

Quantitative Equity Portfolio Management

[Book Link](#)

Author(s): Ludwig and Deahwan

Publish Year: 2006

Tags: [#Portfolio_Management](#) [#Factor_Investing](#) [#idea_factor_modelling](#)

Part One: An Overview of QEPM

- "If return are somewhat predictable over the long run, then stable quantitative models should work more reliably than picking individual stocks intermittently on qualitative information"

Chapter 1: The Power of QEPM

- "Factor tilting can be a very effective way of managing an equity portfolio and is one of the many powerful tools of QEPM"
- "The first step in building a model of stock returns is to choose a mix of explanatory variables for the model"

Chapter 2: The Fundamentals of QEPM

- "Given the returns of the reference instrument, it is possible to [...] decompose the portfolio's returns into two parts—one that is related to the reference instrument and one that is not. The related part is typically called *expected returns* or *consensus returns*, and the second part is known as *residual return* or *the returns not explained by the model*"
- Benchmark α
 - "Given the portfolio returns r_P and the benchmark return r_B , we can estimate:"

$$r_P = \alpha + \beta r_B + \epsilon \quad (2.1)$$

- $\alpha + \epsilon$ is the residual return. "Benchmark α is the expected value of the residual return, and [...] ϵ is the deviation of the residual return from its mean. Equation (2.1) is constructed so that ϵ averages zero"
- CAPM α
 - "Given the portfolio returns r_P and the market return r_M , we can estimate:"

$$r_P = \alpha + \beta r_M + \epsilon \quad (2.2)$$

- "The CAPM α , or α^{CAPM} , is the risk adjusted excess return over the market"
- Multifactor α
 - "Given the portfolio returns r_P and the a series of factor returns f_1, f_2, \dots, f_K , we can estimate:"

$$r_P = \alpha + \beta_1 f_1 + \dots + \beta_K f_K + \epsilon \quad (2.3)$$

- "If all the other factors in the APT model are uncorrelated with the first factor (the benchmark), then the benchmark α is given by

$$\alpha^B = E(\alpha + \beta_1 f_1 + \dots + \beta_K f_K) \quad (2.5)$$

The benchmark α includes the effect of all the other factors. [...] In general, if we allow the benchmark α to be correlated with all other factors, the benchmark α is given by

$$\alpha^B = E(\alpha + \beta_1 f_1 + \dots + \beta_K f_K) - (\beta_2 \gamma_2 + \dots + \beta_K \gamma_K) E(f_1) \quad (2.6)$$

where γ_j is the coefficient in the regression of f_1 on f_j , i.e.:

$$\frac{Cov(f_1, f_j)}{Var(f_1)}$$

"when measuring the performance of a manager with benchmark α , benchmark α can be positive even if multi-factor α equals zero".

- "The *ex-ante* α is the expected α , whereas the *ex-post* α is the realised one. The *ex-ante* α is of interest to the quantitative portfolio manager when he or she is constructing his or her portfolio"
- "The *ex-ante* information ratio (IR) is given by:

$$IR = \frac{\alpha^B}{\omega} \quad (2.7)$$

where α^B is the expected or predicted α of the portfolio manager, and ω is the predicted standard deviation of the residual (i.e., $S(\epsilon)$)"

- **Seven Tenets of QEPM:**

1. *Markets are mostly efficient.*
2. *Pure arbitrage opportunities do not exist.*
3. *Quantitative analysis creates statistical arbitrage opportunities.*
4. *Quantitative analysis combines all the available information in an efficient way.*
5. *Quantitative models should be based on sound economic theories.*
6. *Quantitative models should reflect persistent and stable patterns.*
7. *Deviations of a portfolio from the benchmarks are justified only if the uncertainty is small enough.*

- "the fundamental law [Advances in Active Portfolio Management](#) is for *understanding* QEPM, not for *doing* QEPM. A quantitative portfolio manager does not use the law to actually create portfolios. Rather, the law provides the manager with a useful conceptual framework for analysing the QEPM process"
- "Given the definition of the information ratio in Eq. (2.7):"

$$\frac{\alpha^B}{\omega} = IC \cdot \sqrt{BR} \quad (2.9)$$

- "One of the essential tasks of QEPM is to predict future stock returns by estimating a model that specifies a relationship between the stock return and a list of explanatory variables (a.k.a *factors*). Suppose that the model specifies that the return of stock i at time t , r_{it} , is a linear function of the value of K factor premiums at time t , that is $f_{1t}, f_{2t}, \dots, f_{Kt}$:

$$r_{it} = \alpha_i + \beta_{i1}f_{1t} + \dots + \beta_{iK}f_{Kt} + \epsilon_{it} \quad (2.10)$$

where $\alpha_i, \beta_{i1}, \dots, \beta_{iK}$ are parameters to be estimated, and ϵ_{it} is the random-error term (i.e., the deviation of the stock from its expected value)"

