**Attilio Meucci**

# Linear Factor Models

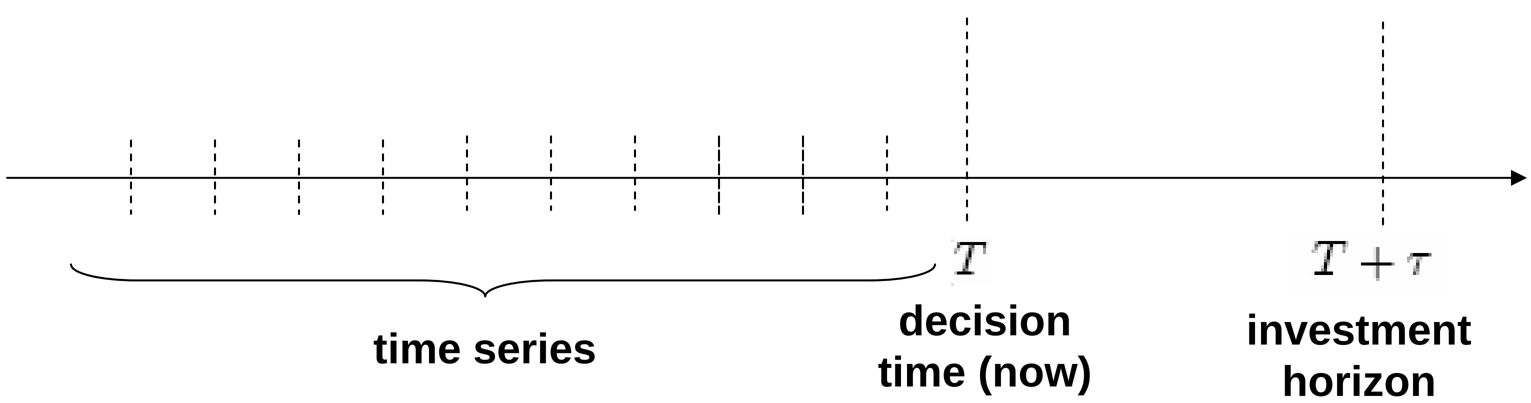
Formulas and figures in this presentation refer to the book **Risk and Asset Allocation**, Springer.

The notation, say, (5.24) refers to Formula 24 in Chapter 5 of the book

The notation, say, (T4.12) refers to Formula 12 in the Technical Appendices for Chapter 4, which can be downloaded from www.symmys.com

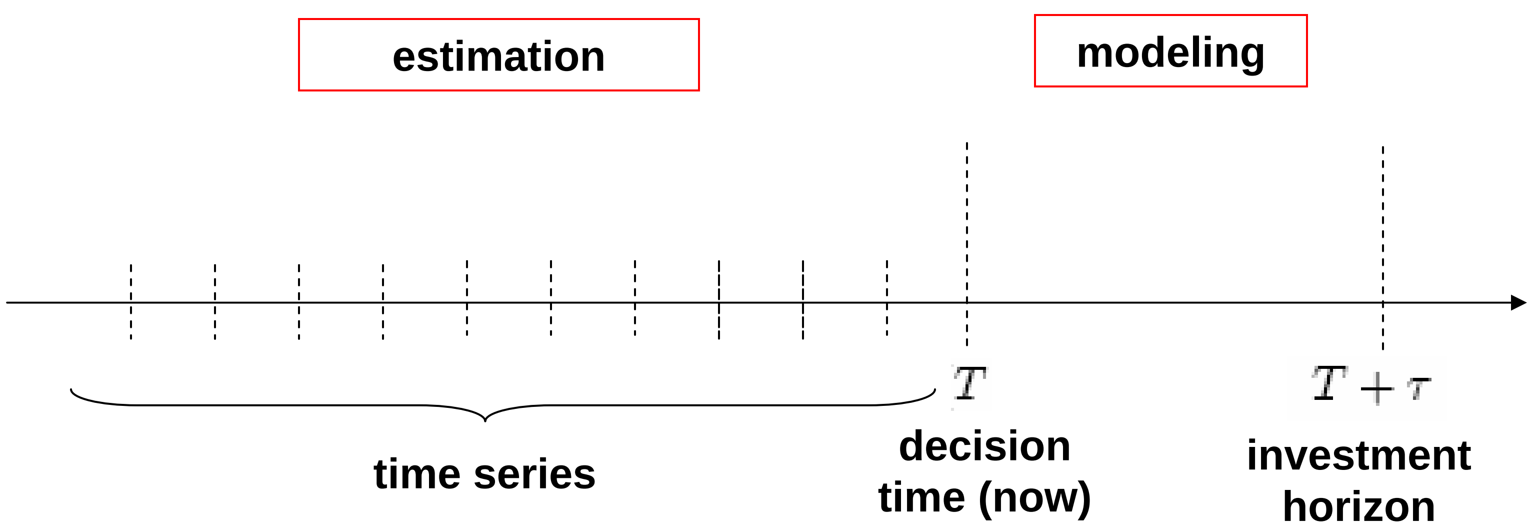
# LINEAR FACTOR MODELS – estimation vs. modeling

