,i = 1,2,\cdots,N$$

- The center economics question is why average returns vary across assets
- The model says that average returns should be proportional to betas. (Figure 12.1)

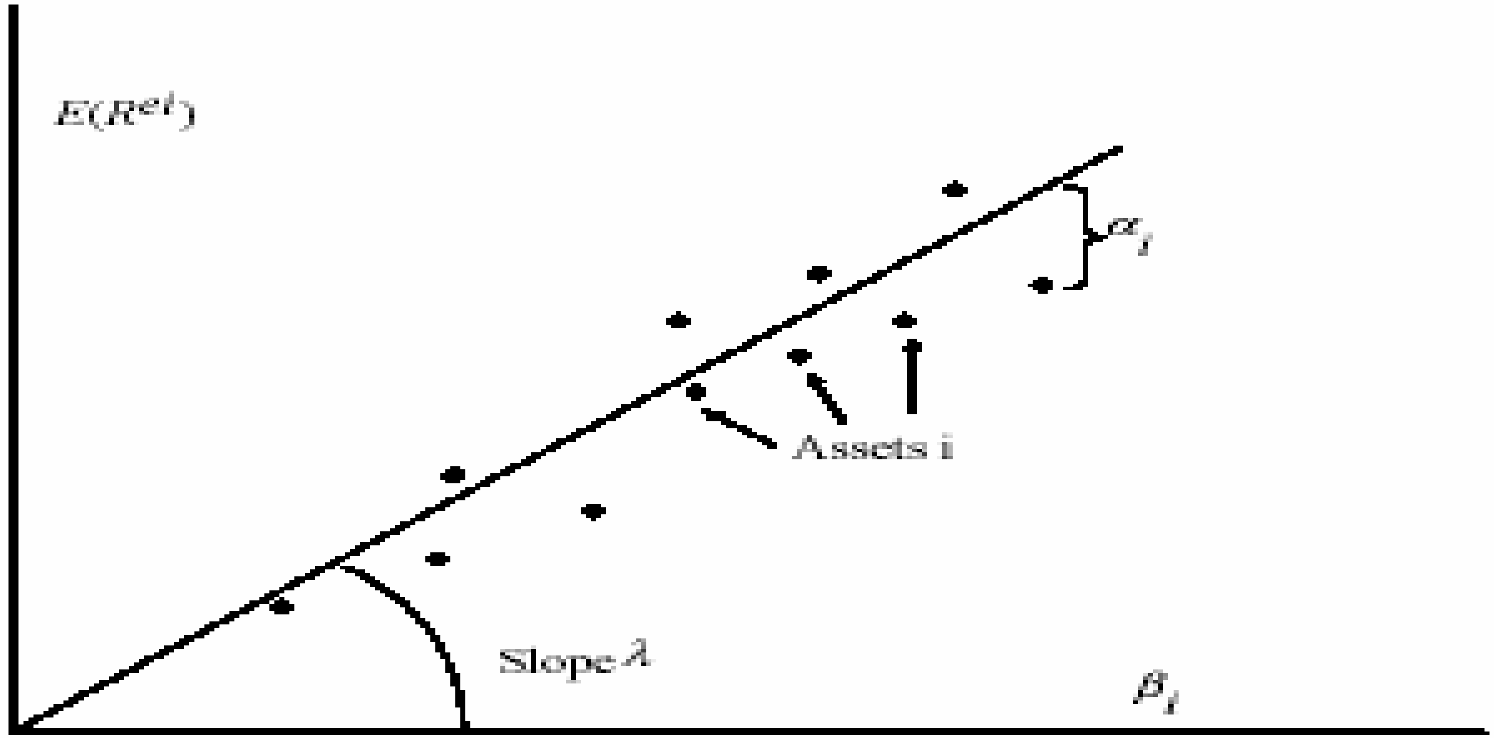

Figure 26. Cross-sectional regression

Run a two-pass regression

- First, find $\hat{\beta}$ from time-series regressions

$$R_t^{ei} = a_i + \beta_i' f_t + \varepsilon_t^i, \qquad t = 1, 2, \cdots, T$$

- Then, estimate $\hat{\lambda}$ from cross-sectional regression.

$$E_T\left(R^{ei}\right) = \beta_i' \lambda + \alpha_i, \qquad i = 1, 2, \cdots, N$$

pricing errors are $\hat{\alpha}_i$

OLS Cross-Sectional Regression

- The OLS cross-sectional estimates are:

$$\hat{\lambda} = (\beta'\beta)^{-1}\beta' E_T(R^e)$$
$$\hat{\alpha} = E_T(R^e) - \hat{\lambda}\beta \qquad (12.11)$$