- "The fundamental law assesses how well Eq. (2.10) explains the stock-return process, and it expresses the equation's goodness of fit as the product of *number of explanatory variables* and each variable's *average contribution*."
- "When constructing the portfolio, a manager should use all the information gathered from the estimation of Eq. (2.10)"
- **Lemma 1 (THE INFORMATION CRITERION)** "*The fundamental law of active portfolio management as expressed in Eq. (2.9) is valid only if the portfolio manager combines all the information in the most efficient way*"
 - Example: portfolio manager following analyst-revision strategy: "Construct an equal-weight portfolio of stocks whose analyst ratings improved in the last month. We can consider 'inclusion in the portfolio' a variable. That is, we may define a variable β_{it} that has a value of 1 if stock i 's rating improved at a time t and a value of 0 otherwise. The manager estimates:

$$r_{i,t+1} = \alpha + \beta_{it}f + \epsilon_{i,t+1} \quad (2.11)$$

where $r_{i,t+1}$ is the return to stock i at time $t + 1$, α and f are the parameters to be estimated, and $\epsilon_{i,t+1}$ is the error (the part of the return that is not explained by the model). If the value of f is positive, the equation suggests that those stocks with $\beta_{it} = 1$ will have higher future returns. Thus the portfolio manager constructs an equal-weighted portfolio of stocks for which $\beta_{it} = 1$, indicating that their ratings improved in the last month". But the portfolio manager is not using all the available information. "Stocks with higher volatility should smaller weights in the eventual portfolio. Since the portfolio manager did not use this piece of information, the information criterion is not satisfied."

- **Data Mining**
 - "Data mining is the practice of running regressions of historical stock returns on so many combinations of factors that one is virtually guaranteed of finding a model or handful of factors that seem significantly correlated with stock returns but are in fact not particularly meaningful."

- "We cannot completely avoid data mining or data snooping when we test a factor unless we use data that have never before been used to test that particular factor"
- "careful empirical study is not enough to avoid the data-mining problem. The best way to generate meaningful statistics is to follow tenet 5 and make sure that the model is based on sound economic theory and common sense"
- Parameter Stability
 - "For example, if we believe that β is generally not constant for more than five years, we should not estimate it from a sample that includes data covering more than five years."
- Parameter Uncertainty
 - "Even if the active portfolio does not look too risky on its surface, it could have a high degree of parameter uncertainty, which is a component of risk"
 - "The standard error, which measures the error of the estimation, depends on two things: the variance of the error in the model and the variance of the explanatory variables"
 - "the portfolio manager should consider not only the estimated value but also the standard error of the estimates. If the standard error is large, then that should be considered part of the portfolio's risk. This is one of the fundamental implications of Bayesian econometrics for portfolio selection". Discussed in Chapter 14.

Chapter 3: Basic QEPM Models

	Fundamental Factor Model	Economic Factor Model
Data type	P/E ratio, Mcap, etc.	GDP, macro variables; other variables
Factor Exposure (β)	directly observed	must be estimated using time series regression
Factor Premium (f)	must be estimated using cross-sectional regression	Directly observed (may be determined by zero-investment portfolio or PCA)
Expected Return	factor exposure x factor premium	factor exposure x factor premium
Risk	factor exposure x risk of factor premium + specific risk	factor exposure x risk of factor premium + specific risk
Criterion for selecting		
Theoretical basis		Better
Factor accommodation		Better
Ease of implementation	Better	
Data requirement	Better	
Intuitive appeal	Depends	Depends
Assumption	stable factor premium	stable factor exposure

• Basic QEPM Models and Portfolio Construction Procedures

- "*Factors* are explanatory variables that represent different types of risk. A factor model shows that the average stock return is proportional to the stock's exposure to the risk that factor represents (the *factor exposure*) and to the payoff for each unit of exposure to risk (the *factor premium*)"
- 1. **Fundamental factor model**
 - "uses fundamental factors, which are stock characteristics such as P/E ratio and market cap"
 - Factor exposure directly observable using financial statements and other data sources.

