Asset pricing I: Pricing Models

Markus K. Brunnermeier

a.y. 2014/2015

Contents

Chapter 1

Introduction

Asset pricing is the study of the value of claims to uncertain future payments. Two components are key to value an asset: the ***timing*** and the ***risk*** of its payments. While time effects are relatively easy to explain, corrections for risk are much more important determinants of many assets' values. For example, over the last 50 years U.S. stocks have given a real return of about 9% on average. Only about 1% of this can be attributed to interest rates; the remaining 8% is a premium earned for holding risk.

This raises the question: what determines the price of financial claims? That is, why do prices move over time, and why do different asset have different prices?[1] There are several approaches that have been used to answer these questions:

- Statistical approaches look at statistical relationships between asset prices
- "Weak" economic approaches look at some basic relations that must hold between asset prices, such as the absence of risk-free profitable strategies[2]
- Economic models derive prices from the fundamental characteristics of an economy[3]

Financial claims are promises of payments at various points in the future: for example, a stock is a claim on future dividends; a bond is a claim over coupons and principal; an option is a claim over the future value of another asset. More formally, suppose that we are at date t, then we can we define payments $x_{t+\tau}$ for $\tau \geq 1$ and expect the price of these payments to be something like $p_t \approx \mathbb{E}_t \sum_{\tau \geq 1} [x_{t+\tau}]$, with some adjustment for time and risk. Another way to think about financial claims is in terms of returns, defined as how much money we make if we hold an asset for a given amount of time: $R_{t+1} = \frac{p_{t+1}+x_{t+1}}{p_{t+1}} - 1$. We call *excess return* the difference between the returns of two assets i and j: $R^e_{t+1} = R_{i,t+1} - R_{j,t+1}$. We can interpret these three representations as follows:

[1] We will see that these questions refer to "time-series" and "cross-sectional" problems, respectively.
[2] This concept is referred to as "no arbitrage".
[3] Such as preferences, technology, etc.

we can invest p_t today and get $\{x_{t+\tau}\}$ in the future, or invest 1 unit today and get R_{t+1} in the future, or yet invest 0 units today and get R^e_{t+1} in the future. What are the properties of returns? Can we predict when assets will have high or low returns? Can we predict which assets have higher or lower returns? Historically, there have been two schools of thought on this subject:

- **The old view** (1970s): Expected returns do not move much over time: stocks returns are unpredictable because prices move with news about future cash-flows. The "classic" model for asset pricing, called CAPM, works pretty well: returns with high covariance with the market return have are higher on average as predicted by the mdoel. The beta parameter in the CAPM model derives from the covariance between asset cash-flows and market cash-flows.

- **The modern view**: Expected returns move a lot over time: stock returns are predictable. Prices move with news about changes in the discount rate used by people to discount assets. We can understand the cross-sectional relation between asset prices with multi-factor models: characteristics other than the beta are associated with returns, and non-market betas matter a lot. Finally, betas derive from the covariance between discount rates and market discount rates.

Asset pricing theory can be used to describe both the way the world *works* and the way the world *should work*. Once we observe the prices, we can use asset pricing theory to understand why prices are what they are, and modify our theory if the predictions are not consistent with the observations; or we can decide that the observed prices are wrong, or *mispriced*, and take advantage of the trade opportunity. Much of asset pricing theory stems from one simple concept:

Price equals expected discounted payoff

The rest is elaboration, special cases and a few tricks. There are two approaches to this elaboration, called ***absolute asset pricing*** and ***relative asset pricing***. In absolute asset pricing we price each asset by reference to its exposure to fundamental macroeconomic risk.[4] This approach is most popular in many academic settings in which we use asset pricing to give an explanation for why prices are what they are in order to predict how prices might change if policy or economic structure changed. In relative pricing we infer an asset's value *given* the prices of some other asset. Black-Scholes option pricing is the classic example of this approach.

The central and unfinished task of asset pricing theory is to understand and measure the sources of aggregate risk that drive asset prices. Of course, this is also the central question of macroeconomics, and in fact a lot of empirical work has documented stylized facts and links between macroeconomics and finance. For example, expected returns vary across time and across assets in ways that are linked to macroeconomic variables: we have learned that the risk premium on stocks[5] is much larger than

[4] The classic examples of this approach are the Walrasian general equilibrium model discussed in Léon Walras' "Elements Of Pure Economics" (1877), and Gerard Debreu's "Theory of Value" (1959).

[5] The difference between the stock return and the risk-free interest rate.

the interest rate, and varies a lot more than interest rates. This means that attempts to line up investments with interest rates are vain, as much of the variation in cost of capital comes from the varying risk premium. Similarly, we have learned that some measure of risk aversion must be quite high, or people would all borrow like crazy to buy stocks. Moreover, while standard macroeconomics theory predicts that agents do not care about business cycles,[6] asset prices reveal that they do: agents forgo substantial return premia to avoid assets whose value falls in recessions. And yet theory still lags behind: we do not yet have a well-described model that fully explains these correlations.

The rest of these notes is organized as follows. We start with frictionless markets and minimal assumptions, adding more structure as the course progresses to obtain more and deeper implications. In the second part of the course we will extend the discussion to multi-period settings, and finally conclude by studying financial markets with frictions.

1.1 Market Efficiency

When is an asset fairly valued? We cannot answer this question without referring to a specific model or assumption about the asset. Consider for instance the following assumption:

Prices incorporate and reflect all publicly available information

A direct consequence of this hypothesis is that stock returns should be unpredictable: for example, when new information is released, stock prices should jump to the new fair level and then keep trading around it. Are these features observed in reality? We will address this question with the next subsections.

A related famous assumption is the Random Walk hypothesis, which states that stock market prices evolve according to a random walk, and therefore cannot be predicted. This is compatible with the Efficient Markets hypothesis outlined before, and it implies that stock price movements can only be attributed to 1) news on future corporate cash flows, 2) changes in "risk premia" - the amount of extra return that investors demand to hold risk (we will come back to this in the next lectures) or 3) shifts in behavioral bias.

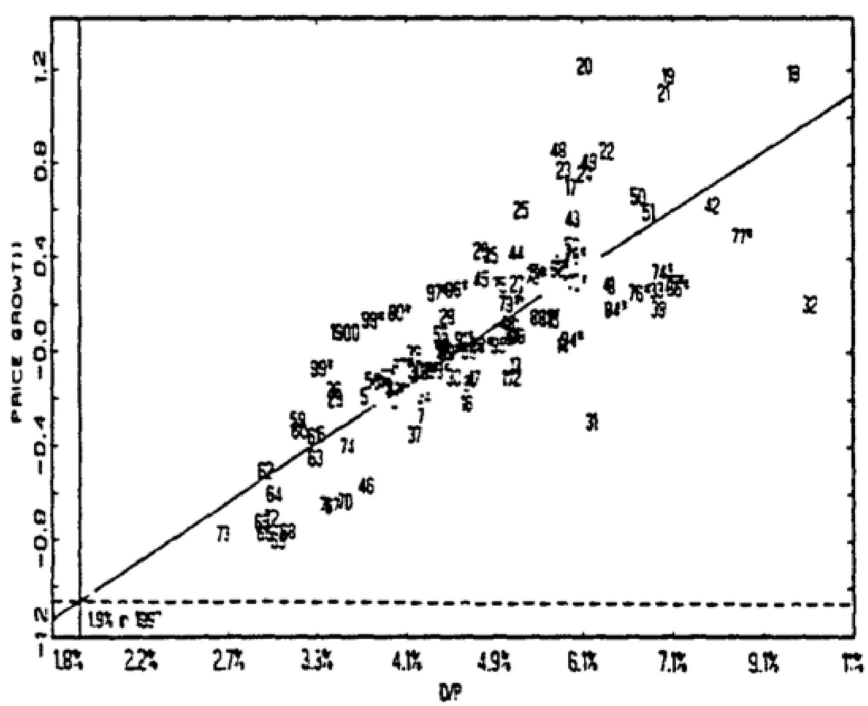
Testing these hypotheses can be done in one of two main ways. One approach is to look at *cross-sectional* data, that is, data collected at the same point in time, or regardless of differences in time. Studies that use this approach look at whether some factors can explain the stock price changes, potentially in contradiction to the efficient markets or random walk hypothesis. Another approach is to look at *time-series* data, that is, data collected as a sequence of data points. Time-series studies look into the existence of trends, seasonalities, or event-specific behaviors that would invalidate the efficient markets or random walk hypothesis.

[6] Lucas (1987).

1.1.1 Dividend/Price Ratio and Stock Prices

A consequence of the efficient markets hypothesis is that the Dividend/Price (D/P) ratio should be a good indicator of future dividend movements, because when the dividend is expected to go up the price also goes up (since it gives right to more cash flows), thus decreasing the D/P ratio until the moment when the dividend is declared. After that moment, the ratio goes back to its previous level.[7] However, it turns out that the D/P ratio is much better at predicting future *stock price* movements than dividend movements: a simple regression of the S&P 500 index price growth on the D/P ratio shows that the latter is a good predictor of future price growth.

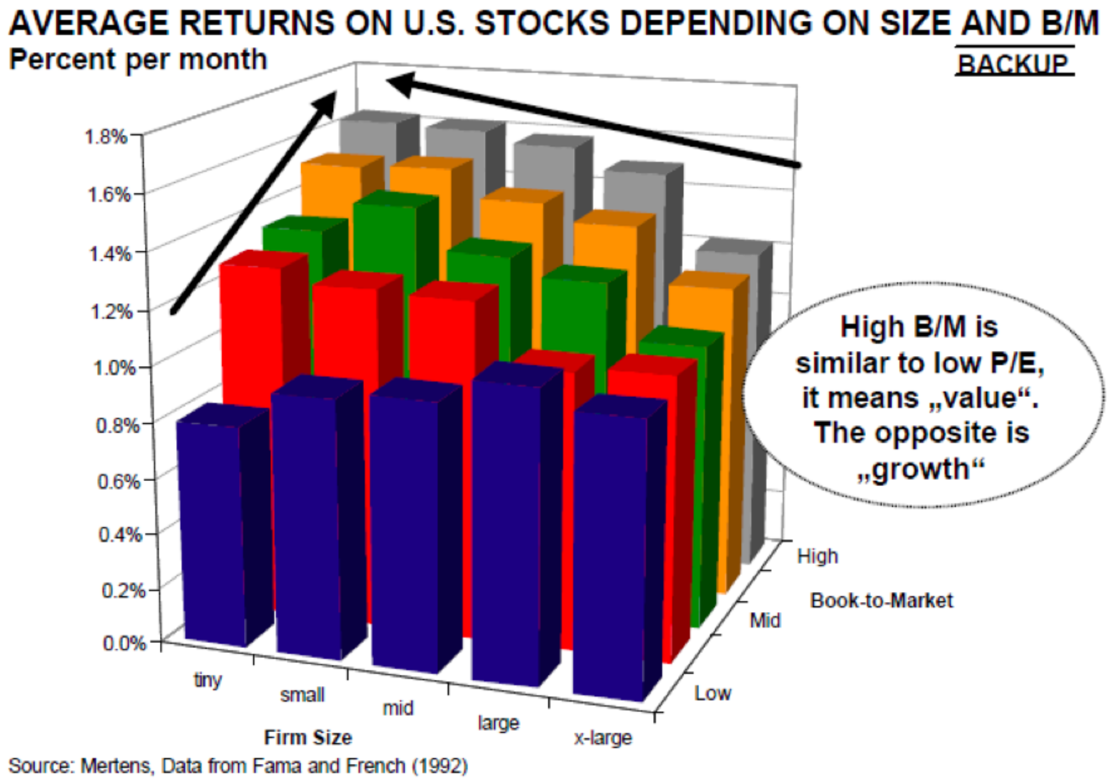
PANEL B. PRICE GROWTH UNTIL NEXT TIME D/P CROSSES ITS MEAN VERSUS D/P



1.1.2 Size and Book to Market as drivers of Stock Returns

Average monthly stock returns appear to be higher the smaller the company size as measured by its assets, and the higher its Book-to-Market, that is, the more a company is perceived as a "value" company by the market (as opposed to "growth" stocks).

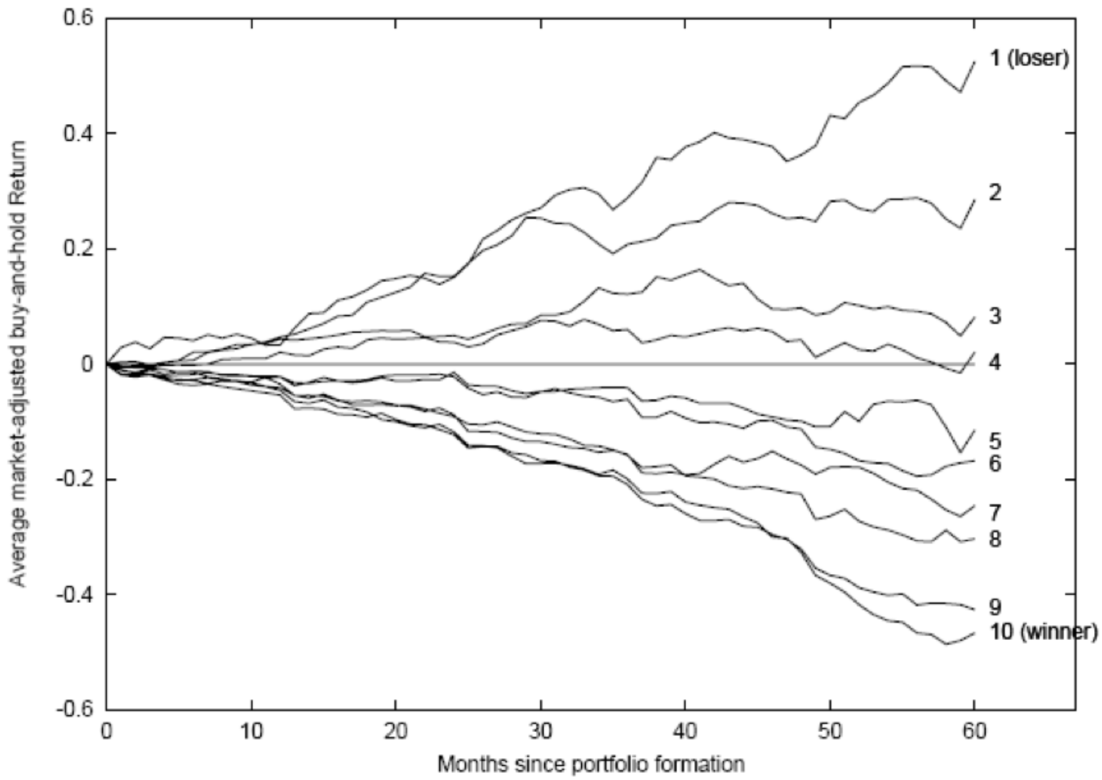
[7] Because the dividend will either be increased, as the market predicts, thus increasing the D/P ratio; or it will not be increased, and consequently the stock price will decrease to the level it was at before the expectation was formed and thus increasing the D/P ratio.



1.1.3 Winners and Losers

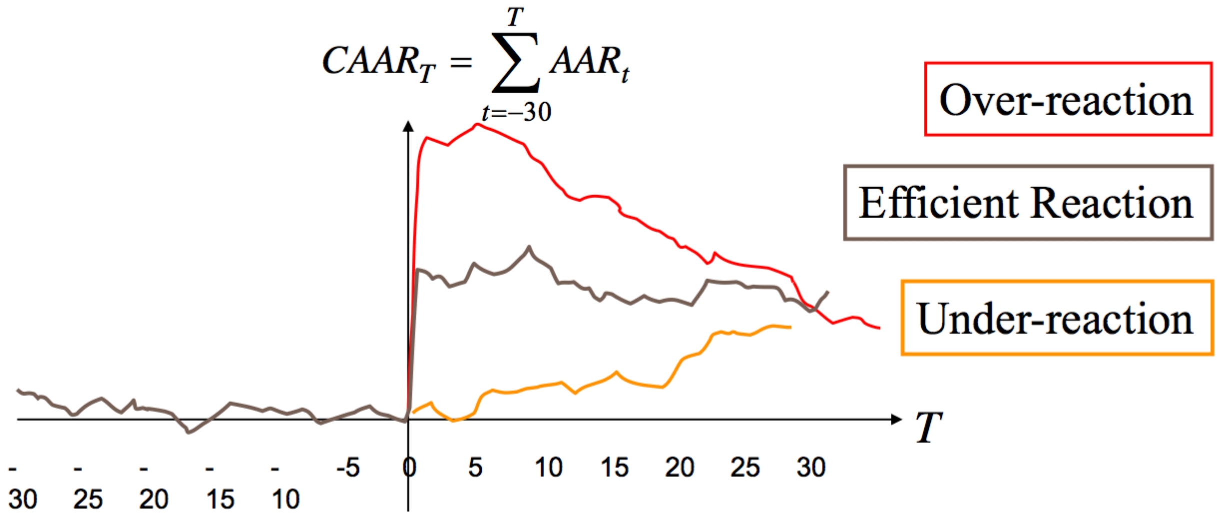
The average market-adjusted return for the strategy "buy and hold the n-th decile performing stocks" shows a remarkable pattern: stocks that perform well over some time tend to underperform in the following period (and vice versa), and this effect is stronger the better the past performance (and vice versa). Over the medium-long term, mean-reversion appears to be a common feature in stocks.[8]

[8]On the contrary, for shorter time-periods, momentum seems to drive prices more than mean-reversion.



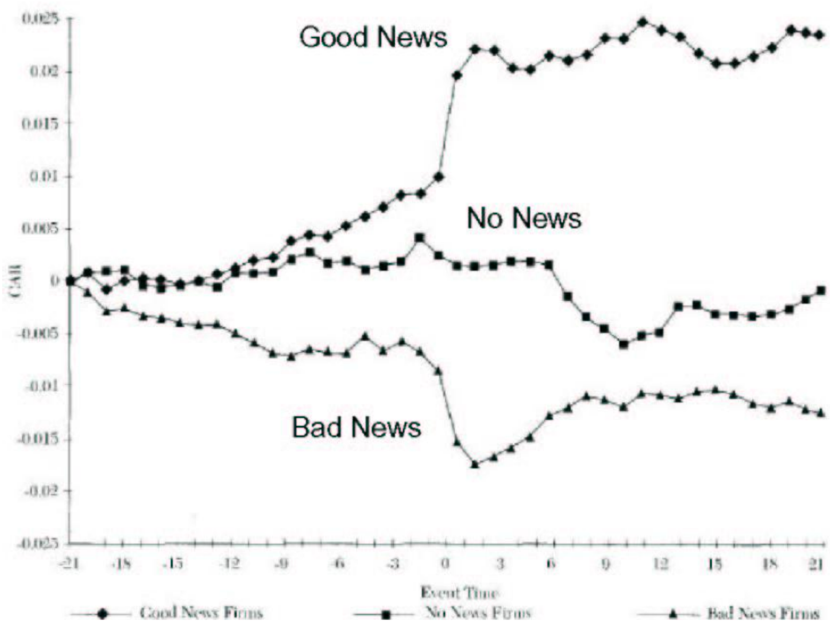
1.1.4 Event Studies

What is the impact on daily stock returns of a company-specific event such as a merger, bankruptcy or special dividend announcement? To find out, suppose we define as date $t = 0$ the date of the anouncement of the financial event and calculate for each firm i the return $R_{i,t}$ for $t = -30, ..., 30$. Then we do the same for the market- or sector-reference group return, $R_{m,t}$ to be used as comparison and define abnormal returns as $AR_{i,t} = R_{i,t} - R_{m,t}$. Finally, we consider the firms' cumulative average returns $CAAR_T = \sum_{t=-30}^{T} AAR_t$, where $AAR_t = \frac{1}{N}\sum_{i=1}^{N} AR_{i,t}$: with no news at all, we expect AAR_t to be close to zero. Below is a plot of the stylized possible reactions: an efficient reaction is one in which the stock price maintains its new level after the event, an under-reaction is one in which the stock price drifts up after the announcement, whereas in an over-reaction the stock price drifts lower after the initial spike.

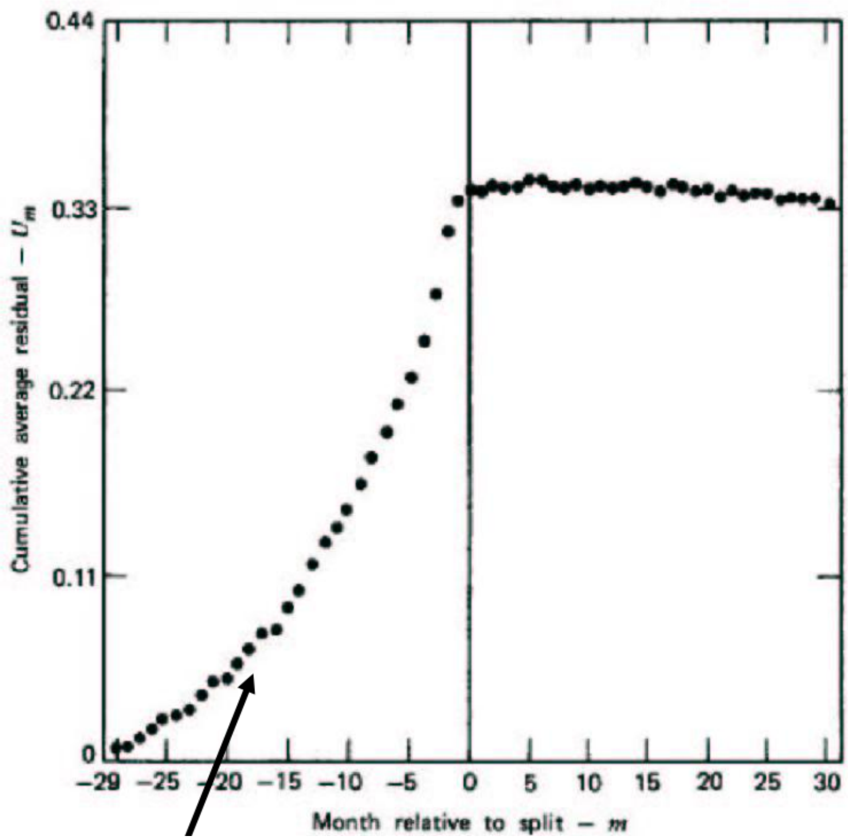


For instance, let us look at earning announcements. Ball and Brown (1968) show that there is not only a pre-announcement drift, arguably due to insider information, but also a post-announcement drift possibly due to the under- or over-reaction of the market to the earnings news released. This constitutes a rejection of the Efficient Market Hypothesis even in its semi-strong form, which states that stock prices reflect all publicly available information.

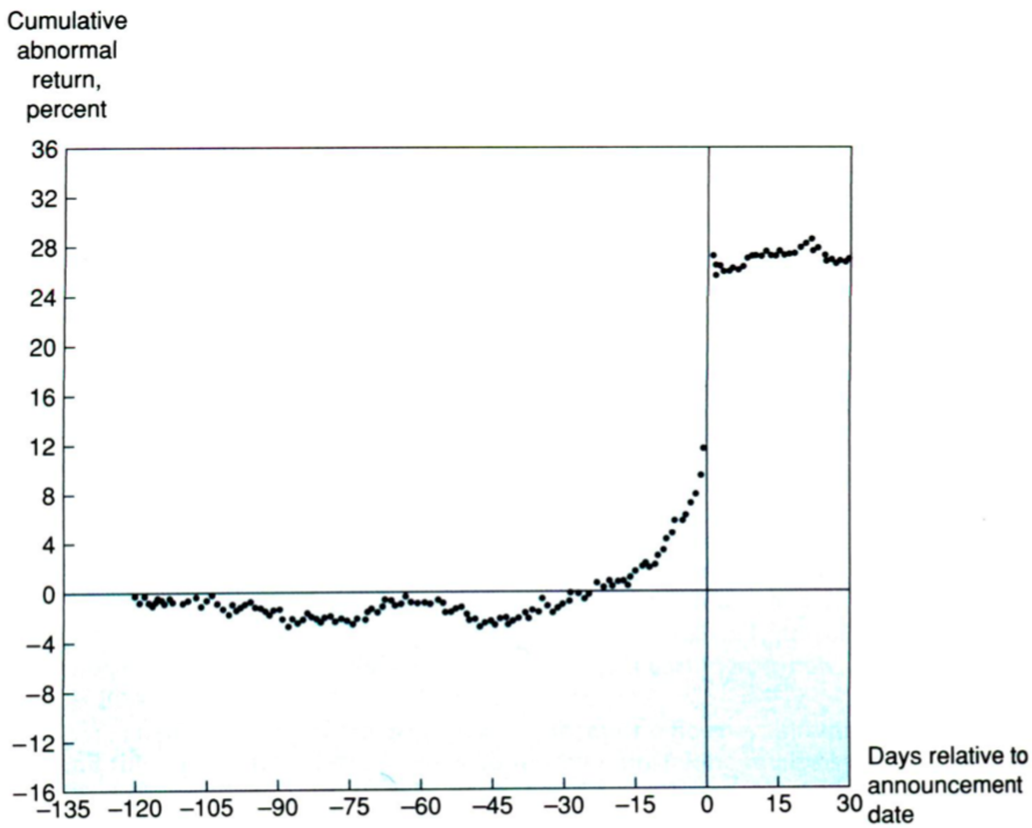
In a related paper, MacKinlay (1997) considers separately firms whose reported earnings were better, equal or worse than expected earnings (as measured by the consensus estimate). Below is a plot of their cumulative average abnormal returns: on average, over- and under-reactions are quite limited.



