We already talked about Principal Components (PCs). There are some rigourous theory available for finding PCs using eigenvectors and eigenvalues, that I can cover on demand and that you should know if you use PCs for more than pure data description.

The followings start from notes that I wrote long ago for math-econ majors at Brown. You can read the first sections quickly, but it gives the notion of how factor models most often are used. Text for graduate students that covers this material is "The Economics of Financial Markets," by Campbell, Lo, and MacKinlay, and "Asset Pricing" (I am looking at the revised edition) by John Cochrane. The later book's treatment is a little clearer to me, but maybe harder if you not mathematically trained, and the first book discusses more the finance literature and you should use that book (both actually) if wanted to do asset pricing.

Multi-index models. We start with another implication of the CAPM model. Look at the CAPM equation

$$R_{it} = \alpha_i + \beta_i * R_{Mt} + u_{it} \; . \qquad (1)$$

For simplicity we assume that the safe rate of interest is constant and that the R_{Mt} is normalized such that $E\{R_{Mt}\} = 0$ (the mean market return as well as the safe return are then absorbed in the constant α_i. We only do this to simplify notation). If the CAPM is true then all "systematic influences" on stock prices goes through the market return – for example, an inflation shock may increase the market return by 1% and the impact on asset 1 (say) will then be $\beta_1 * 1\%$. What does this imply for the joint distribution of the error terms? It implies that the error terms are not only independent of their own value on other dates, and independent of the market return – they are also mutually independent. Formally $E\{u_{it}u_{jt}\} = 0$ if $i \neq j$. IBM and Exxon stock may move together but only through the influence of the market. Formally we have

$$\begin{aligned} COV\{R_{it}, R_{jt}\} &= E\{(\beta_i R_{Mt} + u_{it}) * (\beta_j R_{Mt} + u_{jt})\} \\ &= E\{(\beta_i R_{Mt}) * (\beta_j R_{Mt})\} + E\{\beta_i R_{Mt} u_{jt}\} + E\{\beta_j R_{Mt} u_{it}\} + E\{u_{it} u_{jt}\} \\ &= \beta_i \beta_j VAR(R_{Mt}) + 0 + 0 + 0 \; , \end{aligned}$$

when $i \neq j$. (When $i = j$ the COV is of course the variance of the stock i and the last term is not 0 – in this situation the above is the decomposition of the variance of stock return i). So if the CAPM is true the covariance between two different stocks is proportional to the beta's and to the variance of the market return.

Example: Economic growth might cause high returns on stocks. If the CAPM is true the mechanism will be that high growth makes the market go up and the effect on individual stocks is equal to the stock's beta times the impact on the market. No other impact on individual stock returns are allowed. Notice that this is quite a strong prediction. The market will usually go up when the economy picks up, but all stocks with negative beta's should go down.

In order to test this implication of the CAPM, we could estimate the CAPM relation for each stock and save the error term e_{it} from the regression, since this is an estimate of the true innovation

u_{it}. Then we could perform a test of whether those error terms were uncorrelated (e.g. by regressing one on the other) for each pair of stocks. The problem is that there is such a large number of pairs of stocks that this would become quite impractical. (And it is not clear how you would choose the level of significance when you do a large number of tests of a single hypothesis.) So instead researchers, who suspect that stock returns co-move more than explained by the CAPM, have looked for other "factors" or "indices" that might explain movements in stock returns. You might, for example, expect that stock returns move with inflation, in a way that is not fully captured by the market return (note that we of course expect the market return itself to co-move with inflation). So if we let INF_t denote inflation at time t, we might want to estimate the relation

$$R_{it} \; = \; \alpha + \beta_i * R_{Mt} + \lambda_i INF_t + v_{it} \; . \qquad (2)$$

A test for the CAPM would then be to test if $\lambda_i = 0$. Remember the interpretation of the multiple regression model: λ_i does not capture the influence of inflation on stock returns – it only captures the influence of inflation over and above what already is captured by the market. Note that if model (2) is true and you erroneously estimate model (1), the error term will be $u_{it} = \lambda_i INF_t + v_{it}$ and the covariance between u_i and u_j would be $\lambda_i \lambda_j VAR(INF)$ if the v_{it} terms are uncorrelated (you can derive this in the same fashion as I derived the covariance above).