- "factor premium must be estimated from the historical relationship between stock returns and the factor exposures"

2. Economic factor model

- originally only macroeconomic variables "but is general enough to handle other types of factors as well"
- "stock's exposure to economic (or other) factors is not observable and instead must be estimated from the historical relationship between stock returns and factor premiums"
- "factor premium be determined up to a proportionality without a statistical estimation in certain cases. In other cases, it is determined by constructing *zero-investment portfolios* or through a mathematical method called *principal-component analysis*"
- "fundamental factor model and the economic factor model employ different techniques for modelling stock returns."
- "The specific steps involved in constructing a portfolio depend partly on which model is used"
- Factor choice
 - "Good factors are really the secret sauce of QEPM"
 - "The economic factor model is more flexible, allowing for both economic factors and all fundamental factors"
- Data decision
 - cross-sectional dimension: universe selected; "attributes of the data set will affect the attributes of the ultimate portfolio and number of stocks will affect the ease of estimation"
 - time-series dimension: periodicity and time period of the entire data set; "may affect parameter stability and parameter uncertainty"
 - "decide whether to use the gross return and whether to use the return in excess of a risk-free rate"
- Factor exposure (*factor loading*)
 - Straight-forward in fundamental factor model; "if the factor is the P/B ratio, the factor exposure to the stock i is the latest observed value of the P/B ratio for stock i "
 - economic model: "the factor premiums must be determined first. Then the factor exposure can be estimated from the relationship between returns and factor premiums"
 - given factor premium f_1, \dots, f_K , return of stock i , r_i , can be estimated:

$$r_i = \alpha_i + \beta_{i1}f_1 + \dots + \beta_{iK}f_K + \epsilon \quad (3.1)$$

- "estimated by a time-series regression using observations made at various time periods. [...]. For example, the factor premium f_j might be the real GDP growth. The factor exposure β_{ij} found by estimating the regression shows how sensitive stock returns are to real GDP growth"
- Factor premium
 - fundamental factor model: "factor premium is estimated from the historical relationship between the stock return and factor exposure. [...] by means of either cross-sectional regression, which involves using observations for various stocks at a single point in time, or panel regression, which entails using observations for various stocks at many points in time"
 - Macroeconomic factors: "the values of the variable themselves are taken as factor premiums. For example, if the inflation rate is 3% this month, then the premium for the inflation factor is 3%"
 - "We can determine the factor premium by calculating the return on a zero-investment portfolio of investments that exhibit the factor in question"
 - example for size factor: "A zero-investment portfolio could be created by going long a small-cap sub-portfolio and shorting a equivalent amount of a large-cap sub-portfolio. The factor premium on size would be the difference between the return of the small-cap sub-portfolio and the return of the large-cap sub-portfolio"
 - This is the approach I am currently thinking of taking, i.e. calculate factor premiums first and then calculate the factor loadings. This is the approach Sanskar was talking about 'being more statistical' in nature. Sanskar is using the Fundamental Factor Model, as opposed to the approach I have in mind: Economic Factor Model
- Expected return
 - Estimation of the model "essentially allows us to determine the expected returns of stocks based on their factor exposures and the factor premiums"
- Risk
 - "Non-diversifiable risk is represented by $r_i = \alpha_i + \beta_{i1}f_1 + \dots + \beta_{iK}f_K$ in Eq. (3.1)"