In a stock split, for each share held shareholders receive two shares. This usually signals that the price has gone up far too much for small investors to be able to invest in the company, so firms often decide to split the shares to make the investment more accessible to small investors. Normally a stock split results in a price increase because many small investors can now buy the stock thus driving the price up. Below is a plot of the cumulative average abnormal returns for the event of the stock split: there are two possible explanations for this pattern. One is that inside information is leaked prior to the announcement of the split. Another possibility however is that there is a selection bias: only companies that perform well announce stock splits (those whose price increases to a level unaccessible to small investors) and hence the pattern is due to selection bias.



Finally, below is a plot of cumulative average abnormal returns in the event of a take-over announcement. Clearly, towards the announcement date there are signs of some leaked inside information, and yet this is not completely factored in by the market since the price jumps on the announcement date.



1.1.5 Government Bonds

The Expectation Hypothesis holds that that the long-term yield is determined by the market's expectation for the future short-term yield. According to EH, the term structure of interest rates is fully determined by expectations about future short-term interest rates (and possibly maturity-dependent risk premia). It can be stated in the following way:

Single-period holding returns on bonds of all maturities are equal in expectation

For example, holding a 5-year zero-coupon bond and selling it after 1 year gives you the same return as holding a 1-year zero-coupon bond for 1 year. The one-year continuously compounded return on a 1-year bond is:

$$r_{1,1} = \mathbb{E}_t \left[\ln \frac{1}{Z(t,t+1)} \right] = \ln \frac{1}{e^{-y(t,t+1)\cdot 1}} = y(t,t+1)$$

While the one-year continuously compounded return on a m-year bond is:

$$r_{1,m} = \mathbb{E}_t \left[\ln \frac{Z(t+1,t+m)}{Z(t,t+m)} \right] = \mathbb{E}_t \left[\ln \frac{e^{-y(t+1,t+m)\cdot(m-1)}}{e^{-y(t,t+m)\cdot m}} \right] = m\cdot y(t,t+m) - (m-1)\cdot \mathbb{E}_t \left[y(t+1,t+m) \right]$$

Where $Z(s,t)$ is the price of a zero-coupon bond bought at time s and maturing at time t. So setting $r_{1,1} = r_{1,m}$ and rearranging we get

$$y(t,t+m) - y(t,t+1) = (m-1)\cdot \mathbb{E}_t\left[y(t+1,t+m) - y(t,t+m)\right]$$

Plus a constant, if we allow for time-dependent risk premia. This implies that according to EH the *yield spread* $y(t,t+m) - y(t,t+1)$ forecasts short-term changes in the yield on the long-term bond: a high yield spread predicts a rise in the yield of the long bond, just enough to generate a capital loss to offset that bond's higher yield. This is an empirically testable fact: we can regress the yield spread on the short-run changes in long yields and verify whether the beta coefficient is $(m-1)$.

Table 2
Regression Coefficients

	Long bond maturity (months)						
Dependent variable	*2*	*3*	*6*	*12*	*24*	*48*	*120*
Short-run changes in long yields	0.019 (0.194)	−0.135 (0.285)	−0.842 (0.444)	−1.443 (0.598)	−1.432 (0.996)	−2.222 (1.451)	−4.102 (2.083)

The coefficient is statistically significant and negative: clearly, the Expectation Hypothesis fails to capture some empirical feature. It neglects the risks inherent in investing in bonds, namely 1) interest rate risk, or the risk to re-sell the bond at a lower price when interest rates rise, and 2) reinvestment risk, or the risk that the proceeds from the bond are reinvested at a lower rate.

1.1.6 Corporate Bonds

The Merton Model postulates that the assets of a firm follow a geometric brownian motion process

$$\frac{dA_t}{A_t} = \mu \cdot dt + \sigma \cdot dW_t^P$$

Assuming the firm has debt outstanding with face value D, if at time T the firm is insolvent $(A_T < D)$ the debtholders take possession of the firm. The payoff to a debtholder is $\min(A_T, D) = D - \max(A_T - D, 0) = D - (A_T - D)^+$, and therefore the value of the corporate bond can be priced using the Black-Scholes formula as

$$V_t(A_t, D, T-t) = A_t - C_t^{BS}(A_t, D, T-t) = D\cdot e^{-r(T-t)} - P_t^{BS}(A_t, D, T-t)$$

By put-call parity. However, this model implies a credit spread to Treasuries that is consistently lower than the observed credit spreads:

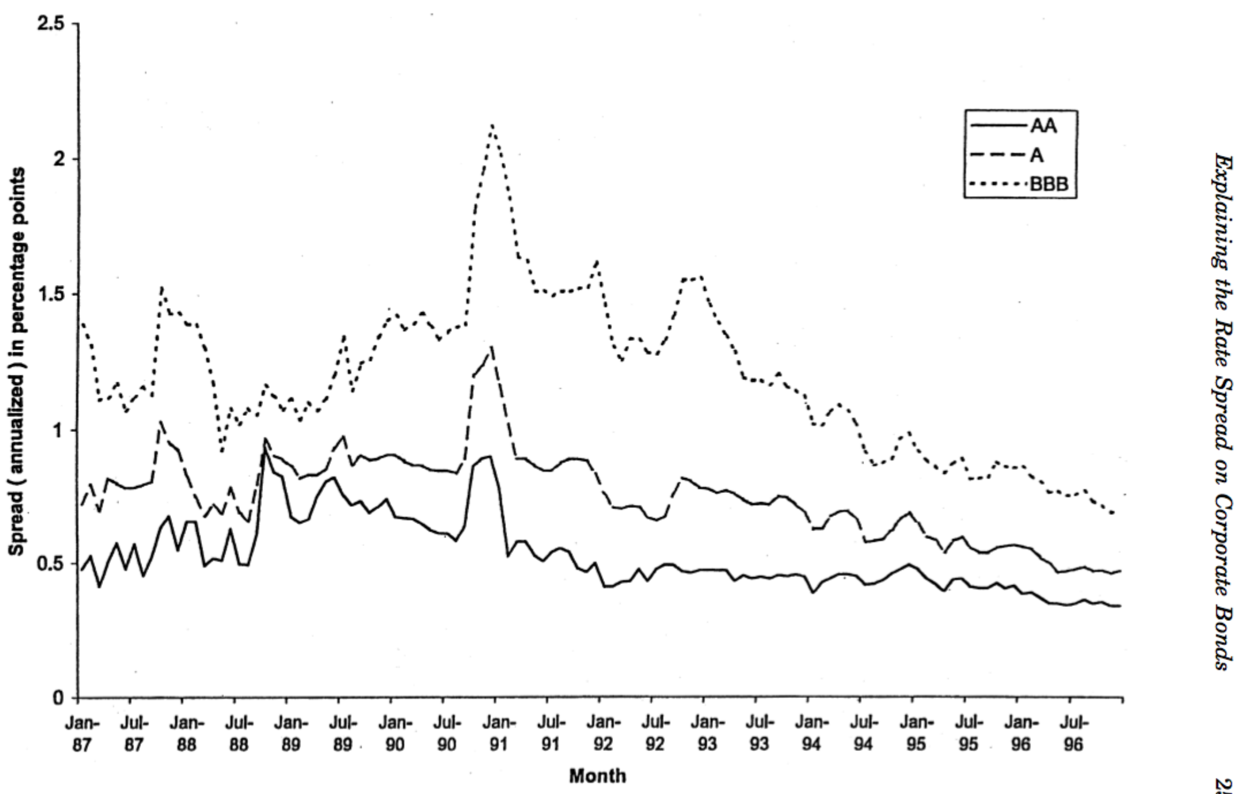


Figure 1. Empirical spreads on industrial bonds of six years maturity.

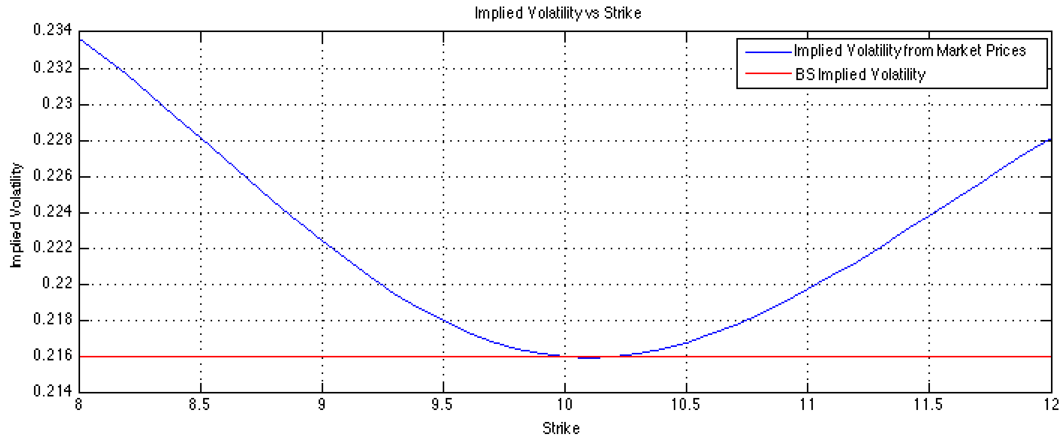
There are several assumptions that do not hold in reality, ranging from log-normally distributed firm assets to the non-standard features of corporate bonds (they may have covenants, variable coupons, etc.)

1.1.7 Derivatives Pricing

In this course we will encounter the Black-Scholes model for options pricing, according to which a European call option on a stock worth S with strike K and maturity T today is worth (assuming zero interest rates)

$$S_0 N\left(\frac{\ln\frac{S_0}{K}+\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right)-Ke^{-rT}N\left(\frac{\ln\frac{S_0}{K}-\frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right)$$

Where σ is the annualized standard deviation, or *volatility*, of stock returns. Assuming that this model is correct, then for a given set of strikes $\{K_i\}_{i=1}^n$ and corresponding option prices we can back out the implied volatility σ_{IV} using (for instance) the bisection method. If the Black-Scholes model was correct, we would see a flat line: by construction, $\sigma_{IV}=\sigma$. However, using option prices as observed in the market gives a different result: strike levels far from S have a *higher implied volatility than implied by Black-Scholes.*

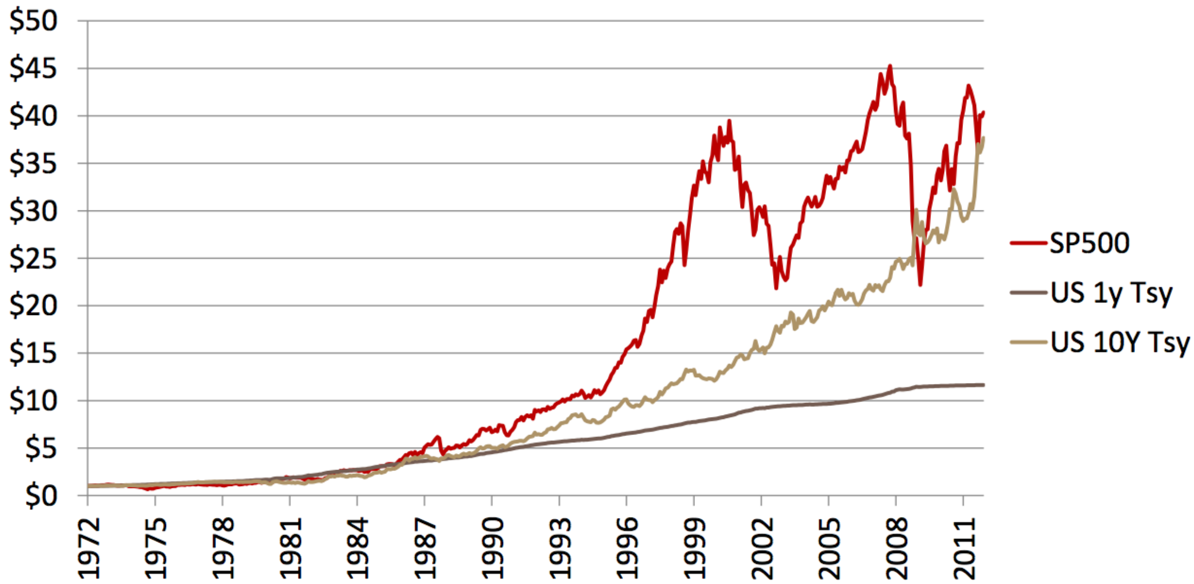


Interestingly, this phenomenon became more evident after the 1987 Black Monday Crash. It has been explained by the fact that rare events (crashes or rallies) occur more frequently in reality than under the Black-Scholes model, hence the market incorporates that real feature of stocks by "bidding the tails".

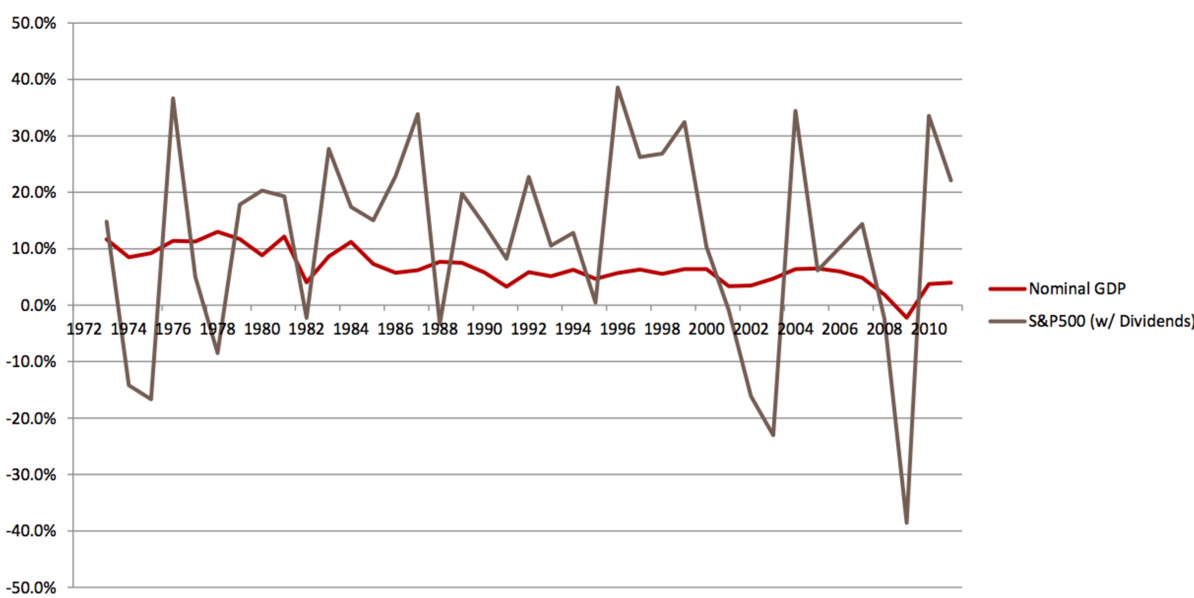
We conclude this section with two empirical facts about stocks, macroeconomic factors, and how risk measurements connect the two.

1.2 Stocks and Macroeconomic Factors

Let us begin by comparing the returns of a stock index and that of riskless bonds. The picture below shows the return from investing \$1 on January 1st 1972 in the S&P 500 index (red), a one-year Treasury Bill (grey) and a 10-year Treasuty Note (brown). The return from the S&P 500 index has been higher on average than both Treasury returns, yet much more volatile. The 1-year Treasury Bill return is much lower and less volatile than the 10-year Treasury, a fact that is accounted for by the relative size of their duration.

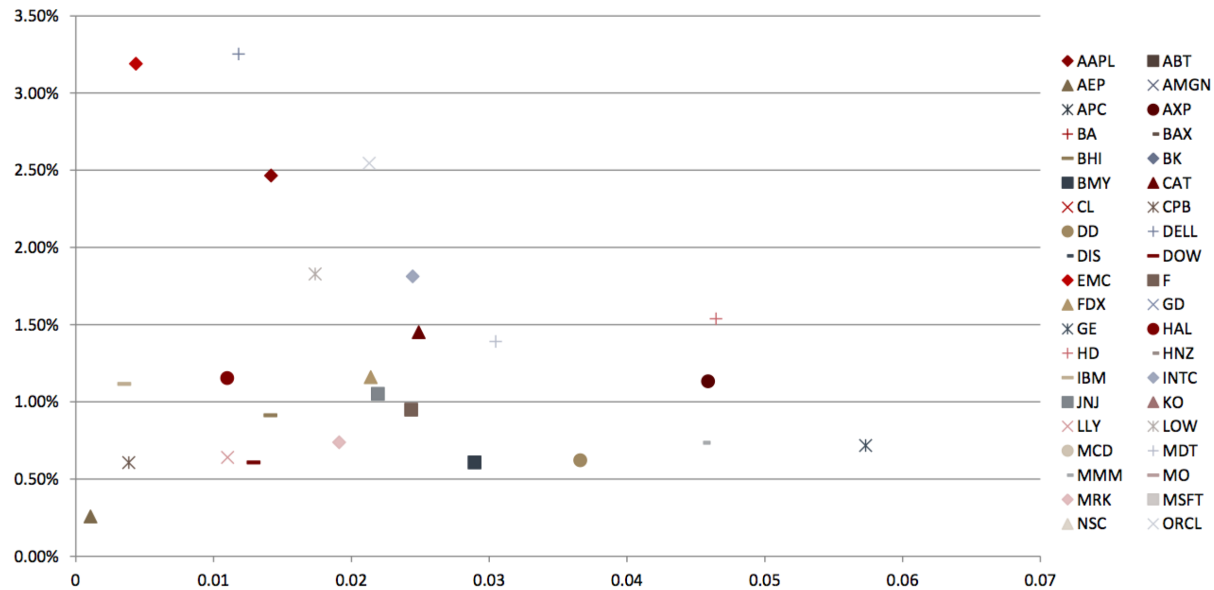


Similarly, looking at the yearly growth of the S&P 500 versus that of nominal GDP (which is mostly comprised of consumption), we see that stocks are far more volatile than consumption.



1.2.1 Measuring Risk with Covariance

We will see that wisely managing a portfolio allows us to eliminate the stock-specific risk exposure. What is left is the systemic risk, or market risk, which is captured by the covariance between the stock and the market, and hence motivates us to use covariance as a risk measure for stocks. Below is a plot of the covariance between some US large companies and the S&P 500.



1.3 The 2013 Nobel Prize in Economics

We conclude this chapter with an excerpt from the scientific background paper for the 2013 Nobel Prize in Economics written by the Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences:

"While prices of financial assets often seem to reflect fundamental values, history provides striking examples to the contrary, in events commonly labeled bubbles and crashes. Mispricing of assets may contribute to financial crises and, as the recent recession illustrates, such crises can damage the overall economy. Given the fundamental role of asset prices in many decisions, what can be said about their determinants?

The 2013 Nobel prize was awarded empirical work aimed at understanding how asset prices are determined. Eugene Fama, Lars Peter Hansen and Robert Shiller have developed methods toward this end and used these methods in their applied work. Although we do not yet have complete and generally accepted explanations for how financial markets function, the research of the Laureates has greatly improved our understanding of asset prices and revealed a number of important empirical regularities as well as plausible factors behind these regularities.

The question of whether asset prices are predictable is as central as it is old. If it is possible to predict with a high degree of certainty that one asset will increase more in value than another one, there is money to be made. More important, such a situation would reflect a rather basic malfunctioning of the market mechanism. In practice, however, investments in assets involve risk, and predictability becomes a statistical concept. A particular asset-trading strategy may give a high return on average, but is it possible to infer excess returns from a limited set of historical data?

Furthermore, a high average return might come at the cost of high risk, so predictability need not be a sign of market malfunction at all, but instead just a fair compensation for risk-taking. Hence, studies of asset prices necessarily involve studying risk and its determinants.

Predictability can be approached in several ways. It may be investigated over different time horizons; arguably, compensation for risk may play less of a role over a short horizon, and thus looking at predictions days or weeks ahead simplifies the task. Another way to assess predictability is to examine whether prices have incorporated all publicly available information. In particular, researchers have studied instances when new information about assets becomes became known in the marketplace, i.e., so-called event studies. If new information is made public but asset prices react only slowly and sluggishly to the news, there is clearly predictability: even if the news itself was impossible to predict, any subsequent movements would be. In a seminal event study from 1969, and in many other studies, Fama and his colleagues studied short-term predictability from different angles. They found that the amount of short-run predictability in stock markets is very limited. This empirical result has had a profound impact on the academic literature as well as on market practices.

If prices are next to impossible to predict in the short run, would they not be even harder to predict over longer time horizons? Many believed so, but the empirical research would prove this conjecture incorrect. Shiller's 1981 paper on stock-price volatility and his later studies on longer-term predictability provided the key insights: stock prices are excessively volatile in the short run, and at a horizon of a few years the overall market is quite predictable. On average, the market tends to move downward following periods when prices (normalized, say, by firm earnings) are high and upward when prices are low.

In the longer run, compensation for risk should play a more important role for returns, and predictability might reflect attitudes toward risk and variation in market risk over time. Consequently, interpretations of findings of predictability need to be based on theories of the relationship between risk and asset prices. Here, Hansen made fundamental contributions first by developing an econometric method – the Generalized Method of Moments (GMM), presented in a paper in 1982 – designed to make it possible to deal with the particular features of asset-price data, and then by applying it in a sequence of studies. His findings broadly supported Shiller's preliminary conclusions: asset prices fluctuate too much to be reconciled with standard theory, as represented by the so-called Consumption Capital Asset Pricing Model (CCAPM). This result has generated a large wave of new theory in asset pricing. One strand extends the CCAPM in richer models that maintain the rational-investor assumption. Another strand, commonly referred to as behavioral finance – a new field inspired by Shiller's early writings – puts behavioral biases, market frictions, and mispricing at center stage.

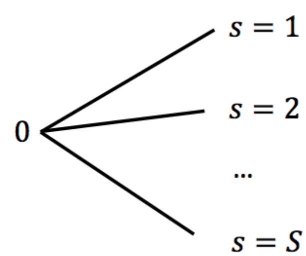
A related issue is how to understand differences in returns across assets. Here, the classical Capital Asset Pricing Model (CAPM) – for which the 1990 prize was given to William Sharpe – for a long time provided a basic framework. It asserts that assets that correlate more strongly with the market as a whole carry more risk and thus require a higher return in compensation. In a large number of studies, researchers have attempted to test this proposition. Here, Fama provided seminal methodological insights and carried out a number of tests. It has been found that an extended model with three factors – adding a stock's market value and its ratio of book value to market value – greatly improves the explanatory power relative to the single-factor CAPM model. Other factors have been found to play a role as well in explaining return differences across assets. As in the case of studying the market as a whole, the cross-sectional literature has examined both rational-investor–based theory extensions and behavioral ones to interpret the new findings.**"**

One last remark: Shiller and Fama's works ultimately discuss the same findings, but interpreting them differently. What Shiller calls irrational bubbles and behavioral biases Fama would call efficient markets and varying risk premia is nothing but the empirical observation that the *stochastic discount factor* varies a lot over time. Whether we label this as "irrational behavior" or "time-varying risk premia" is an hypothesis that is very hard to test empirically.

Chapter 2

One Period Model

We begin our analysis of asset pricing by considering a simple setting in which there are two dates: today ($t = 0$) and some future date ($t = 1$). We know everything about today's state of the world $s = 0$, but we don't know which one of $s = 1, \dots, S$ states will materialize in the future.



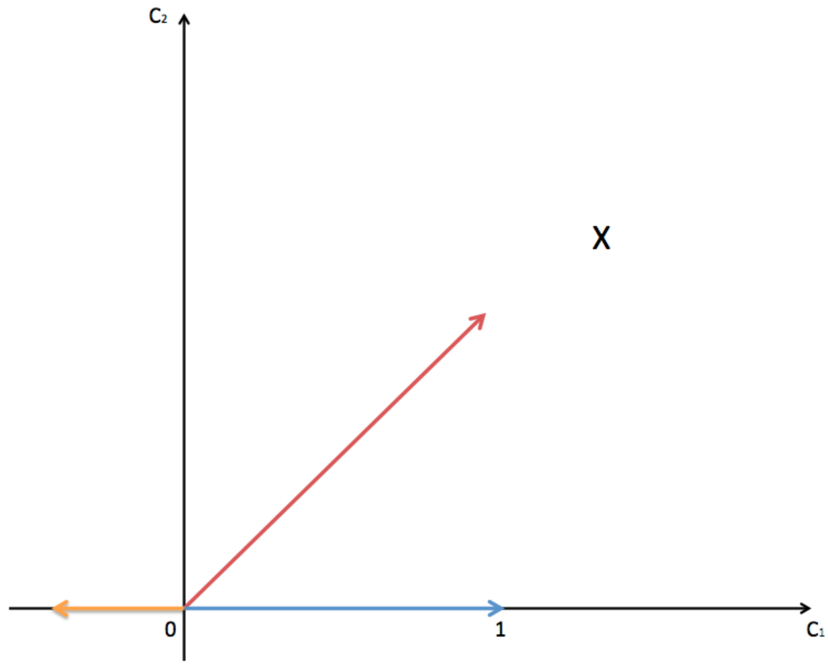
Under this setup, we can represent a security $j \in \{1, \dots, J\}$ as a vector

$$x^j = \begin{bmatrix} x_1^j \\ \vdots \\ x_S^j \end{bmatrix}$$

and define a *security structure* by a matrix X:

$$X = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^{J-1} & x_1^J \\ x_2^1 & x_2^2 & \cdots & x_2^{J-1} & x_2^J \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{S-1}^1 & x_{S-1}^2 & \cdots & x_{S-1}^{J-1} & x_{S-1}^J \\ x_S^1 & x_S^2 & \cdots & x_S^{J-1} & x_S^J \end{bmatrix} = \begin{bmatrix} x^1 & x^2 & \cdots & x^{J-1} & x^J \end{bmatrix}$$

An important example of a security structure is given by securities that are standard basis vectors, which are called *Arrow-Debreu* securities. Consider for example $S = 2$ and $e_1 = (1, 0)'$:



Note that if this is the only security available in the market, we can replicate any security in the horizontal axis (e.g. the orange security above) by buying a quantity $\alpha \in \mathbb{R}$ of e_1, but we cannot replicate any of the securities outside of the horizontal axis (e.g. the red security). When the available securities are not enough to replicate all securities in $\mathbb{R}^S$ we say that markets are *incomplete*.