Note that the multi-index model is a model for the behavior of stock returns over time, so if you find that the inflation affects returns over time (in excess of what is explained by the market) this may or may not lead to rejections of the CAPM in the usual 2-step testing of the CAPM. The topic of the next handout will the APT which impose restrictions on mean returns in multi-index models.

Example: Assume that the covariance between inflation and the market return is 0.5. Also assume that the CAPM model is true. What would be the covariance between inflation and the return to stock i?

$$COV(R_i, INF) \; = \; E\{(\beta_i R_M + u_i) * INF\} \; = \; \beta_i * 0.5 + 0 \; .$$

What would be the estimate λ_i if you estimated relation (2)? It would be zero. Inflation affects returns, but only through the market. (In practice, you would not get exactly 0, but an insignificant parameter).

Multi-index models in general are models of the form

$$R_{it} \; = \; \alpha_i + b_{i1} * I_{1t} + ... + b_{ik} * I_{kt} + u_{it} \; , \qquad (3)$$

where the "factors" or "indices" $I_1, ..., I_k$ are variables that affect stock returns or might influence stock returns. The coefficient b_{ik} to the k'th index I_k measures the sensitivity of the returns on stock i to index k. I will follow the jargon (originating from statistical factor analysis) of referring to the b_{ik} coefficients as "loadings". E.g. b_{i1} measures the "load" of index 1 in stock i. We will assume that all indices have mean 0—this turns out to be most practical and is not restrictive, since it is just a question of normalization (the means being absorbed in the intercept). [The APT model, which we cover on the next handout, is all about the mean return in multi-index models.] Like we did for the CAPM, we assume that the error terms in (3) are independent over time, independent of the indices, and independent across stocks. Multi-index models come in a few different varieties: I. Models where the factors are not observed (and it is not known what they are). IIa. Models where the factors are observed and the market returns is one of the indices (e.g. $I_{1t} = R_{Mt} - R_F$), and IIb: Models where the factors are observed but the market return is NOT one of the factors. We will not estimate models in the group I. This is technically difficult and it is somewhat unsatisfactory not to be able to interpret the factors. Nevertheless, this type of models are used a lot and are among the main competitors to the CAPM model in academic research, although I suspect (it is hard to measure) that practitioners may be more prone to use type II versions of the multi-index models.

The equation (2) is an example of a type IIa. model if the coefficient λ_i is not zero. When we consider the model to be an index model we would use the notation

$$R_{it} \;=\; \alpha_i + b_{Mi} * R_{Mt} + b_{Ii} * INF_t + u_{it} \;.$$

Here b_{Mi} is the load of the market index in portfolio i. NOTE that this is not equal to the market β_i of portfolio i, except in the unlikely situation where inflation is uncorrelated with the market returns. (The coefficient in a multiple regression is only equal to the coefficient in a simple regression if the regressors are uncorrelated. See your 162 notes.)

Models of type IIb. may be used in two different situations. One is the situation where you don't think that market has *any* influence on individual stock returns once other factors have been allowed for. The other situation is one where the market does have an impact, but you are interested in the way that factor returns varies with certain factors. This is easier to understand through an example: Assume that you are the manager of a large pension fund which has promised to pay pensions that are adjusted for cost-of-living increases, so that your payouts will go up with inflation. You would like to know how your portfolio returns are likely to vary with inflation and you would therefore estimate a relation of the form

$$R_{it} \;=\; \alpha_i + b_{Ii} * INF_t + b_{2i} * I_{2t} + ... + b_{ki} * I_{kt} + u_{it} \;,$$

where $I_2, .., I_k$ are some other factors that may influence returns. You will not be interested in including the market, since this will pick up some of the influence of inflation. But the market itself is not predictable and all you care about is hedging inflation, so it is not useful for you to know the loading on the market. This is one important use of multi-index models as discussed by EG pp. 391-394.