- "Diversifiable risk, on the other hand, can be removed from the portfolio by diversifying holdings. It is captured by ϵ_i . often called *stock-specific risk*"
- Stock's total risk can be estimated by "calculating the variance or the standard deviation of each component of risk"
- Forecasting
 - "In all likelihood, the factor exposure will not change in the immediate future, but the factor premium *is* likely to change in a very short amount of time. Thus it is usually necessary to forecast the factor premium. We explain how to forecast in Chapter 8"
- Security Weighting
 - "With the estimates of stock returns and risks based on the new forecasts, the portfolio manager can use optimisation techniques to select the stocks for the portfolio and assign them their respective weights"
- **The Equivalence of The Basic Models** #idea_factor_modelling
 - **Lemma 5 (THE EQUIVALENCE OF FACTOR MODELS)** "*The fundamental factor model and the economic factor model are equivalent to each other if the expected stock return is a linear function of the fundamental factor exposure*"
 - Proof: "First, we assume that expected stock returns are a linear function of fundamental factor premiums, as described by the fundamental factor mode. Then, under this assumption, we show that the expected stock return can be expressed as a linear function of the factor exposure, as described by the economic factor model"
 - "For practitioners, the equivalence of the factor models means that to the extent that the fundamental factor model is a correct model, it does not really matter whether one uses the fundamental factor model or the economic factor model. The two models produce essentially identical results. The only catch is that if the fundamental factor is not correct, then it is best to use the economic factor model because it has a stronger theoretical basis than the fundamental factor model does"
- The Screening and Ranking of Stocks with the z-score
 - "a Z-score is stock's standardised exposure to a fundamental factor"
- Hybrids of the Models and the Information Criterion
 - "Practitioners often try to add extra inputs to the basic QEPM models by combining them with ad hoc models. [...] The most critical problem with hybrids is that they frequently violate the information criterion."
 - Example of the danger of combining a fundamental factor model with a *z*-score analysis, i.e. "calculates the expected return based on the *z*-score method and adds this expected return to the fundamental factor model as a constant"
 - Two stocks, and their returns determined by "P/E ratios of the previous period":

$$r_{A,t+1} = 0.1 \left(\frac{P}{E} \right)_{A,t} + \epsilon_{A,t+1} \quad \epsilon_{A,t+1} \sim N(0, 20) \quad (3.8)$$

$$r_{B,t+1} = 0.1 \left(\frac{P}{E} \right)_{B,t} + \epsilon_{B,t+1} \quad \epsilon_{B,t+1} \sim N(0, 10) \quad (3.9)$$

"We also assume that the covariance between the two random variables ϵ_A and ϵ_B is zero." Suppose P/E for A = 20 and B = 10. "Suppose further that the average and the variance [of P/E] are constant over time and across stocks"

• The Z-Score Model

- Table 3.2 shows the assumed P/E ratios and the subsequent *z*-score.

TABLE 3.2

Price-to-Earnings Ratio and the Z-Score

	Stock A	Stock B	Mean	Variance	SD
$\left(\frac{P}{E}\right)_T$	20	10	15	25	5
z_T	1	-1	0	1	1

- Using the *z*-score, prediction of the expected stock return:

$$r_{i,t+1} = a_i + bz_{it} + v_{i,t+1} \quad i = A, B; t = 1, \dots, T-1 \quad (3.10)$$

where z_{it} is the z -score of the stock i at time t , a_i is a constant, and v_{t+1} is the error. Using Table 3.2 estimates of a_A, a_B , and b ("by demeaning i.e., subtracting the mean from both sides, of the equation and applying the standard ordinary least squares (OLS) formula"):

$$b = \frac{1}{2} \frac{C(r_A, z_A)}{V(z_A)} + \frac{1}{2} \frac{C(r_B, z_B)}{V(z_B)}$$

$$= \frac{1}{2} \sqrt{V\left(\frac{P}{E}\right)_A} \frac{C\left(r_A, \left(\frac{P}{E}\right)_A\right)}{V\left(\frac{P}{E}\right)_A} + \frac{1}{2} \sqrt{V\left(\frac{P}{E}\right)_B} \frac{C\left(r_B, \left(\frac{P}{E}\right)_B\right)}{V\left(\frac{P}{E}\right)_B}$$

$$= \frac{1}{2} \cdot 5 \cdot 0.1 + \frac{1}{2} \cdot 5 \cdot 0.1 = 0.5$$

$$a_i = E(r_i) - bE(z_i) \quad i = A, B$$

$$= 1.5 - 0.5 \cdot 0 = 1.5$$

- $\text{Covariance}(x,y)/\text{Variance}(x)$ gives the linear association between x and y , i.e. if $y = mx + b$, then $C(x,y)/V(x) = m$, and $C(x,y)/V(y) = 1/m$
- "Given the time T value of the Z -score, the preceding estimates imply the following expected returns for stock A and stock B:"

$$E(r_{A,T+1}) = a_A + bz_{A,T} = 1.5 + 0.5 \cdot 1 = 2 \quad (3.14)$$

$$E(r_{B,T+1}) = a_B + bz_{B,T} = 1.5 + 0.5 \cdot (-1) = 1 \quad (3.15)$$

"The Z -score model itself does not create any problem with efficient use of information. The problem arises when the Z -score model is combined incorrectly with other models"