***Risk and Asset Allocation*, Springer - `symmys.com`**



# LINEAR FACTOR MODELS – estimation vs. modeling

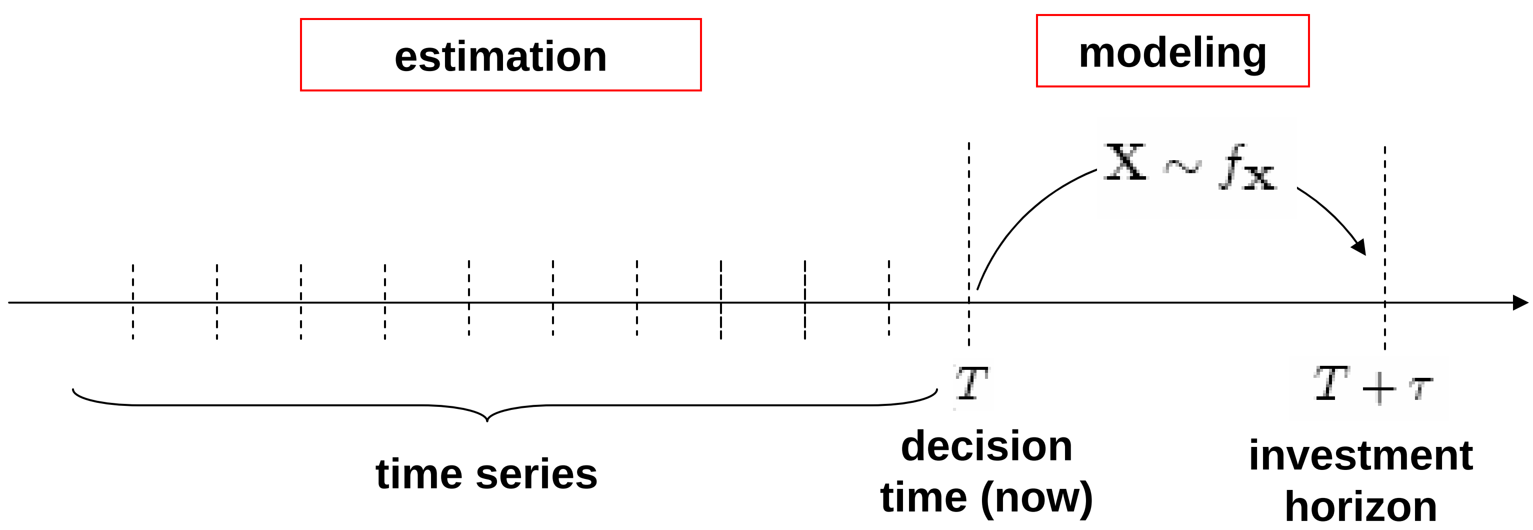
***Risk and Asset Allocation*, Springer - `symmys.com`**



# LINEAR FACTOR MODELS – estimation vs. modeling

***Risk and Asset Allocation*, Springer - `symmys.com`**

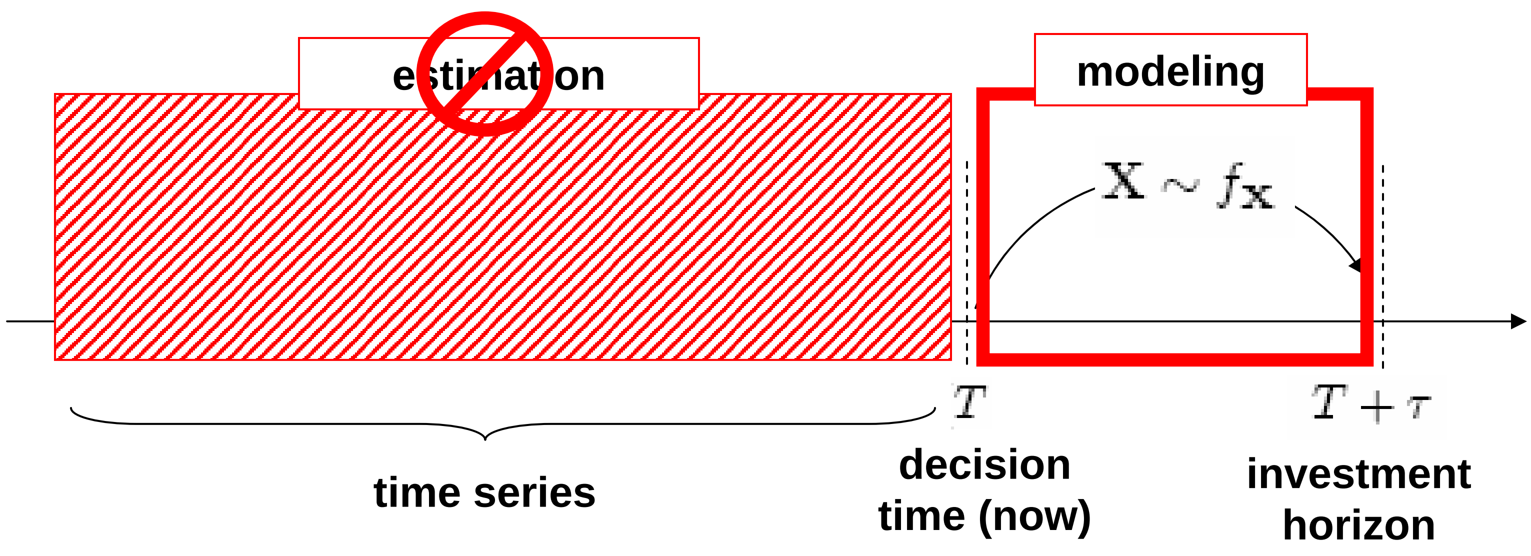
$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**



# LINEAR FACTOR MODELS – estimation vs. modeling

***Risk and Asset Allocation*, Springer - `symmys.com`**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**



# LINEAR FACTOR MODELS

***Risk and Asset Allocation*, Springer - `symmys.com`**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

**Stocks: comp. returns**

$$\mathbf{X} \equiv \begin{pmatrix} \ln\left(P_{T+\tau,1}/P_{T,1}\right) \\ \vdots \\ \ln\left(P_{T+\tau,N}/P_{T,N}\right) \end{pmatrix}$$

N=500: stocks in S&P500

# LINEAR FACTOR MODELS

***Risk and Asset Allocation*, Springer - `symmys.com`**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

**Stocks:**
**comp. returns**

$$\mathbf{X} \equiv \begin{pmatrix} \ln\left(P_{T+\tau,1}/P_{T,1}\right) \\ \vdots \\ \ln\left(P_{T+\tau,N}/P_{T,N}\right) \end{pmatrix}$$

N=500: stocks in S&P500

**Bonds:**
**yield changes**

$$\mathbf{X} \equiv \begin{pmatrix} Y_{T+\tau}^{(\upsilon_1)} - Y_T^{(\upsilon_1)} \\ \vdots \\ Y_{T+\tau}^{(\upsilon_N)} - Y_T^{(\upsilon_N)} \end{pmatrix}$$

N=360: 1m,2m,…,30y points on the curve

# LINEAR FACTOR MODELS

***Risk and Asset Allocation*, Springer - *symmys.com***

$\mathbf{X} \sim f_{\mathbf{X}}$ $\quad N \times 1$ **horizon-specific (random) market drivers with** **known** **distribution**

**Stocks: comp. returns**

$$\mathbf{X} \equiv \begin{pmatrix} \ln\left(P_{T+\tau,1}/P_{T,1}\right) \\ \vdots \\ \ln\left(P_{T+\tau,N}/P_{T,N}\right) \end{pmatrix}$$