Note that you may have a situation where for example

$$R_{it} \;=\; \alpha_i + b_{Ii} * INF_t + b_{Gi} * G_{2t} + u_{it} \;,$$

where G is the rate of growth in GDP. If you find such a relation, is that evidence against the CAPM? No, not in itself. If the market return were given by

$$R_{Mt} \ = \ \alpha_i + b_{IM} * INF_t + b_{GM} * G_{2t} + u_{Mt} \ ,$$

then the CAPM relations $R_{it} = \alpha_i + \beta_i R_{mt} + u_{it}$ would still hold if $B_{Ii} = \beta_i b_{IM}$ and $B_{Gi} = \beta_i b_{GM}$.

Today's topic is **the APT model**. The APT model is a model that tells us what the mean return has to be in the multi-factor model. Consider the multi-factor model

$$R_{it} \ = \ \alpha_i + b_{i1} I_t + b_{i2} G_t + u_{it} \ . \tag{1}$$

In general there will be more than 2 factors, but I limit myself to 2 here to keep it more transparant. Let us think of I_t and G_t as the inflation rate and the growth rate in period t, normalized to have zero mean, but again it could be something else. This material is covered in EG on pp. 368-404 (in particular see page 369-372 for an alternative to this note) and BKM pp.300-302 (the discussion in BKM is very brief).

As mentioned earlier we assume that the error terms u_{it} are independent between different stocks, formally $E\{u_{it}u_{jt}\} = 0$ when $i \neq j$. Now assume that you invest 1/Nth of a dollar in each of N stocks with returns $R_{1t}, ..., R_{Nt}$, where N is a large number. The payoff, R_{pt}, to this portfolio will be:

$$\begin{aligned} R_{pt} &= \Sigma_{i=1}^N R_{it}/N \\ &= \Sigma_{i=1}^N (\alpha_i + b_{i1} I_t + b_{i2} G_t + u_{it})/N \\ &= \Sigma_{i=1}^N \alpha_i/N + (\Sigma_{i=1}^N b_{i1}/N) * I_t + (\Sigma_{i=1}^N b_{i2}/N) * G_t + \Sigma_{i=1}^N u_{it}/N \ . \end{aligned}$$

The important thing to notice is that Var $(\Sigma_{i=1}^N u_{it}/N) \ = \ 1/N^2 \Sigma_{i=1}^N \text{Var}(u_{it})$. If furthermore Var $u_{it} \ \leq \ \sigma^2$ (this is just a mathematical way of expressing that the highest residual variance is σ^2) then Var $(\ \Sigma_{i=1}^N u_{it}/N) \ \leq \ 1/N^2 \Sigma_{i=1}^N \sigma^2 \ = \ \sigma^2/N$. The point is that we can diversify away the variance (the risk) of the portfolio coming from the error terms (often called the "idiosyncratic errors"). When N becomes very large, then $\sigma^2/N \approx 0$.
Now consider an asset A, say, for which the loading of the G_t factor is zero and the loading on the I_t factor is one. Such an asset always exists if you can go short in individual stocks. (Why?) Define $\lambda_I = ER_{At} - R_F$, where R_F is the safe return, i.e. λ_I is the expected excess return on asset A. We have

$$R_{At} \ = \ R_F + \lambda_I + I_t + u_{At} \ ,$$

so λ_I is the risk premium required for an investor to hold the asset. If you had another asset Q, say, for which $R_{Qt} \ = \ R_F + \lambda_I + I_t + u_{Qt}$ then the intercept $(R_F + \lambda_I)$ had to be the same. This follows because we do not care about the error term, since that represents risk that can be diversified away, so in equilibrium the two assets must have the same mean return. Therefore any asset that has a factor loading of 1 for factor I_t will have the same excess return which we refer to as the **"factor risk premium"**.

If I_t is the inflation rate at time t, what is λ_I? Be aware that it is NOT the average inflation rate or anything like that, but rather it is the average excess return that the typical investor requires in order to hold an asset whose return is perfectly correlated with inflation!