• A Hybrid of the Z-Score Model and a Fundamental Factor Model

- "Adam first estimated the fundamental factor model, i.e., Eq. (3.1). Then he did the Z -score analysis. Now he create his hybrid model by setting the term α in the fundamental factor model according to the expected return from the Z -score model. [...] For time $T + 1$, he incorrectly modifies Eq. (3.1) into:"

$$r_{A,t+1} = 1 + 0.1 \left(\frac{P}{E} \right)_{A,t} + \epsilon_{A,t+1} \quad \epsilon_{A,t+1} \sim N(0, 20) \quad (3.16)$$

$$r_{B,t+1} = 1 + 0.1 \left(\frac{P}{E} \right)_{B,t} + \epsilon_{B,t+1} \quad \epsilon_{B,t+1} \sim N(0, 20) \quad (3.17)$$

- "To construct his optimal tangent portfolio, Adam finds" the weight of stock A to be 100% and stock B to be 0%.
- Information Loss
 - "he is actually not achieving the maximum Sharpe ratio because his estimates of the expected return are not correct"
 - "Given the actual Sharpe ratio and the maximum Sharpe ratio, the information loss (IL) is the difference between the two" which is 0.1005 in this case.
- "The moral of the story is clear. The best thing to do is to choose one model and stick with it. Combining basic models is hard to justify [...]. It does no good to tack an ad hoc model onto the one that the software generates, so designing an entire custom model may be the best course of action."
- Choosing the right model
 - "The economic factor model upholds tenet 5 better than the fundamental factor does. [...] An economic factor model can measure the sensitivity of stocks to various economic risk factors."
 - fundamental factor typically assumes "the factor premium is stable for some time. Similarly, in implementing the economic factor model, the assumption is often that the factor exposure is stable for some time".
 - "If a portfolio manager wants to use a mix of all types of factors, the economic factor model is the necessary choice."
 - "For fundamental factor model, only factor premiums need to be estimated. For the economic factor models, factor exposures of individual stocks need to be estimated, and forecasts of factor premiums are usually needed as well."
 - Fundamental factor models do not require a large amount of historical data other than to "estimate the variance of individual stock returns, which is ultimately required to build optimised portfolios." "On the other hand, to estimate the economic factor model, portfolio managers have to gather a relatively large period of historical returns because the estimation of factor exposures requires a time-series analysis of returns."

Part Two: Portfolio Construction and Maintenance

Chapter 4: Factors and Factor Choice

	Univariate Regression Tests	Multiple Regression Tests	Unidimensional Zero-Investment Portfolio	Multidimensional Zero-Investment Portfolio
Used in	Fundamental Factor Model	Fundamental Factor Model	Economic Fac... Model	Economic Factor Model
Equation	$r_{i,t} = \alpha + f\beta_{i,t} + \epsilon_{i,t}$	$r_{i,t} = \alpha + f_1\beta_{i1,t} + \dots + f_K\beta_{iK,t} + \epsilon_{i,t}$		
Pros	simple to perform; to get an early idea of what might be relevant	considers relationship among factors		
Cons	not helpful if the factors are correlated	Data mining is a big risk		
Remarks		ideal for <10 factors		

- "Factors come in many varieties: fundamental, technical, macroeconomic, and alternative".
- "Good factors therefore exhibit relationships with stock returns that not only are stable and persistent but also can be explained by economic theory".
- Fundamental Factors
 - "We group the fundamental factors into six subcategories:"
 1. Valuation factors
 2. Solvency factors
 - "companies with high solvency factor values are more liquid or solvent and less likely to be forced into bankruptcy by payments owed on debt and other liabilities"
 3. Operational efficiency factors
 - "attempt to describe how efficiently a firm operates over both the short and the long term"
 - Inventory turnover ratio, total asset turnover, etc.
 4. Operational profitability factors
 - "One or more of these operating profitability factors (or transformations of these factors) appear in most QEPM models"
 5. Financial risk factors
 - high debt-to-equity ratios are not always bad; "businesses that generate large, stable cash flows can handle relatively large amounts of debt because they can finance the interest payments on the debt"
 - In 2003, IBM's D/E ratio of 0.61 was well above the industry average of 0.35, "but its interest coverage of 63.11 times also was greater than the industry average. Its earning power covered its debt burden."
 - "Low financial risk does not just avert disaster; it is a big plus for expansion and new investment"
 6. Liquidity risk factors
 - "gauge whether a stock's liquidity matches trading demand"
 - Trading Turnover "is calculated as the average daily trading volume (ADV) divided by market capitalisation, measures the percentage of outstanding shares traded in a given period."
- Technical Factors
 - "Portfolio managers typically use technical factors to capture very short-term changes in the relative value of stocks."
- Economic Factors
 - "For macroeconomic models of stock returns, the portfolio manager should choose factors that affect all stock returns to some degree"
- Alternative Factors