N=500: stocks in S&P500

**Bonds: yield changes**

$$\mathbf{X} \equiv \begin{pmatrix} Y_{T+\tau}^{(v_1)} - Y_{T}^{(v_1)} \\ \vdots \\ Y_{T+\tau}^{(v_N)} - Y_{T}^{(v_N)} \end{pmatrix}$$

N=360: 1m,2m,…,30y points on the curve

**Derivatives: log impl. vol. changes**

$$\mathbf{X} \equiv \begin{pmatrix} \ln \sigma_{T+\tau}^{(m_1,v_1)} - \ln \sigma_{T}^{(m_1,v_1)} \\ \vdots \\ \ln \sigma_{T+\tau}^{(m_Q,v_S)} - \ln \sigma_{T}^{(m_Q,v_S)} \end{pmatrix}$$

N=QxS, Q=10 times to expiry and S=10 moneyness levels

# LINEAR FACTOR MODELS

***Risk and Asset Allocation*, Springer - `symmys.com`**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

**Stocks: comp. returns**

$$\mathbf{X} \equiv \begin{pmatrix} \ln\left(P_{T+\tau,1}/P_{T,1}\right) \\ \vdots \\ \ln\left(P_{T+\tau,N}/P_{T,N}\right) \end{pmatrix}$$

N=500: stocks in S&P500

**Bonds: yield changes**

$$\mathbf{X} \equiv \begin{pmatrix} Y_{T+\tau}^{(v_1)} - Y_{T}^{(v_1)} \\ \vdots \\ Y_{T+\tau}^{(v_N)} - Y_{T}^{(v_N)} \end{pmatrix}$$

N=360: 1m,2m,…,30y points on the curve

**Derivatives: log impl. vol. changes**

$$\mathbf{X} \equiv \begin{pmatrix} \ln\sigma_{T+\tau}^{(m_1,v_1)} - \ln\sigma_{T}^{(m_1,v_1)} \\ \vdots \\ \ln\sigma_{T+\tau}^{(m_Q,v_S)} - \ln\sigma_{T}^{(m_Q,v_S)} \end{pmatrix}$$

N=QxS, Q=10 times to expiry and S=10 moneyness levels

**~~portfolio returns~~**

**~~connection with financial theory~~**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

## LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$: $N \times K$ **(deterministic) loadings**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with** **known** **distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

$\mathbf{U}$ **"small"** **?**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

$\mathbf{U}$ "small" ?

- "distance" among random variables

$$\mathrm{E}\left\{\left(\mathbf{X}-\tilde{\mathbf{X}}\right)'\left(\mathbf{X}-\tilde{\mathbf{X}}\right)\right\}$$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

**OPTIMALITY CRITERIA**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

$\mathbf{U}$ **"small"** **?**

- **"distance" among random variables**

$$\frac{\mathrm{E}\left\{\left(\mathbf{X}-\tilde{\mathbf{X}}\right)'\left(\mathbf{X}-\tilde{\mathbf{X}}\right)\right\}}{\operatorname{tr}\left\{\operatorname{Cov}\left\{\mathbf{X}\right\}\right\}}$$

## LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$\mathbf{U}$ **"small"** **?**

- **multivariate r-square:**
  **"distance" among random variables**

$$R^2\left\{\mathbf{X}, \widetilde{\mathbf{X}}\right\} \equiv 1 - \frac{\mathrm{E}\left\{\left(\mathbf{X} - \widetilde{\mathbf{X}}\right)'\left(\mathbf{X} - \widetilde{\mathbf{X}}\right)\right\}}{\mathrm{tr}\left\{\mathrm{Cov}\left\{\mathbf{X}\right\}\right\}} \quad (3.116)$$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

**OPTIMALITY CRITERIA**

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

$\mathbf{U}$ **"small"** **?**

- **"recovered" market** $\widetilde{\mathbf{X}} \equiv \mathbf{B}\mathbf{F}$

- **multivariate r-square:**
  **"distance" among random variables**

$$R^2\left\{\mathbf{X}, \widetilde{\mathbf{X}}\right\} \equiv 1 - \frac{\mathrm{E}\left\{\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)'\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)\right\}}{\operatorname{tr}\left\{\operatorname{Cov}\left\{\mathbf{X}\right\}\right\}} \tag{3.116}$$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

## OPTIMALITY CRITERIA

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

- **"recovered" market** $\widetilde{\mathbf{X}} \equiv \mathbf{BF}$

- **multivariate r-square:**
  **"distance" among random variables**

$$R^2\left\{\mathbf{X}, \widetilde{\mathbf{X}}\right\} \equiv 1 - \frac{\mathrm{E}\left\{\left(\mathbf{X} - \widetilde{\mathbf{X}}\right)'\left(\mathbf{X} - \widetilde{\mathbf{X}}\right)\right\}}{\mathrm{tr}\left\{\mathrm{Cov}\left\{\mathbf{X}\right\}\right\}} \tag{3.116}$$

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$;

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

# LINEAR FACTOR MODELS

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

## APPROACHES

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

## LINEAR FACTOR MODELS – "residual" approach

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $\quad N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $\quad K \times 1$ **(random) risk factors:** $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ **known**

$\mathbf{B}$ $\quad N \times K$ **(deterministic) loadings, known**

$\mathbf{U}$ $\quad N \times 1$ **(random) residuals**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"RESIDUAL" approach**

e.g. $\mathbf{X}$ bond returns

$\mathbf{B}$ key rate durations

$\mathbf{F}$ changes in key rates

# LINEAR FACTOR MODELS – "residual" approach

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known**

$\mathbf{B}$: $N \times K$ **(deterministic) loadings, known**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

## OPTIMALITY CRITERIA

✓ $K \ll N$

✗ $\mathrm{Cor}\{\mathbf{F},\mathbf{U}\} = \mathbf{0}_{K\times N}$:

