Lecture 5: Factor pricing models and arbitrage pricing theory

Bernardino Adao

ISEG, Lisbon School of Economics and Management

March 7, 2025

Overview

- The discount factor of the simple consumption-based model does not perform well empirically
- Linear **factor pricing models** are *more or less ad-hoc* ways to solve that problem
- Are based on empirical observations, unlike the **APT**, which is based on arbitrage pricing
- A k factor model explains the expected excess return on portfolio i according to

$$E\left(R^i\right) - R^f = \beta_{i,1} E\left(f_1\right) + \ldots + \beta_{i,k} E\left(f_k\right)$$

where the $E\left(f\right)$ are expected premiums, and the **factor loadings**, the β's, are the slopes in the time-series regression

$$R_t^i - R_t^f = \alpha_i + \beta_{i,1} f_{1,t} + \ldots + \beta_{i,k} f_{k,t} + \varepsilon_{i,t}$$

- The $E\left(f\right)$'s are obtained by doing the cross section regressions of the $E\left(R^i\right)$'s on the β's.

Overview

- For instance the Fama-Fench 3 factor model considers:
  - the market excess return: $R_m - R_f$
  - the SMB (**S**mall **M**inus **B**ig), the difference between the return on a portfolio of small stocks minus the return on a portfolio of large stocks
  - the HML (**H**igh **M**inus **L**ow), the difference between the return on a portfolio of high-book-to-market stocks minus the return on a portfolio of low-book-to-market stocks.
  - These factors are calculated with combinations of portfolios composed by ranked stocks
  - The Capitalization ranking, the Book-to-Market ranking and the available historical market data may be accessed on Kenneth French's web page

Discount factor and single factor model (or single beta model)

Proposition:

$$\begin{aligned} p &= E(mx) \text{ implies } E(R^i) = \alpha + \beta_{i,m}\lambda_m, \\ \text{where } \lambda_m &= \frac{var(m)}{E(m)},\ \alpha \equiv \frac{1}{E(m)} \text{ and } \beta_{i,m} = -\frac{cov(m, R^i)}{var(m)} \end{aligned}$$

Proof:

$$1 = E(mR^i) = E(m)E(R^i) + cov(m, R^i)$$

$$E(R^i) = \frac{1}{E(m)} - \frac{cov(m, R^i)}{E(m)}$$

$$\begin{aligned} E(R^i) &= \alpha + \left(-\frac{cov(m, R^i)}{var(m)}\right)\left(\frac{var(m)}{E(m)}\right) \\ &= \alpha + \beta_{i,m}\lambda_m \end{aligned}$$

Discount factor and single factor model

- Alternatively, can write the formula for excess returns:

$$E\left(R^{i}\right) = \frac{1}{E\left(m\right)} - \frac{cov\left(m, R^{i}\right)}{var(m)}\left(\frac{var(m)}{E\left(m\right)}\right)$$

$$E\left(R^{i}\right) = R^{f} - \frac{cov\left(m, R^{i} - R^{f}\right)}{var(m)}\left(\frac{var(m)}{E\left(m\right)}\right)$$

$$E\left(R^{ie}\right) = \beta_{i,m}\lambda_{m}$$

where $R^{ie} = R^{i} - R^{f}$, using the fact that

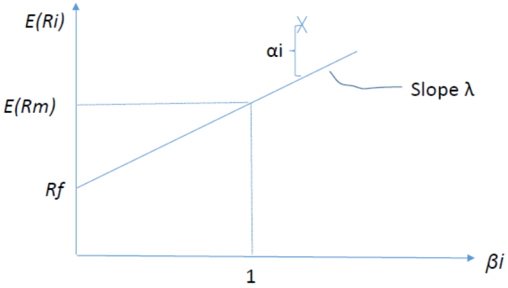
$$R^{f} = \frac{1}{E\left(m\right)} \text{ and } cov\left(m, R^{i} - R^{f}\right) = cov\left(m, R^{i}\right)$$

Discount factor and single beta model

- λ_m is the slope of this cross-sectional relationship and the model implies $\alpha_i = \alpha = R^f$
- The $cov\left(m, R^i\right)$ is in general negative. Positive expected returns are associated with positive correlation with consumption growth, and hence negative correlation with marginal utility growth (m).
- Thus, we expect

$$\beta_{i,m} = -\frac{cov\left(m, R^i - R^f\right)}{var(m)} > 0.$$

Discount factor and single beta model



Graphical Representation: Security Market Line (SML)

- The Security Market Line (SML) is a graphical representation of CAPM.
- It shows the relationship between expected return and beta.
- Assets above the SML are undervalued (offering higher return for their risk).
- Assets below the SML are overvalued (offering lower return for their risk).