- Analyst factors
 - "Wall Street analysts' earnings forecasts, buy-sell recommendations, and related information can be valuable in forecasting stock returns."
 - "many quants look for good neglected (not widely covered by analysts) stocks using a multi-factor model that examines fundamental characteristics and assigns a bonus to neglected status"
- Corporate finance policy factors
 - stock buybacks and insider purchases.
- social responsibility factors
 - "Probably the most useful factor in this arena is the employee relations factor"

• Factor Choice #idea_factor_modelling

- "To choose factors for the fundamental factor model, we suggest using the **univariate** and **multiple regression** techniques. To choose factors for the economic factor model, we suggest using the **unidimensional** and **multidimensional zero-investment portfolio** techniques"

• **Univariate regression Tests**

- "Many portfolio managers [...] performing a series of simple regressions of factors. [...] run panel regressions on each factor versus the stock returns in the universe. Thus"

$$r_{i,t} = \alpha + f\beta_{i,t} + \epsilon_{i,t} \quad (4.1)$$

"estimate f from this panel regression [...]. If a factor has a significant value of f , then the factors is useful in explaining stock returns".

- This is not really helpful if the factors chosen are highly correlated with each other.
- Helpful to provide first level cut if there are too many factors

• **Multiple regression Tests**

- "if the portfolio manager includes too many factors in this regression, there is a chance of *misspecification bias*"
- If there are fewer than 10 factors, "then a multivariate panel regression can be estimated:"

$$r_{i,t} = \alpha + f_1\beta_{i1,t} + \dots + f_K\beta_{iK,t} + \epsilon_{i,t} \quad (4.2)$$

"Once this regression is estimated, all insignificant factors are dropped. The significant factors go into the factor model."

- "One of the biggest dangers here is data mining"
- "Data mining can be a serious problem when one uses the *stepwise regression* technique. This method involves starting with a collection of factors and then running a regression on every possible combination of the factors and picking the set of factors with the largest R^2 ."

• **Unidimensional Zero-Investment Portfolio**

- "Typically, a manager will split the universe of stocks conditional on a particular factor exposure into thirds, quintiles, or deciles. [...] a portfolio is created from the first division, and another portfolio is created from the last division. The returns of the top division are subtracted from those of the last division"
- Steps for P/B ratio factor:
 1. "rank the universe of stocks by their P/B ratio exposure".
 2. "create an equal-weighted portfolio stocks in the first quintile and an equal-weighted portfolio of stocks in the fifth quintile".
 3. "compute the returns of the two portfolios for each month"
 4. "calculate statistics on the historical returns of the first-quintile portfolio minus the fifth-quintile portfolio"
 - "repeated for every factor"
- One can do a statistical test of whether the average return of the zero-investment portfolio "is significantly different from zero". Calculate portfolio returns for T periods and obtained r_1, \dots, r_T . Then the t statistic:

$$\hat{t} = \frac{\bar{r}}{\hat{S}(\bar{r})} = \frac{\bar{r}\sqrt{T}}{\sqrt{\frac{1}{T-1} \sum_{s=1}^T (r_s - \bar{r})^2}} \quad (4.3)$$

where \bar{r} is the average return and $\hat{S}(\bar{r})$ is the standard error of the mean.

- If t statistic > 2 , "the average portfolio return is significantly different from zero, and the factor is statistically significant."