If we introduce another asset $e_2 = (0, 1)'$, all securities in $\mathbb{R}^2$ become replicable. In this case, we say that markets are *complete*. Note that adding another asset $x^3 \in \mathbb{R}^2$ does not benefit us in any way, as we were already able to span the whole set $\mathbb{R}^2$ with the previous two. The security structure generated by e_1 and e_2 is

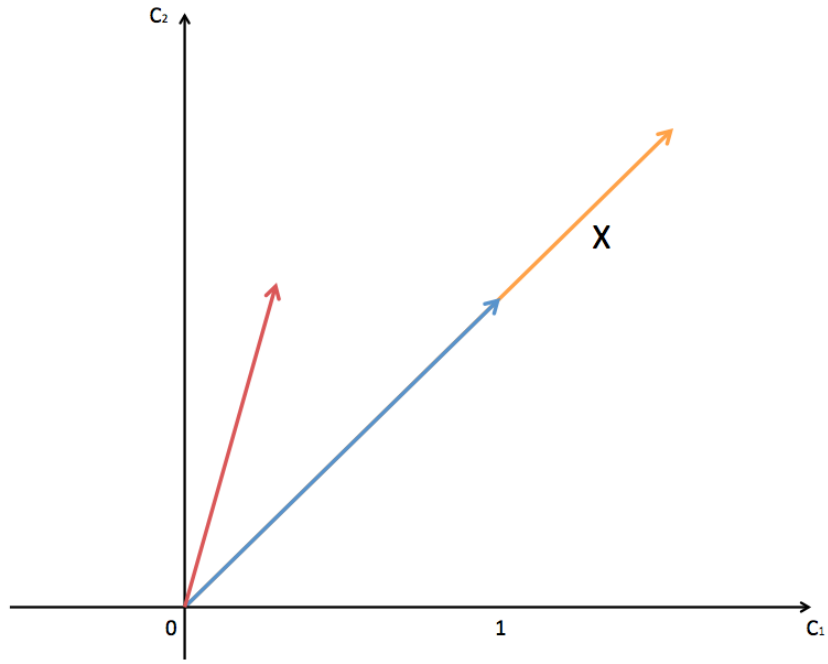
$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In general, given S possible states in $t = 1$ and $J = S$ securities we call *Arrow-Debreu security structure* the matrix

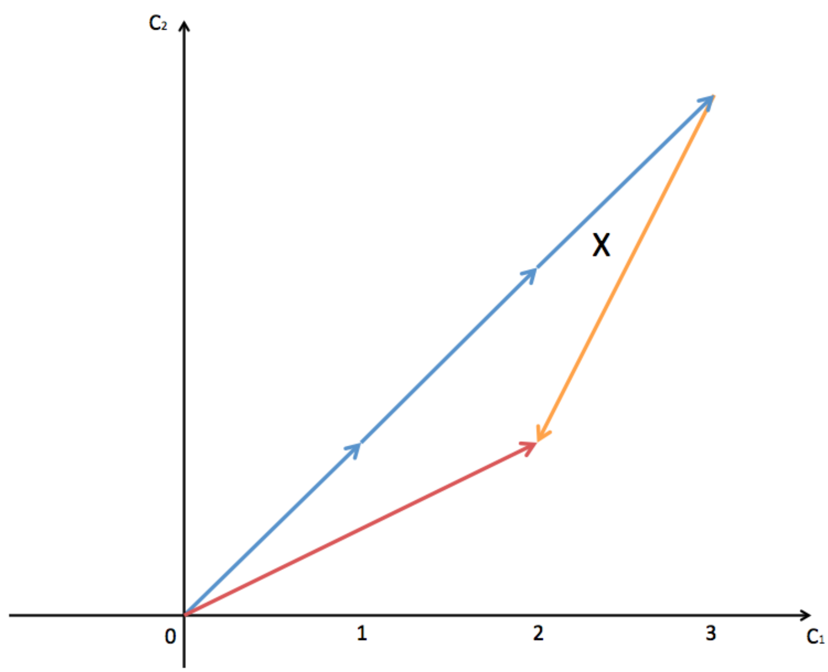
$$X^{AD} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Note that 1) all payoffs (securities) are linearly independent vectors in $\mathbb{R}^S$ 2) markets are complete by construction (since also $J = S$).

Consider now a general security structure in $\mathbb{R}^2$. Suppose there is only a riskless bond that pays 1 in each state of the world: $b = (1, 1)$.



Under this security structure, all riskless securities on the 45 degree line are replicable (e.g. orange security above) but none of the securities outside of it are replicable (e.g. 45 degree line). Adding a risky security, say $c = (2, 1)'$, allows us to span the whole $\mathbb{R}^2$ set. For instance, suppose we wish to replicate the security $d = (1, 2)'$. We can buy 3 securities b (in blue below) and sell short one unit of security c (in orange), thus obtaining security d (in red):



Finally, notice that in the market structure considered here

$$X = \begin{bmatrix} b & c \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

b and c are linearly independent vectors in $\mathbb{R}^2$, and therefore the payoffs that can be obtained with their linear combinations coincide with those that can be obtained with linear combinations of the Arrow-Debreu securities e_1 and e_2 considered before: therefore, markets are complete.

Before going further, it is worthwile two three important remarks.

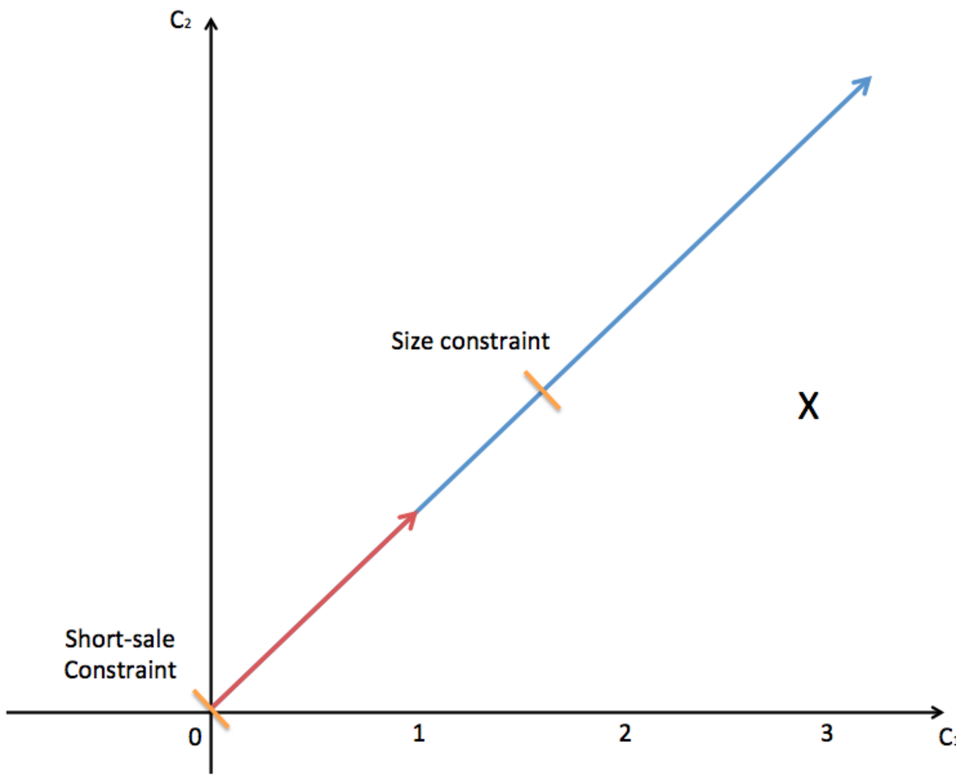
1) The state space itself is an important modeling choice. The relevance of this can hardly be overstated, because market completeness depends directly on how we specify the state space. Suppose for example that the security structure looks like:

$$X = \begin{bmatrix} 1 & 0 & 9 & 4 & 2 \\ 2 & 1 & 5 & 2 & 2 \\ 3 & 1 & 8 & 4 & 9 \\ 4 & 2 & 1 & 2 & 2 \\ 4 & 2 & 1 & 2 & 2 \end{bmatrix}$$

Clearly markets are incomplete, but if the state $s = 5$ is a state of the world equivalent to $s = 4$ in all aspects with the only exception that "at noon my cat is on the third floor", then state $s = 5$ probably isn't a relevant one and for modeling purposes should be disregarded. And after eliminating state $s = 5$ markets are complete (in the relevant states $s = 1, 2, 3, 4$). Another example where the state space choice is important is in derivatives pricing, where we assume that the state space corresponds to the space of possible values of the underlying asset.

2) Although we said we would not deal with frictions until the third part of the course, market incompleteness *is* a market friction: an incomplete market is equivalent to a market in which there are infinite transaction costs to trade assets whose payoffs are inearly independent from the existing assets, for which instead transaction costs are zero. From this point of view, we are dealing with a very "black and white" setup: either we have no frictions at all (with market completeness) or an extreme friction (with market incompleteness). Later in this course we will explore the "grey" area in between.

Finally, trade limitations are another kind of market friction we will encounter in this part of the course. Suppose there are only two states of the world and the security structure includes only a riskless bond, which we can neitherbuy in large amounts nor *short-sell* (assume there is a law that prohibits these practices):



Therefore, we can only achieve payoffs on the 45 degree line between the two yellow segments.

2.1 General Security Structure

We now have all the tools to formalize our analysis. We can define:

- A **Portfolio** as a vector $h \in \mathbb{R}^J$, whose entries represent a quantity for each asset in the security structure.
- The **Portfolio Payoff** is given by $\sum_{j=1}^{J} h_j x^j = Xh$
- The **Asset Span** is $\langle X \rangle = \left\{ z \in \mathbb{R}^S : \exists h \in \mathbb{R}^J \textit{such that } z = Xh \right\}$

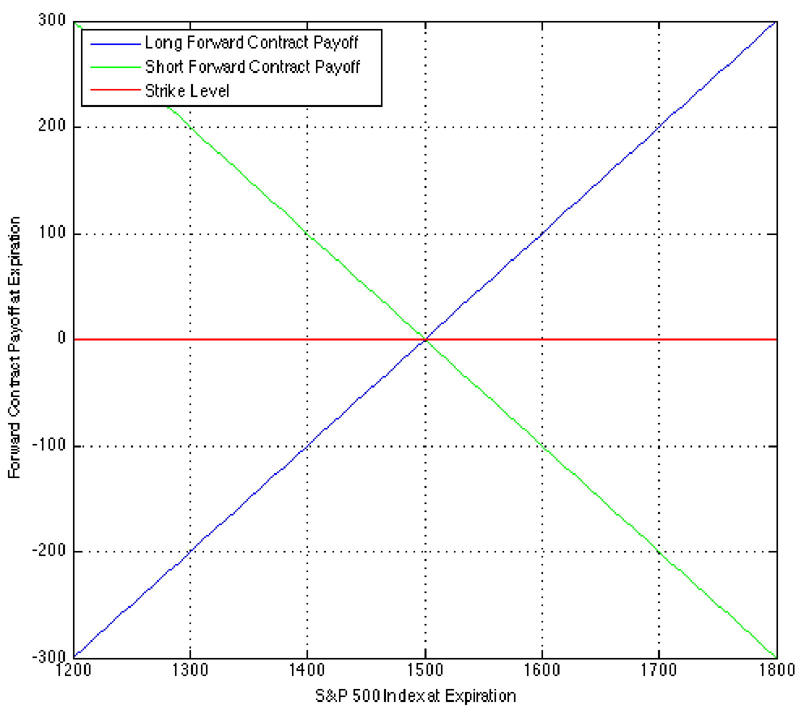
Note that it is always the case that $\langle X \rangle \subseteq \mathbb{R}^S$, and we have market completeness if and only if $\langle X \rangle = \mathbb{R}^S$, that is, if and only if $rank(X) = S$. In other words, market completeness refers to a market structure in which there are at least S *linearly independent* assets. We say that security j is redundant if there exists $h \in \mathbb{R}^J$ such that $x^j = Xh$ and $h_j = 0$: when markets are complete and $J > S$ there are $J - S$ redundant assets. When $rank(X) < S$ markets are incomplete, and if there are $J < S$ linearly independent assets then $S - J$ linearly independent assets are needed to achieve market completeness.

2.2 Derivatives

While securities represent property rights (hence contracts), derivatives *derive* their value from an underlying security. The most popular derivatives are Swaps, Futures and Options. A natural question to ask is: since a derivative's value is a function of the underlying security, are derivatives always redundant assets? We will see that this is not the case in general, and the answer depends on the fact that *functional dependency* does not imply *linear dependency*.

2.2.1 Forward Contracts

Forward contracts are binding agreements to buy or sell a given security at a specified price, quantity, time and delivery logistics. Futures contracts are the same from a payoff point of view, but differ from forwards because they are traded on exchanges that require collateral posting and hence are more liquid.



The strike level is normally set so that the value of the contract at initiation is zero. Compared to an outright long position in the underlying asset, it requires no cash outlay at initiation and gives the same synthetic exposure to the underlying (that is, if the underlying ends up higher by 10 at expiration, the forward will pay 10 as well). Assuming that a bond pays \$1 at maturity, the relation between a unit of the underlying and the corresponding forward contract is given by

$$Forward\ Contract\ Value = Value\ of\ Underlying - Strike \times Bond$$

To see this, suppose the forward contract on a stock S is agreed at time 0 to buy an asset at a future date T at a specified price K, so that the payoff is $(S_T - K)$. Now we ask: what is the fair price of this contract today? If we can replicate this payoff with a combination of stocks and bonds whose price is known today, we know that the price of the forward is going to be equal to the price of that combination of stocks and bonds.

Suppose today we buy the stock at price S_0, and sell K bonds maturing at T at a price of e^-rT each. What do we get at time T? The stock will be worth S_T and each bond will be worth 1, so K bonds will be worth K - since we sold them, we will have to *pay* K. In total, we have $S_T - K$. Therefore we replicated the payoff of the forward contract, and hence the price of the forward contract today will be equal to the value of the portfolio of long 1 stock and short K bonds today, that is, $S_0 - Ke^{-rT}$, that is, the value of the underlying asset minus the strike price times the bond price.

The forward contract can settle in one of two ways: 1) in cash, i.e. parties exchange the difference between the underlying at maturity and the strike price in dollar value, which is less costtly and more practical than 2) by physical delivery, i.e. parties exchange the underlying at the agreed price. Credit risk can be an issue for over-the-counter forwards, for which credit checks and bank letters of credit may be required other than collateral postings, while for exchange-based transaction counterparty risk is reduced since the clearing house guarantees the transactions.

2.2.2 Options

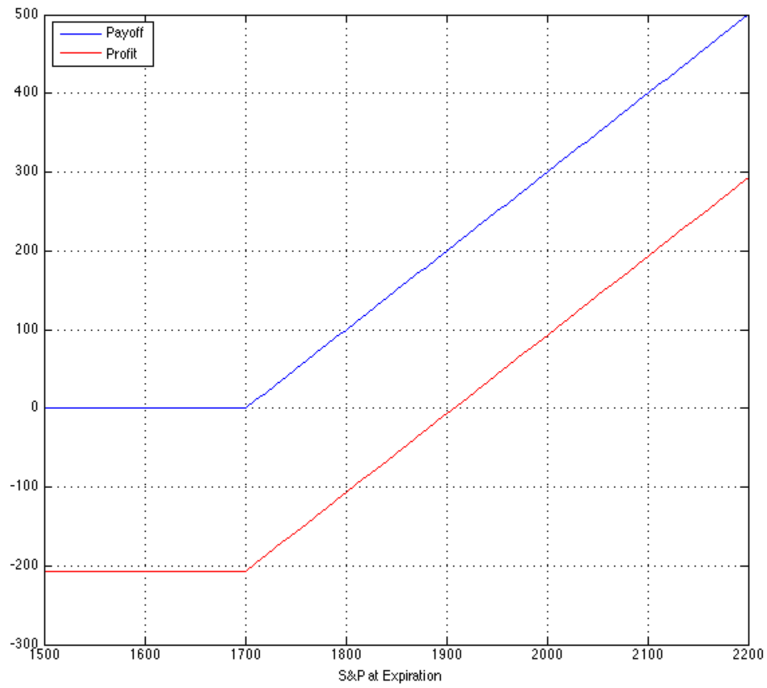
A *call option* is a contract that gives the right (not the obligation) to buy an asset at a specified price on a future date. From the point of view of the buyer, it preserves the upside potential from owning an asset without the downside risk. The seller has an obligation to sell if the buyer chooses to buy. The strike price is the price at which the parties agree to exchange the underlying. When the buyer chooses to buy the asset, we say he *exercises* the option. The expiration date is the date by which the buyer has to decide whether to exercise his right. Classified by exercise style, there are three main classes of options: 1) European options can only be exercised at expiration date 2) American options can be exercised anytime between inception and expiry and 3) Bermudan options can be exercised during some specified dates before or on expiration date.

For European options, because the buyer only exercises when the spot price at maturity is higher than the strike price (otherwise he would rationally buy the security in the market for less). Therefore the payoff at expiration is given by

$$\max\{S - K, 0\} \equiv (S - K)^+$$

and the profit is given by the payoff minus the future value of the option price, called *option premium*.

Example: suppose you buy a 1-year call option on the S&P 500, currently trading at 1680, with strike price 1700 and premium 197.34 dollars. Assuming a 5% one-year risk-free rate, if the index value in one year is 2000 then the profit is given by 2000 - 1700 - 197.34×(1 + 5%) = 92.79, but if it turns out to be 1600 then the loss is just 197.34×(1 + 5%) = 207.21.



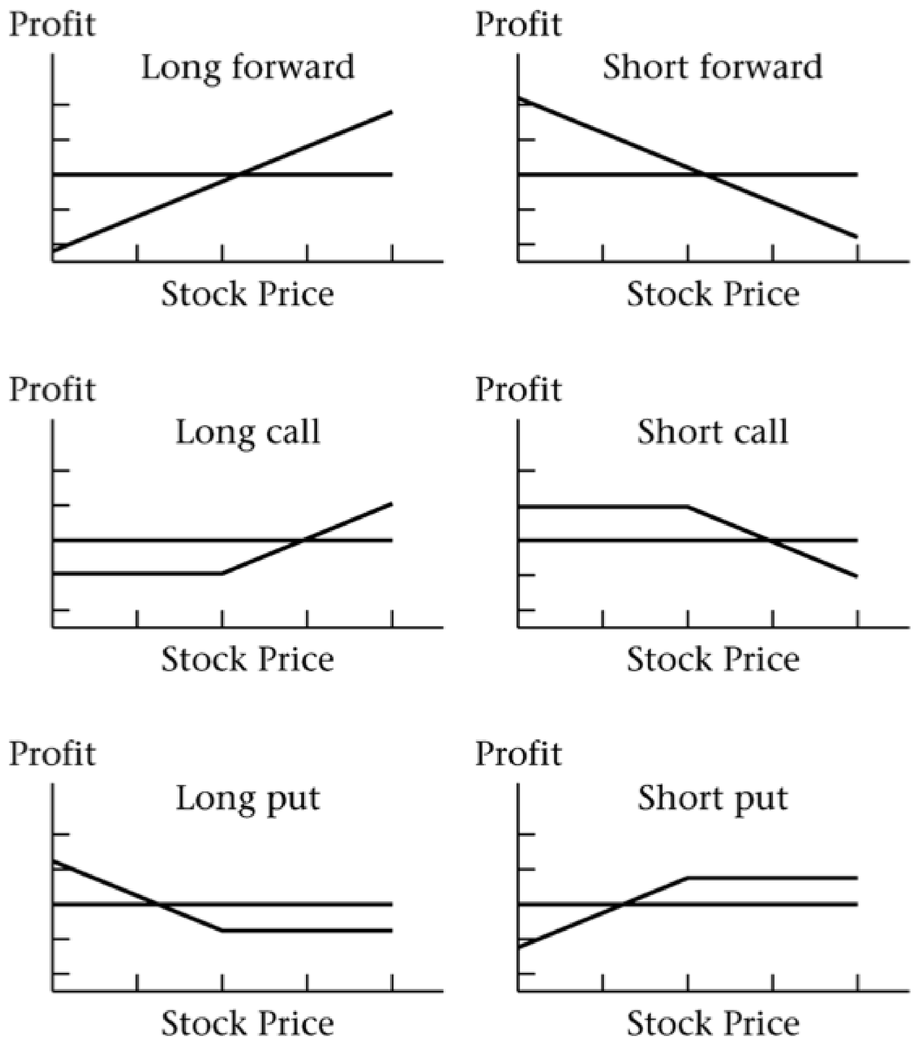
Similarly, a *put option* gives the buyer the right but not the obligation to sell an asset at a determined price on a future date. The seller of a put is obligated to buy if called upon to do so by the buyer. Similarly to a call option, for a buyer the payoff is $\max(K-S,0) \equiv (K-S)^+$ and the profit is given by the payoff minus the future value of the option price. It is worthwhile to remark that while a call option increases in value when the underlying rises in value, the opposite is true for a put option.

We say that an option is *in the money (out of the money)* if it would have a positive (negative) payoff if exercised immediately. If the spot equals the strike we say the option is *at the money*. There are two interesting examples of derivatives different from options which turn out to be valued just like options:

1. **Homeowner insurance**: because insurance pays only in the case of a damage to the house, it can be thought of as a put option on the value of the house.

2. **Equity-Linked Certificates of Deposit**: this contract pays the invested amount plus 70% of the gain in the S&P 500 index. For instance, suppose we invest $10,000 when the S&P 500 is at 1700, then the payoff would be $10,000 \times \left(1 + 0.7 \max\left\{\frac{S\&P500_{Final}}{1700} - 1, 0\right\}\right)$.

Below is a short recap of the profit functions of the derivatives discussed so far:



2.3 Back to Security Structures

We may now address the question: are derivatives useful to make markets complete? To make things concrete, let us consider a specific example in which the possible stock value at time $t = 1$ are equal to the index of the state of the world: $s = (1, 2, \ldots, S)$. We can introduce $S - 1$ call options with payoff $(s - k)^+$ for $k = 1, \ldots, S - 1$: we obtain the securities

$$c_1 = (0, 1, 2, \ldots, S - 2, S - 1)'$$

$$c_2 = (0, 0, 1, \ldots, S - 3, S - 2)'$$

$$\vdots$$

$$c_{S-1} = (0, 0, 0, \dots, 0, 1)'$$

Which together with the stock give rise to the security structure

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ S & S-1 & \cdots & 1 \end{pmatrix}$$

This is an upper triangular $S \times S$ matrix whose determinant is one (the product of the terms on the diagonal). Therefore X is full rank and markets are complete.

2.3.1 Prices

Let $p \in \mathbb{R}^J$ be the vector of prices for each asset. Then the cost of portfolio h is given by

$$p \cdot h \equiv \sum_{j=1}^{J} p_j h_j$$

and if $p_j \neq 0$, the (gross) return vector for asset j is given by $R_j = \frac{x^j}{p_j}$.

Exercises

1) Using the Laplace expansion of the determinant of an $S \times S$ matrix, prove that

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ S & S-1 & \cdots & 1 \end{pmatrix}$$

Has determinant equal to 1.

2) Now repeat exercise 1 by adding to the security structure put options on the stock, following the same steps as explained in the text (for call options). Are markets complete?

3) Suppose there exist only a risk-free asset $x^1 = (1, 1, \ldots, 1)'$ and a risky asset $x^2 \neq x^1$ and S states of the world. Let p_1 and p_2 be the prices of these two assets. A forward contract on the stock is an agreement to pay an amount F at a future date $t = T$ in exchange for the payment x_s^j when the state $s \in \{1, 2, \ldots, S\}$ realizes, with no cash flow exchange at time $t = 0$. Assuming arbitrage opportunities are ruled out, find the fair value of F.

Chapter 3

Pricing in the One Period Model

Let us begin with some notation: for $x, y \in \mathbb{R}^n$ we write

- $y \geq x$ if for each $i = 1, \dots, n$ $y_i \geq x_i$
- $y > x$ if $y \geq x$ and $y \neq x$
- $y \gg x$ if for each $i = 1, \dots, n$ $y_i > x_i$
- $y \cdot x$ for the inner product $\sum_{i=1}^{n} x_i y_i$

A fundamental concept in Asset Pricing is that of *no arbitrage*. In our setup it has three forms: given any two portfolios $h, k \in \mathbb{R}^J$and a security structure $X \in \mathbb{R}^{S \times J}$,

1. **Law of One Price**: if $Xh = Xk$ then $p \cdot h = p \cdot k$
2. **No Strong Arbitrage**: if $Xh \geq 0$ then $p \cdot h \geq 0$
3. **No Arbitrage**: if $Xh > 0$ then $p \cdot h > 0$

These definitions are related through the following lemmas:

Lemma 1: Law of One Price (LOOP) implies that *every portfolio with zero payoff has zero price.*

Proof: grouping terms we can equivalently write $X(h - k) = 0 \Rightarrow p \cdot (h - k) = 0$, and therefore portfolio $w \equiv h - k$, which has zero payoff by construction, also has zero price.

Lemma 2: No Arbitrage (NA) implies No Strong Arbitrage (NSA).

Proof: trivial.

Lemma 3: NSA implies LOOP.