✗ $\mathbf{U}$ **"small" ⇔ $R^2\{\mathbf{X},\mathbf{BF}\}$ large**

✗ $\mathbf{U}$ **idiosyncratic**

## APPROACHES

**1 - $\mathbf{F}$ $\mathbf{B}$: exogenous**

**2 - $\mathbf{F}$ exogenous $\mathbf{B}$: from optimality criteria**

**3 - $\mathbf{B}$: exogenous $\mathbf{F}$ from optimality criteria**

**4 - $\mathbf{F}$ $\mathbf{B}$: from optimality criteria**

## "RESIDUAL" approach

e.g. $\mathbf{X}$ bond returns

$\mathbf{B}$: key rate durations

$\mathbf{F}$ changes in key rates

# LINEAR FACTOR MODELS – exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors:** $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ **known**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"TIME SERIES" approach (MISNOMER)**

e.g. $\mathbf{X}$ stock compounded returns

$\mathbf{B}$ "betas"

$\mathbf{F}$ S&P index return, …

# LINEAR FACTOR MODELS – exogenous factors

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors:** $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ **known**

$\mathbf{B}$: $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$;

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$: **exogenous**

$$\boxed{\textbf{2 - } \mathbf{F} \textbf{ exogenous } \mathbf{B} \textbf{: from optimality criteria}}$$

**3 -** $\mathbf{B}$: **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$: **from optimality criteria**

**"TIME SERIES" approach (MISNOMER)**

e.g. $\mathbf{X}$ stock compounded returns

$\mathbf{B}$: "betas"

$\mathbf{F}$ S&P index return, …

## LINEAR FACTOR MODELS – exogenous factors: Fama - French



$$C_{t,\tau}^{(n)} \equiv \ln\left(\frac{P_t^{(n)}}{P_{t-\tau}^{(n)}}\right) \qquad (3.183)$$

$C^M$ **broad stock market index**

$SmB$ **“Small minus Big” market capitalization**

$HmL$ **“High minus Low” book to market value**

# LINEAR FACTOR MODELS – exogenous factors: Fama - French

***Risk and Asset Allocation*, Springer - `symmys.com`**

$$C^{(n)}_{t,\tau} \equiv \ln\left(\frac{P^{(n)}_{t}}{P^{(n)}_{t-\tau}}\right) \qquad (3.183)$$

$C^{M}$ **broad stock market index**

$SmB$ **“Small minus Big” market capitalization**

$HmL$ **“High minus Low” book to market value**

$$\begin{aligned} C^{(n)}_{t,\tau} \equiv \mathrm{E}\left\{C^{(n)}_{t,\tau}\right\} &+ \beta^{(n)}_{\boxed{\phantom{M}}}\left(C^{M}_{t,\tau} - \mathrm{E}\left\{C^{M}_{t,\tau}\right\}\right) \\ &+ \gamma^{(n)}_{\boxed{\phantom{M}}}\left(SmB_{t,\tau} - \mathrm{E}\left\{SmB_{t,\tau}\right\}\right) \\ &+ \zeta^{(n)}_{\boxed{\phantom{M}}}\left(HmL_{t,\tau} - \mathrm{E}\left\{HmL_{t,\tau}\right\}\right) + U^{(n)}_{t,\tau} \end{aligned} \qquad (3.184)$$

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \quad (3.120)$$

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

$\boxed{\text{2 - } \mathbf{F} \text{ exogenous } \quad \mathbf{B} \text{ from optimality criteria}}$

3 - $\mathbf{B}$ exogenous $\quad$ $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

## "TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1} \quad (3.121)$$

## LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

### OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

✗ $\mathbf{U}$ idiosyncratic

### APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1} \quad (3.121)$$

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

## "TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{BF}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{XF}'\}\,\mathrm{E}\{\mathbf{FF}'\}^{-1} \quad (3.121)$$

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F} \mapsto \begin{pmatrix} 1 \\ \mathbf{F} \end{pmatrix}$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$: $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

### OPTIMALITY CRITERIA

$K \ll N$

✓ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ large

✗ $\mathbf{U}$ idiosyncratic

### APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$: exogenous

$\boxed{\text{2 - } \mathbf{F} \text{ exogenous} \quad \mathbf{B} \text{: from optimality criteria}}$

3 - $\mathbf{B}$: exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$: from optimality criteria

### "TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{BF}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{XF}'\}\,\mathrm{E}\{\mathbf{FF}'\}^{-1} \quad (3.121)$$

## LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

**"TIME SERIES" approach (MISNOMER)**

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1} \quad (3.121)$$

## LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$$\mathrm{Cov}\{\mathbf{F}\} \equiv \mathbf{E}\Lambda\mathbf{E}' \qquad (3.133)$$

$$\mathbf{C}_{XF} \equiv \mathrm{Cor}\{\mathbf{X}, \mathbf{E}'\mathbf{F}\} \qquad (3.139)$$

$$R^2 = \frac{\mathrm{tr}\left(\mathbf{C}_{XF}\mathbf{C}'_{XF}\right)}{N}.$$

~ $\mathbf{U}$ **"small"** $\Leftrightarrow R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

**"TIME SERIES" approach (MISNOMER)**

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\mathrm{argmax}}\, R^2\{\mathbf{X}, \mathbf{BF}\} \qquad (3.120)$$

$$= \mathrm{E}\left\{\mathbf{XF}'\right\}\mathrm{E}\left\{\mathbf{FF}'\right\}^{-1} \qquad (3.121)$$

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F} \mapsto \begin{pmatrix} 1 \\ \mathbf{F} \end{pmatrix}$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$: $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

✓ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

~ $\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

✗ $\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$: exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$: from optimality criteria

3 - $\mathbf{B}$: exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$: from optimality criteria

## "TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \quad (3.120)$$

$$= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\, \mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1} \quad (3.121)$$