- Assuming the true errors are I.I.d. over time and independent of the factors.

$$\sigma^2\left(\hat{\lambda}\right) = \frac{1}{T}(\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1} \qquad (12.12)$$

$$\operatorname{cov}\left(\hat{\alpha}\right) = \frac{1}{T}(I-\beta(\beta'\beta)^{-1}\beta')\Sigma(I-\beta(\beta'\beta)^{-1}\beta') \qquad (12.13)$$

- Since,

$$E(\alpha\alpha') = E(\begin{pmatrix}\alpha_1 \\ \vdots \\ \alpha_N\end{pmatrix}\begin{pmatrix}\alpha_1 & \cdots & \alpha_N\end{pmatrix}) = \frac{1}{T^2}E(\begin{pmatrix}\sum_{t=1}^{T}\varepsilon_t^1 \\ \vdots \\ \sum_{t=1}^{T}\varepsilon_t^N\end{pmatrix}\begin{pmatrix}\sum_{t=1}^{T}\varepsilon_t^1 & \cdots & \sum_{t=1}^{T}\varepsilon_t^N\end{pmatrix}$$

$$= \frac{1}{T}E(\begin{pmatrix}\varepsilon_t^1 \\ \vdots \\ \varepsilon_t^N\end{pmatrix}\begin{pmatrix}\varepsilon_t^1 & \cdots & \varepsilon_t^N\end{pmatrix}) = \frac{1}{T}\Sigma$$

$$\sigma^2\left(\hat{\lambda}\right) = (\beta'\beta)^{-1}\beta'\mathrm{cov}[E_T(R^e)]\beta(\beta'\beta)^{-1}$$
$$= (\beta'\beta)^{-1}\beta'\mathrm{cov}[\beta'\lambda + \alpha]\beta(\beta'\beta)^{-1}$$
$$= (\beta'\beta)^{-1}\beta' E(\alpha\alpha')\beta(\beta'\beta)^{-1}$$

- The test statistic is :

$$\hat{\alpha}' \operatorname{cov}(\hat{\alpha})^{-1} \hat{\alpha} \sim \chi^2_{N-1} \qquad (12.14)$$

The degree of freedom is N-1 not N, and N-K for K-factors model

A test of residuals is unusual in OLS regressions.

GLS Cross-Sectional Regression

- Since $E[\alpha\alpha'] = \frac{1}{T}\Sigma$ as the error covariance matrix ,GLS estimates are:

$$\hat{\lambda} = \left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\Sigma^{-1}E_T\left(R^e\right),$$

$$\hat{\alpha} = E_T\left(R^e\right) - \hat{\lambda}\beta$$

- The variance of these estimates are:

$$\sigma^2\left(\hat{\lambda}\right) = \frac{1}{T}\left(\beta'\Sigma^{-1}\beta\right)^{-1}, \qquad (12.15)$$

$$\operatorname{cov}\left(\hat{\alpha}\right) = \frac{1}{T}\left(\Sigma - \beta(\beta'\Sigma^{-1}\beta)^{-1}\beta'\right). \qquad (12.16)$$

- A GLS regression can be understand as a transformation of the space of returns
- The test statistic for the pricing errors

$$T\hat{\alpha}'\Sigma^{-1}\hat{\alpha} \sim \chi^2_{N-1} \qquad (12.17)$$

Correction for the fact that β are estimated

- Since the betas are estimated, the asymptotic standards errors should be corrected (Shanken 1992)

$$\sigma^2\left(\hat{\lambda}_{OLS}\right) = \frac{1}{T}[(\beta'\beta)^{-1}\beta'\Sigma\beta(\beta'\beta)^{-1}(1+\lambda'\Sigma_f^{-1}\lambda)+\Sigma_f]$$

$$\sigma^2\left(\hat{\lambda}_{GLS}\right) = \frac{1}{T}[(\beta'\Sigma^{-1}\beta)^{-1}(1+\lambda'\Sigma_f^{-1}\lambda)+\Sigma_f]$$

(12.18)

Compare to (12.12)(12.15)

- The asymptotic covariance matrix of the pricing errors is :

$$\operatorname{cov}\left(\hat{\alpha}_{OLS}\right) = \frac{1}{T}(I_N - \beta(\beta'\beta)^{-1}\beta')\Sigma(I_N - \beta(\beta'\beta)^{-1}\beta') \times \left(1 + \lambda'\Sigma_f^{-1}\lambda\right) \quad (12.19)$$

$$\operatorname{cov}\left(\hat{\alpha}_{GLS}\right) = \frac{1}{T}\left(\Sigma - \beta\left(\beta'\Sigma^{-1}\beta\right)^{-1}\beta'\right)\left(1 + \lambda'\Sigma_f^{-1}\lambda\right) \quad (12.20)$$