Assumptions of CAPM

- Investors are rational and risk-averse.
- Markets are efficient (all available information is reflected in prices).
- Investors hold diversified portfolios (only systematic risk matters).
- No transaction costs or taxes.
- Investors can borrow and lend at the risk-free rate.
- Typically these assumptions do not hold but CAPM can still be a useful benchmark

Connecting SDF to CAPM

- Under some assumptions (such as mean-variance preferences or log-normal returns), the SDF takes the form:

$$m_{t+1} = a + bR^m_{t+1}$$

  where a and b are constants related to risk aversion and consumption preferences.

  - Exercise: using general pricing $E_t\left(m_{t+1}\left(R^i_{t+1} - R^f_{t+1}\right)\right) = 0$ and replacing m_{t+1} obtain $E\left(R^i\right) - R^f = \beta_i E\left(R^m - R^f\right)$, where $\beta_i = -\frac{cov\left(R^m, R^i\right)}{var\left(R^m\right)}$

Multiple factor pricing and multiple beta models

- The single beta model does not work well in practice
- Average excess returns rise from growth (low book-to-market, "high price") to value (high book-to-market, "low price").
- Figure below (for 10 portfolios) includes the results of multiple regressions on the market excess return and Fama and French's hml factor,

$$R^e_{i,t} = a_i + b_i \times rmrf_t + h_i \times hml_t + \varepsilon_{i,t}$$

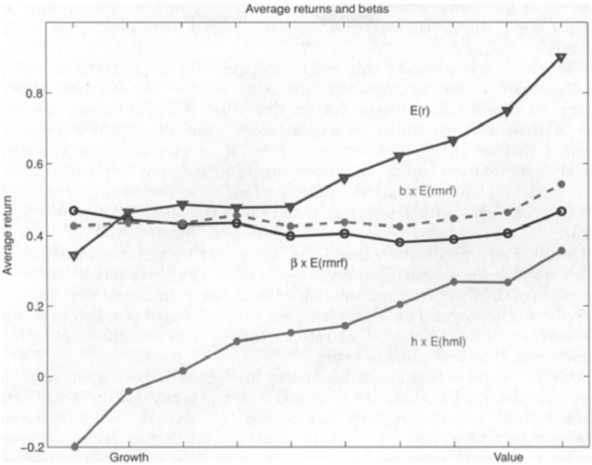


Figure 6. Average returns and betas. 10 Fama–French book-to-market portfolios. Monthly data, 1963–2010.

TABLE 1
Summary Statistics and Three-Factor Regressions for Simple Monthly Percent Excess Returns on 25 Portfolios Formed on Size and BE/ME: 7/63–12/93, 366 Months

	Book-to-market equity (BE/ME) quintiles									
Size	Low	2	3	4	High	Low	2	3	4	High
	Panel A: Summary statistics									
	Means					Standard deviations				
Small	0.31	0.70	0.82	0.95	1.08	7.67	6.74	6.14	5.85	6.14
2	0.48	0.71	0.91	0.93	1.09	7.13	6.25	5.71	5.23	5.94
3	0.44	0.68	0.75	0.86	1.05	6.52	5.53	5.11	4.79	5.48
4	0.51	0.39	0.64	0.80	1.04	5.86	5.28	4.97	4.81	5.67
Big	0.37	0.39	0.36	0.58	0.71	4.84	4.61	4.28	4.18	4.89

FIGURE 1 Fama and French (1996), Table 1.