# LINEAR FACTOR MODELS - exogenous factors

***Risk and Asset Allocation*, Springer - `symmys.com`**

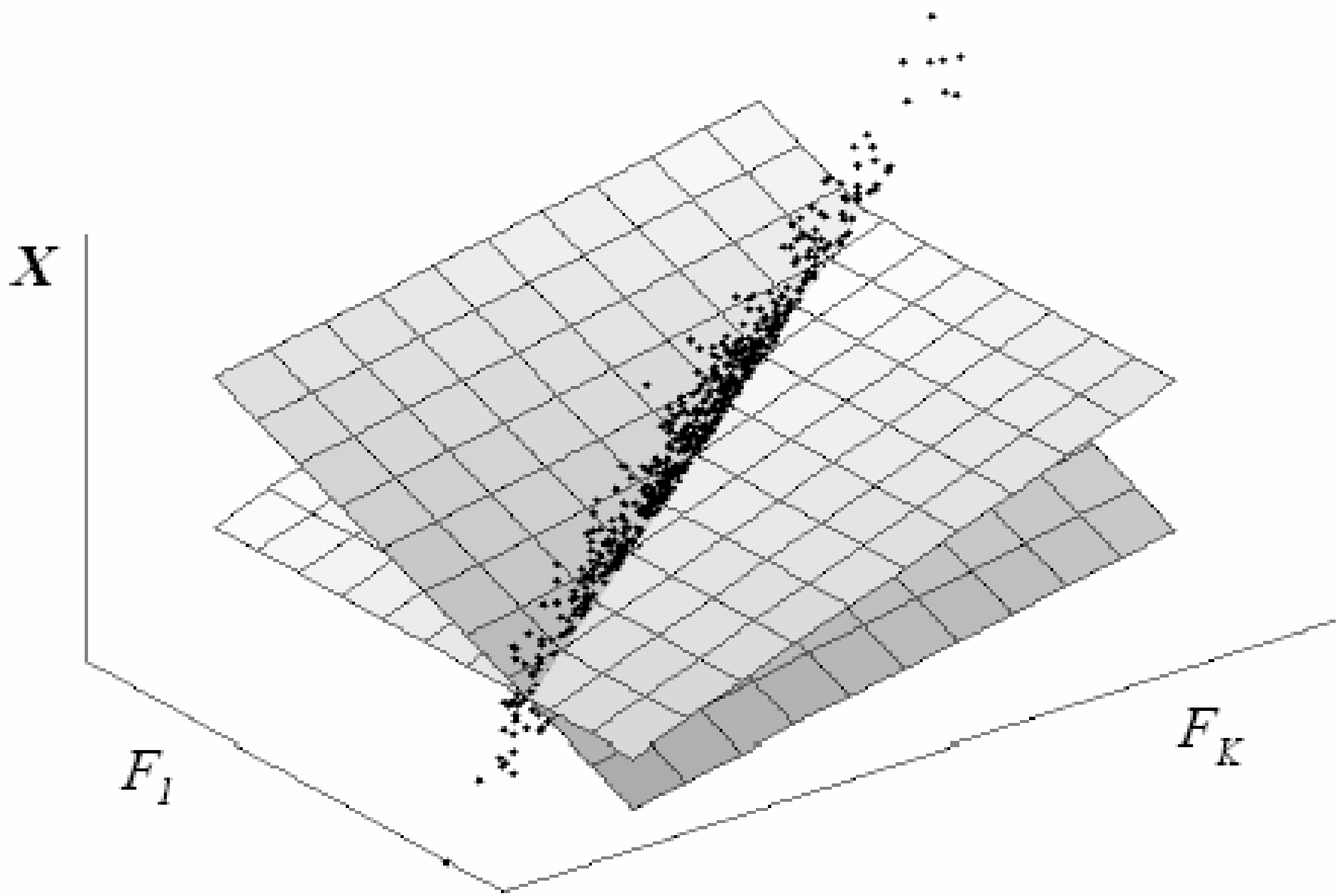



Fig. 3.13. Collinearity: the regression plane is not defined

# LINEAR FACTOR MODELS - exogenous factors selection routine

## *Risk and Asset Allocation*, Springer - `symmys.com`

$$I_N \equiv \{1, \ldots, N\} \quad (3.187)$$

$$I_K \equiv \{n_1, \ldots, n_K\} \quad (3.188)$$

$$I_K^* = \underset{I_K \subset I_N}{\operatorname{argmax}}\, \mathcal{O}\left(I_K\right) \quad (3.191)$$

Step 0. Set $K \equiv N$, and consider the initial set $I_K \equiv \{1, \ldots, N\}$

Step 1. Consider the $K$ sets obtained from $I_K$ by dropping the generic $k$-th element:

$$I_K^k \equiv \{n_1, \ldots, n_{k-1}, n_{k+1}, \ldots n_K\}, \quad k = 1, \ldots, K. \tag{3.198}$$

Step 2. Evaluate the above sets:

$$k \mapsto v_K^k \equiv \mathcal{O}\left(I_K^k\right), \quad k = 1, \ldots, K. \tag{3.199}$$

Step 3. Determine the worst element in $I_K$:

$$k^* \equiv \underset{k \in \{1, \ldots, K\}}{\operatorname{argmax}} \left\{v_K^k\right\}. \tag{3.200}$$

Step 4. Drop the worst element in $I_K$:

$$I_{K-1} \equiv I_K^{k^*}. \tag{3.201}$$

Step 5. If $K = 2$ stop. Otherwise set $K \equiv K - 1$ and go to Step 1.

# LINEAR FACTOR MODELS - exogenous factors

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F} \mapsto \begin{pmatrix} 1 \\ \mathbf{F} \end{pmatrix}$ (random) risk factors: $f_{\mathbf{F}}$ $f_{\mathbf{X},\mathbf{F}}$ known

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

✓ $K \ll N$

✓ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

~ $\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

✗ $\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

$\boxed{\text{2 - } \mathbf{F} \text{ exogenous} \quad \mathbf{B} \text{ from optimality criteria}}$

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"TIME SERIES" approach (MISNOMER)

$$\mathbf{B}_r \equiv \underset{\mathbf{B}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\} \qquad (3.120)$$

$$= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1} \qquad (3.121)$$

# LINEAR FACTOR MODELS – exogenous loadings

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings, known**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"CROSS SECTION" approach**

e.g. $\mathbf{X}$ stock compounded returns

$\mathbf{B}$ GICS 1/0 industry partition

$\mathbf{F}$ industry factors

# LINEAR FACTOR MODELS – exogenous loadings

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings, known

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ large

$\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"CROSS SECTION" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

# LINEAR FACTOR MODELS – exogenous loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings, known

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

$\boxed{\text{3 - } \mathbf{B} \text{ exogenous } \quad \mathbf{F} \text{ from optimality criteria}}$

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"CROSS SECTION" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$\mathbf{F}_c \equiv \underset{\mathbf{F} \equiv \mathbf{A}'\mathbf{X}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$$

$$= (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{X}$$

# LINEAR FACTOR MODELS – exogenous loadings

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings, known

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

✓ $K \ll N$

✗ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

~ $\mathbf{U}$ "small" ⇔ $R^2\{\mathbf{X}, \mathbf{BF}\}$ large

✗ $\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"CROSS SECTION" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$\mathbf{F}_c \equiv \underset{\mathbf{F} \equiv \mathbf{A}'\mathbf{X}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{BF}\}$$

$$= (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{X}$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"PCA" approach**

e.g. $\mathbf{X}$ yield curve changes

$\mathbf{B}$ market / slope / butterfly

$\mathbf{F}$ parallel shift / tilt / twist

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\boxed{\mathbf{F}\ \mathbf{B}\ \text{from optimality criteria}}$