Similarly, consider an asset B, say, such that

$$R_{Bt} \;=\; R_F + \lambda_G + G_t + u_{Bt} \; .$$

Here λ_G is the factor risk premium for the factor G.

We now want to find the mean return to asset i in terms of the factor risk premiums. In other words, we want to find out what α_i should be in terms of the safe rate of interest and the factor risk premiums.

In order to derive an expression for the mean return to asset i, we consider a portfolio, p, where you invest $1/(1 - b_{i1} - b_{i2})$\$ in asset i, $-b_{i1}(1 - b_{i1} - b_{i2})$\$ in asset A and $-b_{i2}(1 - b_{i1} - b_{i2})$\$ in asset B. Let us consider the return to such a portfolio:

$$R_{pt} \;=\; \frac{1}{1 - b_{i1} - b_{i2}} R_{it} - \frac{b_{i1}}{1 - b_{i1} - b_{i2}} R_{At} - \frac{b_{i2}}{1 - b_{i1} - b_{i2}} R_{Bt} \; .$$

I have chosen the fractions invested in asset i and assets A and B, such that the factor loadings for asset p are all 0. We can check this by plugging in the expressions for R_{it}, R_{At}, and R_{Bt}:

$$\begin{aligned} R_{pt} &= (R_{it} - b_{i1}R_{At} - b_{i2}R_{Bt})/(1 - b_{i1} - b_{i2}) \\ &= [\alpha_i + b_{i1}I_t + b_{i2}G_t + u_{it} - b_{i1}(R_F + \lambda_I + I_t + u_{At}) - b_{i2}(R_F + \lambda_G + G_t + u_{Bt})]/(1 - b_{i1} - b_{i2}) \\ &= [\alpha_i - b_{i1}(R_F + \lambda_I) - b_{i2}(R_F + \lambda_G)]/(1 - b_{i1} - b_{i2}) + (u_{it} - b_{i1}u_{At} - b_{i2}u_{Bt})/(1 - b_{i1} - b_{i2}) \; . \end{aligned}$$

Notice that this portfolio only has *diversifiable* risk. (The portfolio weights were chosen such that the factors cancel out.) It therefore must the be case that the portfolio has mean return equal to the safe rate of interest R_F, since otherwise one could make a positive return on average on a portfolio with essentially no risk. A positive return on an essentially riskless portfolio is called an **arbitrage opportunity**, and the APT (Arbitrage Pricing Theory) is derived from the assumption that NO arbitrage opportunities exists. In an efficient market one can not make arbitrage profits so it must be the case that the mean return to portfolio p: $[\alpha_i - b_{i1}(R_F + \lambda_I) - b_{i2}(R_F + \lambda_G)]/(1 - b_{i1} - b_{i2}) = R_F$, or

$$\alpha_i = R_F + b_{i1}\lambda_I + b_{i2}\lambda_G \; .$$

This is the content of the Arbitrage Pricing Theory (the APT).

You use it in the following way: Assume that you know that there are 2 non-diversifiable factors affecting stock returns, e.g., the inflation rate I_t and the growth rate G_t. (For brevity, I may sometimes say "two factors in the economy" or something like this). Assume that you know that an asset which has a unit loading of I has a factor risk premium of λ_I and an asset with a unit loading of G has a factor risk premium of λ_G. Now assume that you observe that a stock has a loading of b_{i1} on the I factor and a loading of b_{i2} for the G factor. The expected return to this stock then has to be $ER_{it} = R_F + b_{i1}\lambda_I + b_{i2}\lambda_G$.

