Attilio Meucci

REVIEW of FACTORS MODELS

http://ssrn.com/abstract=1635495

Definition of factor models

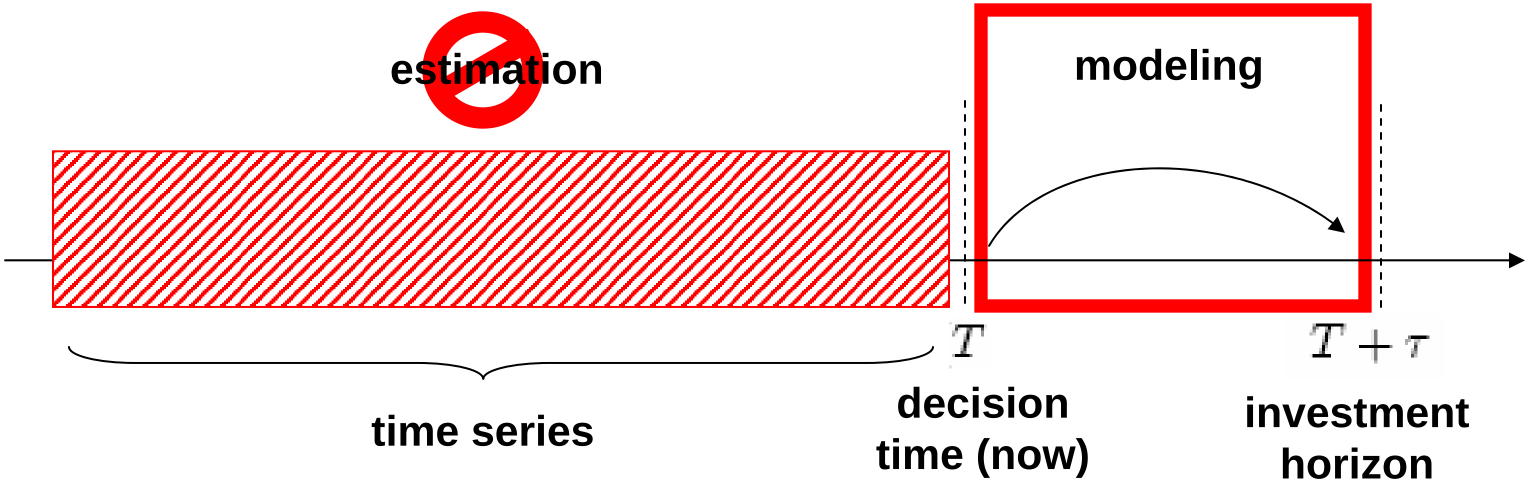
$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U} \quad \begin{cases} \mathbf{X} & N \times 1 & \text{“Market”} \\ \mathbf{a} & N \times 1 & \text{Constant} \\ \mathbf{B} & N \times K & \text{Exposures} \\ \mathbf{F} & K \times 1 & \text{Risk factors} \\ \mathbf{U} & N \times 1 & \text{Residuals} \end{cases}$$

Prices, price changes, returns, spreads, spread changes, impl. vol. surf., etc.

$$\underbrace{\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix}}_{\text{“Market”}} = \underbrace{\begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \vdots \\ a_N \end{pmatrix}}_{\text{Constant}} + \underbrace{\begin{pmatrix} b_{1,1} & \cdots & b_{1,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,k} & \cdots & b_{N,K} \end{pmatrix}}_{\text{Exposures of market to factors}} \underbrace{\begin{pmatrix} F_1 \\ \vdots \\ F_k \\ \vdots \\ F_K \end{pmatrix}}_{\text{Factors}} + \underbrace{\begin{pmatrix} U_1 \\ \vdots \\ U_n \\ \vdots \\ U_N \end{pmatrix}}_{\text{Residuals}}$$

Definition of factor models

$$
\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}
\quad
\left\{
\begin{array}{lll}
\mathbf{X} & N \times 1 & \text{“Market” with given distribution } f_{\mathbf{X}} \\
\mathbf{a} & N \times 1 & \text{Constant} \\
\mathbf{B} & N \times K & \text{Exposures} \\
\mathbf{F} & K \times 1 & \text{Risk factors} \\
\mathbf{U} & N \times 1 & \text{Residuals}
\end{array}
\right.
$$



Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{a}^*, \mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{a},\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} \mathcal{F}\{\mathbf{a} + \mathbf{B}\mathbf{F}, \mathbf{X}\}$$

fitness

Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**“Market”**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

“Dominant + residual” factor model

$$(\mathbf{a}^*, \mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{a},\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} \mathcal{F}\{\mathbf{a} + \mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathcal{F}\{\mathbf{Y}, \mathbf{X}\} \equiv R^2_{\mathbf{W}}\{\mathbf{Y}, \mathbf{X}\} \equiv -\operatorname{tr}(\operatorname{Cov}\{\mathbf{W}(\mathbf{Y} - \mathbf{X})\})$$

Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{a}^*, \mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{a},\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} \mathcal{F}\{\mathbf{a} + \mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathcal{F}\{\mathbf{Y}, \mathbf{X}\} \equiv R^2_{\mathbf{W}}\{\mathbf{Y}, \mathbf{X}\} \equiv -\operatorname{tr}(\operatorname{Cov}\{\mathbf{W}(\mathbf{Y} - \mathbf{X})\})$$

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}}\{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U} \quad \begin{cases} \mathbf{X} & N \times 1 & \text{“Market”} \\ \mathbf{a} & N \times 1 & \text{Constant} \\ \mathbf{B} & N \times K & \text{Exposures} \\ \mathbf{F} & K \times 1 & \text{Risk factors} \\ \mathbf{U} & N \times 1 & \text{Residuals} \end{cases}$$

“Systematic + idiosyncratic” factor model

not correlated

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \vdots \\ a_N \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,k} & \cdots & b_{N,K} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} F_1 \\ \vdots \\ F_k \\ \vdots \\ F_K \end{matrix}} \end{pmatrix} + \begin{pmatrix} \boxed{\begin{matrix} U_1 \\ \vdots \\ U_n \\ \vdots \\ U_N \end{matrix}} \end{pmatrix}$$

$$
\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}
\quad
\left\{
\begin{array}{lll}
\mathbf{X} & N \times 1 & \text{“Market”} \\
\mathbf{a} & N \times 1 & \text{Constant} \\
\mathbf{B} & N \times K & \text{Exposures} \\
\mathbf{F} & K \times 1 & \text{Risk factors} \\
\mathbf{U} & N \times 1 & \text{Residuals}
\end{array}
\right.
$$

“Systematic + idiosyncratic” factor model

$$
\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix}
=
\begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \vdots \\ a_N \end{pmatrix}
+
\begin{pmatrix}
b_{1,1} & \cdots & b_{1,k} & \cdots & b_{n,K} \\
\vdots & & \vdots & & \vdots \\
b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,K} \\
\vdots & & \vdots & & \vdots \\
b_{N,1} & \cdots & b_{N,k} & \cdots & b_{N,K}
\end{pmatrix}
\begin{pmatrix} F_1 \\ \vdots \\ F_k \\ \vdots \\ F_K \end{pmatrix}
+
\begin{pmatrix} \boxed{U_1} \\ \vdots \\ \boxed{U_n} \\ \vdots \\ U_N \end{pmatrix}
$$

not correlated

$$\mathrm{Var}\{\sum_{n=1}^{N} w_m X_n\} = \begin{cases} \underbrace{\sum_{m,n=1}^{N} \sum_{j,k=1}^{K} w_n w_m B_{n,k} B_{m,j} \operatorname{Cov}\{F_k, F_j\}}_{\text{factors}} \\ \underbrace{+2 \sum_{m,n=1}^{N} \sum_{k=1}^{K} w_m w_n B_{m,k} \operatorname{Cov}\{F_k, U_n\}}_{\text{cross-term}} \\ \underbrace{+ \sum_{m,n=1}^{N} w_n w_m \operatorname{Cov}\{U_n, U_m\}.}_{\text{residual}} \end{cases}$$