	Book-to-Market Equity (BE/ME) Quintiles									
Size	Low	2	3	4	High	Low	2	3	4	High
Panel B: Regressions: $R_i - R_f = a_i + b_i(R_M - R_f) + s_i SMB + h_i HML + e_i$										
	a					t(a)				
Small	−0.45	−0.16	−0.05	0.04	0.02	−4.19	−2.04	−0.82	0.69	0.29
2	−0.07	−0.04	0.09	0.07	0.03	−0.80	−0.59	1.33	1.13	0.51
3	−0.08	0.04	−0.00	0.06	0.07	−1.07	0.47	−0.06	0.88	0.89
4	0.14	−0.19	−0.06	0.02	0.06	1.74	−2.43	−0.73	0.27	0.59
Big	0.20	−0.04	−0.10	−0.08	−0.14	3.14	−0.52	−1.23	−1.07	−1.17
	b					t(b)				
Small	1.03	1.01	0.94	0.89	0.94	39.10	50.89	59.93	58.47	57.71
2	1.10	1.04	0.99	0.97	1.08	52.94	61.14	58.17	62.97	65.58
3	1.10	1.02	0.98	0.97	1.07	57.08	55.49	53.11	55.96	52.37
4	1.07	1.07	1.05	1.03	1.18	54.77	54.48	51.79	45.76	46.27
Big	0.96	1.02	0.98	0.99	1.07	60.25	57.77	47.03	53.25	37.18
	s					t(s)				
Small	1.47	1.27	1.18	1.17	1.23	39.01	44.48	52.26	53.82	52.65
2	1.01	0.97	0.88	0.73	0.90	34.10	39.94	36.19	32.92	38.17
3	0.75	0.63	0.59	0.47	0.64	27.09	24.13	22.37	18.97	22.01
4	0.36	0.30	0.29	0.22	0.41	12.87	10.64	10.17	6.82	11.26
Big	−0.16	−0.13	−0.25	−0.16	−0.03	−6.97	−5.12	−8.45	−6.21	−0.77
	h					t(h)				
Small	−0.27	0.10	0.25	0.37	0.63	−6.28	3.03	9.74	15.16	23.62
2	−0.49	0.00	0.26	0.46	0.69	−14.66	0.34	9.21	18.14	25.59
3	−0.39	0.03	0.32	0.49	0.68	−12.56	0.89	10.73	17.45	20.43
4	−0.44	0.03	0.31	0.54	0.72	−13.98	0.97	9.45	14.70	17.34
Big	−0.47	0.00	0.20	0.56	0.82	−18.23	0.18	6.04	18.71	17.57
	R^2					s(e)				
Small	0.93	0.95	0.96	0.96	0.96	1.97	1.49	1.18	1.13	1.22
2	0.95	0.96	0.95	0.95	0.96	1.55	1.27	1.28	1.16	1.23
3	0.95	0.94	0.93	0.93	0.92	1.44	1.37	1.38	1.30	1.52
4	0.94	0.92	0.91	0.88	0.89	1.46	1.47	1.51	1.69	1.91
Big	0.94	0.92	0.87	0.89	0.81	1.19	1.32	1.55	1.39	2.15

Multiple Factor pricing and multiple beta models

- The table shows that small stocks tend to have higher returns than big stocks
- high book to market stocks tend to have higher returns than low book to market stocks
- the estimated intercepts say that the model leaves:
  - a large negative unexplained part for the portfolios of the smallest size and lowest book to market quintiles
  - and a large positive unexplained return for the portfolio of stocks in the largest size and lowest book to market quintiles
  - otherwise the intercepts are close to zero

Multiple Factor pricing and multiple beta models

- Derived from observed patterns in asset returns
- Uses specific well-known factors
- Examples of Factor Price Models:
  - Fama-French Three-Factor Model: Includes market risk, size (SMB), and value (HML).
  - Carhart Four-Factor Model: Adds momentum as a fourth factor.
  - Five-Factor Models (e.g., Fama-French 5-Factor Model): Adds investment and profitability factors.

Summing-up

- CAPM (single beta model) works when the stocks are grouped by size only
- Does not work when stocks are grouped by book to market ratio: does not price well value and growth stocks
- This observation motivates efforts to tie the discount factor m to other data
- Linear factor pricing models are the most popular models of this sort in finance

Multiple Factor pricing and multiple beta models

- **Factor pricing models** replace the consumption-based expression for marginal utility growth with a linear model of the form

$$m_{t+1} = a + \mathbf{b}'\mathbf{f}_{t+1}$$

a and $\mathbf{b}$ are free parameters and $\mathbf{f}_{t+1}$ are the factors.