Proof: we prove the contrapositive “if LOOP does not hold, then NSA does not hold either”. If the LOOP does not hold, then $Xh = Xk$ and $p \cdot h \neq p \cdot k$. Assume $p \cdot h < p \cdot k$, then we have $X(h-k) = 0 \geq 0$ and $p \cdot (h-k) < 0$, which is a violation of NSA for the portfolio $w \equiv h - k$. Similarly, if $p \cdot h > p \cdot k$ then $X(k-h) = 0 \geq 0$ and $p \cdot (k-h) < 0$, again a violation of NSA for the portfolio $q \equiv k - h$.

3.1 Forwards Revisited

Consider the following payment and payoff timing combinations:

1. Outright purchase
2. Fully leveraged purchase: you borrow the money to execute the purchase
3. Prepaid forward: you pay today to receive shares in the future
4. Forward contract: you agree to a price now, which you pay when you receive the shares in the future

3.1.1 Prepaid Forwards

Suppose we wish to price a prepaid forward for a stock with no dividends. Clearly the timing of delivery is irrelevant, and the price of the prepaid forward $F^p_{0,T}$ for a stock delivered at $t = T$ is just equal to the current stock price S_0 at $t = 0$. This reasoning is called *pricing by analogy.*

Another way of getting to the same result is *pricing by arbitrage.* Suppose that at $t = 0$ we observe that $F^p_{0,T} > S_0$, then we could buy the stock today at S_0, sell the forward at F^p_0 and pocket the difference $F^p_{0,T} - S_0 > 0$. In $t = T$ our stock is worth S_T, and we owe S_T from the forward contract we sold in $t = 0$: as a result, in $t = T$ we receive 0. This would constitute an arbitrage. A similar argument can be used for the case when $F^p_{0,T} < S_0$. Therefore, absence of arbitrage requires that $F^p_{0,T} = S_0$.

When there are dividends the two pricing arguments do not hold any longer: the holder of the stock, unlike that of the forward, will not receive dividends in the period $[0, T]$. As a consequence, it must be that $F^p_{0,T} < S_0$ since $F^p_{0,T} = S_0 - PV(\textit{Dividend Payments in } [0,T])$. In particular:

- With discrete dividends D_{t_i} for $t_1, \dots, t_n \in [0, T]$ and assuming reinvestment at the risk-free rate, we have $F^p_{0,T} = S_0 - \sum_{i=1}^{n} PV_0(D_{t_i}) = S_0 - \sum_{i=1}^{n} D_{t_i} e^{-rt_i}$
- With continuous dividends with annualized dividend yield δ, we have $F^p_{0,T} = S_0 e^{-\delta T}$

Note that this only applied to deterministic dividends: when dividends are stochastic the securities structure turns from a matrix into a cube and we can no longer use our one period model.

3.1.2 Forwards

Obviously the forward price is just the future value of the prepaid forward:

- With no dividends, $F_{0,T} = S_0 e^{rT}$
- With discrete dividends, $F_{0,T} = S_0 e^{rT} - \sum_{i=1}^{n} FV_T(D_{t_i}) = S_0 e^{rT} - \sum_{i=1}^{n} D_{t_i} e^{r(T-t_i)}$
- With continuous dividends, $F_{0,T} = S_0 e^{(r-\delta)T}$

Indexes are an example of assets with continuous dividends. We call *forward premium* the quantity $\frac{F_{0,T}}{S_0}$, which can be used to infer the current stock price from the forward price. The *annualized forward premium* is $\pi = \frac{1}{T} \ln \frac{F_{0,T}}{S_0}$.

We can also use a no-arbitrage argument to price a forward: assuming continuous dividends with rate δ, we can buy $e^{-\delta T}$ units of the stock worth S_0 for total price of $S_0 e^{-\delta T}$ by borrowing the full amount. At $t = 0$ there is no cash outlay; however, at $t = T$ the portfolio is worth $S_T - S_0 e^{(r-\delta)T}$. Since the long forward payoff is $S_T - F_{0,T}$, this implies that to exclude arbitrage it must be the case that $F_{0,T} = S_0 e^{(r-\delta)T}$.

TABLE 5.3 Demonstration that borrowing $S_0 e^{-\delta T}$ to buy $e^{-\delta T}$ shares of the index replicates the payoff to a forward contract, $S_T - F_{0,T}$.

	Cash Flows	
Transaction	**Time 0**	**Time *T* (expiration)**
Buy $e^{-\delta T}$ Units of the Index	$-S_0 e^{-\delta T}$	$+S_T$
Borrow $S_0 e^{-\delta T}$	$+S_0 e^{-\delta T}$	$-S_0 e^{(r-\delta)T}$
Total	0	$S_T - S_0 e^{(r-\delta)T}$

It follows that

$$\textit{Forward} = \textit{Stock} - \textit{Zero-Coupon Bond}$$

An interesting application of this fact is the so-called *cash and carry arbitrage*: a market maker can make a riskless profit by (for instance) selling short a forward contract and going long a synthetic forward: the payoff from this strategy is $F_{0,T} - S_0 e^{(r-\delta)T}$.

A natural question at this point is: is the forward price a market prediction of the future price? The answer is no: the formula $F_{0,T} = S_0 e^{(r-\delta)T}$ shows clearly that the forward price provides no more information than r, δ and S_0.

3.2 Options Revisited

Consider two European options, one call and one put, with the same strike K and time to expiry T. The *put-call parity* relation requires that

$$C(K,T) - P(K,T) = PV_0\,(Forward\ Price - Strike) = e^{-rT}\,(F_{0,T} - K)$$

Note that if $F_{0,T} = K$ the long call short put portfolio above is equivalent to a synthetic forward, and in fact will have zero price.[1] With a dividend stream $\{D_{t_i}\}_{i=1}^{n}$ we can rewrite the above relation as

$$C(K,T) - P(K,T) = S_0 - PV_0\left(\{D_{t_i}\}_{i=1}^{n}\right) - e^{-rT}K$$

While for an index (with continuous dividends) we have

$$C(K,T) - P(K,T) = S_0 e^{-\delta T} - e^{-rT}K$$

3.2.1 Option Price Boundaries

Because an American option can be exercised at any time, while a European option can only be exercised at maturity, it must be the case that

$$C_A\,(K,T) \geq C_E\,(K,T)$$

$$P_A\,(K,T) \geq P_E\,(K,T)$$

In general, the American call option price cannot exceed the stock price (otherwise you would never buy the option but just buy the stock). Moreover, the European call option cannot be lower than 1) the price implied by put-call parity by setting to zero the put price, or 2) zero, whichever is highest. That is,[2]

$$S_0 > C_A\,(K,T) \geq C_E\,(K,T) > e^{-rT}\,(F_{0,T} - K)^{+}$$

Similarly, the American put option price cannot exceed the strike price (otherwise you would never buy the option but just buy the bond). Moreover, the European put option cannot be lower than 1) the price implied by put-call parity by setting to zero the call price, or 2) zero, whichever is highest. That is,

$$K > P_A\,(K,T) \geq P_E\,(K,T) > e^{-rT}\,(K - F_{0,T})^{+}$$

[1] An alternative definition of "At The Money" option is to say that an option is at the money when the forward price equals the strike price. Under this definition, a "long call short put" portfolio replicates a forward when the two options are at the money.

[2] Since $\max\left\{e^{-rT}\,(F_{0,T} - S_0)\,, 0\right\} = e^{-rT}\,(F_{0,T} - S_0)^{+}$ and $C_E\,(K,T) > 0$.

Rationally, we never exercise an American call option on a stock with no dividends: this is because[3]

$$C_A(K,T) \geq C_E(K,T) = S_0 - Ke^{-rT} + P_E(K,T) =$$

$$= S_0 - K + \underbrace{K(1-e^{-rT}) + P_E(K,T)}_{>0} > S_0 - K$$

That is, for a holder of an American option it is always best to sell the option rather than exercise it early. By LOOP, the price of an American call option on a stock with no dividends is the same as that of an European option. Note that this is not true for a dividend-paying stock, as well as for an American put on a non-dividend-paying stock.

3.2.2 Time to Expiration

An American option (both put and call) with more time to expiration is at least as valuable as an American option with less time to expiration. This is because the longer option can easily be converted into the shorter option by exercising it early. European call options on dividend-paying stock and European puts may be less valuable than an otherwise identical option with less time to expiration. A European call option on a non-dividend paying stock will be more valuable than an otherwise identical option with less time to expiration.

3.2.3 Strike Price

Let $K_1 < K_2$, then we know that $C(K_1) \geq C(K_2)$ and $P(K_1) \leq P(K_2)$ (since their payoff is more likely to be positive at $t = T$). A less obvious fact is that $C(K_1) - C(K_2) \leq K_2 - K_1$, because the maximum payoff for a *collar* (long option with low strike, short option with high strike) with strikes $K_1 < K_2$ is $K_2 - K_1$, and the price of the collar cannot exceed its maximum payoff. In the same way, for put options we have $P(K_2) - P(K_1) \leq K_2 - K_1$. Finally, the option price is *convex* with respect to its strike: for $K_1 < K_2 < K_3$,

$$\frac{C(K_2) - C(K_1)}{K_2 - K_1} \leq \frac{C(K_3) - C(K_2)}{K_3 - K_2}$$

Below is a brief recap of the put-call parity relations examined so far:

[3] Remember that with no dividends $F_{0,T} = S_0 e^{rT}$.

TABLE 9.9 Versions of put-call parity. Notation in the table includes the spot currency exchange rate, x_0; the risk-free interest rate in the foreign currency, r_f; and the current bond price, B_0.

Underlying Asset	Parity Relationship
Futures Contract	$e^{-rT} F_{0,T} = C(K,T) - P(K,T) + e^{-rT} K$
Stock, No-Dividend	$S_0 = C(K,T) - P(K,T) + e^{-rT} K$
Stock, Discrete Dividend	$S_0 - PV_{0,T}(Div) = C(K,T) - P(K,T) + e^{-rT} K$
Stock, Continuous Dividend	$e^{-\delta T} S_0 = C(K,T) - P(K,T) + e^{-rT} K$
Currency	$e^{-r_f T} x_0 = C(K,T) - P(K,T) + e^{-rT} K$
Bond	$B_0 - PV_{0,T}(Coupons) = C(K,T) - P(K,T) + e^{-rT} K$

To conclude, it is worthwile to note that many of these results on option price bounds can be derived within our (very simple) two period model: they are incredibly robust. Once we adopt a more specific settings (like the famous Black-Scholes model) we can use more sophisticated tools such as needs dynamic replication, and the results become deeper but at the same time more hinging on the specific model used.

3.3 Back to the One Period Model

So far we used prices of existing assets to directly derive the price (or price bounds) on other assets. Now we will go along an indirect route: first we will derive the price of each individual state - called *state price* - and then we will use this theoretical tool to derive the price of other assets.

For a given price vector $p \in \mathbb{R}^J$ and $z \in \langle X \rangle$ define the set v as

$$v(z) \equiv \{p \cdot h : z = Xh\}$$

If LOOP holds then v is a *linear functional,* that is, a function mapping $\langle X \rangle$ onto $\mathbb{R}$ such that:

1. v is single-valued, because if $Xh = z$ then by LOOP there exist only one $p \in \mathbb{R}^J$ such that $p \cdot h = v$, and therefore v is a singleton. This means that for any $h \in \mathbb{R}^J$ such that $Xh \in \langle X \rangle$ we can write we can write $v\,(Xh) = p \cdot h$.

2. v is linear on $\langle X \rangle$,[4] since for any $\alpha, \beta \in \mathbb{R}$, $z_1 = Xh \in \langle X \rangle$ and $z_2 = Xk \in \langle X \rangle$ we have $\alpha z_1 + \beta z_2 \in \langle X \rangle$ and

$$\alpha v(z_1) + \beta v(z_2) = \alpha v(Xh) + \beta v(Xk) = \alpha p \cdot h + \beta p \cdot k = p \cdot (\alpha h + \beta k) =$$

[4] That is, for all $z_1, z_2 \in \langle X \rangle$ and $\alpha, \beta \in \mathbb{R}$ such that $\alpha z_1 + \beta z_2 \in \langle X \rangle$ it holds that $v(\alpha z_1 + \beta z_2) = \alpha v(z_1) + \beta v(z_2)$.

$$= v\left(X(\alpha h + \beta k)\right) = v\left(\alpha Xh + \beta Xk\right) = v\left(\alpha z_1 + \beta z_2\right)$$

3. $v(0) = 0$, since $Xh = 0$ always has the solution $h = 0$ and because v is single-valued (by LOOP) it must be that $p \cdot h = p \cdot 0 = 0$ is the only price of the payoff $z = 0$.

The converse is also true: if there exists a linear functional v defined in $\langle X \rangle$, then LOOP holds.

3.3.1 State Prices

Definition: a *vector of state prices* is a vector $q \in \mathbb{R}^S$ such that $p = X'q$.[5]
Definition: a linear functional $V : \mathbb{R}^S \to \mathbb{R}$ is a valuation function if

1. $V(z) = v(z)$ for every $z \in \langle X \rangle$
2. For every $z \notin \langle X \rangle$, $V(z) = q \cdot z$ for $q \in \mathbb{R}^S$ with $q_s = V(e_s)$: V extends v from $\langle X \rangle$ to $\mathbb{R}^S$.

Recall that e_s is the standard basis introduced in chapter 2: $e_s \in \mathbb{R}^S$ is a vector with the s^{th} entry equal to 1 and all other entries equal to zero. The next proposition addresses the relationship between v,V and q.

Proposition: if LOOP holds and q is a vector of state prices, then $V(z) = q \cdot z$ for all $z \in \langle X \rangle$.

To see this we only need to show that also for $z \in \langle X \rangle$ we have $V(z) = q \cdot z \,(= v(z))$. Suppose that q is a vector of state prices and LOOP holds, then for $z \in \langle X \rangle$ we have

$$v(z) = p \cdot h = \sum_{j=1}^{J} p_j h_j = \sum_{j=1}^{J} \left(x^j \cdot q\right) h_j = \sum_{j=1}^{J} \left(\sum_{s=1}^{S} x_s^j \cdot q_s\right) h_j =$$

$$= \sum_{s=1}^{S} \left(\sum_{j=1}^{J} x_s^j \cdot h_j\right) q_s = \sum_{s=1}^{S} z_s q_s = q \cdot z$$

Moreover the converse is also true, and therefore the valuation function $V(z) = q \cdot z$ is a linear functional for all $z \in \mathbb{R}^S$ *if and only if* q is a vector of state prices and LOOP holds.

Below is a graphical example of state prices. Given the securities structure (red arrows)

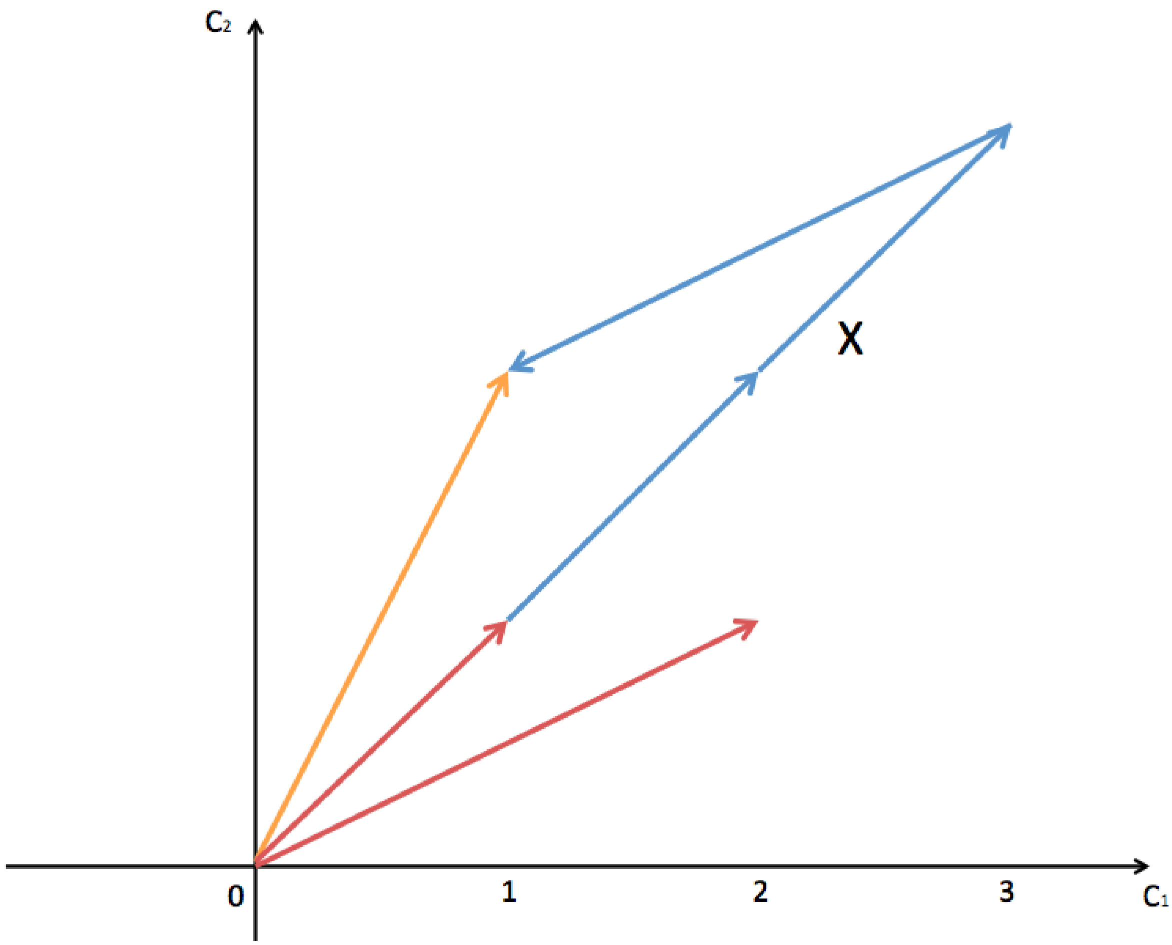
$$X = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

[5] That is for $j = 1, \ldots J$ we have $p_j = x^j \cdot q$.

we know that $p_1 = q_1 + q_2$ and $p_2 = 2q_1 + q_2$. As a consequence, we can value the security $x_3 = (1, 2) = 3x_1 - x_2$ (in yellow) simply as

$$3p_1 - p_2 = q_1 + 2q_2$$

by LOOP.



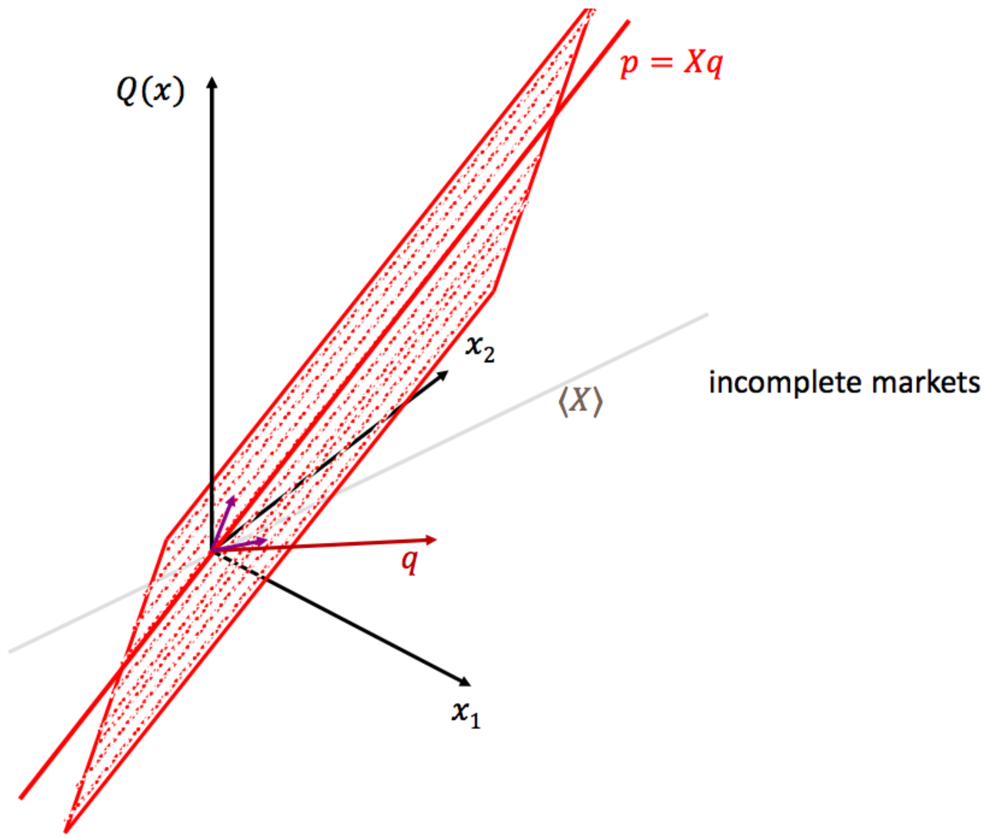
3.3.2 The Fundamental Theorem of Finance

- **Proposition 1**: Security prices exclude arbitrage if and only if there exists a valuation functional with $q \gg 0$.
- **Proposition 2**: Let X be a $S \times J$ matrix, and $p \in \mathbb{R}^J$. There is no $h \in \mathbb{R}^J$ satisfying $h \cdot p \leq 0$, $Xh \geq 0$ and at least one strict inequality if and only if there exists a vector $q \in \mathbb{R}^S$ with $q \gg 0$ and $p = X'q$.

It is hard to overstate the importance of this theorem: the *absence* of arbitrage is equivalent to the *existence* of a vector of positive state prices.

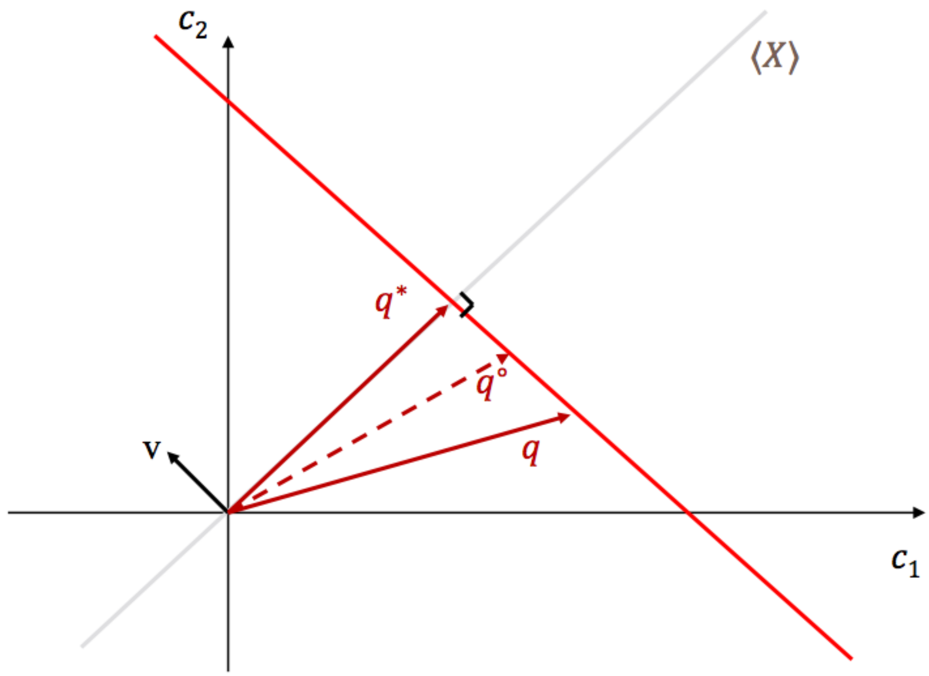
3.3.3 State Prices and Incomplete Markets

We have established an astonishing equivalence between arbitrage (the absence of arbitrage) and state prices (the existence of positive state prices). A natural follow-up question is whether there is a similar equivalence between state prices and market completeness. Suppose for instance that there are two states of the world and only one bond $x_1 = (1,1)'$ with price p_1. What are the state prices consistent with this incomplete market structure? We know that any $p_1 = q_1 + q_2$ would work, hence the state prices consistent with no arbitrage are all $q \in \mathbb{R}^2$ such that $q_1 \in (0, p_1)$ and $q_2 = p_1 - q_1$.



In the picture above, the red plane is the set of q's consistent with prices and the security structure

Of all $q \in \mathbb{R}^2$ consistent with no arbitrage however, there is a very special one that also belongs to $\langle X \rangle$: it is the unique *projection* of any state price vector q on $\langle X \rangle$. In our example, this would be achieved for $q_1 = \frac{p_1}{2}$ since then $q = \left(\frac{p_1}{2}, \frac{p_1}{2}\right)' = \frac{p_1}{2} \times (1,1)' \in \langle X \rangle$. We call this state price ***pricing kernel*** and denote it as q^*.



Now we are ready to state the important relation between state prices and market completeness:

- **Proposition 3**: Markets are complete and there is no arbitrage if and only if there exists a **unique** valuation functional.

An intuitive proof of this proposition is the following: if markets are complete, then for a given $X$and p the system $X'q = p$ has a unique solution $q \in \mathbb{R}^S_+$ (positivity follows from the assumption of no arbitrage). If markets are not complete, then there exists a vector $v \in \mathbb{R}^S$ such that $v \neq 0$ and $Xv = 0$. If there is no arbitrage, then there is a $q \gg 0$ and some $\alpha \in \mathbb{R}$ such that $q + \alpha v \gg 0$ and $X(q + \alpha v) = 0$. Since this is also true for any fraction of α, it follows that there are infinitely many state price vectors of the type $q + \alpha v$.