The test statistic in GLS is :

$$T(1 + \lambda'\Sigma_f^{-1}\lambda)\hat{\alpha}'_{GLS}\Sigma^{-1}\hat{\alpha}_{GLS} \sim \chi^2_{N-K}$$

- How important the corrections ?

  in CAPM , $\lambda = E(R^{em})$, so in annual data

$$\lambda^2 / \sigma^2 (R^{em}) \approx (0.08 / 0.16)^2 = 0.25$$

It is too large to ignore,but in monthly interval

$$\lambda^2 / \sigma^2 (R^{em}) \approx 0.25 / 12 \approx 0.02$$

It makes little difference.

- The additive term can be very important, given some assumption ,write (12.19)as follow

$$\sigma^2\left(\hat{\lambda}\right) = \frac{1}{T}[\frac{1}{N}\sigma^2(\varepsilon) + \sigma^2(f)]$$

Even with N=1,most factor model have fairly high R^2 ,so

$$\sigma^2(\varepsilon) < \sigma^2(f)$$

Typical assets numbers N=10 to 50

even if the the Shanken correction can be ignored , it should be included in $\sigma\left(\hat{\lambda}\right)$

Derivation and formulas that do not require I.I.d. errors

- Derivation by GMM, map all effects into GMM framework. The moments are

$$g_T(b) = \begin{bmatrix} E\left(R_t^e - a - \beta f_t\right) \\ E\left[\left(R_t^e - a - \beta f_t\right) f_t\right] \\ E\left(R^e - \beta\lambda\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

It is overidentified, since N extra moment

The ingredients of the recipe:

the parameter vector is :

$$b' = [a', \beta', \lambda]$$

The a matrix is :
$$a = \begin{bmatrix} I_{2N} & \\ & \gamma' \end{bmatrix}$$

The d matrix is

$$d = \frac{\partial g_T(b)}{\partial b'} = \begin{bmatrix} -I_N & -I_N E(f) & 0 \\ -I_N E(f) & -I_N E(f^2) & 0 \\ 0 & -\lambda I_N & -\beta \end{bmatrix}$$

- The S matrix is:

$$S=\sum_{j=-\infty}^{\infty} E\left(\begin{bmatrix} R_t^e - a - \beta f_t \\ \left(R_t^e - a - \beta f_t\right) f_t \\ R_t^e - \beta\lambda \end{bmatrix}\begin{bmatrix} R_{t-j}^e - a - \beta f_{t-j} \\ \left(R_{t-j}^e - a - \beta f_{t-j}\right) f_{t-j} \\ R_{t-j}^e - \beta\lambda \end{bmatrix}'\right)$$

$$=\sum_{j=-\infty}^{\infty} E\left(\begin{bmatrix} \varepsilon_t \\ \varepsilon_t f_t \\ \beta\left(f_t - Ef\right)+\varepsilon_t \end{bmatrix}\begin{bmatrix} \varepsilon_{t-j} \\ \varepsilon_{t-j} f_{t-j} \\ \beta\left(f_{t-j} - Ef\right)+\varepsilon_{t-j} \end{bmatrix}'\right)$$

- Calculate the standard error formula (11.4)(11.5) (pricing errors is last N moment)

- Recover classic formula (12.18) (12.19) (12.20) by adding assumption:

1. the errors are I.I.d. and independent of the factors
2. the factors are uncorrelated over time

Thus:

$$S = \begin{bmatrix} \Sigma & E(f)\Sigma & \Sigma \\ E(f)\Sigma & E\left(f^2\right)\Sigma & E(f)\Sigma \\ \Sigma & E(f)\Sigma & \beta\beta'\sigma^2(f)+\Sigma \end{bmatrix}$$

- Plug into (11.4) get the (3,3) element

$$\operatorname{var}\left(\hat{b}\right) = \frac{1}{T}(ad)^{-1} aSa'(ad)^{-1}' \qquad (11.4)$$

- Plug the items into (11.5) to get the sample distribution of the pricing errors

$$\sqrt{T} g_T\left(\hat{b}\right) \to N[0, \left(I - d(ad)^{-1} a\right) S \left(I - d(ad)^{-1} a\right)'] \qquad (11.5)$$

The result is the same as (12.20)