- This specification is equivalent to a **multiple-beta model**

$$E(\mathbf{R}) = \alpha + \boldsymbol{\beta}'\boldsymbol{\lambda}_{\mathbf{f}}$$

- **Procedure:** Get the $\boldsymbol{\beta}^i$ by running the regression

$$R^i_{t+1} = a + \boldsymbol{\beta}^{i\prime}\mathbf{f}_{t+1} + \varepsilon_{it+1}.$$

After that, the $\boldsymbol{\lambda}_{\mathbf{f}}$ is obtained by running the regression of $E(R^i_{t+1})$ on the the $\boldsymbol{\beta}^{i\prime}$

Multiple Factor pricing and multiple beta models

Theorem:

$$m_{t+1} = a + \mathbf{b}'\mathbf{f}_{t+1} \Longleftrightarrow E(\mathbf{R}) = \alpha + \boldsymbol{\beta}'\boldsymbol{\lambda}_{\mathbf{f}}$$

- It is easier to prove for excess returns.
- In this case

$$E\left(mR^{e}\right) = 0$$

and we do not get the value for $E\left(m\right)$. Thus, we can normalize it to any constant for instance $E\left(m\right) = 1$ or

$$m = 1 + \mathbf{b}'[\mathbf{f} - E(\mathbf{f})]$$

Multiple Factor pricing and multiple beta models

Theorem: Given the model

$$m = 1 + \mathbf{b}'[\mathbf{f} - E(\mathbf{f})] \text{ with } E\left(mR^e\right) = 0 \tag{1}$$

one can find $\boldsymbol{\lambda}_\mathbf{f}$ such that

$$E(\mathbf{R}^e) = \boldsymbol{\beta}'\boldsymbol{\lambda}_\mathbf{f} \tag{2}$$

where $\boldsymbol{\beta}$ are the multiple regression coefficients of excess returns $\mathbf{R}^e$ on the factors.
Conversely, given $\boldsymbol{\lambda}_\mathbf{f}$ in (2), we can find **b** such that (1) holds.

Multiple Factor pricing and multiple beta models

Proof: From (1),

$$0 = E\left(mR^e\right) = E\left(R^e\right) + cov(R^e, \mathbf{f}')\mathbf{b}$$

Thus,

$$E\left(R^e\right) = -cov(R^e, \mathbf{f}')\mathbf{b}$$

Divide and multiply by $var(\mathbf{f})$

$$E(R^e) = -cov(R^e, \mathbf{f}')var(\mathbf{f})^{-1}var(\mathbf{f})\mathbf{b} = \boldsymbol{\beta}'\boldsymbol{\lambda}_{\mathbf{f}}$$

where $\boldsymbol{\lambda} = -var(\mathbf{f})\mathbf{b}$.

Multiple Factor pricing and multiple beta models

- **What should one use for factors $\mathbf{f}$** ?
  - Factor pricing models look for variables that are good proxies for aggregate marginal utility growth, i.e., variables for which

$$\frac{\beta u'(c_{t+1})}{u'(c_t)} \approx a + \mathbf{b}'\mathbf{f}_{t+1}$$

-
  - Consumption is related to: (i) returns on broadbased portfolios, (ii) interest rates, (iii) GDP growth, (iv) investment, (v) other macroeconomic variables, and (vi) variables that forecast income in the future like: term premium, dividend/price ratio, stock returns, etc.

Multiple Factor pricing and multiple beta models

- **Conclusion:** Factors should be thought as proxies for marginal utility growth
- **Important:** All factor models are derived as specializations of the consumption-based model
- **The idea:**
  - (i) Start with a general equilibrium model which produces relations that express the determinants of consumption from exogenous variables and other endogenous variables; equations of the form

$$c_t = g(f_t).$$

-
  - (ii) use this kind of equation to substitute out for consumption in the basic first order conditions.

Capital Asset Pricing Model (CAPM)

- It was independently developed by Lintner (1965), Mossin (1964) and Sharpe (1964)
- The CAPM is

$$m = a + bR_w$$

R_w = wealth portfolio return

CAPM

- The CAPM is the first, most famous and was the most widely used model in asset pricing
- The values for the parameters a and b are found by requiring the discount factor m price any two assets
  - For instance with
    $$1 = E(mR_W)$$
    and
    $$1 = E(m)R_f$$
    get 2 equations and 2 unknowns
  - It is conventional to proxy R_W by the return on a broad-based stock portfolio such as a NYSE index or a S&P500 index.