3.4 Asset Pricing Formulas

We now present four asset pricing formulas. They are effectively equivalent, but each one can be interpreted in a particular way and derived from the one period model.

3.4.1 State Price Model

This is just the pricing formula seen above: $p_j = \sum_{s=1}^{S} q_s x_s^j$

3.4.2 Stochastic Discount Factor

Analogously to the state price model, we can write $p_j = \sum_{s=1}^{S} \pi_s \frac{q_s}{\pi_s} x_s^j$ where π is the physical probability distribution of states. Defining the random variable *Stochastic Discount Factor* as $m_s \equiv \frac{q_s}{\pi_s}$, we get $p_j = \sum_{s=1}^{S} \pi_s m_s x_s^j = \mathbb{E}\left[m \cdot x^j\right]$.

Note that $p_j = \mathbb{E}\left[m \cdot x^j\right] = \mathbb{E}\left[x^j\right]\mathbb{E}\left[m\right] + Cov\left[m, x^j\right]$, and since for a risk-free bond $x_s^b = 1$ for all$s$, we have $p_b = \mathbb{E}\left[m\right] = \frac{1}{R^f}$ where R^f is the gross risk-free return. Therefore, for any asset j, $p_j = \frac{\mathbb{E}\left[x^j\right]}{R^f} + Cov\left[m, x^j\right]$. Typically, $Cov\left[m, x^j\right] < 0$.

Defining $R^j \equiv \frac{x^j}{p_j}$, we get $\mathbb{E}\left[m \cdot R^j\right] = 1$. Since for a risk-free bond $R^f = \frac{1}{\mathbb{E}[m]}$, we can write $\mathbb{E}\left[m \cdot \left(R^j - R^f\right)\right] = 0$, or

$$\mathbb{E}\left[m \cdot \left(R^j - R^f\right)\right] = \mathbb{E}\left[m\right]\left(\mathbb{E}\left[R^j\right] - R^f\right) + Cov\left(m, R^j\right) = 0$$

That is,

$$\mathbb{E}\left[R^j\right] - R^f = -\frac{Cov\left(m, R^j\right)}{\mathbb{E}\left[m\right]}$$

Which implies that the *excess return* for a generic asset j is determined solely by the covariance with the stochastic discount factor. This also means that an investor is only compensated (with a higher return) for holding *systematic* risk, not *idiosyncratic* risk.

Consider the stochastic discounf factor obtained from the pricing kernel $m^* \equiv \begin{bmatrix} \frac{q_1^*}{\pi} \\ \vdots \\ \frac{q_S^*}{\pi} \end{bmatrix}$. Note that m^* is the projection of any stochastic discount factor m on $\langle X \rangle$, that is,

$$m^* = proj\left(m | \langle X \rangle\right) \in \langle X \rangle$$

Which means that there exists a vector $h^* \in \mathbb{R}^J$ such that $m^* = Xh^*$. Therefore for any asset j we can write

$$p_j = \mathbb{E}\left[m^* \cdot x^j\right]$$

So for all assets we have

$$p = \mathbb{E}\left[X' m^*\right] = \mathbb{E}\left[X' X h^*\right] = \mathbb{E}\left[X' X\right] h^*$$

$\mathbb{E}\left[X' X\right]$ is a second order moment: assuming it is invertible we can write

$$h^* = \left(\mathbb{E}\left[X' X\right]\right)^{-1} p$$

And therefore plugging this in $m^* = Xh^*$ we obtain

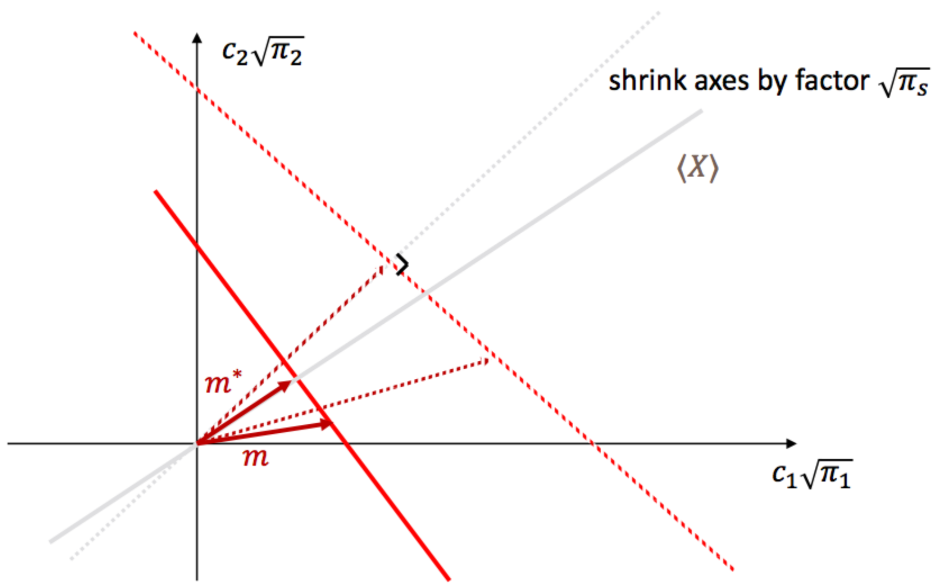
$$m^* = X\left(\mathbb{E}\left[X'X\right]\right)^{-1} p$$

This is similar to the language of *linear regressions*. When we run a regression of p on

$$y = Xb + \varepsilon$$

We find the linear combination of X that is “closest” to y by minimizing the varianc of the residual ε. We do this by forcing the residual to be “orthogonal” to X, $\mathbb{E}\left[X\varepsilon\right] = 0$. The projection of y onto X is defined as the fitted value $Xb = X\left(\mathbb{E}\left[X'X\right]\right)^{-1}\mathbb{E}\left[X'y\right]$. This idea is often illustrated by a residual vector ε that is perpendicular to a plane defined by the variable X. Thus, when the inner product is defined by a second moment, the operation “project y onto X” is a regression.

Finally, note that we can represent our previous discussion for state prices by shrinking the axes by a factor $\sqrt{\pi}$:



3.4.3 Equivalent Martingale Measure

Starting again from $p_j = \sum_{s=1}^{S} q_s x_s^j$, for a riskless bond we have $p_b = \sum_{s=1}^{S} q_s = \frac{1}{1+r^f}$, where r^f is the risk-free net return. Thus we can write $p_j = \frac{1}{1+r^f}\sum_{s=1}^{S}\frac{q_s}{\sum_{s=1}^{S} q_s}x_s^j = \frac{1}{1+r^f}\sum_{s=1}^{S}\hat{\pi}_s x_s^j = \frac{1}{1+r^f}\mathbb{E}^{\mathbb{Q}}\left[x^j\right]$, where $\hat{\pi}_s \equiv \frac{q_s}{\sum_{s=1}^{S} q_s}$.[6] Compared to the stochastic discount factor approach, we simply used a different

[6]The $\mathbb{Q}$ notation comes from the literature about risk-neutral valuation.

probability measure to discount future states. We will see that the significance of this probability measure is market-determined and its importance is paramount in options pricing theory.

3.4.4 State-Price Beta Model

Consider the stochastic discounf factor obtained from the pricing kernel $m^* \equiv \begin{bmatrix} \frac{q_1^*}{\pi} \\ \vdots \\ \frac{q_S^*}{\pi} \end{bmatrix}$, and define its return as $R^* = \frac{m^*}{p_{m^*}} \equiv \alpha m^*$ for $\alpha > 0$. Then we can write

$$\mathbb{E}\left[R^j\right] - R^f = -\frac{Cov\left(R^*, R^j\right)}{\mathbb{E}\left[R^*\right]}$$

Defining $\beta_j \equiv \frac{Cov\left(R^*, R^j\right)}{Var(R^*)}$ we can write for the asset j:

$$\mathbb{E}\left[R^j\right] - R^f = -\beta_j \frac{Var\left(R^*\right)}{\mathbb{E}\left[R^*\right]}$$

While for security x^*

$$\mathbb{E}\left[R^*\right] - R^f = -\frac{Var\left(R^*\right)}{\mathbb{E}\left[R^*\right]}$$

Therefore, for security j we have

$$\mathbb{E}\left[R^j\right] - R^f = \beta_j \left(\mathbb{E}\left[R^*\right] - R^f\right)$$

Which, if we assume a linear model for R^j and R^* can be specified and tested empirically as

$$R_k^j - R^f = \beta_j \left(R_k^* - R^f\right) + \varepsilon_k$$

with $Cov\left(R^*, \varepsilon\right) = \mathbb{E}\left[\varepsilon\right] = 0$.

In summary, the four equivalent pricing relations are:

1. **State Price Model**: $p_j = \sum_{s=1}^{S} q_s x_s^j$
2. **Stochastic Discount Factor**: $p_j = \mathbb{E}\left[m x^j\right]$
3. **Equivalent Martingale Measure**: $p_j = \frac{1}{1+r^f} \mathbb{E}^{\mathbb{Q}}\left[x^j\right]$
4. **State-Price Beta Model**: $\mathbb{E}\left[R^j\right] - R^f = \beta_j \left(\mathbb{E}\left[R^*\right] - R^f\right)$

As a last remark, note that whenever markets are incomplete, the multiplicity of state price vectors q translates directly into the multiplicity of stochastic discount factors m and of equivalent martingale measures $\hat{\pi}$.

3.5 Recovering State Prices from Option Prices

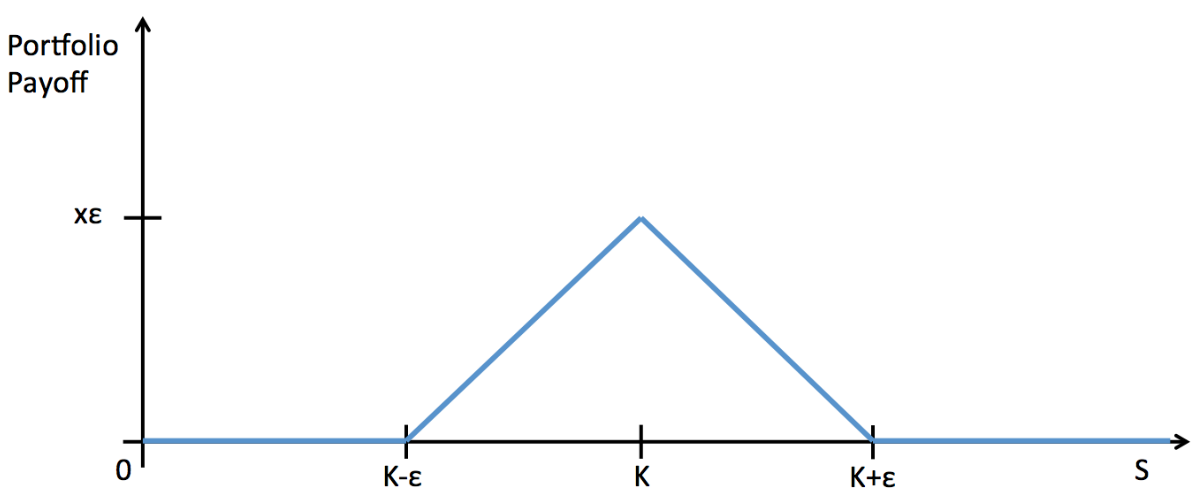
Let's assume for a moment that S_T, the value of the asset at expiration, can take on a *continuum* of values: at time $t = T$ the stock price can have any value $S_T \in \mathbb{R}_+$. In this section we will use the law of one price to derive the price of an Arrow-Debreu security for a continuum of states, which is a function $q : \mathbb{R}_+ \to \mathbb{R}$ called *state price density*. In ths context, a state price density is the price of an asset that pays one dollar in a particular state $x \in \mathbb{R}_+$ and zero in all others (here each value of the price for the underlying asset at time $t = T$ corresponds to a different state). Assuming further that there exist call options offered in the market that cover a continuum of strike prices (one for each possible outcome at expiration), the result of section 2.2.3 extends to this case: markets are complete.

Fix a strike $K > 0$, and a (small) number $\varepsilon > 0$. Consider the following portfolio:

- Buy x call options with strike $K - \varepsilon$
- Sell $2x$ call options with strike K
- Buy x call options with $K + \varepsilon$

State	Payoff
$S_T \in [0, K - \varepsilon)$	0
$S_T \in [K - \varepsilon, K)$	$x\,(S_T - (K - \varepsilon)) = x\,(S_T - K) + x\varepsilon$
$S_T \in [K, K + \varepsilon)$	$x\,(S_T - (K - \varepsilon)) - 2x\,(S_T - K) = -x\,(S_T - K) + x\varepsilon$
$S_T \in [K + \varepsilon, +\infty)$	$x\,(S_T - (K - \varepsilon)) - 2x\,(S_T - K) + x\,(S_T - (K + \varepsilon)) = 0$

The payoff of this portfolio (symmetric butterfly) looks like the following:



This portfolio corresponds the payoff of an Arrow-Debreu security for state $S_T = K$ if two conditions are satisfied:

1. The area under the payoff diagram (the sum of all possible payoffs) must be equal to one.
2. The option portfolio must yield a payoff of zero for all values of the underlying asset different from K.

The area under the payoff diagram is equal to $\frac{(x\varepsilon)\times(2\varepsilon)}{2} = x\varepsilon^2$: to achieve (1), we set $x = \frac{1}{\varepsilon^2}$; to achieve (2), we take the limit for $\varepsilon \to 0$. Now that we have replicated the payoff of an Arrow-Debreu security, we can use the law of one price to determine the state prices: for a fixed $\varepsilon > 0$ the price of the portfolio is

$$V(K,\varepsilon) = \frac{C(K+\varepsilon,T) - 2C(K,T) + C(K-\varepsilon,T)}{\varepsilon^2}$$

and therefore

$$q(K) = \lim_{\varepsilon\to 0} V(K,\varepsilon) = \lim_{\varepsilon\to 0} \frac{C(K+\varepsilon,T) - 2C(K,T) + C(K-\varepsilon,T)}{\varepsilon^2} = \frac{\partial^2 C(K,T)}{\partial K^2}$$

We established an important fact: because $q(K) > 0$ if and only if $\frac{\partial^2 C(K,T)}{\partial K^2} > 0$, by the fundamental theorem of finance there is no arbitrage if and only if for all $K > 0$ we have $\frac{\partial^2 C(K,T)}{\partial K^2} > 0$, that is, if option prices are convex.

There is another important relationship related to state prices and $\frac{\partial^2 C(K,T)}{\partial K^2}$. We can write the option price using the equivalent martingale measure representation as follows:

$$C(K,T) = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right]$$

Taking the first derivative with respect to K we get

$$\frac{\partial C(K,T)}{\partial K} = -e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[I\{S_T - K \geq 0\}\right] = -e^{-rT}P^{\mathbb{Q}}\{S_T \geq K\} = -e^{-rT}\left(1 - F^{\mathbb{Q}}_{S_T}(K)\right)$$

Where $F^{\mathbb{Q}}_{S_T}$ is the cumulative distribution function for the random variable S_T under the equivalent martingale measure $\mathbb{Q}$.[7] Differentiating again with respect to K we get

$$\frac{\partial^2 C(K,T)}{\partial K^2} = e^{-rT} f^{\mathbb{Q}}_{S_T}(K)$$

Together with our previous result,

$$\frac{\partial^2 C(K,T)}{\partial K^2} = q(K) = e^{-rT} f^{\mathbb{Q}}_{S_T}(K)$$

[7] We define $I\{S_T - K \geq 0\}$ as a function that is equal to one if $S_T - K \geq 0$ and zero otherwise. Note that we can take the partial derivative of $C(K,T)$ with respect to Ksince $\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)^+\right] = \int_{\mathbb{R}} f^{\mathbb{Q}}_{S_T}(x)(x-K)^+ dx = \int_K^\infty f^{\mathbb{Q}}_{S_T}(x)(x-K)\,dx$ is smooth in K.

From the findamental theorem of finance, it follows that there exists a (positive) equivalent martingale measure density function $f_{S_T}^{\mathbb{Q}}$ if and only if no arbitrage holds.

Note that this implies that we can evaluate *any* future payoff $h(S_T)$ in 3 ways:

$$p = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[h\left(S_T\right)\right] = e^{-rT}\int_{\mathbb{R}} f_{S_T}^{\mathbb{Q}}\left(x\right)h\left(x\right)dx = \int_{\mathbb{R}} \frac{\partial^2 C\left(K,T\right)}{\partial K^2}\left(x\right)h\left(x\right)dx = \int_{\mathbb{R}} q(x)h\left(x\right)dx$$

Where the last term is the continuous-states equivalent of the formula $p = Xq$.

This means the following -*apparently unrelated*- conditions are equivalent:

- Absence of arbitrage
- Existence of a positive equivalent martingale measure density function
- Convexity of options

Note that in the market we only directly observe the option prices for some discrete and finite set of strikes: but if all the option prices $C(K,T)$ for the strikes $K \in \{0, \Delta, 2\Delta, \ldots, N\Delta\}$ are observable (for N large and $\Delta > 0$ small enough) we can approximate $\frac{\partial^2 C(K,T)}{\partial K^2}$ in the following way

$$\frac{\partial^2 C\left(K,T\right)}{\partial K^2} \approx \frac{C\left(K+\Delta,T\right) - 2C\left(K,T\right) + C\left(K-\Delta,T\right)}{\Delta^2}$$

and therefore we can also calculate the empirical market-implied probability distribution of S_T , $f_{S_T}^{\mathbb{Q}}$, and the empirical state price density q.

Going back to our finite-state one-period model, if all the option prices $C(s,T)$ for the strikes $s \in \{0, 1, 2, \ldots, S\}$ are observable, we have

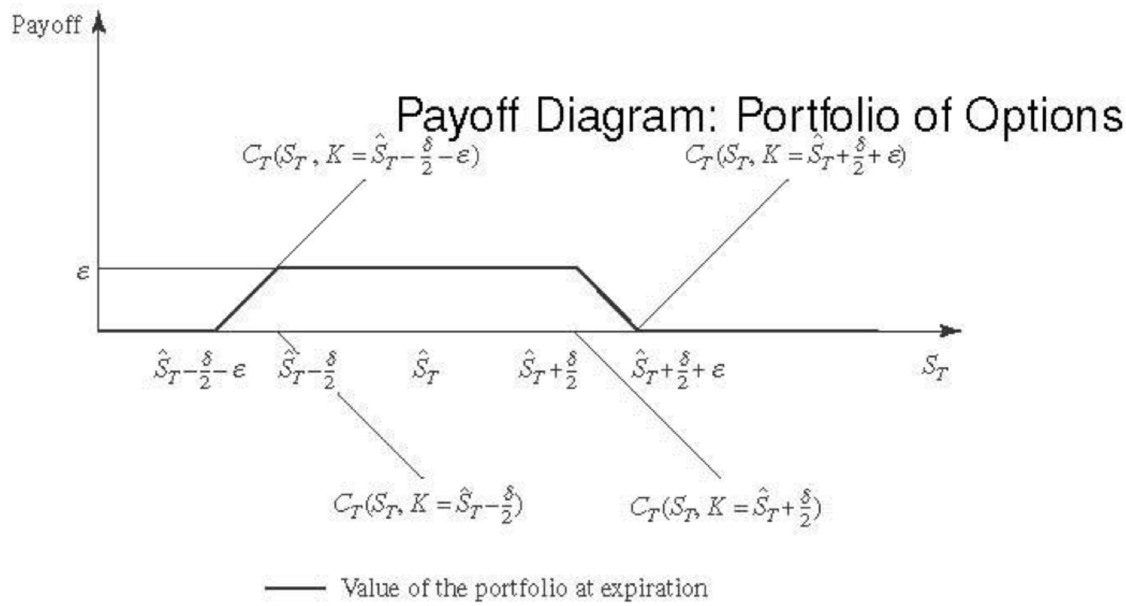
$$\triangle^2 C\left(s,T\right) \equiv C\left(s+1,T\right) - 2C\left(s,T\right) + C\left(s-1,T\right) = q_s = \frac{\hat{\pi}_s}{\left(1+r^f\right)}$$

Appendix

The following is an alternative derivation from option prices of the state price density function. We can construct the following portfolio: for some $\varepsilon > 0$ and $\delta > 0$ and a fixed $\hat{S}_T$

- Buy one call with strike $\hat{S}_T - \frac{\delta}{2} - \varepsilon$
- Sell one call with $\hat{S}_T - \frac{\delta}{2}$
- Sell one call with $\hat{S}_T + \frac{\delta}{2}$

- Buy one call with $\hat{S}_T + \frac{\delta}{2} + \varepsilon$



If we buy $\frac{1}{\varepsilon}$ units of this portfolio, when $S_T \in \left[\hat{S}_T - \frac{\delta}{2}, \hat{S}_T + \frac{\delta}{2}\right]$ the total payoff is equal to 1. The total value of this portfolio is

$$\frac{1}{\varepsilon}\left[C\left(K = \hat{S}_T - \frac{\delta}{2} - \varepsilon, T\right) - C\left(\hat{S}_T - \frac{\delta}{2}\right) - C\left(\hat{S}_T + \frac{\delta}{2}\right) + C\left(\hat{S}_T + \frac{\delta}{2} + \varepsilon\right)\right]$$

Letting $\varepsilon \to 0$ this boils down to

$$-\frac{\partial C\left(K = \hat{S}_T - \frac{\delta}{2}, T\right)}{\partial K} + \frac{\partial C\left(K = \hat{S}_T + \frac{\delta}{2}, T\right)}{\partial K}$$

Finally, dividing by δ and letting $\delta \to 0$, we obtain the continuum-states version of a vector of state prices, the state price density function $\frac{\partial^2 C(K = \hat{S}_T, T)}{\partial K^2}$.

Suppose we want to evaluate a one-year "wedding cake option" of the type

$$Payoff = \begin{cases} \$1,000,000 & if\ S_T \in [1700, 1750] \\ \$0 & otherwise \end{cases}$$

We now have the technology to price this. Its value will be equal to the integral of the state price density over the interval $[1700, 1750]$, that is, for $T = 1$,

$$\int_{1700}^{1750} \frac{\partial^2 C(K,T)}{\partial K^2} dK = \frac{\partial C\left(K = 1750, 1\right)}{\partial K} - \frac{\partial C\left(K = 1700, 1\right)}{\partial K}$$

Finally, note that this is equivalent to a portfolio comprising a long position in a binary 1700 call and a short position in a 1750 binary call.

Exercises

1) Determine whether the following statements are true or false. Provide a proof or a counter-example.

1. Law of one price and complete markets imply no strong arbitrage.
2. Law of one price and complete markets imply no arbitrage.
3. No strong arbitrage and complete markets imply no arbitrage.

2) Suppose there exist 3 states of the world $s = 1, 2$ and 2 assets x^1, x^2.

1. Suppose $x^1 = (2, 1, 0)'$ and $x^2 = (0, 1, 0)'$. Describe the asset span. Are markets complete?
2. Suppose $p_1 = 4$ and $p_2 = 3$. What type of no-arbitrage requirements does this market satisfy?
3. What are the restrictions on p_1 and p_2 such that this market satisfies LOOP, NSA and NA? (Write each restriction separately)
4. Repeat 1), 2) and 3) for $x^1 = (1, 1, 0)'$ and $x^2 = (0, 2, 0)'$.
5. Repeat 1), 2) and 3) for $x^1 = (1, 1, 0)'$, $x^2 = (0, 2, 0)'$ and $x^3 = (0, 1, 1)'$.

3) Suppose a stock index is currently trading at \$300, the dividend yield on the index is 3% per annum, and the risk-free interest rate is 8% per annum. What is the lower bound for the price of a 6-month European call option on the index when the strike price is \$290?

Now assume a stock currently sells for \$32. A 6-month call option with a strike of \$30 has a premium of \$4.29, and a 6 month put with the same strike has a premium of \$2.64. Assume a 4% continuously compounded risk-free rate. What is the present value of dividends payable over the next 6 months?

Finally, suppose a stock is priced at \$23 per share. The interest rate is 7% per annum and the stock pays no dividend. A three-month European call option with a strike price of \$30 has a price of \$0.3 What is the value of a European put with the same underlying asset, same strike price and same time to expiration?

4) Suppose there are 3 call options traded on a stock with strike prices equal to 40, 50 and 60 and with prices $C(40) = 8$, $C(50) = 6$ $C(60) = 2$.

1. Show that prices allow for arbitrage and provide an arbitrage portfolio with initial price equal to zero. What kind of strategy is this?
2. Is it possible to exploit the arbitrage with a symmetric butterfly spread with zero initial price?

5) Suppose there exist 3 states of the world $s = 1, 2, 3$ and 2 assets $x^1 = (2, 1, 0)'$ and $x^2 = (0, 1, 0)'$.

1. Suppose $p_1 = 1$ and $p_2 = 0.3$. What state prices are consistent with these prices?
2. Solve for the unique pricing kernel q^*.
3. Use the pricing kernel to value a third asset $x^3 = (0, 1, 1)'$. What other state prices (different from q^*) are consistent with no arbitrage?
4. Now suppose $p_3 = 0.6$. Solve for the state price vector. Does this market permit arbitrage?
5. Solve for the stochastic discount factor assuming the physical probability is such that $\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$.
6. Solve for the distribution under the equivalent martingale measure.

6) Suppose a stock index is currently trading at \$25, and there are 5 possible states of the world in $t = T$ such that $S_T \in \{15, 20, 25, 30, 35\}$.

1. Given a zero risk-free interest rate, describe a valid equivalent martingale measure.
2. Under this measure, price call options at $K = 15, 20, 25, 30, 35$.
3. Use this information to recover state prices.