"PCA" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

$\boxed{\text{4 - } \mathbf{F}\ \mathbf{B} \text{ from optimality criteria}}$

"PCA" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$(\mathbf{B}_p, \mathbf{A}_p) \equiv \underset{\mathbf{B},\mathbf{A}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{A}'\mathbf{X}\} \tag{3.147}$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

$\boxed{\text{4 - } \mathbf{F}\ \mathbf{B} \text{ from optimality criteria}}$

## "PCA" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$(\mathbf{B}_p, \mathbf{A}_p) \equiv \underset{\mathbf{B},\mathbf{A}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{B}\mathbf{A}'\mathbf{X}\} \qquad (3.147)$$

$$\mathbf{A} \equiv \mathbf{B} = \mathbf{E}_K$$

$$\mathbf{E}_K \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}\right)$$

$$\mathrm{Cov}\{\mathbf{X}\} \equiv \mathbf{E}\mathbf{\Lambda}\mathbf{E}'.$$

## LINEAR FACTOR MODELS – principal component analysis

***Risk and Asset Allocation*, Springer - `symmys.com`**

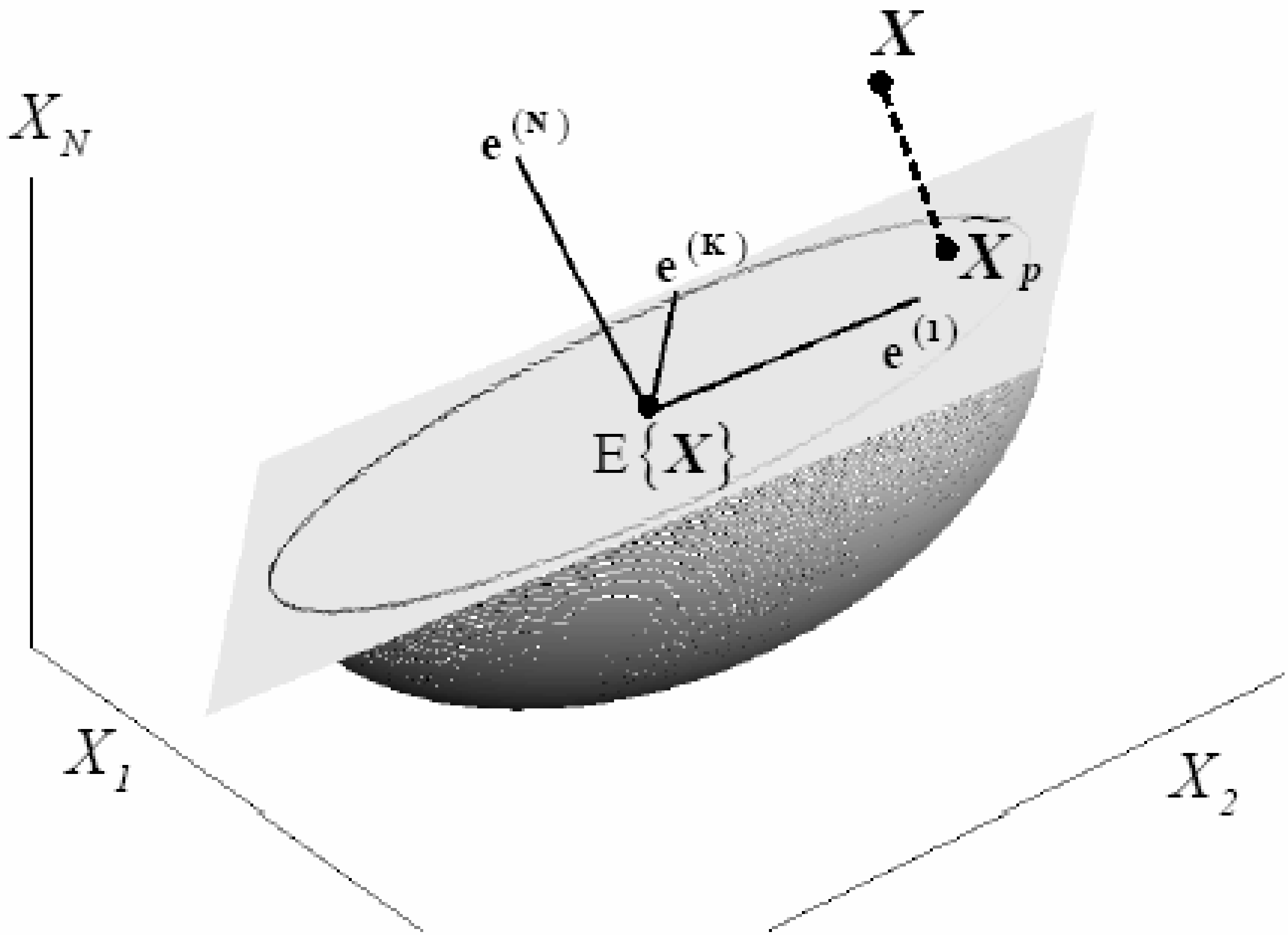



**Fig. 3.14.** Hidden factor dimension reduction: PCA

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{U}$ **"small"** $\Leftrightarrow R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

**"PCA" approach**

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$(\mathbf{B}_p, \mathbf{A}_p) \equiv \underset{\mathbf{B},\mathbf{A}}{\operatorname{argmax}}\, R^2\{\mathbf{X}, \mathbf{B}\mathbf{A}'\mathbf{X}\} \quad (3.147)$$

$$\mathbf{A} \equiv \mathbf{B} = \mathbf{E}_K$$

$$\mathbf{E}_K \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}\right)$$

$$\operatorname{Cov}\{\mathbf{X}\} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'.$$

$$R^2 = \frac{\sum_{n=1}^{K} \lambda_n}{\sum_{n=1}^{N} \lambda_n}. \quad (3.162)$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

## OPTIMALITY CRITERIA

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

✓ $\mathbf{U}$ "small" $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ large

$\mathbf{U}$ idiosyncratic

## APPROACHES

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

## "PCA" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$(\mathbf{B}_p, \mathbf{A}_p) \equiv \underset{\mathbf{B},\mathbf{A}}{\operatorname{argmax}} R^2\left\{\mathbf{X}, \mathbf{B}\mathbf{A}'\mathbf{X}\right\} \qquad (3.147)$$

$$\mathbf{A} \equiv \mathbf{B} = \mathbf{E}_K$$

$$\mathbf{E}_K \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}\right)$$

$$\mathrm{Cov}\{\mathbf{X}\} \equiv \mathbf{E}\Lambda\mathbf{E}'.$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

## OPTIMALITY CRITERIA

✓ $K \ll N$

✓ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$.