CAPM

- The CAPM is more often expressed in its beta representation

$$E(R_i) = \alpha + \beta_{i,R_w}\left[E\left(R_w\right) - \alpha\right]$$

- There are many derivations of the CAPM

1) Two period quadratic utility;

$$U(c_t, c_{t+1}) = -(c_t - c^*)^2 - \beta E[(c_{t+1} - c^*)^2]$$

- the quadratic utility assumption means marginal utility is *linear* in consumption

CAPM

- the constraints are

$$c_{t+1} = W_{t+1}$$

$$W_{t+1} = R^{W}_{t+1}(W_t - c_t)$$

$$R^{W} \equiv \sum_{i=1}^{N} \theta_i R_i$$

$$\sum_{i=1}^{N} \theta_i = 1.$$

CAPM

$$\Rightarrow m_{t+1} = \beta \frac{c_{t+1} - c^*}{c_t - c^*} = \beta \frac{R^W_{t+1}(W_t - c_t) - c^*}{c_t - c^*}$$

$$= -\frac{\beta c^*}{c_t - c^*} + \frac{\beta (W_t - c_t)}{c_t - c^*} R^W_{t+1}$$

$$\Longleftrightarrow m_{t+1} = a_t + b_t R^W_{t+1}$$

CAPM

2) One period, exponential utility $u(c) = -e^{-\theta c}$, and normal returns;

$$E[u(c)] = E\left(-e^{-\theta c}\right)$$

θ is the coefficient of absolute risk aversion.

- If consumption is normally distributed, we have

$$Eu(c) = -e^{-\theta E(c) + \frac{1}{2}\theta^2 \sigma^2(c)}$$

the budget constraint is

$$c = y^f R_f + \mathbf{y}^T \mathbf{R}$$

$$W = y^f + \mathbf{y}^T \mathbf{1}$$

where $\left(y^f_, \mathbf{y}^T\right)$ is the vector of investments in the riskless and risky assets

CAPM

$$\Rightarrow Eu(c) = -e^{-\theta E(y^f R_f + \mathbf{y}^T \mathbf{R}) + \frac{1}{2}\theta^2 \sigma^2 (y^f R_f + \mathbf{y}^T \mathbf{R})}$$

$$\Rightarrow Eu(c) = -e^{-\theta(y^f R_f + \mathbf{y}^T E\mathbf{R}) + \frac{1}{2}\theta^2 \mathbf{y}^T \Sigma \mathbf{y}}$$

Lagrangian:

$$L = -e^{-\theta(y^f R_f + \mathbf{y}^T E\mathbf{R}) + \frac{1}{2}\theta^2 \mathbf{y}^T \Sigma \mathbf{y}} + \lambda \left[W - y^f - \mathbf{y}^T \mathbf{1} \right]$$

Maximization of this expression w.r.t. $\left(y^f_, \ \mathbf{y}^T \right)$

$$-\theta R_f e^{-\theta(y^f R_f + \mathbf{y}^T E\mathbf{R}) + \frac{1}{2}\theta^2 \mathbf{y}^T \Sigma \mathbf{y}} - \lambda = 0$$

$$\left(-\theta E R_i + \theta^2 \sum_{j=1}^{N} y_j cov\left(R_j, R_i\right) \right) e^{-\theta(y^f R_f + \mathbf{y}^T E\mathbf{R}) + \frac{1}{2}\theta^2 \mathbf{y}^T \Sigma \mathbf{y}} - \lambda = 0$$

CAPM

- Rewriting the second equation

$$\left(-\theta E\mathbf{R}+\theta^2\mathbf{y}^T\Sigma\right)e^{-\theta(y^fR_f+\mathbf{y}^TE\mathbf{R})+\frac{1}{2}\theta^2\mathbf{y}^T\Sigma\mathbf{y}}-\lambda\mathbf{1}=0$$

- Taking the ratio of the 2 equations

$$-\theta R_f\mathbf{1}=\left(-\theta E\mathbf{R}+\theta^2\mathbf{y}^T\Sigma\right)$$

$$\Rightarrow\frac{E\mathbf{R}-R_f\mathbf{1}}{\theta}\Sigma^{-1}=\mathbf{y}^T$$

Conclusion: investors invest more in risky assets if their expected return is higher, less if the risk aversion coefficient is higher, and less if assets are riskier

CAPM

the expression above is

$$\Leftrightarrow E\mathbf{R} - R_f = \theta \mathbf{y}^T \Sigma$$

$$\Rightarrow ER_i - R_f = \theta cov(R_i, \mathbf{y}^T \mathbf{R})$$

$$\Rightarrow ER_i - R_f = \theta cov(R_i, y^f R_f + \mathbf{y}^T \mathbf{R})$$

Define the market rate of return

$$R_m = y^f R_f + \mathbf{y}^T \mathbf{R}$$

CAPM

$$\Rightarrow ER_i = R_f + \theta cov(R_i, R_m)$$

$$\Rightarrow \frac{(ER_m - R_f)}{var(R_m)} = \theta$$

$$\Rightarrow ER_i = R_f + \frac{cov(R_i, R_m)}{var(R_m)} (ER_m - R_f)$$

$$\Leftrightarrow ER_i = \alpha + \beta_{R_i,R_m} (ER_m - \alpha)$$

Conclusion: the coeficient of absolute risk aversion is proportional to the price of risk