7) Suppose there are S possible states of the world in $t = T$ and each has a (physical) probability of occurrence $\eta_s > 0$ with $\sum_{s=1}^{S} \eta_s = 1$. Consider the vector $\mu \in \mathbb{R}^S$ with for $s = 1, \ldots, S$ $\mu_s = \frac{q_s}{\eta_s}$, where q is a state-price vector. Write $\mathbb{E}(y) = \sum_{s=1}^{S} \eta_s y_s$ for any $y \in \mathbb{R}^S$.

1. Consider an asset with payoff $x = (x_1, \ldots, x_S)$. Show that the price of this asset must be $\mathbb{E}(z)$, where $z_s = \mu_s x_s$. Interpret this result.
2. Let the rate of return for an asset with price $p > 0$ in state s be $r_s = \frac{x_s}{p}$ and let $w_s = r_s \mu_s$. Show that $\mathbb{E}(w) = 1$. Is there some function of the excess return $r - r^f$ of the asset such that $\mathbb{E}\left[f(r - r^f)\right] = 0$?

8) Show in detail how to retrieve state prices using put options both in a continuous and discrete states setup.

Chapter 4

Risk Preferences and Expected Utility Theory

So far in this course we dealt with relative asset pricing: we derived asset prices through information available on other, pre-existing assets. In this chapter we will study absolute asset pricing, a task for which we have to specify the agents' preferences towards *risk*. We will see that risk involves knowledge of the probabilities of uncertain events, and agents' attitude towards risk determines how much they would be willing to pay for an asset - that is the core of absolute asset pricing. A concept related to risk that we will not see in detail is that of *uncertainty*, which relates to the agents' preferences over events whose probability distribution is unknown.

It is important to stress that preferences over risk refer and are defined over "final payoff gambles", that is, after combining any random payoffs, such as the payoff of securities, investments, insurance, and even non-random endowments.

We will start this chapter trying to answer the follwoing question: is there a criterion that allows us to compare random payoffs?

4.1 State-by-State Dominance

A first attempt to answer this question is given by the principle of *state-by-state dominance.* Suppose we have to choose one of the following three investments:

	$t = 0$	$t = 1$	
	Initial Cost	Probabilities $\pi_1 = \pi_2 = \frac{1}{2}$	
		$s = 1$	$s = 2$
Investment 1	-1000	1050	1200
Investment 2	-1000	500	1600
Investment 3	-1000	1050	1600

It is hard to tell which one is better between investment 1 and 2: the two possible states they have the same probability and in state 1 investment 1 pays more, while in state 2 investment 2 pays more. However, investment 3 pays at least as much as investment 1 and 2 in any state of the world! Clearly from this perspective investment 3 is more desirable, and we say that it state-by-state dominates investments 1 and 2.

> **Definition (State-by-State Dominance)**: given two random variables X and Y defined over the state space $(\Omega, \mathcal{F}, P)$, we say that Y state-by-state dominates X if $\forall \omega \in \Omega \; X(\omega) \leq Y(\omega)$.

We have already developed the notation needed to express state-by-state dominance: recall that for $x, y \in \mathbb{R}^n$ we write

1. $y \geq x$ if for each $i = 1, \ldots, n$ $y_i \geq x_i$
2. $y > x$ if $y \geq x$ and $y \neq x$
3. $y \gg x$ if for each $i = 1, \ldots, n$ $y_i > x_i$

Assuming that x and y are scaled to have the same price, under any of the three conditions above we would say that y state-by-state dominates x (you will notice that in the list above we see an "increasing degree" of dominance).

A problem with state-by-state dominance is that it is an incomplete ranking: in fact, it is as incomplete as $\mathbb{R}^S$. Another popular criterion which has been widely used is the *mean-variance dominance*: assume that we like more expected return and less volatility. Consider the following example:

	$t = 0$	$t = 1$			
	Initial Cost	Probabilities $\pi_1 = \pi_2 = \frac{1}{2}$			
		$s = 1$	$s = 2$	$\mathbb{E}[R]$	$\sigma(R)$
Investment 1	-1000	+5%	+20%	+12.5%	7.5%
Investment 2	-1000	-50%	+60%	+5%	55%
Investment 3	-1000	+5%	+60%	32.5%	27.5%

Investment 1 has a higher expected return and lower volatility than investment 2. However, it is not the case that investment 1 state-by-state dominates investment 2! Moreover, while investment 3 state-by-state dominates both investment 1 and 2, it only mean-variance dominates investment 2. Evidently, mean-variance dominance and state-by-state dominance are somehow unrelated orderings. However, while mean-variance dominance is an incomplete ranking, it does simplify the problem of ranking investments to two dimensions only (it is as incomplete as $\mathbb{R}^2$). It looks like, potentially, a way of getting round the incompleteness problem is to find a criterion that amounts to some number, for instance, in $\mathbb{R}$, an ordered set. For example, suppose we rank investments according to which *Sharpe Ratio* is highest:

$$\textit{Sharpe Ratio} = \frac{\mathbb{E}[R] - r^f}{\sigma(R)}$$

In the example above, assuming $r^f = 3\%$, we get 1.27, 0.04 and 1.07 for investment 1, 2 and 3 respectively, and therefore we would select investment 1 (which, remember, we would eliminate based on both state-by-state dominance and mean-variance dominance!)

Before going on, it is worthwile to remark that we can write the "payoff per state" representation of random payoffs:

State	s_1	s_2	s_3	s_4	s_5
Probability	π_1	π_2	π_3	π_4	π_5
Payoff X	10	10	20	20	20
Payoff Y	10	20	20	20	30

in terms of "probability lotteries":

Payoff	10	20	30
Probability X	$p_{10} = \pi_1 + \pi_2$	$p_{20} = \pi_3 + \pi_4 + \pi_5$	$p_{30} = 0$
Probability Y	$q_{10} = \pi_1$	$p_{20} = \pi_2 + \pi_3 + \pi_4$	$p_{30} = \pi_5$

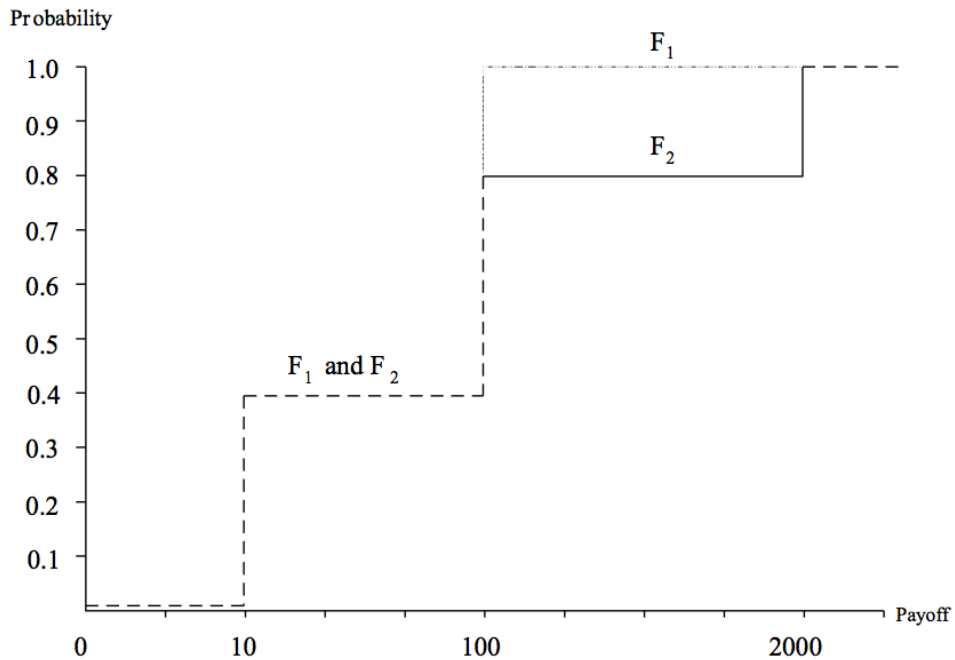
This allows us to define preferences $p_x \succ p_y$ over probability distributions over states (rather than on payoffs $x \succ y \in \mathbb{R}^S$) and will turn out to be useful when thinking of *final payoffs*. Note that we can always go from the"payoff per state" to the "probability lotteries" representation, but the reverse is not true.

4.2 Stochastic Dominance

Another criterion to rank investment is the *stochastic dominance.* It is related to state-by-state dominance, and is still an incomplete ordering of investments. Suppose there are two investment opportunities 1 and 2:

Event	e_1	e_2	e_3
Payoff	10	100	2000
Probability 1	40%	60%	0%
Probability 2	40%	40%	20%

In this case, every investment pays the same amount in each event, but the probabilities of the events happening vary. To reconcile this with our previous "states of the world" discussion, it is enough to note that the above example is such that investment 1 pays off 100 in the states s_2 and s_3 so that $e_2 = \{s_2, s_3\}$, while investment 2 pays off 100 in state 2 and 2000 in state 3 (and obviously $e_1 = s_1$).



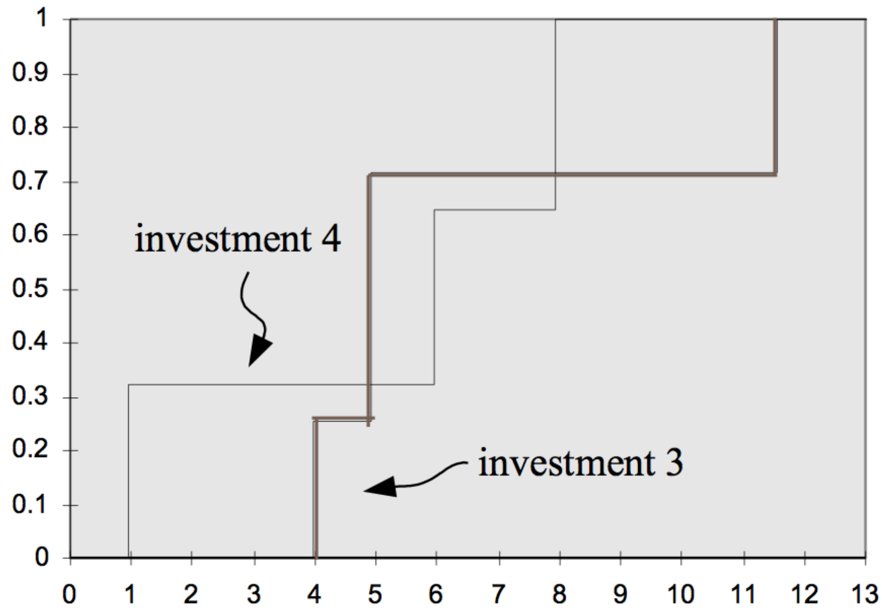
Definition (First Order Stochastic Dominance): let F_A and F_B represent, respectively, the cumulative distribution functions of two random variables (investments payoff) defined in the interval $[a, b]$. We say that F_A first-order stochastically dominates (FSD) F_B if $\forall x \in [a, b]$ $F_A(x) \leq F_B(x)$.

Visually, FSD occurs whenever F_A "stays below" F_B for the whole domain. State-by-State Dominance implies First Order Stochastic dominance.

Next, consider the following example:

Payoff	1	4	5	6	8	12
Probability 3	0%	25%	50%	0%	0%	25%
Probability 4	33%	0%	0%	33%	33%	0%

It is easy to verify that FSD does not hold for any investment. However, investment 3 is somewhat special because its CDF stays below investment 4's "most of the time".

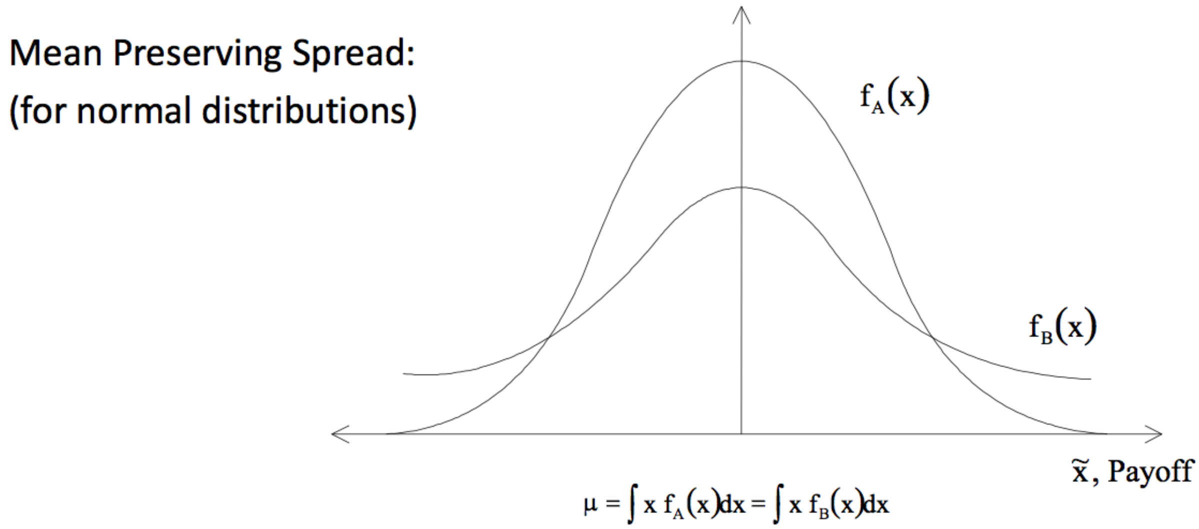


Definition (Second Order Stochastic Dominance): let F_A and F_B represent, respectively, the cumulative distribution functions of two random variables (investments payoff) defined in the interval $[a, b]$. We say that F_A second-order stochastically dominates (SSD) F_B if $\forall x \in [a, b]$ $\int_a^x [F_B(t) - F_A(t)]\, dt \geq 0$.

First Order Stochastic Dominance implies Second Order Stochastic Dominance. Clearly, in our example, investment 3 SSD investment 4. Another related concept is that of *mean-preserving spread*:

Definition (Mean-Preserving Spread): we say that the random variable x_A is a mean-preserving spread of the random variable x_B if $x_A = x_B + \varepsilon$, where the random variable ε is independent of x_A and x_B, and has zero mean and positive variance.

For normally distributed x_A, x_B and ε the picture looks like:



The concept of mean preserving spread and that of second-order stochastic dominance are connected by the following proposition:

> **Proposition**: Let F_A and F_B be the CDFs of two random variables x_A and x_B defined on the same space with identical means. Then F_A SSD F_B if and only if x_A is a mean-preserving spread of x_B.

4.3 Von Neumann Morgenstern Expected Utility Theory

So far we mostly encountered criteria to rank uncertain payoffs which are incomplete, in the sense that we can encounter two investments that we are unable to compare and decide which one is better. We saw that this is linked to the fact that the criteria examined require to compare finite-dimensional vectors (in the case of state-by-state dominance and mean-variance dominance) or even functions, which can be considered as infinite-dimensional vectors (in the case of first and second-order stochastic dominance). In this section we will elaborate another method which aims at simplifying the problem by assigning to each investment a number $u \in \mathbb{R}$, so that every two investments are comparable. To do this by avoiding arbitrary choices of this function that maps investments onto numbers, we will assume the existence of *regular preferences* over investments.

Let's start with an example. Suppose you can enter the following bet: we flip a fair coin once. If tails comes up, I'll give you \$1. If heads comes up, we'll flip again, and if tails comes up I'll pay you \$2, while if heads comes up we'll flip again. Every time we flip again, the prize for a tails up doubles. The question is: what is the fair price for this bet?

Since for the n-th flip the probability of a tails is $\left(\frac{1}{2}\right)^n$ and the payoff is 2^{n-1}, and the number of flips ranges from 1 to infinity, it is straightforward to see that the expected payoff is

$$Expected\ Payoff = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \times 2^{n-1} = \sum_{n=1}^{\infty} \frac{1}{2} = \infty$$

However this is the correct expected value, it does not make sense to pay all the money in the world for this coin-flipping game (investment), since it pays you less that \$64 in 99% of cases. Clearly this criterion is inappropriate to value uncertain payoffs. This problem is was named the St. Petersburg paradox and was first proposed by Daniel Bernoulli, who proposed the following solution: because when you are rich an additional dollar is less valuable to you than when you are poor, we should weight less the largest payoffs. He proposed to discount large payoffs using the natural logarithm: introduce the parameter p for the denominator of the probability, in the previous example equal to 2. Then we can write:

$$Bernoulli\ Value = \sum_{n=1}^{\infty} \left(\frac{1}{p}\right)^n \times \ln\left(2^{n-1}\right) = \ln 2 \sum_{n=1}^{\infty} (n-1) p^{-n} =$$

$$1 = -\ln 2 \sum_{n=1}^{\infty} (1-n) p^{-n} = -\ln 2 \sum_{n=1}^{\infty} \frac{d}{dp}\left[p^{(1-n)}\right] = -\ln 2 \frac{d}{dp}\left[\sum_{n=1}^{\infty} p^{(1-n)}\right] =$$

$$= -\ln 2 \frac{d}{dp}\left[\sum_{n=0}^{\infty} \frac{1}{p^n}\right] = -\ln 2 \frac{d}{dp}\left[\frac{1}{1-p}\right] = \ln 2 \frac{1}{(1-p)^2}$$

Which for $p = 2$ is just $\ln 2$. If the value is $\ln 2$ and the function used to discount large payoffs is the natural logarithm, we deduce that the price we are willing to pay for the coin-flipping game is just x such that $\ln x = \ln 2$, that is, \$2. Let's dig deeper into this idea of specifying a function that maps uncertain payoffs (investments) to a single number (utility).

4.4 Representation of Preferences

Suppose all we care about is consumption in each state. Then, the choice set would consist of vectors in $\mathbb{R}^{S+1}$ where the first element represents consumption at $t = 0$, and all other elements represent consumption in the possible S states at $t = T$. To reconcile this definition with our one-period model, suppose an investor has an initial wealth W (constant across states and time) and is considering buying a security j with price p_j and payoff x^j. The vector $\left(W - p_j, (W - p_j) \times I + x^j\right) = (c_0, c_1, ... c_S)^T \in \mathbb{R}^{S+1}$ (where I is a $S \times 1$ vector where each element is equal to one) would be the corresponding consumption profile. Moreover, given two consumption profiles $c, c' \in \mathbb{R}^{S+1}$, if an investor chose c over c' we say that that investor *prefers* c to c' and write $c \succ c'$.

Representation Theorem: Suppose the preference ordering is i) complete ii) transitive iii) continuous (and satisfies some further regularity conditions that we do not discuss here). Then the preference ordering can be represented by a utility function, that is, $c \succ c'$ if and only if there exists a utility function U such that $U(c) > U(c')$.

A related but different approach was developed by John Von Neumann and Oskar Morgenstern. In this setup, we define the (finite) set of outcomes X (in the previous setup for instance, this was the uncertain consumption profile at time $t = T$) and we define a *lottery* as a probability measure over X, that is a function $p : X \to [0,1]$ such that 1) for $x \in X$ $p(x) \geq 0$ and 2) $\sum_{x \in X} p(x) = 1$. The set of all possible lotteries is

$$P \equiv \left\{ p : X \to [0,1] \,|\forall x \in X \; p(x) \geq 0, \; \sum_{x \in X} p(x) = 1 \right\}$$

Assume there exists a preference relation $\succ$ over P and consider the following three axioms:

Von Neumann - Morgenstern Axioms

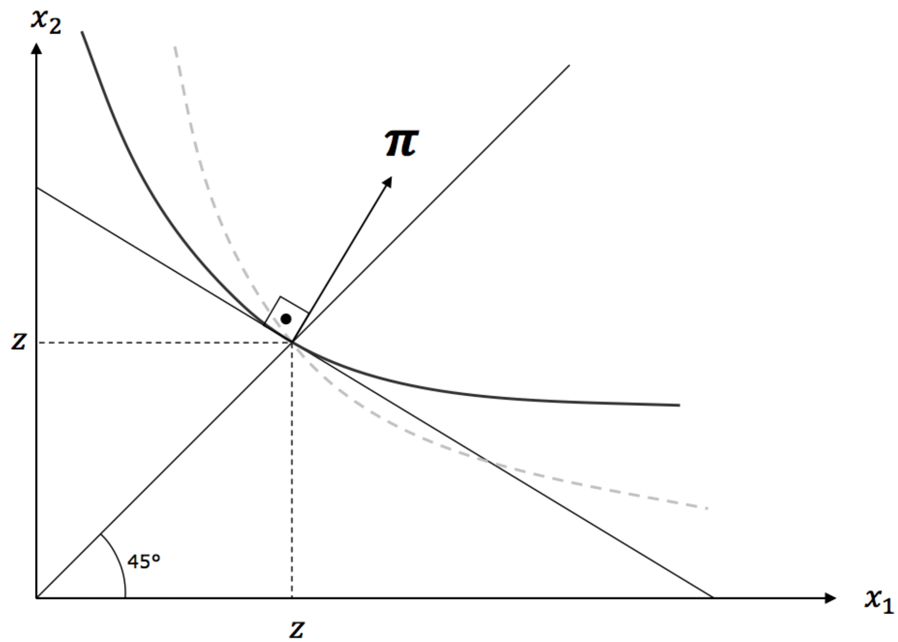
1. (*Regularity*) Agents have complete and transitive preferences over P
2. (*Independence*) For any three lotteries p, q, $r \in P$ and $\alpha \in (0,1]$, $p \succeq q \Leftrightarrow \alpha p + (1-\alpha) r \succeq \alpha q + (1-\alpha) r$
3. (*Continuity*) For any three lotteries p, q, $r \in P$ such that $p \succ q \succ r$ there exist $\alpha, \beta \in (0,1)$ such that $\alpha p + (1-\alpha) r \succ q \succ \beta p + (1-\beta) r$

Then the following theorem holds:

Theorem (Expected Utility Representation). If a preference relation $\succ$ over P satisfies the Regularity, Independence and Continuity axioms, then there exists a function $u : X \to \mathbb{R}$ such that

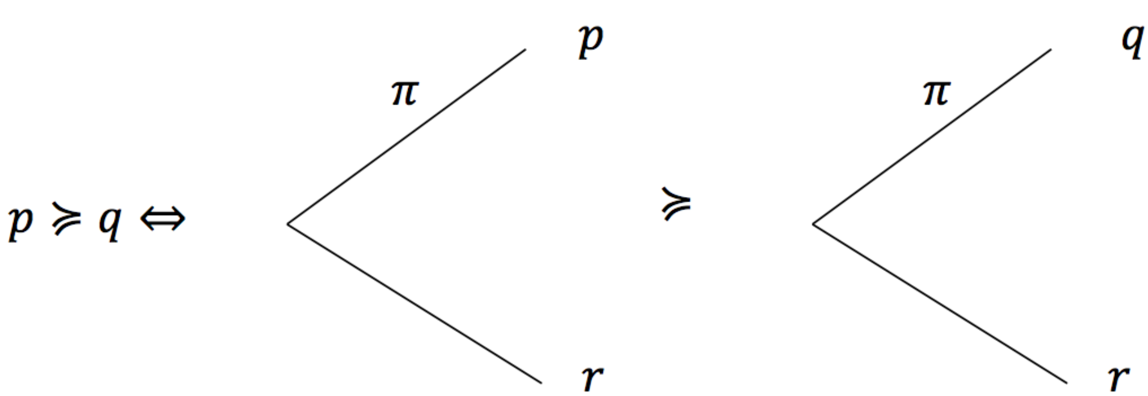
$$p \succeq q \Leftrightarrow \sum_{x \in X} p(x) u(x) \geq \sum_{x \in X} q(x) u(x)$$

That is, preferences over lotteries correspond to the *expected utility* of the lotteries. Below is a graphical representation of an expected utility function over a lottery defined on a two-dimensional state space:



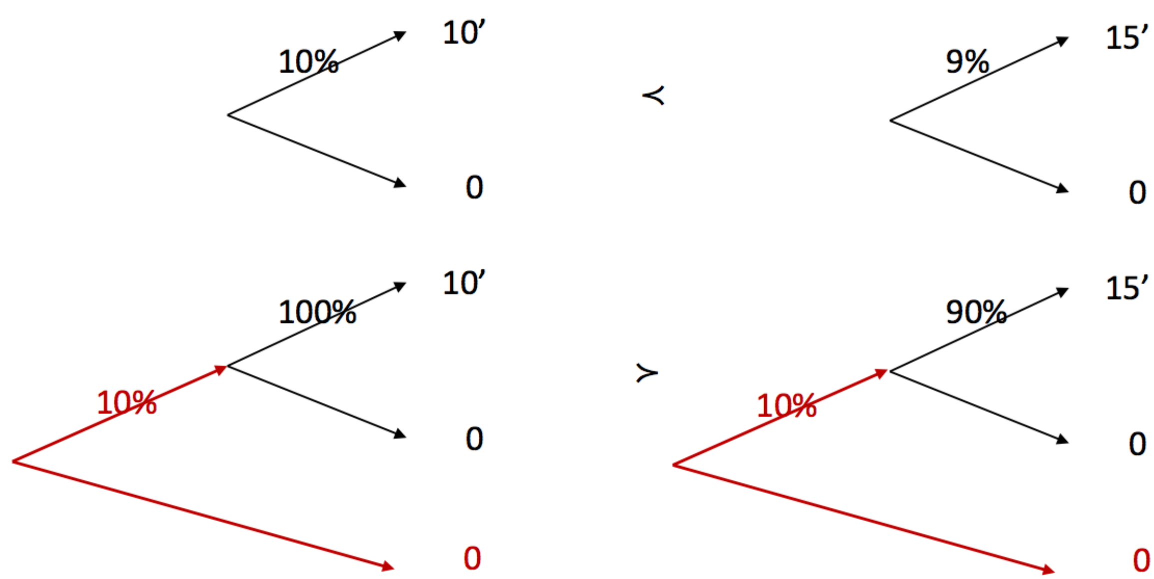
The 45° line represents the set of lotteries that pay off the same amount in both states. The negatively sloped straight line is the set of lotteries in the (x_1, x_2) plane such that the expectation of the lottery is equal to a value $c > 0$, that is, if π is the probability of the state 1 occurring, $c = \pi x_1 + (1 - \pi) x_2$, or $x_2 = \frac{\pi}{1-\pi} x_1 + \frac{c}{1-\pi}$. The convex curve represents the indifference curve for the agent: to give up a certain consumption amount in state 2, he is willing to take some amount in state 1 to keep his utility level constant. The curve is the locus of lotteries in the (x_1, x_2) plane such that $k = \pi u(x_1) + (1 - \pi) u(x_2)$, $k \in \mathbb{R}$.