✓ $\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

✗ $\mathbf{U}$ **idiosyncratic**

## APPROACHES

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

## "PCA" approach

$$\mathbf{F} \equiv \mathbf{A}'\mathbf{X}$$

$$(\mathbf{B}_p, \mathbf{A}_p) \equiv \underset{\mathbf{B},\mathbf{A}}{\operatorname{argmax}} R^2\{\mathbf{X}, \mathbf{B}\mathbf{A}'\mathbf{X}\} \quad (3.147)$$

$$\mathbf{A} \equiv \mathbf{B} = \mathbf{E}_K$$

$$\mathbf{E}_K \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}\right)$$

$$\mathrm{Cov}\{\mathbf{X}\} \equiv \mathbf{E}\Lambda\mathbf{E}'.$$

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}.$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{B}\mathbf{F}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"FACTOR ANALYSIS" approach**

e.g. $\mathbf{X}$ stock compounded returns

$\mathbf{B}$ statistical loadings

$\mathbf{F}$ N/A

## LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ **horizon-specific (random) market drivers with known distribution**

$\mathbf{F}$ $K \times 1$ **(random) risk factors**

$\mathbf{B}$ $N \times K$ **(deterministic) loadings**

$\mathbf{U}$ $N \times 1$ **(random) residuals**

**OPTIMALITY CRITERIA**

$K \ll N$

$\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

$\mathbf{U}$ **"small"** $\Leftrightarrow$ $R^2\{\mathbf{X}, \mathbf{BF}\}$ **large**

$\mathbf{U}$ **idiosyncratic**

**APPROACHES**

**1 -** $\mathbf{F}$ $\mathbf{B}$ **exogenous**

**2 -** $\mathbf{F}$ **exogenous** $\mathbf{B}$ **from optimality criteria**

**3 -** $\mathbf{B}$ **exogenous** $\mathbf{F}$ **from optimality criteria**

**4 -** $\mathbf{F}$ $\mathbf{B}$ **from optimality criteria**

**"FACTOR ANALYSIS" approach**

$$\mathrm{Cov}\{\mathbf{X}\} \equiv \mathbf{BB}' + \mathbf{D}$$

**diagonal**

# LINEAR FACTOR MODELS – endogenous factors and loadings

$$\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$$

$\mathbf{X} \sim f_{\mathbf{X}}$ $N \times 1$ horizon-specific (random) market drivers with known distribution

$\mathbf{F}$ $K \times 1$ (random) risk factors

$\mathbf{B}$ $N \times K$ (deterministic) loadings

$\mathbf{U}$ $N \times 1$ (random) residuals

**OPTIMALITY CRITERIA**

✓ $K \ll N$

✓ $\mathrm{Cor}\{\mathbf{F}, \mathbf{U}\} = \mathbf{0}_{K \times N}$

✗ $\mathbf{U}$ "small" ⇔ $R^2\{\mathbf{X}, \mathbf{BF}\}$ large

✓ $\mathbf{U}$ idiosyncratic

**APPROACHES**

1 - $\mathbf{F}$ $\mathbf{B}$ exogenous

2 - $\mathbf{F}$ exogenous $\mathbf{B}$ from optimality criteria

3 - $\mathbf{B}$ exogenous $\mathbf{F}$ from optimality criteria

4 - $\mathbf{F}$ $\mathbf{B}$ from optimality criteria

"FACTOR ANALYSIS" approach

$$\mathrm{Cov}\{\mathbf{X}\} \equiv \mathbf{BB}' + \mathbf{D}$$

diagonal

## LINEAR FACTOR MODELS – estimation vs interpretation

### estimation

**invariants** $\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

# LINEAR FACTOR MODELS – estimation vs interpretation

## estimation

**invariants** $\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

## interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$ "basis"

# LINEAR FACTOR MODELS – estimation vs interpretation

## estimation

**invariants** $\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

## interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$ "basis"

systematic idiosyncratic

**~~CAPM~~:** $\mathbf{R} = \beta R_M + \mathbf{I}$

**~~APT~~:** $\mathbf{R} = \mathbf{DZ} + \mathbf{I}$

**financial theory**

# LINEAR FACTOR MODELS – estimation vs interpretation

## estimation

**invariants** $\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

## interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$ "basis"

$R^f_{t,\tau} \equiv \left( \dfrac{1}{Z^{(t)}_{t-\tau}} - 1 \right)$ (3.181)

systematic idiosyncratic

**CAPM:** **if** $\mathbf{R} = \beta R_M + \mathbf{I}$ $\Rightarrow$ $\mathrm{E}\{\mathbf{R}\} = \beta E\{R_M\} + (1-\beta) R_f$ (3.180)

**APT:** **if** $\mathbf{R} = \mathbf{DZ} + \mathbf{I}$ $\Rightarrow$ $\mathrm{E}\{\mathbf{R}\} = \xi_0 \mathbf{1} + \mathbf{D}\xi$ (3.186)

**financial theory**

# LINEAR FACTOR MODELS – estimation vs interpretation

## estimation

**invariants** $\mathbf{X} \equiv \mathbf{BF} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

## interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$

"basis"

systematic idiosyncratic

**CAPM:** **if** $\mathbf{R} = \beta R_M + \mathbf{I}$ $\Rightarrow$ $\mathrm{E}\{\mathbf{R}\} = \beta E\{R_M\} + (1-\beta) R_f$ (3.180)

**APT:** **if** $\mathbf{R} = \mathbf{DZ} + \mathbf{I}$ $\Rightarrow$ $\mathrm{E}\{\mathbf{R}\} = \xi_0 \mathbf{1} + \mathbf{D}\xi$ (3.186)