EXAMPLE: Assume that when the inflation rate goes up by 1% the return to IBM goes down by 0.5%, and when the growth rate picks up by 1% the return to IBM goes up by 2%. (This means that IBM has a loading of -0.5 for inflation and a loading of 2 for growth). Assume that you know that an asset which has a loading of one for inflation and a zero loading for growth has an expected return of 6%. Further assume that an asset with a unit loading for growth and a zero loading for inflation has an expected return of 20%. Further assume that the safe rate of interest is 5%. Now what would be the expected return to IBM stock? First we see that $\lambda_I = 6\% - 5\% = 1\%$ and $\lambda_G = 20\% - 5\% = 15\%$. The return to IBM stock will therefore be

$$R_{IBM} \;=\; 5\% - 0.5 * 1\% + 2 * 15\% \;=\; 34.4\% \;.$$

We will continue with the APT model. The APT model claims that stock returns have the form

$$R_{it} \;=\; R_F + b_{i1}\lambda_1 + ... + b_{ik}\lambda_k + b_{i1}I_{1t} + ... + b_{ik}I_{kt} + u_{it} \;,$$

where $I_1, ..., I_k$ are factors (or indices) normalized to have mean 0, and $\lambda_1, ..., \lambda_k$ are the factor risk premiums corresponding to factor 1 to factor k, respectively.

First: What are likely factors? The market index of course is important, but often you might not want to include this if you want to look for an *explanation* of what affects stock returns. So what else? Here are some suggestions (but you might yourself think of other potential factors):
a) Consumption Growth. By the logic of the CCAPM (so expect positive risk premium).
b) Production growth. Maybe want to insure against production risk (and therefore also expect positive risk premium).
c) Inflation. Inflation may have pervasive effects, reshuffling wealth. (Expected sign not obvious, if pension funds dominant, you might expect a negative risk premium, since pension funds may accept a lower return in order to obtain inflation insurance).
d) Term structure—the difference between the yields to long versus short term bonds. Remember that R_F is the short term interest rate. Long term investors may care about long yields and may prefer stocks with low exposure to the long rate. (This is more speculative, so it not obvious what sign to expect).
e) Risk Aversion. The difference between bond rated as very safe and junk bonds. (Expected sign of risk premium not obvious).

It may also be the case that *unexpected* changes in, say, inflation has a different effect than expected changes in inflation. I will not get into theory about how one can identify unexpected changes, but for those with a good econometrics background this is a feasible topic for a term paper.

Having identified a set of POTENTIAL factors, how do you proceed? Follow the following steps:
A: Regress (let's call it the 1st step regression) your sample of stock returns (or portfolio returns) on the full set of potential factors. Run regressions of the form

$$R_{it} \;=\; \alpha_i + b_{i1} I_{1t} + ... + b_{ik} I_{kt} + u_{it} \;,$$

for each stock over a period of time. This gives you estimates $\hat{\alpha}_i$, and $\hat{b}_{i1}, ..., \hat{b}_{ik}$ for each asset i.
B: Select the factors that are "usually" significant. E.g. the factors that are significant in more than 50% of the regressions. The cut-off depends on how long a data series you have and whether you are more nervous to erroneously leave out a factor than to include a factor that should not be there. More advanced methods exist and ideally you would like to make ONE single test for all the 1st step regression together, but we will not do that in this class.
C: Now repeat the regressions for each stock, but this time only regressing stock returns on the factors that you decided were important. (Call this the 2nd step regression).
D: Now we want to find the factor risk premiums related to each factor. Calculate the average excess return to each stock. Then regress this average excess return ($\hat{\alpha}_i - R_F$) on the b-coefficients ($\hat{b}_{i1}, ..., \hat{b}_{ik}$) that you estimated in part C. (We will call this the 3rd step regression). I.e. run the regression

$$\hat{\alpha}_i - R_F = \hat{b}_{i1}\lambda_1 + ... + \hat{b}_{ik}\lambda_k + \epsilon_i \;.$$

The estimated λs will be your estimate of the factor risk premiums. (If the estimated α- and b-parameters we equal to the true parameters the error terms ϵ_i would be identically 0. The error term captures that the parameters are estimates and that the model isn't literally true.) NOTE that this way of estimating the factor risk premium is similar to the way you estimated the second pass regression for the CAPM. For the APT the regression does not give us a *test* a test since the theory does not suggest in advance what the risk premiums should be. Of course, ways to test the APT exist; but these are more complicated and goes beyond this class.

If you had k stocks and k factors the 3rd step regression would choose the values of the factor risk premiums such that $R^2 = 1$, i.e. a perfect fit. For 2 stocks and 2 factors we can find the factor risk premiums by hand if we know the expected excess returns and the factor loadings.