Going back to the axioms, consider the Independence axiom: it basically states that if we "dilute" p and q with some third lottery r in the same way, but we prefer p to q, then our preference ranking should not change. It can be visualized using trees:



However, this axiom rarely holds in experiments. For instance, suppose you are given the choice between lottery p that pays \$10 with probability 10% (and zero otherwise) and lottery q that pays

\$15 with probability 9% (and zero otherwise). Because the prize is 50% higher in lottery q than in lottery p and the probability of winning a positive amount is only 10% lower in lottery q than lottery p, most people in experiments choose lottery q. However, between a lottery u that pays \$10 with 100% probability and a lottery v that pays \$15 with 90% probability, most people choose the sure bet implied by lottery u. However, it turns out that lotteries u and v can be mixed with a third lottery r that pays zero with probability 100%, using a parameter $\alpha = 10\%$, to obtain lotteries p and q again(see picture below): the preference relation between p and q is now the opposite of the initial preference, and therefore this constitutes an empirical violation of the independence axiom.[1]

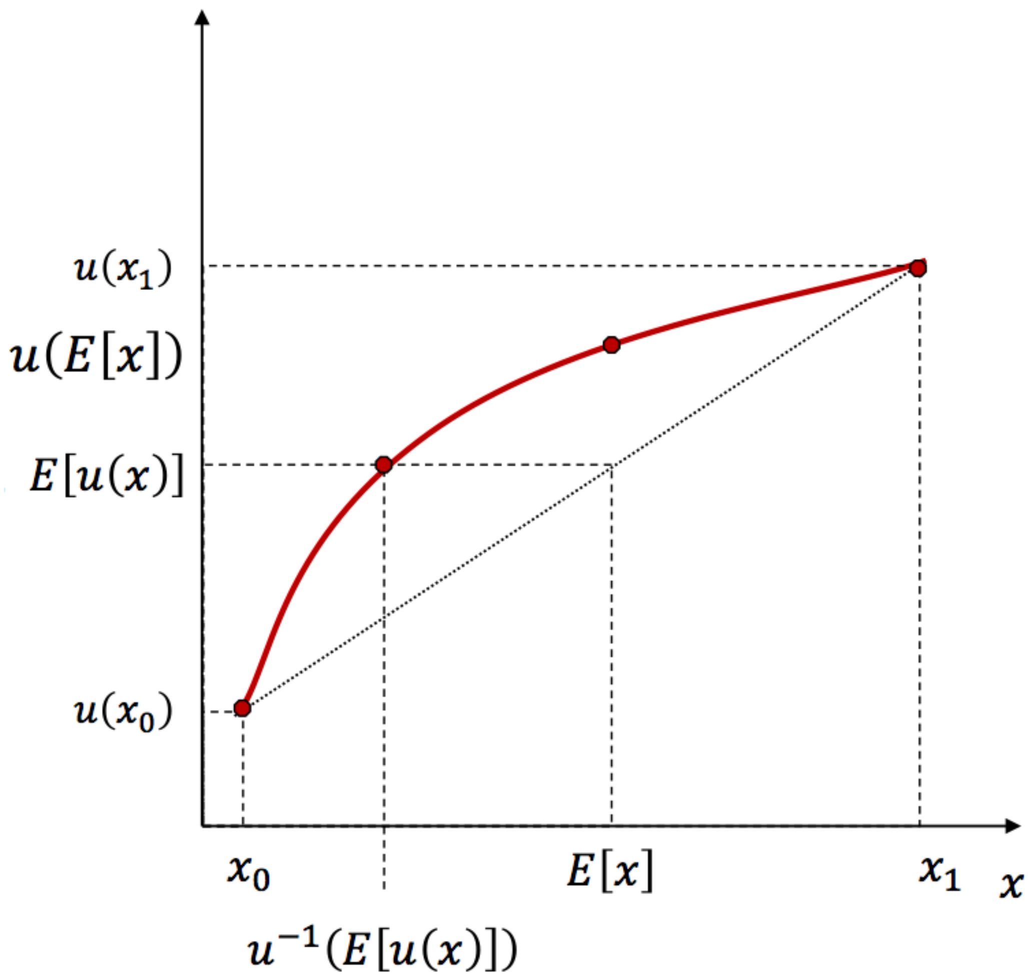


4.5 Risk Aversion, Concavity, Certainty Equivalent

Consider a lottery p with possible outcomes $x_0 < x_1$. We define the *Certainty Equivalent* as the certain payoff which gives the same expected utility as the uncertain lottery p. Clearly, if $\mathbb{E}[u(x)]$ is the expected utility of lottery p, $u^{-1}(\mathbb{E}[u(x)])$ will be the certainty equivalent of p. Now suppose that the agent is risk-averse: clearly, for him the certainty equivalent will be *lower* than the expected value of the lottery (he will be willing to "give up" some money off the expected value of the lottery to have a certain payoff). As it turns out, this property is equivalent to the Von Neumann-Morgenstern function u being *concave*.

$$u^{-1}(\mathbb{E}[u(x)]) < \mathbb{E}[x] \Leftrightarrow u(\cdot) \textit{ is concave}$$

[1] Top left: lottery p. Top right: lottery q. Bottom left, in black: lottery u. Bottom right, in black: lottery v. In red: with $\alpha = 10\%$ and $r = 0$ we can transform u and v into p and q.



And the difference $\mathbb{E}[x] - u^{-1}(\mathbb{E}[u(x)])$ is called *Risk Premium.* This fact derives directly from *Jensen's Inequality:*

Theorem (Jensen's Inequality): Let g be concave over $[a, b]$ and let x be a random variable such that $P[x \in [a, b]] = 1$. If the expectations $\mathbb{E}[x]$ and $\mathbb{E}[g(x)]$ exist, then $\mathbb{E}[g(x)] \leq g(\mathbb{E}[x])$. Furthermore, if $g(\cdot)$ is concave then the inequality is strict.

The following theorems establish a relationship between First and Second Order Stochastic Dominance and Expected Utility:

Theorem 1: Let F_A and F_B be two CDFs for the random payoffs $x \in [a, b]$. Then F_A FSD F_B if and only if $\mathbb{E}_A[U(x)] \geq \mathbb{E}_B[U(x)]$ for all *nondecreasing* utility functions $U(\cdot)$.

Theorem 2: Let F_A and F_B be two CDFs for the random payoffs $x \in [a, b]$. Then F_A SSD F_B if and only if $\mathbb{E}_A[U(x)] \geq \mathbb{E}_B[U(x)]$ for all *nondecreasing and concave* utility functions $U(\cdot)$.

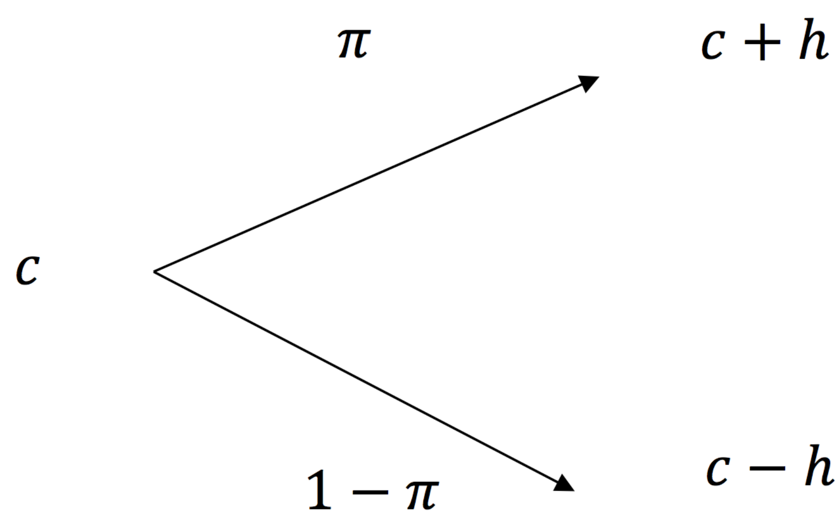
We conclude this subsetion with a parallel between (non-expected) utility representation and expected utility representation. In the former case, suppose $U(c) = U(c_0, c_1, ... c_S)$ represents a complete, transitive and continuous preference ordering between consumption profiles. Then also $V(c) = f(U(c))$, for f strictly increasing, represents the same preference ordering. In the latter case, suppose that $\mathbb{E}[u(c)]$ represents a preference ordering satisfying the Von Neumann-Morgenstern axioms. Then for $a, b \in \mathbb{R}$ the affine function $v(c) = a + bu(c)$ represents the same preference ordering.

4.6 Measures of Risk Aversion

We define the following measures of risk aversion defined over the wealth Y of the agent:

1. Absolute Risk Aversion: $R_A(Y) = -\frac{u''(Y)}{u'(Y)}$
2. Relative Risk Aversion: $R_R(Y) = -Y \cdot \frac{u''(Y)}{u'(Y)} = Y \cdot R_A$
3. Risk Tolerance: $R_T(Y) = \frac{1}{R_A(Y)}$

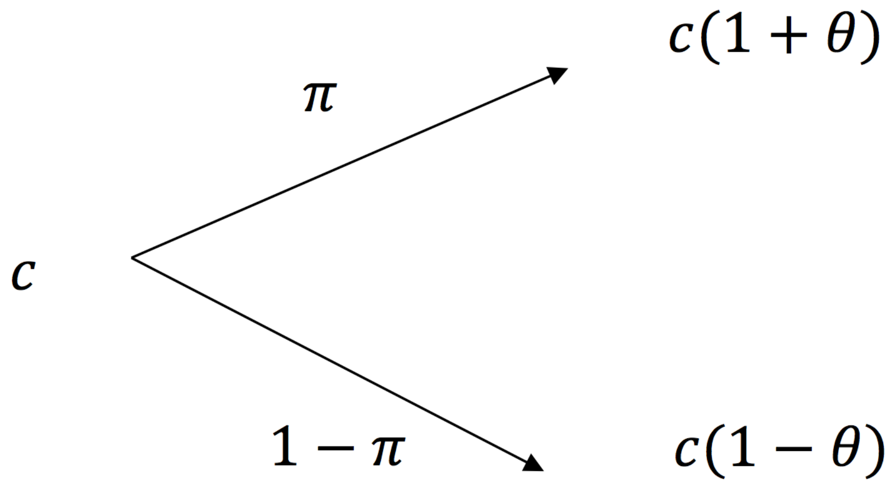
There is an interesting link between expected utility and these measures of risk aversion: consider the following lottery, where Y is the initial wealth of the agent:



What is the change in π needed for an agent to be indifferent when h rises? We can Taylor-expand to the second degree the expected utility in a neighbourhood of Y to obtain:

$$\pi(Y, h) = \frac{1}{2} + \frac{1}{4} h R_A(Y) + \textit{higher order terms}$$

Similarly, consider the lottery



Again, the change in π needed for an agent to be indifferent when h rises is obtained by Taylor-expanding to the second degree the expected utility in a neighbourhood of Y to obtain:

$$\pi(Y,h) = \frac{1}{2} + \frac{1}{4}\theta R_R(Y) + \textit{higher order terms}$$

Similarly, suppose we add a small risk (lottery) x to the initial lottery w. Since $\mathbb{E}\left[u(w+x)\right] = u\left(c^{CE}\right)$, we can use a second order Taylor expansion to show that

$$w - c^{CE} \approx R_A(w)\frac{Var(x)}{2}$$

Where $w - c^{CE}$ is the risk premium, and it is approximately linear in the variance of the additive risk, with slope equal to half the coefficient of absolute risk aversion. Similarly, suppose we have a *multiplicative* risk. Suppose that $\mathbb{E}\left[u(gw)\right] = u\left(kw\right)$, where g is a positive random variable with unit mean and k is is the certainty equivalent growth rate. This time we get

$$1 - k \approx R_R(w)\frac{Var(g)}{2}$$

Therefore the coefficient of *absolute* risk aversion is relevant for additive risk, while the *relative* risk aversion is relevant for multiplicative risk.

We will now introduce two specific functional forms for the Von Neumann-Morgenstern function U:

- Constant Absolute Risk Aversion (CARA) utility function: $U(x) = -e^{-\rho Y}$, for some $\rho \in \mathbb{R}_+$
- Constant Relative Risk Aversion (CRRA) utility function: $U(x) = \begin{cases} \frac{Y^{1-\gamma}}{1-\gamma} & if\, \gamma \neq 1 \\ \ln Y & if\, \gamma = 1 \end{cases}$

For example, consider the CRRA utility and a lottery that adds \$50,000 or \$100,000 to the initial wealth Y. The certainty equivalent is defined as the number $CE \in \mathbb{R}$ such that

$$\frac{(Y+CE)^{1-\gamma}}{1-\gamma}=\frac{1}{2}\times\frac{(Y+50,000)^{1-\gamma}}{1-\gamma}+\frac{1}{2}\times\frac{(Y+100,000)^{1-\gamma}}{1-\gamma}$$

For $Y=0$ we get

$$\begin{array}{cc}
\gamma=0 & CE=75,000 \text{ (risk neutrality)}\\
\gamma=1 & CE=70,711\\
\gamma=2 & CE=66,246\\
\gamma=5 & CE=58,566\\
\gamma=10 & CE=53,991\\
\gamma=20 & CE=51,858\\
\gamma=30 & CE=51,209
\end{array}$$

While for $Y=100,000$ and $\gamma=5$ the certainty equivalent is already 66,530 (close to $\gamma=2$ for $Y=0$!)

4.7 Risk Aversion and Portfolio Allocation

Suppose that all assets are consumed in $t=T$, so that there are no savings in $t=T$. Moreover, assume there is a riskless asset with net return r^f and a risky asset with a random net return r. We want to maximize the expected utility by choosing the risky asset allocation parameter $a\in\mathbb{R}$:

$$\max_{a\in\mathbb{R}} E\left[U\left(Y_0\left(1+r^f\right)+a\left(r-r^f\right)\right)\right]$$

For some initial wealth Y_0. The problem has first order conditions (FOC)

$$E\left[U'\left(Y_0\left(1+r^f\right)+a\left(r-r^f\right)\right)\left(r-r^f\right)\right]=0$$

We can characterize the solution to the problem with the following theorem:

Theorem: assume $U'>0$, $U''<0$ and let $\hat{a}$ denote the solution to the problem above. Then

$$\hat{a}>0\Leftrightarrow\mathbb{E}\left[r\right]>r^f$$

$$\hat{a}=0\Leftrightarrow\mathbb{E}\left[r\right]=r^f$$

$$\hat{a}<0\Leftrightarrow\mathbb{E}\left[r\right]<r^f$$

Proof: define $W(a)\equiv\mathbb{E}\left[U\left(Y_0\left(1+r^f\right)+a\left(r-r^f\right)\right)\right]$, then the FOC can br written as $W'(a)=0$. By risk aversion (that is, since $U'>0$ and $U''<0$) we have $W''(a)=\mathbb{E}\left[U''\left(Y_0\left(1+r^f\right)+a\left(r-r^f\right)\right)\left(r-r^f\right)^2\right]<0$, so $W'(a)$ is everywhere decreasing in a.

This implies that $\hat{a} > 0$ if and only if $W'(0) > 0$, and since $U' > 0$ it follows that $\hat{a} > 0$ if and only if $\mathbb{E}[r] > r^f$. The other assertions follow similarly.

How do measures of risk aversion depend on the initial wealth Y_0? This is the subject of Arrow's theorem (1971):

Theorem 1: Let $\hat{a} = \hat{a}(Y_0)$ be the solution to the problem above. Then

$$\frac{\partial R_A}{\partial Y_0} < 0 \Rightarrow \frac{\partial \hat{a}}{\partial Y_0} > 0$$

$$\frac{\partial R_A}{\partial Y_0} = 0 \Rightarrow \frac{\partial \hat{a}}{\partial Y_0} = 0$$

$$\frac{\partial R_A}{\partial Y_0} > 0 \Rightarrow \frac{\partial \hat{a}}{\partial Y_0} < 0$$

Theorem 2: If, for all wealth levels Y:

$$\frac{\partial R_R}{\partial Y_0} < 0 \Rightarrow \frac{da/a}{dY/Y} > 1$$

$$\frac{\partial R_R}{\partial Y_0} = 0 \Rightarrow \frac{da/a}{dY/Y} = 1$$

$$\frac{\partial R_R}{\partial Y_0} > 0 \Rightarrow \frac{da/a}{dY/Y} < 1$$

In the special case $U(Y) = \ln Y$ the FOC is

$$\mathbb{E}\left[\frac{r - r^f}{Y_0\left(1 + r^f\right) + a\left(r - r^f\right)}\right] = 0$$

and assuming that r can take on two possible values $r_1$and r_2 with $r_1 < f^f < r_2$, it is possible to show that

$$\frac{a}{Y_0} = \frac{\left(1 + r^f\right)\left(E[r] - r^f\right)}{\left(r_2 - r^f\right)\left(r^f - r_1\right)} > 0$$

That is, a constant fraction of the initial wealth is invested in the risky asset. How does all this generalize to the J assets case? The answer is provided by the following theorem:

Theorem (Cass and Stiglitz): Let the vector

$$\begin{bmatrix} \hat{a}_1(Y_0) \\ \vdots \\ \hat{a}_J(Y_0) \end{bmatrix}$$

denote the amount optimally invested in the J risky assets if the initial wealth is Y_0. Then

$$\begin{bmatrix} \hat{a}_1(Y_0) \\ \vdots \\ \hat{a}_J(Y_0) \end{bmatrix} = \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_J \end{bmatrix} \cdot f(Y_0)$$

if and only if either i) $U'(Y_0) = (\theta Y_0 + \kappa)^\Delta$ or ii) $U'(Y_0) = \xi e^{-vY_0}$.

In other words, it is sufficient to offer a *mutual fund*: the optimal proportion to invest in each asset is given by the vector $[\hat{a}_1, \ldots, \hat{a}_J]^T$, and agents would then invest an amount depending on their initial wealth level according to the function $f(Y_0)$.

To conclude this subsection, we generalize the class of utility functions by starting from the absolute risk aversion (and then backing out the resulting utility function).

We call *linear risk tolerance* (or *hyperbolic risk aversion*) any utility function such that

$$R_A = -\frac{u''(c)}{u'(c)} = \frac{1}{A + B \cdot c}$$

Clearly, for $B = 0$ and $A \neq 0$ we have a CARA utility function. If $B \neq 0$ we obtain a generalized power function

$$u(c) = \frac{1}{B-1} \left(A + B \cdot c\right)^{\frac{B-1}{B}}$$

For which if $B \to 1$ we obtain the log utility $u(c) = \ln(A + B \cdot c)$, for $B = -1$ we obtain the quadratic utility $-\frac{1}{2}(A - c)^2$, and for $A = 0$ we obtain the CRRA utility function $u(c) = \frac{1}{B-1}(B \cdot c)^{\frac{B-1}{B}}$.

4.8 Alternative Theories

There are a number of theories alternative to Von Neumann-Morgenstern that have been proposed. Savage's expected utility theory (1954) combines the idea of utility function with *subjective* probabilities, showing that under a set of axioms it is possible to derive subjective probabilities that agents attach to events starting from their preferences over lotteries. This is considered an extension of the Von Neumann-Morgenstern paradigm, but like in vN-M it still suffers from the fact that empirically its axioms are violated.

An interesting experiment that opened the door to *ambiguity theory* is Ellsberg's paradox. Suppose there are 10 balls in an urn and you have to make a choice between lottery 1 which pays you \$100 if you pick a blue ball, and lottery 2 which pays you \$100 if you pick a red ball, *without knowing* how many balls are red or blue. Which one would you choose? Unlike preferences over risk, in this case with *uncertainty* there are no probability distributions: it turns out that empirically people are also uncertainty averse.

Daniel Kahneman used insights from psychology to show that people first decide on a *reference point* to compare outcomes with, and then consider lesser outcomes as losses and greater outcomes as gains. In particular, he showed that empirically they are risk-averse in the gains domain, and risk-loving in the loss space.

4.9 Savings

Within our one period model, consider an asset stucture composed of one risk-free asset with constant gross return R^f. The agent has to make a decision on how much to consume in each period $t = 0$ and $t = T$. If R^f increases, we will see two effects: one is the incentive to save more at time $t = 0$, since by doing so we can now consume more in $t = T$ -this is called s*ubstitution effect*- and the incentive to consume more (save less) at time $t = 0$, since the agent is effectively richer when the gross riskless return increases -this is called *income effect*. In the log utility case, these two effects cancel each other; in other cases ($\gamma \neq 1$) one effect can prevail over the other.

Suppose now there is a risk-free assets but the endowment in the future period is random. How does the behavior of the agent change when we increase his exposure to risk? An old idea (going back to J. M. Keynes) is that people save more when they face greater uncertainty. This phoenomenon is called *precautionary savings* and it generally arises under two circumstances, one linked to the shape of the utility function and one to the presence of borrowing constraints.

Let's jump forward a little and suppose that in a multi-period setting agents maximize the expected utility

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t u\left(c_t\right)\right]$$

subject to the intertemporal budget constraint

$$c_{t+1} = e_{t+1} + (1+r)\left(e_t - c_t\right)$$

The standard Euler equation is

$$u'\left(c_t\right) = \beta\left(1+r\right)\mathbb{E}_t\left[u'\left(c_{t+1}\right)\right]$$

And if $u''' > 0$, Jensen's inequality implies

$$\frac{1}{\beta\left(1+r\right)} = \frac{\mathbb{E}_t\left[u'\left(c_{t+1}\right)\right]}{u'\left(c_t\right)} > \frac{u'\left(\mathbb{E}_t\left[c_{t+1}\right]\right)}{u'\left(c_t\right)}$$

Which shows that the *marginal rate of intertemporal substitution* is higher in the presence of uncertainty in c_{t+1}. The difference between the two marginal rates, with and without uncertainty, is attributed to precautionary savings. to see this, suppose the variance of e_{t+1} increases (in a mean preserving fashion). Since numerator $\mathbb{E}_t\left[u'\left(c_{t+1}\right)\right]$ is increasing in the variance of c_{t+1}, in order for the above equality to hold c_t must decrease, that is, savings increase due to precautionary savings.

Suppose now there are no risk-free assets and one risky asset with random gross return R.

> **Theorem (Rothschild and Stiglitz)**: Let R_A and R_B be two return distributions with identical means such that $R_B = R_A + \varepsilon$ where ε is a white noise, and let s_A and s_B be the savings out of Y_0 corresponding to the return distributions of $R_A$and R_B respectively. Then

- If $R_{R'}\left(Y\right) \leq 0$ and $R_R\left(Y\right) > 1$, then $s_A < s_B$
- If $R_{R'}\left(Y\right) > 0$ and $R_R\left(Y\right) \leq 1$, then $s_A > s_B$

Define *Absolute Prudence* as the quantity

$$P_A(w) = -\frac{u'''(w)}{u''(w)}$$

and Relative Prudence as

$$P_R(w) = -c \cdot \frac{u'''(w)}{u''(w)}$$

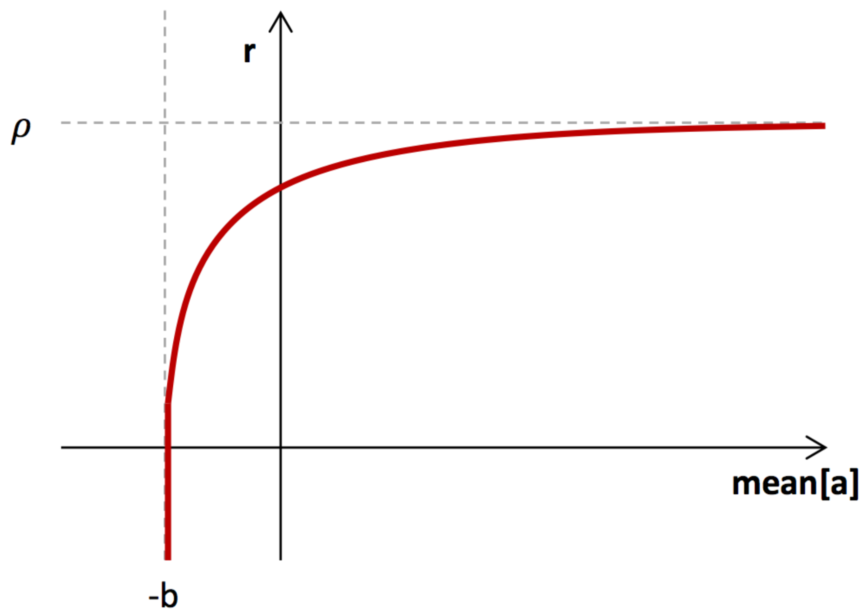
Note that this depends directly on the curvature of u' ($u''' > 0$ implies that u' is convex) and does not follow directly from risk aversion. It turns out that precautionary savings occur if $P(w) > 0$. Moreover, we have the following result:

> Theorem: Let R_A and R_B be two return distributions such that R_A SSDR_B, and let s_A and s_B (respectively) be the savings out Y_0. Then

$$s_A \geq s_B \Leftrightarrow cP(c) \leq 2$$

$$s_A < s_B \Leftrightarrow cP(c) > 2$$

Finally, agents might save precautionarily also because they are concerned that they will face borrowing constraints in some state in the future: Bewley (1977) showed that with idiosyncratic income shocks (borrowing constraints) negative mean asset holdings across individuals (mean[a] below) result from simple individual optimization:



4.10 Mean-Variance Preferences

A different approach is to specify the utility function over the mean-variance space of investment returns. This is a much simpler approach than the Von Neumann-Morgenstern, so it is natural to ask under which conditions the two approaches are equivalent. For instance, if all lotteries have normally distributed prizes (not necessarily independent), because any linear combination of jointly normal random variables is also jointly normal and therefore described by its first two moments, expected utility can be expressed as a function of just mean and variance as well.