CAPM

3) Infinite horizon, log utility and normally distributed returns.
Suppose the investor has log utility

$$u(c_t) = \log c_t$$

$$\Leftrightarrow u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma} \text{ with } \gamma = 1$$

CAPM

the wealth portfolio is the claim to all consumption dividends

$$p_t\left(W\right) = E_t \sum_{j=1}^{\infty}\left(m_{t+j}c_{t+j}\right) = E_t \sum_{j=1}^{\infty} \frac{\beta^j u'\left(c_{t+j}\right)}{u'\left(c_t\right)} c_{t+j}$$

$$= E_t \sum_{j=1}^{\infty} \beta^j \left(\frac{c_t}{c_{t+j}}\right) c_{t+j} = \frac{\beta}{1-\beta} c_t$$

$$R_{w,t+1} = \frac{p_{t+1}\left(W\right) + c_{t+1}}{p_t\left(W\right)} = \frac{\left(\frac{\beta}{1-\beta} + 1\right) c_{t+1}}{\left(\frac{\beta}{1-\beta}\right) c_t} = \frac{1}{\beta}\frac{c_{t+1}}{c_t} = \left[\frac{\beta u'\left(c_{t+1}\right)}{u'\left(c_t\right)}\right]^{-1}$$

conclusion: the return on the wealth portfolio is proportional to consumption growth and discount factor equals the inverse of the wealth portfolio return

$$m_{t+1} = \left(R_{w,t+1}\right)^{-1}$$

Arbitrage Pricing Theory (APT)

- The APT was introduced by Ross (1976) as an alternative to the **FPM**
- Explains asset returns through arbitrage
- APT is

$$m_{t+1} = a + \mathbf{b}^T \mathbf{f}_{t+1}$$

"there is a discount factor linear in the vector **f** that prices returns"

- The APT is "more general" than the **FPM** as the factors are not predefined
- Any factors can be used
  - statistical factor identification

Arbitrage Pricing Theory (APT)

- **Approximate APT**
- Consider a **statistical** characterization for the payoff of asset i

$$x_i = E(x_i) + \sum_{j=1}^{M} \beta_{i,j} \tilde{f}_j + \varepsilon_i, i = 1, 2, ... N$$

where

$$\tilde{f} \equiv f - E(f)$$

- the factor decomposition can be regarded as a regression equation with

$$E(\varepsilon_i) = cov(\varepsilon_i, \tilde{f}_j) = 0, \text{ all } i \text{ and } j$$

Arbitrage Pricing Theory (APT)

- **The APT assumes** that

$$E(\varepsilon_i \varepsilon_k) = 0, \text{ for } i \neq k$$

This imposes a **restriction on the covariance matrix** of the payoffs $\mathbf{x}$

$$\begin{aligned} cov(x_i, x_k) &= E\left[(x_i - Ex_i)(x_k - Ex_k)\right] \\ &= E\left(\sum_{j=1}^{M} \beta_{i,j}\tilde{f}_j + \varepsilon_i\right)\left(\sum_{j=1}^{M} \beta_{k,j}\tilde{f}_j + \varepsilon_k\right) \\ &= \left(\sum_{j=1}^{M} \beta_{i,j}\beta_{k,j} E\left(\tilde{f}_j\right)^2 + \sum_{j\neq l}^{M}\sum_{j=1}^{M} \beta_{i,j}\beta_{k,l} E\left(\tilde{f}_j\tilde{f}_l\right) + E\left(\varepsilon_i\varepsilon_k\right)\right) \end{aligned}$$

where $E(\varepsilon_i \varepsilon_k) = 0$ for $i \neq k$ and $= \sigma^2(\varepsilon_i)$ for $i = k$

- That is the idiosyncratic terms ε_i must be uncorrelated

Arbitrage Pricing Theory (APT)