• **Multidimensional Zero-Investment Portfolio**

- "Zero-investment portfolio also can be created by considering many factor simultaneously".
- If there are two factors, each stock is assigned a ranking based on each factor. "Based on these rankings, the portfolio manager can group stocks into a number of joint quintiles or deciles". Say the portfolio manager wants to create 10 groups out of each factor, "then there will be 100 groups eventually [...] By taking intersections of these portfolios"
- "The method to create the zero-investment portfolio depends on what the portfolio manager is interested in. If he or she is interested in whether small size and low P/B ratio together influence stock returns, he or she could create the zero-investment portfolio by taking a long position on a small-low portfolio a short position on a large-high portfolio"
- Techniques to Reduce the Number of Factors [#idea_factor_modelling](#)
 1. *within-correlation coefficient*
 - "correlation coefficient adjusted for the panel structure of the data":

$$\rho(X, Y) = \frac{\sum_t C_t(X, Y)}{\sum_s \sqrt{V_s(X) V_s(Y)}} \quad (4.4)$$

2. Kendall's τ

- "shows the rank correlation (i.e., whether the ranking by one variable is related to the ranking by another variable)"
- "We calculate the numerator (the number of pairs for which the rankings by two variable are the same minus the number of pairs for which the rankings by two variables are the opposite of each other) and the denominator (the total number of pairs) for each period and then aggregate over time"

• Appendix 4A [link to the tables](#)

Chapter 5: Stock Screening and Ranking

- skipped this chapter for now

Chapter 6: Fundamental Factor Models

- In the fundamental factor model, factor exposure "it some observable fundamental characteristic of the stock. [...] The factor premium, on the other hand, is not known. It is the proportionality between the average stock return and the factor exposure, and it must be estimated empirically."
- "The average stock return is the payoff for taking nondiversifiable risk"

$$\text{Nondiversifiable risk} = \text{factor exposure}^2 \cdot \text{factor risk} \quad (6.9)$$

where factor risk is the variance of "one unit of exposure".

- Total risk of a stock:

$$V(r_i) = V(\alpha_i + \beta_{i1}f_1 + \dots + \beta_{iK}f_K) + V(\epsilon_i) \quad (6.10)$$

"The diversifiable component, [...] is the term $V(\epsilon_i)$ "

• Preliminary work

- "The primary decision if the choice of factors"
- "In practice, financial economists tend to prefer a monthly interval and, to a lesser degree, a weekly interval. Daily and annual intervals are rather uncommon"
- "Financial economists tend to include between 36 and 60 time intervals in their analyses when using a monthly interval [...]. When using a weekly interval, the number of time intervals tends to be greater than 60"

• Benchmark and α

- "If the portfolio manager aims to outperform a benchmark while minimising the portfolio's tracking error, then the benchmark must play a role in the model. One approach to incorporating the benchmark into the fundamental factor model is to use the model to predict only the residual return rather than the entire stock return". Estimate:

$$r_i = \tilde{\alpha}_i + \tilde{\beta}_i r_B + \tilde{\epsilon}_i \quad (6.14)$$

using a time-series regression. "After estimating $\tilde{\alpha}$ and $\tilde{\beta}$, we define the residual return \tilde{r}_i as"

$$\tilde{r}_i \equiv \tilde{\alpha}_i + \tilde{\epsilon}_i \quad (6.15)$$

"simply substitute the residual return for the stock return given in the fundamental factor model. [...] An alternative way to account for the benchmark, rather than predicting the residual stock return, is to add the benchmark to the model as one of the factors."

• Factor Exposure

- "once the factors are chosen, determining the factor exposure is rarely a challenging task"
- "After all the factor exposures are determined, we collect the numbers and save them as a set of vectors $\{(\beta_{11}, \dots, \beta_{N1}), \dots, (\beta_{1T}, \dots, \beta_{NT})\}$ where β_{it} is the factor exposure of stock i for time t "

• The Factor Premium

- "the factor premium is estimated from the pooled cross-sectional regression (i.e., panel regression) of the stock return on the factor exposure"
- "While the OLS estimators is simple to obtain, it may not be the most reliable estimator. We suggest that portfolio managers do a number of robustness checks on it"
- OLS estimator of the factor premium
 - "Given the returns of N stocks over T time periods, $\{(r_{11}, \dots, r_{N1}), \dots, (r_{1T}, \dots, r_{NT})\}$, and the factor exposures of N stocks over T time periods, $\{(\beta_{11}, \dots, \beta_{N1}), \dots, (\beta_{1T}, \dots, \beta_{NT})\}$, we can estimate the following:"

$$r_{it} = \beta'_{it} f + \epsilon_{it} \quad (6.17)$$

- "OLS estimator of the factor premium f is given as:"

$$\hat{f} = \frac{\sum_{t=1}^T \sum_{i=1}^N (\beta_{it} - \bar{\beta})(r_{it} - \bar{r})}{\sum_{t=1}^T \sum_{i=1}^N (\beta_{it} - \bar{\beta})(\beta_{it} - \bar{\beta})'} \quad (6.18)$$

where

$$\bar{\beta} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \beta_{it} \quad (6.19)$$

$$\bar{r} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N r_{it} \quad (6.20)$$