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \vdots \\ a_N \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,k} & \cdots & b_{N,K} \end{pmatrix} \begin{pmatrix} F_1 \\ \vdots \\ F_k \\ \vdots \\ F_K \end{pmatrix} + \begin{pmatrix} U_1 \\ \vdots \\ U_n \\ \vdots \\ U_N \end{pmatrix}$$

$$\mathrm{Var}\{\sum_{n=1}^{N} w_m X_n\} = \left\{ \begin{array}{l} \underbrace{\sum_{m,n=1}^{N} \sum_{j,k=1}^{K} w_n w_m B_{n,k} B_{m,j} \operatorname{Cov}\{F_k, F_j\}}_{\text{factors}} \\ +2 \underbrace{\sum_{m,n=1}^{N} \sum_{k=1}^{K} w_m w_n B_{m,k} \operatorname{Cov}\{F_k, U_n\}}_{\text{cross-term}} \\ + \underbrace{\sum_{m,n=1}^{N} w_n w_m \operatorname{Cov}\{U_n, U_m\}}_{\text{residual}}. \end{array} \right.$$

$$\left\{ \begin{array}{l} \underbrace{\sum_{m,n=1}^{N} \sum_{j,k=1}^{K} w_n w_m B_{n,k} B_{m,j} \operatorname{Cov}\{F_k, F_j\}}_{\text{systematic}} \\ + \underbrace{\sum_{n=1}^{N} w_n^2 \operatorname{Var}\{U_n\}}_{\text{idiosyncratic}}. \end{array} \right.$$

not correlated

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \vdots \\ a_N \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,k} & \cdots & b_{n,K} \\ \vdots & & \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,k} & \cdots & b_{N,K} \end{pmatrix} \begin{pmatrix} F_1 \\ \vdots \\ F_k \\ \vdots \\ F_K \end{pmatrix} + \begin{pmatrix} U_1 \\ \vdots \\ U_n \\ \vdots \\ U_N \end{pmatrix}$$

not correlated

Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$$\begin{cases} \mathbf{X} & N \times 1 & \text{“Market”} \\ \mathbf{a} & N \times 1 & \text{Constant} \\ \mathbf{B} & N \times K & \text{Exposures} \\ \mathbf{F} & K \times 1 & \text{Risk factors} \\ \mathbf{U} & N \times 1 & \text{Residuals} \end{cases}$$

“Systematic + idiosyncratic” factor model

$$\begin{aligned} \mathrm{Cor}\{U_n, F_k\} &= 0 \\ \mathrm{Cor}\{U_n, U_m\} &= 0 \end{aligned}$$

Definition of factor models

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**“Market”**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

“Dominant + residual” factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

“Systematic + idiosyncratic” factor model

$$\mathrm{Cor}\{U_n, F_k\} \cancel{=} 0$$

$$\mathrm{Cor}\{U_n, U_m\} \cancel{=} 0$$

All factor models are dominant + residual

No factor model is systematic + idiosyncratic

A. MEUCCI – Review of factor models

Linear factor model $\mathbf{X} \equiv \mathbf{a} + \mathbf{BF} + \mathbf{U}$		dominant + residual $(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{BF}, \mathbf{X}\}$	idiosyncratic $\operatorname{Cor}\{U_n, U_m\} = 0$	systematic $\operatorname{Cor}\{U_n, F_k\} = 0$
"Pure residual" (3)		$\mathbf{B}$ fully constrained $\mathbf{F}$ fully constrained	X	X
"Time-series"	"OLS" (4.2)	$\mathbf{B}$ free $\mathbf{F}$ fully constrained	X	✓
	Generalized (4.3)	$\mathbf{B}$ partly constrained $\mathbf{F}$ fully constrained	X	X
"Cross-section"	"w-OLS" (5.2)	$\mathbf{B}$ fully constrained $\mathbf{F}$ free	X	X
	Generalized (5.3)	$\mathbf{B}$ fully constrained $\mathbf{F}$ partly constrained	X	X
"Statistical"	"PCA" (6.2)	$\mathbf{B}$ free $\mathbf{F}$ free	X	✓
	"FA" (6.3)	$\mathbf{B}$ over-constrained $\mathbf{F}$ over-constrained	X	✓
Hybrid (7)		$\mathbf{B}$ partly constrained $\mathbf{F}$ partly constrained	X	X

Time series models

Application example

$\mathbf{X}$ stock returns

$\mathbf{B}$ "betas"