- The intuition behind the APT is that the completely idiosyncratic movements in asset returns should not carry any risk prices, since investors can diversify idiosyncratic returns away by holding diversified portfolios
- Therefore, risk prices or expected returns on a security should be related to the security's covariance with the common components or "factors" only
- It is important to explore under what conditions the idiosyncratic components have zero (or small) risk prices, so that only the common components matter to asset pricing
- If there were no residual, then we could price securities from the factors by arbitrage (by the law of one price)

Arbitrage Pricing Theory (APT)

- Can estimate a factor structure by running regressions if the factors are known.
  - For instance, the market (value-weighted portfolio), industry portfolios, size and book/market portfolios, small minus big portfolios, momentum portfolios, etc
- However, most of the time do not know the identities of the factor portfolios ahead of time.
  - In this case we have to use one of several statistical techniques under the broad heading of factor analysis (that is where the word “factor” came from) to estimate the factor model

Arbitrage Pricing Theory (APT)

- With multiple (orthogonalized) factors, we obtain

$$cov(x_i, x_k) = \left(\sum_{j=1}^{M} \beta_{i,j} \beta_{k,j} E\left(\tilde{f}_j\right)^2 + E\left(\varepsilon_i \varepsilon_k\right) \right)$$

Exact APT

- Suppose there was no idiosyncratic term i.e. we have an *exact factor model*

$$x_i = E\left(x_i\right) 1 + \boldsymbol{\beta}_i^T \tilde{\mathbf{f}}$$

then the price can only depend on the price of factors

$$p\left(x_i\right) = E\left(x_i\right) p\left(1\right) + \boldsymbol{\beta}_i^T p(\tilde{\mathbf{f}})$$

Arbitrage Pricing Theory (APT)

- with exact factor pricing and $x_i = R_i$

$$1 = E\left(R_i\right)\frac{1}{R_f} + \boldsymbol{\beta}_i^T p(\tilde{\mathbf{f}})$$

since $p\left(1\right) = E\left(m \cdot 1\right)$

$$\Rightarrow E\left(R_i\right) = R_f + \boldsymbol{\beta}_i^T\left[-p(\tilde{\mathbf{f}})R_f\right] = R_f + \boldsymbol{\beta}_i^T\boldsymbol{\lambda}$$

where $\boldsymbol{\lambda} = \left[-p(\tilde{\mathbf{f}})R_f\right]$

- expected returns are linear in the betas, and the constants $(\boldsymbol{\lambda})$ are related to the prices of the factors

Arbitrage Pricing Theory (APT)

In Practice

- Actual returns do not display an exact factor structure
- There is always some idiosyncratic or residual risk
- *But*, factor model regressions often have very high R^2, i.e. the idiosyncratic risks are small
- Thus, there is reason to hope that the APT holds approximately, especially for reasonably large portfolios

Arbitrage Pricing Theory (APT)

Formally:

- Assume:

$$x_i = E(x_i) + \boldsymbol{\beta}_i^T \tilde{\mathbf{f}} + \varepsilon_i$$

$$p(x_i) = E(x_i)\, p(1) + \boldsymbol{\beta}_i^T p(\tilde{\mathbf{f}}) + p(\varepsilon_i)$$

- what is the value of $p(\varepsilon_i)$? Is it small?

Arbitrage Pricing Theory (APT)

- Next we state 2 theorems that can be interpreted to say that the APT holds approximately for portfolios that either have high R^2, or well-diversified portfolios

Arbitrage Pricing Theory (APT)

Theorem: Fix a discount factor m that prices the factors, implying that $p(x_i) = E(mx_i)$, and $\sigma^2(m) \leq A$. Then, as $var(\varepsilon_i) \to 0$, $p(x_i - E(x_i)) \to p(\boldsymbol{\beta}_i^T \mathbf{f})$.
By doing some algebra:

$$\begin{aligned} var(x_i) &= var\left(E(x_i) + \boldsymbol{\beta}_i^T \tilde{\mathbf{f}} + \varepsilon_i\right) \\ &= var(\boldsymbol{\beta}_i^T \mathbf{f}) + var(\varepsilon_i) \end{aligned}$$

which is related to the regression R^2. By definition:

$$\frac{var(\varepsilon_i)}{var(x_i)} = 1 - R^2$$

The theorem says that as $R^2 \to 1$, $var(\varepsilon_i) \to 0$. But $var(\varepsilon_i) \to 0$ means that $\varepsilon_i \to 0$, i.e. ε_i is a random variable that takes values almost always very close to 0.
If m is bounded then $p(\varepsilon_i) = E(m\varepsilon_i) \to 0$.