- "The standard error for the factor premium is the square root of (the diagonal elements of) the following variance:"

$$\hat{V}(\hat{f}) = \frac{\hat{\sigma}^2}{\sum_{t=1}^T \sum_{i=1}^N (\beta_{it} - \bar{\beta})(\beta_{it} - \bar{\beta})'} \quad (6.21)$$

where $\hat{\sigma}^2$ is the estimated variance of ϵ_{it} :

$$\hat{\sigma}^2 = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \left(r_{it} - \beta'_{it} \hat{f} \right)^2 \quad (6.22)$$

- "Table 6.4 present the factor premium estimates and their standard errors. We made the estimations for Standard & Poor's (S&P) 500 stocks in the period from January 2000 to December 2003"

TABLE 6.4

The Factor Premium

P/E	P/B	D/E	Size	M12M
−0.0006	−0.0334	0.0341	−0.0129	−0.0814
(0.0005)	(0.0075)	(0.0125)	(0.0021)	(0.0229)

Note: Estimated for the period from January 2000 to December 2003. Standard errors are inside parentheses. Factors are as in Table 6.2.

- "To check the robustness of the estimation, split the data set into a few subsets and see whether the estimates are very different across the subsets"
 - Subsets can be created along the time dimension or along the cross-sectional dimension.
- "The weakness of the OLS method is that in trying to minimise the sum of the squared residuals, it is highly sensitive to outliers."
 - Minimum Absolute Deviation (a.k.a. *median estimation*): "minimises the sum of the absolute value of residuals rather than the squared residuals."
- "The OLS is the best estimation method only if the variance of ϵ_{it} is constant across different stocks"
 - "In econometrics, the problem of non-constant variance is called *heteroscedasticity*. When the variance of ϵ_{it} is not constant across stocks, the best estimation procedure is the generalised least square (GLS) method. The GLS estimator of f :"

$$\hat{f} = \frac{\sum_{t=1}^T \sum_{i=1}^N \left(\frac{\beta_{it}}{\hat{\sigma}_i} - \tilde{\beta} \right) \left(\frac{r_{it}}{\hat{\sigma}_i} - \tilde{r} \right)}{\sum_{t=1}^T \sum_{i=1}^N \left(\frac{r_{it}}{\hat{\sigma}_i} - \tilde{r} \right)^2} \quad (6.23)$$

where

$$\tilde{\beta} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \frac{\beta_{it}}{\hat{\sigma}_i} \quad (6.24)$$

$$\tilde{r} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \frac{r_{it}}{\hat{\sigma}_i} \quad (6.25)$$

$$\hat{\sigma}_i^2 = \frac{1}{T} \sum_{t=1}^T \left(r_{it} - \beta_{it}' \hat{f} \right)^2 \quad (6.26)$$

- Decomposition of Risk
 - "In the fundamental factor model, the risk of the stock return has two components: non-diversifiable risk captured by $\beta_i' \hat{f}$ and diversifiable risk captured by ϵ_i . The total risk of the stock return is simply the sum of these two risks."
 - Variance of the factor premium:

$$\hat{V}(\hat{f}) = \left(\sum_{t=1}^T \sum_{i=1}^N \beta_{it} \beta_{it}' \right)^{-1} \left(\sum_{t=1}^T \sum_{i=1}^N \hat{\sigma}_i^2 \beta_{it} \beta_{it}' \right) \left(\sum_{t=1}^T \sum_{i=1}^N \beta_{it} \beta_{it}' \right)^{-1} \quad (6.28)$$

where $\hat{\sigma}_i^2$ is an estimate.

- Non-diversifiable risk of stock i is a quadratic equation of the factor exposure

$$\beta_{iT}' \hat{V}(\hat{f}) \beta_{iT} \quad (6.29)$$

"Note that we used the last value of the factor exposure to obtain the non-diversifiable risk of stock i "

Chapter 7: Economic Factor Models

•