$\mathbf{F}$ - S&P index return, - industry indices, …

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

- $\mathbf{X}$ $N \times 1$ "Market"
- $\mathbf{a}$ $N \times 1$ Constant
- $\mathbf{B}$ $N \times K$ Exposures
- $\mathbf{F}$ $K \times 1$ Risk factors exogenous
- $\mathbf{U}$ $N \times 1$ Residuals

Application example

$\mathbf{X}$ **stock returns**

$\mathbf{B}$ **"betas"**

$\mathbf{F}$ **- S&P index return, - industry indices, …**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{B}^* \equiv \underset{\mathbf{B} \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

Application example

$\mathbf{X}$ **stock returns**

$\mathbf{B}$ **"betas"**

$\mathbf{F}$ **- S&P index return, - industry indices, …**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{B}^* \equiv \underset{\mathbf{B}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

(no constraints)

$$\mathbf{B}^* \equiv \mathrm{Cov}\{\mathbf{X}, \mathbf{F}\} \, \mathrm{Cov}\{\mathbf{F}\}^{-1}$$

Time series models

Application example

- $\mathbf{X}$ **stock returns**
- $\mathbf{B}$ **"betas"**
- $\mathbf{F}$ **- S&P index return, - industry indices, …**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures**
$\mathbf{F}$	$K \times 1$	**Risk factors exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F})\in\mathcal{C}}{\operatorname{argmax}}\, R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{B}^* \equiv \underset{\mathbf{B}\in\mathcal{C}}{\operatorname{argmax}}\, R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

(no constraints)

$$\mathbf{B}^* \equiv \operatorname{Cov}\{\mathbf{X}, \mathbf{F}\} \operatorname{Cov}\{\mathbf{F}\}^{-1}$$

"Systematic + idiosyncratic" factor model

$$\operatorname{Cor}\{U_n, F_k\} \cancel{\neq} 0$$

$$\operatorname{Cor}\{U_n, U_m\} \cancel{\neq} 0$$

Cross-section models

Application example

$\mathbf{X}$ **stock returns**

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

$\mathbf{F}$ **industry factors**

$$\mathbf{X} = \mathbf{a} + \mathbf{BF} + \mathbf{U}$$

$\mathbf{X}$ $N \times 1$ **"Market"**

$\mathbf{a}$ $N \times 1$ **Constant**

$\mathbf{B}$ $N \times K$ **Exposures exogenous**

$\mathbf{F}$ $K \times 1$ **Risk factors**

$\mathbf{U}$ $N \times 1$ **Residuals**

Application example

$\mathbf{X}$ stock returns

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

$\mathbf{F}$ industry factors

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F})\in\mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X},$$

Cross-section models

Application example

$\mathbf{X}$ stock returns

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

$\mathbf{F}$ industry factors

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$ $N \times 1$ **“Market”**

$\mathbf{a}$ $N \times 1$ **Constant**

$\mathbf{B}$ $N \times K$ **Exposures exogenous**

$\mathbf{F}$ $K \times 1$ **Risk factors**

$\mathbf{U}$ $N \times 1$ **Residuals**

“Dominant + residual” factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X},$$

$$\mathbf{G}^* \equiv \underset{\mathbf{G} \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

Cross-section models

Application example

$\mathbf{X}$ **stock returns**

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

$\mathbf{F}$ **industry factors**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$ $N \times 1$ **"Market"**

$\mathbf{a}$ $N \times 1$ **Constant**

$\mathbf{B}$ $N \times K$ **Exposures exogenous**

$\mathbf{F}$ $K \times 1$ **Risk factors**

$\mathbf{U}$ $N \times 1$ **Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X}.$$

$$\mathbf{G}^* \equiv \underset{\mathbf{G}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

(no constraints)

$$\mathbf{G}^* = (\mathbf{B}'\mathbf{\Phi}\mathbf{B})^{-1} \mathbf{B}'\mathbf{\Phi}, \leftarrow \mathbf{\Phi} \equiv \mathbf{W}'\mathbf{W}.$$

$$\mathbf{F} \equiv (\mathbf{B}'\mathbf{\Phi}\mathbf{B})^{-1} \mathbf{B}'\mathbf{\Phi}\mathbf{X}.$$

Application example

$\mathbf{X}$ **stock returns**

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

$\mathbf{F}$ **industry factors**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X},$$

$$\mathbf{G}^* \equiv \underset{\mathbf{G} \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