EXAMPLE: Assume that there are 2 factors I_t and G_t. Assume that you know (or have estimated) the relations $R_1 = 15\% + 2 * I_t + 3 * G_t$ and $R_2 = 25\% + 4 * I_t + 3 * G_t$, the safe rate of interest is 5%. (This is only an example to illustrate the principle, in practise you would estimate the APT for a large number of stocks in order to get a precise estimate of the factor risk premiums.) What is the factor risk premiums? Write down the APT relations $ER_i - R_F = b_{i1}\lambda_I + b_{i2}\lambda_G$, here

$$15\% - 5\% \;=\; 2 * \lambda_I + 3\lambda_G \qquad (2)$$

$$25\% - 5\% \;=\; 4 * \lambda_I + 3\lambda_G \qquad (3)$$

This is just 2 equations in 2 unknowns. For example, subtracting (1) from (2) we get $2\lambda_I = 10\%$, so the factor risk premium for the I-factor is 5%. Then substitute in, e.g., (1) and get 10% =

$2 * 5\% + 3 * \lambda_G$ from which we see that the factor risk premium for factor G is 0. (A factor risk premium of 0 is not ruled out by the APT model, although we usually expect a non-zero — positive or negative — value).

Literature: Chen, Roll, and Ross (1986), is one of the first papers to estimate APT-models of the form discussed here. Another paper is: L.K. Chan, J. Karceski, and J. Lakonihok: "The Risk and Return from Factors." *Journal of Financial and Quantitative Analysis,*" vol. 33, June 1998.

Unobserved Factors:
In the case of unobserved factors, we have the model

$$R_{It} \quad = \quad \alpha_i + b_{i1} f_t^1 + b_{i2} f_t^2 + u_{it} \, ,$$

where we assume f^1 and f^2 are mean zero, variance one, orthogonal vectors and the u terms in white noise terms, with variance σ_u^2, independent of each other. (The first two of these assumption are mere normalizations.) We refer to the b terms as factor loadings. The mean is

$$[\alpha_i = r^f + b_{i1}\lambda^1 + b_{i2}\lambda^2 \ ,$$

where λ^k, $k = 1, 2$ are risk premiums associated with the factors.

For now, ignore the mean restrictions. If we define $R_t = \{R_{1t}, ..., R_{Nt}\}$ for a sample of N stocks or portfolios,

$$B \ = \begin{pmatrix} b_{11} & b_{12} \\ . & . \\ . & . \\ b_{N1} & b_{N2} \end{pmatrix} .$$

We have

$$var(R_t) = BB' + \sigma_u I_N \ .$$

Ignoring the mean for now, we can estimate the parameters of B and σ_u by (non-linear) ML estimation (or least-squares). The variance matrix has $N * (N + 1)/2$ separate values and B has $2\,N$ parameters, so for $N = 4$, we have 4*5/2=10 different elements in the variance matrix and 2*4+1 parameters so the model is identified. We would typically use the model for larger values of N.

Notice that if we let Σ be an unrestricted diagonal matrix we would have $var(R_t) = B\Sigma B' + \sigma_u I_N$, but for this to be identified you would need to, e.g., set one of the coefficients in each of the columns of B to unity. (Even if the number of unknown parameters is larger than the number of variances and co-variances, you can see that by multiplying out the matrices.) You could also allow for the unobserved factors to be correlated; i.e., no constrain Σ to be diagonal. For the APT application this makes it hard to interpret, though. For a statistician, like me, this is all the content of a basic factor model: as set of restrictions on the variance covariance matrix.

For the APT application, we have N mean terms, but the model implies that (after subtracting the save return), we have (for $a = \{a_1, ..., aN\}$) and $\Lambda = \{\lambda_1, \lambda_2\ \}$'):

$$a = B\Lambda\ .$$

which gives us more degrees of freedom for $N > 2$ as we add two parameters to fit N mean terms. You could now write down the likelihood function in terms a sum of T terms involving the $N-$ dimensional multivariate normal distribution.