- Once again, it is no need to make assumptions.
- It is quite easy to estimate an S matrix that does not impose the conditions
- If one is really interesting in efficiency, The GLS cross-sectional estimate should use the spectral density matrix as weighting matrix rather Σ^{-1}

Time series vs. Cross Section

- The time series requires factors that are also returns. So that you can estimate factor risk premia by $\hat{\lambda} = E_T(f)$
- The asset pricing model does predict a restrict in the time-series. If imposed $E\left(R^{ei}\right) = \beta_i'\lambda$, you can write time-series regression as

$$R_t^{ei} = \beta_i'\lambda + \beta_i'\left(f_t - E(f)\right) + \varepsilon_t^i, t = 1, 2, \cdots, T$$

- The intercept restriction is

$$a_i = \beta_i^{'}\left(\lambda - E\left(f\right)\right)$$

The restriction makes sense. You can see how $\lambda = E\left(f\right)$ result in a zero intercept

- without an estimate of λ ,you can not check this intercept restriction

- When the factor is a return, so that we can compare the two methods:

1. The time-series regression describes the cross section of expected returns by drawing a line as in figure 12.1 that runs through the origin and through the factor

$$E_T\left(R^{ei}\right) = \beta_i^{'}\lambda + \alpha_i$$

2. The OLS cross-sectional regression picks the slope and intercept to best fit the points

- In including the factor as a test factor The GLS cross-sectional regression =the time-series regression

The time-series regression for f is :

$$f_t = 0 + 1 f_t + 0$$

The residual covariance matrix of the returns is:

$$E\left(\begin{bmatrix} R^e - a - \beta f \\ f - 0 - 1 f \end{bmatrix} [\bullet]'\right) = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$$

- Since f have 0 residual, GLS puts all its weight on that asset .
- Therefore $\hat{\lambda} = E_T(f)$,just as time-series regression.
- The degree of freedom in χ^2 test is N!
- Why ignore the pricing errors of the other asset in estimating the factor risk premium?

$$R_t^e = a + \beta f_t + \varepsilon_t$$

12.3 Fama-MacBeth Procedure

- Fama-MacBeth (1973) procedure:

1. find beta estimates with a time-series regression. $\hat{\beta}_i$
2. Run a cross-sectional regression at each time period:

$$R_t^{ei} = \beta_i^{'} \lambda_t + \alpha_{it}, \qquad i = 1, 2, \cdots, N$$

for each time t $\hat{\lambda}_t, \hat{\alpha}_{it}$

3. Estimate λ, α_i by:

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

4. Generate the sampling errors for these estimates:

$$\sigma^2\left(\hat{\lambda}\right) = \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{\lambda}_t - \hat{\lambda}\right)^2, \sigma^2\left(\hat{\alpha}_i\right) = \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{\alpha}_{it} - \hat{\alpha}_i\right)^2$$

5.when the time series is autocorrelated

$$\sigma^2\left(\hat{\lambda}\right) = \frac{1}{T}\sum_{j=-\infty}^{\infty} \operatorname{cov}_T\left(\hat{\lambda}_t, \hat{\lambda}_{t-j}\right)$$

6. Testing whether all the pricing errors are jointly zero:

Write the parameters in vector from:

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

$$\operatorname{cov}\left(\hat{\alpha}\right) = \frac{1}{T^2}\sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})(\hat{\alpha}_t - \hat{\alpha})'$$

The test statistic is :

$$\hat{\alpha}' \operatorname{cov}\left(\hat{\alpha}\right)^{-1} \hat{\alpha} \sim \chi^2_{N-1}$$

Fama-MacBeth in depth

- Consider a regression:

$$y_{it} = \beta' x_{it} + \varepsilon_{it}, i = 1, 2, \cdots N, t = 1, 2, \cdots, N$$

- *Pooled time-series cross-section* estimate:

stack the i and t observations together and estimate β by OLS

(contemporaneous correlation)

In an expected return-beta asset pricing model, the x_{it} is the $beta_i$ and beta is the lamda

- Take time-series averages and run a *pure cross-sectional regression*:

$$E_T(y_{it}) = \beta' E_T(x_{it}) + u_i, i = 1, 2, \cdots, N$$

- *Fama-MacBeth procedure*: run a cross-sectional regression at each point in time. Then get the estimates.

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

$$\operatorname{cov}\left(\hat{\beta}\right) = \frac{1}{T^2}\sum_{t=1}^{T} (\hat{\beta}_t - \hat{\beta})(\hat{\beta}_t - \hat{\beta})'$$

proposition

- If the x_{it} variables do not vary over time,and if the errors are cross-sectionally correlated but not correlated over time, Then:

The Fama-MacBeth estimate =

The pure cross-sectional OLS estimate =

The pooled time-series cross-sectional OLS estimate

So to the standard errors ,corrected for residual correlation.