(no constraints)

$$\mathbf{G}^* = (\mathbf{B}'\mathbf{\Phi}\mathbf{B})^{-1} \mathbf{B}'\mathbf{\Phi}, \leftarrow \mathbf{\Phi} \equiv \mathbf{W}'\mathbf{W}.$$

$$\mathbf{F} \equiv (\mathbf{B}'\mathbf{\Phi}\mathbf{B})^{-1} \mathbf{B}'\mathbf{\Phi}\mathbf{X}.$$

"Systematic + idiosyncratic" factor model

$$\operatorname{Cor}\{U_n, F_k\} \cancel{\neq} 0$$

$$\operatorname{Cor}\{U_n, U_m\} \cancel{\neq} 0$$

Application example

- $\mathbf{X}$ **yield curve changes**
- $\mathbf{B}$ **market / slope / butterfly**
- $\mathbf{F}$ **parallel shift / tilt / twist**

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures not exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors not exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

Application example

$\mathbf{X}$ **yield curve changes**

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

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

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures not exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors not exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X},$$

Application example

$\mathbf{X}$ yield curve changes

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

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

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures not exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors not exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X}$$

$$(\mathbf{B}^*, \mathbf{G}^*) \equiv \underset{(\mathbf{B},\mathbf{G}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

Application example

$\mathbf{X}$ **yield curve changes**

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

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

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures not exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors not exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X}$$

$$(\mathbf{B}^*, \mathbf{G}^*) \equiv \underset{(\mathbf{B},\mathbf{G})}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

(no constraints)

$$\operatorname{Cov}\{\mathbf{W}\mathbf{X}\} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}' \begin{cases} \mathbf{E} \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(N)}\right) \\ \boldsymbol{\Lambda} \equiv \operatorname{diag}\left(\lambda_1^2, \ldots, \lambda_N^2\right) \end{cases}$$

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

$$\mathbf{B}^* = \mathbf{W}^{-1}\mathbf{E}_K \qquad \mathbf{G}^* \equiv \mathbf{E}'_K \mathbf{W}$$

$$\mathbf{F} \equiv \mathbf{E}'_K (\mathbf{X} - \mathrm{E}\{\mathbf{X}\})$$

Application example

$\mathbf{X}$ **yield curve changes**

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

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

$$\mathbf{X} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U}$$

$\mathbf{X}$	$N \times 1$	**"Market"**
$\mathbf{a}$	$N \times 1$	**Constant**
$\mathbf{B}$	$N \times K$	**Exposures not exogenous**
$\mathbf{F}$	$K \times 1$	**Risk factors not exogenous**
$\mathbf{U}$	$N \times 1$	**Residuals**

"Dominant + residual" factor model

$$(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}}\, R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{F}, \mathbf{X}\}$$

$$\mathbf{F} \equiv g(\mathbf{X}) \equiv \mathbf{d} + \mathbf{G}\mathbf{X}$$

$$(\mathbf{B}^*, \mathbf{G}^*) \equiv \underset{(\mathbf{B},\mathbf{G}) \in \mathcal{C}}{\operatorname{argmax}}\, R^2_{\mathbf{W}} \{\mathbf{B}\mathbf{G}\mathbf{X}, \mathbf{X}\}$$

"Systematic + idiosyncratic" factor model

$$\operatorname{Cor}\{U_n, F_k\} \cancel{\neq} 0$$

$$\operatorname{Cor}\{U_n, U_m\} \cancel{\neq} 0$$

(no constraints)

$$\operatorname{Cov}\{\mathbf{W}\mathbf{X}\} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}' \begin{cases} \mathbf{E} \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(N)}\right) \\ \boldsymbol{\Lambda} \equiv \operatorname{diag}\left(\lambda_1^2, \ldots, \lambda_N^2\right) \end{cases}$$