**financial theory**

**portfolio return** $R_{\mathbf{w}} \equiv \sum_{n=1}^{N} w_n R_n$

## LINEAR FACTOR MODELS – estimation vs interpretation

***Risk and Asset Allocation*, Springer - `symmys.com`**

### estimation

**invariants** $\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

### interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$ “basis”

$$R^f_{t,\tau} \equiv \left(\frac{1}{Z^{(t)}_{t-\tau}} - 1\right) \quad (3.181)$$

systematic idiosyncratic

**CAPM:** **if** $\mathbf{R} = \beta R_M + \mathbf{I} \quad \Rightarrow \quad \mathrm{E}\{\mathbf{R}\} = \beta E\{R_M\} + (1-\beta) R_f \quad (3.180)$

**APT:** **if** $\mathbf{R} = \mathbf{D}\mathbf{Z} + \mathbf{I} \quad \Rightarrow \quad \mathrm{E}\{\mathbf{R}\} = \xi_0 \mathbf{1} + \mathbf{D}\xi \quad (3.186)$

**financial theory**

**portfolio return** $R_{\mathbf{w}} \equiv \sum_{n=1}^{N} w_n R_n \qquad R_{\mathbf{w}} = \mathbf{d}'_{\mathbf{w}} \mathbf{Z} + \eta_{\mathbf{w}}$

# LINEAR FACTOR MODELS – estimation vs interpretation

***Risk and Asset Allocation*, Springer - `symmys.com`**

## estimation

**invariants** $\mathbf{X} \equiv \mathbf{B}\mathbf{F} + \mathbf{U}.$ **horizon prices** $P_{n,T+\tau} = g_n(\mathbf{X})$

## interpretation

return $R_n \equiv \dfrac{P_{n,T+\tau} - P_{n,T}}{B_{n,T}}$ "basis"

$$R^f_{t,\tau} \equiv \left( \frac{1}{Z^{(t)}_{t-\tau}} - 1 \right) \quad (3.181)$$

systematic idiosyncratic

**CAPM:** **if** $\mathbf{R} = \beta R_M + \mathbf{I} \Rightarrow \mathrm{E}\{\mathbf{R}\} = \beta E\{R_M\} + (1-\beta) R_f \quad (3.180)$

**APT:** **if** $\mathbf{R} = \mathbf{D}\mathbf{Z} + \mathbf{I} \Rightarrow \mathrm{E}\{\mathbf{R}\} = \xi_0 \mathbf{1} + \mathbf{D}\boldsymbol{\xi} \quad (3.186)$

**financial theory**

**portfolio return** $R_{\mathbf{w}} \equiv \sum_{n=1}^{N} w_n R_n$ $\quad R_{\mathbf{w}} = \mathbf{d}'_{\mathbf{w}} \mathbf{Z} + \eta_{\mathbf{w}}$

$$\mathbf{d}_{\mathbf{w}} \equiv \underset{\mathbf{d} \in \mathcal{C}}{\operatorname{argmax}} \left\{ \mathcal{R}^2 \left( R_{\mathbf{w}}, \mathbf{d}'\mathbf{Z} \right) \right\}$$

- **hedging**
- **portfolio replication**
- **style analysis**

# LINEAR FACTOR MODELS – pitfalls

***Risk and Asset Allocation*, Springer - *`symmys.com`***

**linear models on invariants for estimation**

**financial theory (CAPM, APT) on linear returns for interpretation**

## LINEAR FACTOR MODELS – pitfalls

***Risk and Asset Allocation*, Springer - `symmys.com`**

**linear models on invariants for estimation**

**residual not idiosyncratic**

**residual correlated with factors**

**financial theory (CAPM, APT) on linear returns for interpretation**

# LINEAR FACTOR MODELS – pitfalls

***Risk and Asset Allocation*, Springer - *symmys.com***

**linear models on invariants for estimation**

- **residual not idiosyncratic**
- **residual correlated with factors**

**financial theory (CAPM, APT) on linear returns for interpretation**

- **OK for non-equity products (e.g. options)**
- **OK  for autocorrelated processes**

# LINEAR FACTOR MODELS – pitfalls

***Risk and Asset Allocation*, Springer - `symmys.com`**

**linear models on invariants for estimation**

**residual not idiosyncratic**

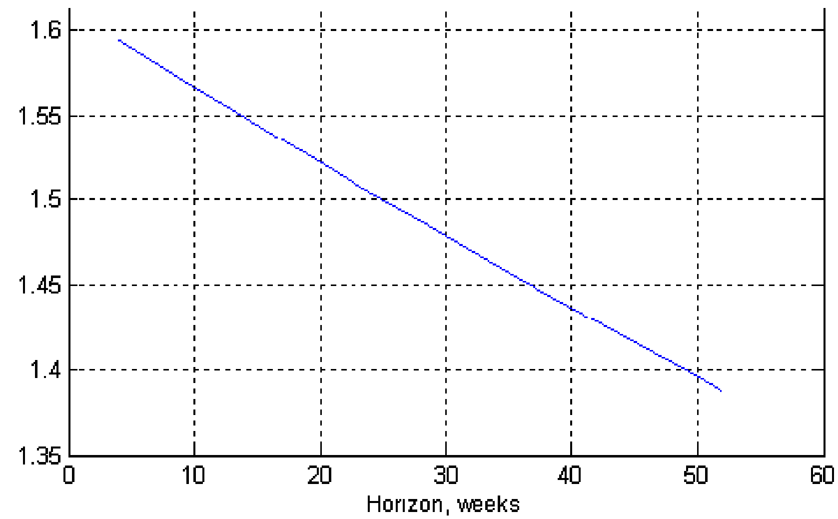
**residual correlated with factors**

**financial theory (CAPM, APT) on linear returns for interpretation**

**OK for non-equity products (e.g. options)**

**OK  for autocorrelated processes**

**betas depend on the horizon**



# LINEAR FACTOR MODELS – pitfalls

***Risk and Asset Allocation*, Springer - `symmys.com`**

**linear models on invariants for estimation**

- **residual not idiosyncratic**
- **residual correlated with factors**

**financial theory (CAPM, APT) on linear returns for interpretation**

- **OK for non-equity products (e.g. options)**
- **OK  for autocorrelated processes**
- **betas depend on the horizon**
- **NOT estimation model**