Arbitrage Pricing Theory (APT)

Theorem: As the number of primitive assets increases, the R^2 of well-diversified portfolios increases to 1.
Proof:
Consider an equally weighted portfolio (in fact it does not need to be equally weighted...)

$$x_p = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\Rightarrow x_p = \frac{1}{N}\sum_{i=1}^{N}\left(a_i + \boldsymbol{\beta}_i^T \mathbf{f} + \varepsilon_i\right)$$

Arbitrage Pricing Theory (APT)

$$\Rightarrow x_p = \frac{1}{N}\sum_{i=1}^{N} a_i + \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\beta}_i^T \mathbf{f} + \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i$$

$$\Rightarrow x_p = a_p + \boldsymbol{\beta}_p^T \mathbf{f} + \varepsilon_p$$

where

$$var(\varepsilon_p) = var\left(\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i\right)$$

since $E(\varepsilon_i \varepsilon_k) = 0$, for $i \neq k$ and if $var(\varepsilon_i)$ is bounded then

$$\lim_{N\to\infty} var(\varepsilon_p) = 0.$$

Arbitrage Pricing Theory (APT)

- These two theorems can be interpreted to say that the APT holds approximately (in the usual limiting sense) for either portfolios that have high R^2, or well-diversified portfolios
- These 2 theorems say that if you fix m and take limits over N or ε we get a good approximation to an exact APT
- However in practice, you fix N or ε and look for an m that can price the portfolios
- **Important**: It may be possible that the approximate APT does not work because we can always choose an m sufficiently "far out" to generate an arbitrarily large price for an arbitrarily small ε_i

Arbitrage Pricing Theory (APT)

Solution: impose ad-hoc restrictions on the m
Example: impose a bound on $\sigma^2(m)$
consider only the $m \in [\underline{m}, \overline{m}]$, where

$$\underline{m} = \arg\min_{m} \left\{ p\left(x_i\right) = E\left(m x_i\right), \text{ s.t. } E(mf) = p(f), m > 0, \sigma^2(m) \leq A \right\}$$

$$\overline{m} = \arg\max_{m} \left\{ p\left(x_i\right) = E\left(m x_i\right), \text{ s.t. } E(mf) = p(f), \; m > 0, \sigma^2(m) \leq A \right\}$$

Arbitrage Pricing Theory (APT)

- If we impose a restriction on the volatility of m then we have an APT limit theorem
- **Theorem:** As $\varepsilon_i \longrightarrow 0$ and $R^2 \longrightarrow 1$, the price $p(x_i - Ex_i)$ assigned by any discount factor m that satisfies $E(mf) = p(f)$, $m > 0, \sigma^2(m) \leq A$ approaches $p(\boldsymbol{\beta}_i^T \mathbf{f})$

Principal Component Analysis (APT)

- Collect asset returns

$$\mathbf{R} = \begin{bmatrix} R_{1,1} & \dots & \dots & R_{1,T} \\ \dots & & & \dots \\ & & & \\ \dots & & & \dots \\ R_{N,N} & \dots & \dots & R_{N,T} \end{bmatrix}$$

Principal Component Analysis (APT)

- Normalize the Data (Mean-Centering & Standardization)
- Subtract the mean of each column to center the data.
- Optionally, divide by the standard deviation to standardize the data.

Principal Component Analysis (APT)

- Compute the Covariance Matrix
- The covariance matrix captures the relationships between assets:
- $$\Sigma = \frac{1}{T}\mathbf{R}\mathbf{R}'$$
- Compute Eigenvalues & Eigenvectors:

$$\Sigma \mathbf{v} = \lambda \mathbf{v}$$

  - v (eigenvectors) = Principal Components (PCs) (new uncorrelated factors)
  - λ (eigenvalues) = Variance explained by each PC.

Principal Component Analysis (APT)

- Rank Principal Components by Variance Explained
- Sort the PCs in descending order of their eigenvalues.
- The first few PCs typically explain most of the variance, allowing us to reduce dimensionality.
- Each PC is a linear combination of the original variables.

$$PC_1 = 0.5R_1 + 0.45R_2 + 0.3R_3 + \ldots$$

Principal Component Analysis (APT)

- Use these PCs as factors in an APT regression:

$$R_i = \alpha + \beta_{i,1} PC_1 + \beta_{i,2} PC_2 + \beta_{i,3} PC_3 + \varepsilon_i$$

- If the factors are economically interpretable, they can be linked to macroeconomic variables (e.g., interest rates, inflation).