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

$$\mathbf{B}^* = \mathbf{W}^{-1}\mathbf{E}_K \qquad \mathbf{G}^* \equiv \mathbf{E}_K'\mathbf{W}$$

$$\mathbf{F} \equiv \mathbf{E}_K' (\mathbf{X} - \mathrm{E}\{\mathbf{X}\})$$

A. MEUCCI – Review of factor models

Linear factor model $\mathbf{X} \equiv \mathbf{a} + \mathbf{BF} + \mathbf{U}$	**dominant + residual** $(\mathbf{B}^*, \mathbf{F}^*) \equiv \underset{(\mathbf{B},\mathbf{F}) \in \mathcal{C}}{\operatorname{argmax}} R^2_{\mathbf{W}} \{\mathbf{BF}, \mathbf{X}\}$	**idiosyncratic** $\operatorname{Cor}\{U_n, U_m\} = 0$	**systematic** $\operatorname{Cor}\{U_n, F_k\} = 0$
"Pure residual" (3)	$\mathbf{B}$ fully constrained $\mathbf{F}$ fully constrained	X	X
"Time-series" — "OLS" (4.2)	$\mathbf{B}$ free $\mathbf{F}$ fully constrained	X	✓
"Time-series" — Generalized (4.3)	$\mathbf{B}$ partly constrained $\mathbf{F}$ fully constrained	X	X
"Cross-section" — "w-OLS" (5.2)	$\mathbf{B}$ fully constrained $\mathbf{F}$ free	X	X
"Cross-section" — Generalized (5.3)	$\mathbf{B}$ fully constrained $\mathbf{F}$ partly constrained	X	X
"Statistical" — "PCA" (6.2)	$\mathbf{B}$ free $\mathbf{F}$ free	X	✓
"Statistical" — "FA" (6.3)	$\mathbf{B}$ over-constrained $\mathbf{F}$ over-constrained	X	✓
Hybrid (7)	$\mathbf{B}$ partly constrained $\mathbf{F}$ partly constrained	X	X

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}};$$

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}};$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

$$\begin{cases} \mathrm{E}\{R_{\mathbf{w}}\} = \mathbf{w}'\,\mathrm{E}\{\mathbf{R}\} \\ \mathrm{Sd}\{R_{\mathbf{w}}\} = \sqrt{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}. \end{cases}$$

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^N u_m P_{m,T}};$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

$$\begin{cases} \mathrm{E}\{R_{\mathbf{w}}\} = \mathbf{w}'\,\mathrm{E}\{\mathbf{R}\} \\ \mathrm{Sd}\{R_{\mathbf{w}}\} = \sqrt{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}. \end{cases}$$

max. Sharpe ratio portfolio

$$\mathbf{w}_{SR} \equiv \underset{\mathbf{w}'\mathbf{1}\equiv 1}{\operatorname{argmax}} \left\{ \frac{\mathrm{E}\{R_{\mathbf{w}}\}}{\mathrm{Sd}\{R_{\mathbf{w}}\}} \right\}$$

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}}.$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

$$\begin{cases} \mathrm{E}\{R_{\mathbf{w}}\} = \mathbf{w}'\,\mathrm{E}\{\mathbf{R}\} \\ \mathrm{Sd}\{R_{\mathbf{w}}\} = \sqrt{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}. \end{cases}$$

max. Sharpe ratio portfolio

$$\mathbf{w}_{SR} \equiv \underset{\mathbf{w}'\mathbf{1}\equiv 1}{\operatorname{argmax}} \left\{ \frac{\mathrm{E}\{R_{\mathbf{w}}\}}{\mathrm{Sd}\{R_{\mathbf{w}}\}} \right\}$$

risk-free return

$$\mathrm{E}\{\mathbf{R}\} = r + \mathbf{b}_{\mathbf{w}_{SR}}\left(\mathrm{E}\{R_{\mathbf{w}_{SR}}\} - r\right)$$

$$\mathbf{b}_{\mathbf{w}} \equiv \frac{\mathrm{Cov}\{\mathbf{R}, R_{\mathbf{w}}\}}{\mathrm{Var}\{R_{\mathbf{w}}\}} = \frac{\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}$$