None of them holds if the x_{it} vary through time

Proof:

Having assuming that the x variable do not vary over time, the regression is :

$$y_{it} = x_i^{'}\beta + \varepsilon_{it}$$

stack up the regressions in vector from:

$$y_t = x\beta + \varepsilon_t$$

The error assumptions mean

$$E(\varepsilon_t \varepsilon_t ') = \Sigma$$

Pooled OLS: to run pooled OLS, stack up the time series and cross sections:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, \qquad X = \begin{bmatrix} x \\ x \\ \vdots \\ x \end{bmatrix}, \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

and then:

$$Y = X\beta + \varepsilon$$

with

$$E(\varepsilon\varepsilon') = \Omega = \begin{bmatrix} \Sigma & & \\ & \ddots & \\ & & \Sigma \end{bmatrix}$$

The estimate and its standard error are:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'Y$$

$$\operatorname{cov}\left(\hat{\beta}_{OLS}\right) = (X'X)^{-1} X'\Omega X (X'X)^{-1}$$

Simplified:

$$\hat{\beta}_{OLS} = (x'x)^{-1} x' E_T(y_t)$$

$$\operatorname{cov}\left(\hat{\beta}_{OLS}\right) = \frac{1}{T}(x'x)^{-1} x'\sum x (x'x)^{-1}$$

Estimate this sampling variance with

$$\hat{\sum} = E_T\left(\hat{\varepsilon}_t \hat{\varepsilon}_t'\right), \qquad \hat{\varepsilon}_t \equiv y_t - x\hat{\beta}_{OLS}$$

Pure cross-section: take the time-series averages,

$$E_T\left(y_y\right) = x\beta + E_T\left(\varepsilon_t\right)$$

Having assumed I.I.d. errors ,so

$$E\left(E_T\left(\varepsilon_t\right)E_T\left(\varepsilon_t{}'\right)\right) = \frac{1}{T}\Sigma$$

the cross-sectional estimate and standard errors are:

$$\hat{\beta}_{XS} = \left(x'x\right)^{-1}x'E_T\left(y_t\right)$$

$$\sigma^2\left(\hat{\beta}_{XS}\right) = \frac{1}{T}\left(x'x\right)^{-1}x'\Sigma x\left(x'x\right)^{-1}$$

Fama-MacBeth: run cross-sectional regression at each moment in time

$$\hat{\beta}_t = (x'x)^{-1} x' y_t$$

then the estimate is the average of the cross-sectional regression estimates,

$$\hat{\beta}_{FM} = E_T\left(\hat{\beta}_t\right) = (x'x)^{-1} x' E_T(y_t)$$

the standard error is :

$$\operatorname{cov}(\hat{\beta}_{FM}) = \frac{1}{T} \operatorname{cov}_T\left(\hat{\beta}_t\right) = \frac{1}{T} (x'x)^{-1} x' \operatorname{cov}_T(y_t) x (x'x)^{-1}$$

with

$$y_t = x\beta_{FM} + \hat{\varepsilon}_t$$

we have

$$\mathrm{cov}_T\left(y_t\right) = E_T\left(\hat{\varepsilon}_t, \hat{\varepsilon}_t'\right) = \hat{\Sigma}$$

and finally

$$\mathrm{cov}\left(\hat{\beta}_{FM}\right) = \frac{1}{T}\left(x'x\right)^{-1} x'\hat{\Sigma}x\left(x'x\right)^{-1}$$

End proof

Varying x

- none of the three procedures are equal anymore, since it ignore the time-series variance in the x_{it}
- an extreme example:

$$y_{it} = \alpha + x_t\beta + \varepsilon_{it}, i = 1,2,\cdots,N, t = 1,2,\cdots,T$$

the grand OLS regression is : $\tilde{x} = x - E_T(x)$

$$\hat{\beta}_{OLS} = \frac{\sum_{it}\tilde{x}_t y_{it}}{\sum_{it}\tilde{x}_t^{\,2}} = \frac{\sum_t \tilde{x}_t \sum_i y_{it}}{N\sum_t \tilde{x}_t^{\,2}} = \frac{\sum_t \tilde{x}_t (1/N)\sum_i y_{it}}{\sum_t \tilde{x}_t^{\,2}}$$

Advantage of Fama-MacBeth Method

- Allows changing betas.
- The standard errors are easily to compute.
- Could be easily modified to consider the estimated beta, Shanken (1992).