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^N u_m P_{m,T}};$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

$$\begin{cases} \mathrm{E}\{R_{\mathbf{w}}\} = \mathbf{w}'\,\mathrm{E}\{\mathbf{R}\} \\ \mathrm{Sd}\{R_{\mathbf{w}}\} = \sqrt{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}. \end{cases}$$

max. Sharpe ratio portfolio

$$\mathbf{w}_{SR} \equiv \underset{\mathbf{w}'\mathbf{1}\equiv 1}{\operatorname{argmax}} \left\{ \frac{\mathrm{E}\{R_{\mathbf{w}}\}}{\mathrm{Sd}\{R_{\mathbf{w}}\}} \right\}$$

risk-free return

$$\mathrm{E}\{\mathbf{R}\} = r + \mathbf{b}_{\mathbf{w}_{SR}} \left(\mathrm{E}\{R_{\mathbf{w}_{SR}}\} - r\right)$$

$$\mathbf{b}_{\mathbf{w}} \equiv \frac{\mathrm{Cov}\{\mathbf{R}, R_{\mathbf{w}}\}}{\mathrm{Var}\{R_{\mathbf{w}}\}} = \frac{\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}{\mathbf{w}'\,\mathrm{Cov}\{\mathbf{R}\}\,\mathbf{w}}$$

Capital Asset Pricing Theorem (CAPM):

If all investors trade off expectation and variance and the distribution is known

market portfolio

$$\mathbf{w}_M = \mathbf{w}_{SR}$$

$$\mathrm{E}\{\mathbf{R}\} = r\mathbf{1}_N + \mathbf{b}_{\mathbf{w}_M} \left(\mathrm{E}\{R_{\mathbf{w}_M}\} - r\right)$$

Capital Asset Pricing Theorem (CAPM):

- Only refers to returns, not generic variables X

- Refers to yet-to-be realized returns, no regression

- Refers to returns of any asset (exotic derivatives, etc.), not only stocks

- Is a constraint on expectations only

- Is not a “time series” factor model with one explicit factor, the market return

- Has identification issues

- In factor modes the “beta” is arbitrary

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}};$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}};$$

$$\mathbf{w}'\mathbf{1} \equiv 1.$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

Arbitrage Pricing Theory (APT):

If R is a "systematic + idiosyncratic" model

$$\mathbf{R} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U};$$

$$\mathrm{Cor}\{U_n, F_k\} = 0$$

$$\mathrm{Cor}\{U_n, U_m\} = 0$$

risk-free return

$$\mathrm{E}\{\mathbf{R}\} \approx r\mathbf{1} + \mathbf{b}_1\left(\mathrm{E}\left\{R_{\mathbf{w}_{(1)}}\right\} - r\right) + \cdots + \mathbf{b}_K\left(\mathrm{E}\left\{R_{\mathbf{w}_{(K)}}\right\} - r\right)$$

Relationships with asset pricing

$$R_n \equiv \frac{P_{n,T+\tau} - P_{n,T}}{P_{n,T}};$$

$$w_n \equiv \frac{u_n P_{n,T}}{\sum_{m=1}^{N} u_m P_{m,T}}:$$

$$\cancel{\mathbf{w}'\mathbf{1} \equiv 1}$$

$$R_{\mathbf{w}} \equiv \mathbf{w}'\mathbf{R}$$

portfolio return

Arbitrage Pricing Theory (APT):

If R is a "systematic + idiosyncratic" model

$$\mathbf{R} = \mathbf{a} + \mathbf{B}\mathbf{F} + \mathbf{U};$$

$$\mathrm{Cor}\{U_n, F_k\} = 0$$

$$\mathrm{Cor}\{U_n, U_m\} = 0$$

risk-free return

$$\mathrm{E}\{\mathbf{R}\} \approx r\mathbf{1} + \mathbf{b}_1\left(\mathrm{E}\left\{R_{\mathbf{w}_{(1)}}\right\} - r\right) + \cdots + \mathbf{b}_K\left(\mathrm{E}\left\{R_{\mathbf{w}_{(K)}}\right\} - r\right)$$

1st factor-mimicking pseudo portfolio

$$\mathbf{w}'_{(1)}\mathbf{B} \equiv (1, 0, \ldots, 0)$$

K-th factor-mimicking pseudo portfolio

$$\mathbf{w}'_{(K)}\mathbf{B} \equiv (0, 0, \ldots, 1)$$

Arbitrage Pricing Theory (APT):

- Only refers to returns, not generic variables X

- Refers to yet-to-be realized returns, no regression

- Refers to returns of any asset (exotic derivatives, etc.), not only stocks

- Is a constraint on expectations only

- Is not a “time series” factor model with K explicit factor, the mimicking portfolios

- Has identification issues