Another example in which mean variance and vN-M are equivalent is when agents have quadratic utility:

$$u(y) = ay - by^2$$

Then the expected utility is

$$\mathbb{E}\left[u(y)\right] = a\mathbb{E}\left[y\right] - b\mathbb{E}\left[y^2\right] = a\mathbb{E}\left[y\right] - b\mathbb{E}\left[y\right]^2 - bVar\left(y\right)$$

Therefore, the expected utility is a function of the mean $\mathbb{E}\left[y\right]$ and variance $Var\left(y\right)$ only. This is also true in another case: suppose all lotteries have normally distributed outcomes (thus the state space now is infinite). Then, because any linear combination of normally distributed random variables is also normal, the distribution of any lottery (or linear combination thereof) will be completely described by its first two moments, mean and variance. Therefore also expected utility can be expressed as a function of these two numbers only.

Exercises

1) Using the definition in the chapter,

1. Show that state-by-state dominance implies first order stochastic dominance.
2. Provide a simple example of discrete, finite-state space random variables (that is, a table of payoffs, states and probabilities) that can be ranked by first order stochastic dominance, but not by state-by-state dominance.

2) Consider the random variables $X \sim N(0,1)$ and $Y \sim N(1,1)$.

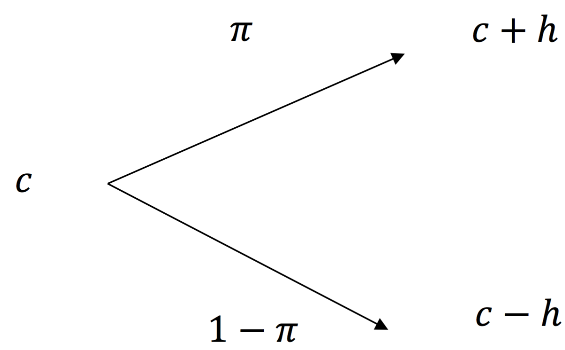
1. Suppose $Y = X + 1$. What is the state-by-state ordering of X and Y (if any)? What is the first-order stochastic dominance ordering of X and Y (if any)?
2. Suppose X and Y are independent random variables. What is the state-by-state ordering of X and Y (if any)? What is the first-order stochastic dominance ordering of X and Y (if any)?
3. What conclusion can we draw about state-by-state and first order stochastic dominance in relation to the probability distribution of the random variables?

3) Suppose an agent has an income of \$10. He has the possibility to buy for \$2 a lottery ticket which pays the winner \$19 (and zero otherwise). Suppose the agent has a von-Neumann Morgenstern utility $u(x) = \ln(x)$ and believes that the probability of winning π is $\frac{1}{3}$.

1. Write the agent's expected utility.
2. Should the agent buy the ticket?
3. Suppose the agent buys the ticket, but before a winner is selected he decides to sell it. What is the minimum certain amount that the agent would accept in exchange for the lottery ticket?

4) To resolve the St.Petersburg Paradox, Bernoulli proposed to give less weight to large payouts in the computation of the expected value of the game, to account for the fact that people value money less and less the richer they become (though this certainly does not apply to MFins). This property is shared by any increasing and concave function: can you guess why Bernoulli (a pretty smart guy) proposed exactly $u(x) = \ln(x)$? (Hint: consider the increase in utility u given by a small increase in the payoff x)

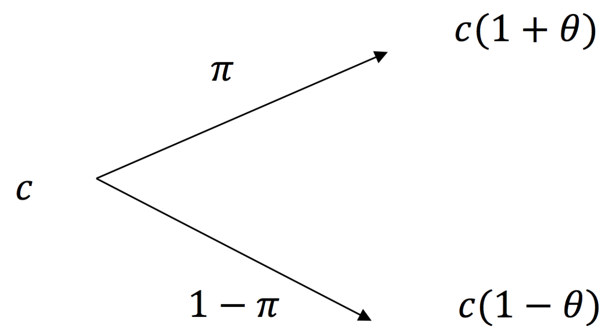
5) Consider a general Von Neumann-Morgenstern utility function $u(x)$ and a lottery which adds an amount h to your personal wealth c in case of victory (which occurs with probability π) and subtracts the same amount in case of loss:



Show that the following relation holds:

$$\pi(c,h)=\frac{1}{2}+\frac{1}{4}hR_A(c)+o\left(h^2\right)$$

Similarly, consider a lottery which increases your personal wealth c by $\theta\%$ in case of victory (which occurs with probability π) and decreases your personal wealth by the same percentage amount in case of loss:



Show that the following relation holds:

$$\pi(c,h)=\frac{1}{2}+\frac{1}{4}\theta R_R(c)+o\left(\theta^2\right)$$

6) Consider adding a lottery x to lottery w. We know that teh certainty equivalent consumption is defined by $\mathbb{E}\left[u(w+x)\right]=u\left(c^{CE}\right)$. Show that

$$w-c^{CE}\approx R_A(w)\frac{Var(x)}{2}$$

Similarly, for a positive random variable g with unit mean we have $\mathbb{E}\left[u(gw)\right]=u\left(kw\right)$, where k is is the certainty equivalent growth rate. Show that

$$1-k\approx R_R(w)\frac{Var(g)}{2}$$

7) Suppose an agent is given the opportunity to enter, for free, a lottery that pays him \$50,000 with probability $\frac{1}{2}$ and \$100,000 with probability $\frac{1}{2}$. Suppose the agent has CRRA utility

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

and $\gamma = 7$. Find the certainty equivalent of this lottery for the initial wealth Y equal to \$0, \$100,000 and \$300,000. What happens to the certainty equivalent when the initial wealth increases? Interpret this result.

8) Show that the solution to the portfolio optimization problem with one risky and one riskless asset is such that agents invest a constant fraction of their wealth on the risky asset, and that this results holds for any CRRA utility function and any distribution of the risky asset return r (please refer to slide 51 in the slides set 4).

9) Show that the Ellsberg Paradox is a violation of the independence axiom (hint: build a compound lottery).

10) Show that an agent with DARA preferences will choose to save precautionarily (hint: show that the absolute prudence coefficient is positive).

11) Show that even if X and Y have positive skewness $Z = X + Y$ can have negative skewness. You can make any additional assumptions to prove your point. What does this fact depend on? This shows that the negative skewness of the S&P 500 is consistent with single-name stocks having positive skewness.

Chapter 5

General Equilibrium, Efficiency and the Equity Premium Puzzle

In section 4.3 we introduced the representation of preferences over a consumption profile $c = (c_0, c_1, \ldots, c_S) \in \mathbb{R}_+^{S+1}$. Suppose that, at birth, agent i is given an endowment $e = (e_0, e_1, \ldots, e_S) \in \mathbb{R}_+^{S+1}$ and has a utility function $U^i : \mathbb{R}_+^{S+1} \to \mathbb{R}$. We say that U^i is:

- Quasiconcave, if the sets $C = \{c : U(c) \geq v\}$ are convex for each $v \in \mathbb{R}$
- Concave, if for any $c, c' \in \mathbb{R}^{S+1}$ and $\alpha \in [0,1]$ we have $U^i(\alpha c + (1-\alpha)c') \geq \alpha U^i(c) + (1-\alpha)U^i(c')$

Moreover, concavity implies quasiconcavity. If U^i is also differentiable, a standard requirement is that $\frac{\partial U^i}{\partial c_s} > 0$ for each $s = 0, 1, \ldots S$. We are now ready to formulate the *portfolio consumption problem*:

$$\max_{c,h} U^i(c_0, c_1, \ldots c_S)$$

Subject to the constratints

$$0 \leq c_0 \leq e_0 - p \cdot h$$

$$0 \leq c_s \leq e_s + \sum_{j=1}^{J} x_s^j h_j \;,\; s = 1, \ldots, S$$

If we define $c_T \equiv (c_1, \ldots, c_S) \in \mathbb{R}^S$ and $e_T \equiv (e_1, \ldots, e_S) \in \mathbb{R}^S$ the latter constraint can be rewritten more compactly as follows:

$$0 \leq c_T \leq e_T + Xh$$

The lagrangean for this problem is

$$\mathcal{L}(c; \lambda, \mu) = U^i(c_0, c_1, \ldots c_S) - \lambda[c_0 - e_0 + p \cdot h] - \mu[c_T - e_T - Xh]$$

Note that $X'h = h'X$ and $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}^S$. Suppose c^* solves the problem above, then there exist $\lambda \in \mathbb{R}$ and $\mu \in \mathbb{R}^S$ such that:

$$\frac{\partial U^i}{\partial c_0}(c^*) = \lambda \tag{5.1}$$

$$\frac{\partial U^i}{\partial c_s}(c^*) = \mu_s \tag{5.2}$$

$$\lambda p = \mu X \tag{5.3}$$

The third FOC implies that

$$p_j = \sum_{s=1}^{S} \frac{\mu_s}{\lambda} x_s^j$$

Plugging in the first and second FOCs we have

$$p_j = \sum_{s=1}^{S} \frac{\partial U^i / \partial c_s}{\partial U^i / \partial c_0} x_s^j$$

Assume that utility is *time separable:*

$$U^i(c) = u(c_0) + \delta \tilde{u}(c_T)$$

And also that the utility over uncertain states $\tilde{u}$ is a Von Neumann-Morgenstern utility:

$$U^i(c) = u(c_0) + \delta \mathbb{E}[u(c_T)]$$

Then we can rewrite our pricing relation as

$$p_j = \sum_{s=1}^{S} \pi_s \delta \frac{\partial u^i / \partial c_s}{\partial u^i / \partial c_0} x_s^j$$

We can read our findings as follows:

- $\delta \frac{\partial u^i / \partial c_s}{\partial u^i / \partial c_0}$ is the stochastic discount factor m_s
- $\pi_s \delta \frac{\partial u^i / \partial c_s}{\partial u^i / \partial c_0} = \pi_s m_s$ is the state price q_s

Thus, we are back to our equivalent asset pricing formulas

$$p_j = x^j \cdot q = \mathbb{E}\left[m \cdot x^j\right]$$

Finally, note that the vector of marginal rates of substitutions $\frac{\partial U^i/\partial c_s}{\partial U^i/\partial c_0}$ is the state price vector (note that this is positive by our assumption on U^i).

5.1 Pareto Efficiency

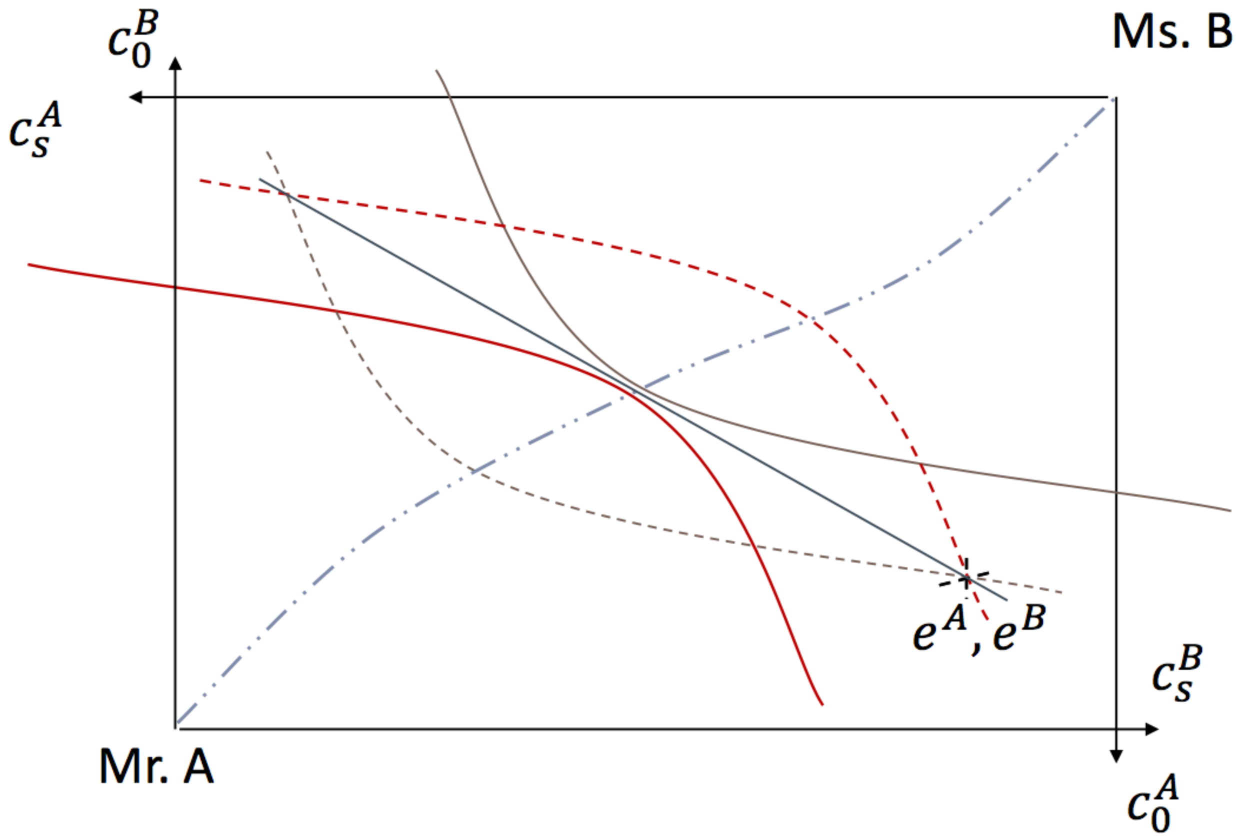
Consider the set of allocations that yield agent i the same utility, $S = \left\{c \in \mathbb{R}^{S+1} : U^i\left(c\right) = k\right\}$ for a given $k \in \mathbb{R}$. This corresponds to a curve in the (c_0, c_s) space whose total differential is

$$\frac{\partial U^i\left(c\right)}{\partial c_0}dc_0 + \frac{\partial U^i\left(c\right)}{\partial c_s}dc_s = 0$$

called *indifference curve* of agent i. Equivalently,

$$\frac{\partial U^i/\partial c_s}{\partial U^i/\partial c_0} = -\frac{dc_0}{dc_s} \equiv MRS^i_{s,0}$$

That is, the negative of the inverse of the slope $\frac{dc_s}{dc_0}$ of the indifference curve is the marginal rate of substitution $MRS^i_{s,0}$ between consumption in state $s = S$ and consumption in state $s = 0$ of agent i. Moreover, we also know from the previous section that the marginal rate of substitution $MRS_{s,0}$ equals the state price q_s. We can represent this with an *Edgeworth box*: fix the consumption axes $\left(c^A_0, c^A_s\right)$ for agent A with the origin in the lower left corner of the box, and draw a box that has width equal to the sum of the endowments $e^A_0 + e^B_0$ in state 0 and height the sum of endowments $e^A_s + e^B_s$ in state s. We fix theconsumption axes $\left(c^B_0, c^B_s\right)$ in the upper right corner of the box. The endowments of the two agents corespond to a point $e^* = e^A$ for the agent A axes, and $e^* = e^B$ for the agent B axes.



In the picture above, the indifference curves of agent A are in grey, while those of agent B are in red. The dashed lines are the indifference curves that pass through the endowment e^A (or e^B if we look at the agent B axes). The dash-dotted blue line represents the set of points in the Edgeworth box for which $MRS_{s,0}^A = MRS_{s,0}^B$, and is called *contract curve*. Allocations on this line are called *Pareto Optimal*, that is, they are allocations with the property that there exists no other allocation in the box such that at least one agent can be made better off without making someone else worse off. It is important to remark that allocative efficiency (Pareto Efficiency) is not the same thing as informational efficiency or fairness. The solid blue straight line has slope $\frac{1}{q_s}$ and is the only straight line passing from the endowment point for which $MRS_{s,0}^A = MRS_{s,0}^B = q_s$. The point c^e where the straight blue line intersects the contract curve is called *Competitive Equilibrium*. The relation between Pareto Optimality and Competitive Equilibrium is captured by the two *Welfare Theorems*:

1. **First Welfare Theorem**: If markets are complete, then the Competitive Equilibrium allocation is Pareto Optimal.

2. **Second Welfare Theorem**: Any Pareto Efficient allocation can be decentralized as a Competitive Equilibrium.

The First Welfare Theorem also implies that there exists a unique state price vector q and that for any two agents $i \neq j$ the vectors MRS^i and MRS^j (whose elements $MRS_{s,0}^i$ for $s = 1, \dots S$) are equal and coincide with q. However, in multi-period (or even multi-good) settings and markets are incomplete, the equilibrium allocation is not Pareto Optimal.

We conclude the chapter with two interesting results about the aggregation of preferences:

Aggregation Theorem 1: Suppose markets are complete. Then asset prices in an economy with K agents are identical to an economy with a single agent (social planner) whose utility is

$$U(c) = \sum_{k=1}^{K} \alpha_k u^k(c)$$

Where α_k is a weight given to agent k, u^k is the utility of agent k and the single agent consumes the aggregate endowment.

Aggregation Theorem 2: Suppose that i) a riskless annuity and the endowments are tradable, ii) agents have common beliefs, iii) agents have a common rate of time preference, iv) agents have Linear Risk Tolerance (LRT or Hyperbolic Risk Aversion) preferences (see section 4.6) with

$$R_A^k(c) = \frac{1}{A_k + Bc}$$

Then asset prices in an economy with many agents are identical to a single agent economy with LRT preferences such that

$$R_A(c) = \frac{1}{\sum_{k=1}^{K} A_k + Bc}$$

5.2 The Sharpe Ratio, Bonds and the Equity Premium Puzzle

Consider a security with price p_t at time t that pays off an amount x_{t+1} at time $t+1$. Using the SDF representation of the price we can write

$$p_t = \mathbb{E}_t\left[m_{t+1}x_{t+1}\right] = \mathbb{E}_t\left[m_{t+1}\right]\mathbb{E}_t\left[x_{t+1}\right] + Cov_t\left[m_{t+1}, x_{t+1}\right]$$

Now, for a given (generic) m_{t+1} we write $R_t^f = \frac{1}{\mathbb{E}_t[m_{t+1}]}$. If there exists a risk-free portfolio, R_t^f will be unique and independent of the particular choice of m_{t+1}. Therefore we have

$$p_t = \frac{\mathbb{E}_t\left[x_{t+1}\right]}{R_t^f} + Cov_t\left[m_{t+1}, x_{t+1}\right]$$

Where the first term is the discounted expected payoff, and the second term is a risk adjustment. Clearly, positive correlation of the payoff with the discount factor adds to the price (or, for a given price, decreases the return on the asset). In returns we have

$$\mathbb{E}_t\left[m_{t+1}R_{t+1}\right]=1$$

And using $R_t^f=\frac{1}{\mathbb{E}_t[m_{t+1}]}$ we get

$$\mathbb{E}_t\left[m_{t+1}\left(R_{t+1}-R_t^f\right)\right]$$

that is, the m-discounted expected return on *any* asset is equal to zero. This also implies

$$Cov_t\left[m_{t+1},R_{t+1}-R_t^f\right]=-\mathbb{E}_t\left[m_{t+1}\right]\mathbb{E}_t\left[R_{t+1}-R_t^f\right]$$

That is

$$\mathbb{E}_t\left[R_{t+1}-R_t^f\right]=-\frac{Cov_t\left[m_{t+1},R_{t+1}\right]}{\mathbb{E}_t\left[m_{t+1}\right]}$$

so that we can see that the expected excess return, the risk premium, is determined by its covariance with the stochastic discount factor. This also holds for a generic portfolio $h\in\mathbb{R}^J$:

$$\mathbb{E}_t\left[\left(R_{t+1}-R_t^f\right)h\right]=-\frac{Cov_t\left[m_{t+1},R_{t+1}h\right]}{\mathbb{E}_t\left[m_{t+1}\right]}=-\frac{\rho_t\left(m_{t+1},R_{t+1}h\right)\sigma_t\left(m_{t+1}\right)\sigma_t\left(R_{t+1}h\right)}{\mathbb{E}_t\left[m_{t+1}\right]}$$

Note that all results hold also for the unconditional expectation $\mathbb{E}\left[\cdot\right]$, so we can drop the expectation subscript:

$$\mathbb{E}\left[\left(R_{t+1}-R_t^f\right)h\right]=-\frac{Cov\left[m_{t+1},R_{t+1}h\right]}{\mathbb{E}\left[m_{t+1}\right]}=-\frac{\rho\left(m_{t+1},R_{t+1}h\right)\sigma\left(m_{t+1}\right)\sigma\left(R_{t+1}h\right)}{\mathbb{E}\left[m_{t+1}\right]}$$

Rearranging in terms of the portfolio Sharpe Ratio,

$$-\frac{\sigma\left(m_{t+1}\right)}{\mathbb{E}\left[m_{t+1}\right]}\rho\left(m_{t+1},R_{t+1}h\right)=\frac{\mathbb{E}\left[\left(R_{t+1}-R_t^f\right)h\right]}{\sigma\left(R_{t+1}h\right)}$$

taking the absolute value of both sides,

$$\frac{\sigma\left(m_{t+1}\right)}{\mathbb{E}\left[m_{t+1}\right]}\left|\rho\left(m_{t+1},R_{t+1}h\right)\right|=\left|\frac{\mathbb{E}\left[\left(R_{t+1}-R_t^f\right)h\right]}{\sigma\left(R_{t+1}h\right)}\right|$$

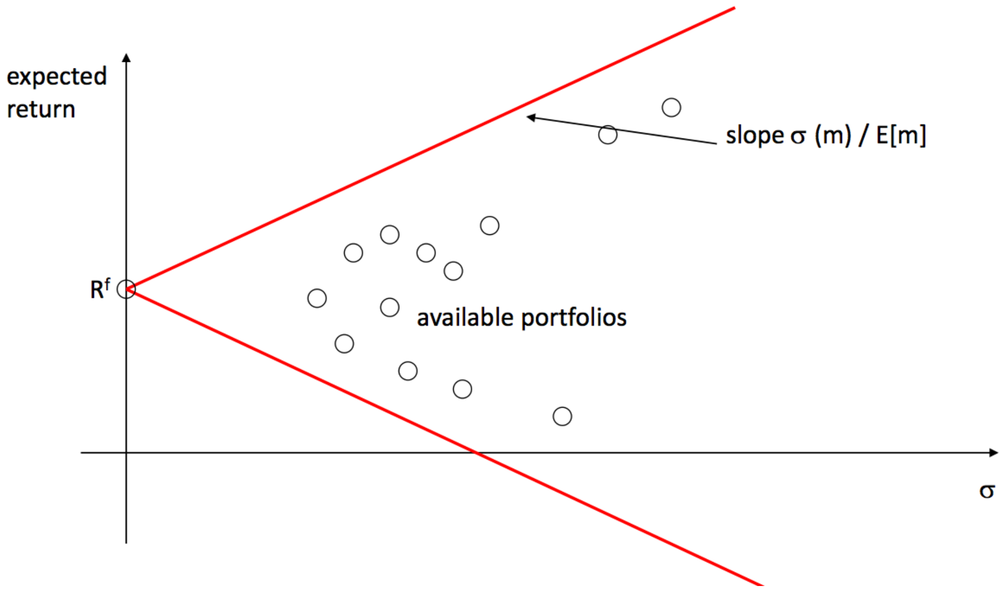
but since for any $h\in\mathbb{R}^J$ $\rho\left(\cdot\right)\in\left[-1,1\right]$,

$$\frac{\sigma\left(m_{t+1}\right)}{\mathbb{E}\left[m_{t+1}\right]}\geq\sup_{h\in\mathbb{R}^J}\left|\frac{\mathbb{E}\left[\left(R_{t+1}-R_t^f\right)h\right]}{\sigma\left(R_{t+1}h\right)}\right|$$

we just proved the following theorem:

Theorem (Hansen-Jagannathan Bound): The ratio of the standard deviation of a stochastic discount factor to its mean exceeds the Sharpe Ratio attained by any portfolio.

This theorem can be used to easily check the viability of a proposed stochastic discount factor. Or, for a given stochastic discount factor, it establishes the maximum Sharpe Ratio any portfolio can attain. Note that the theorem also holds for the expectation conditional at time t.



5.3 Adding Expected Utility

Now assume that agents maximize their expected utility. For $c_0 \in \mathbb{R}$ and $c_1 \in \mathbb{R}^S$,

$$U(c_0, c_1) = \sum_{s=1}^{S} \pi_s u(c_0, c_{1,s})$$

so that

$$\partial_0 u = \left(\frac{\partial u(c_0, c_{1,1})}{\partial c_0}, \ldots, \frac{\partial u(c_0, c_{1,S})}{\partial c_0} \right)$$

$$\partial_1 u = \left(\frac{\partial u(c_0, c_{1,1})}{\partial c_1}, \ldots, \frac{\partial u(c_0, c_{1,S})}{\partial c_S} \right)$$

and the stochastic discount factor can be shown to be $m = \frac{MRS}{\pi} = \frac{\partial_1 u}{\mathbb{E}[\partial_0 u]} \in \mathbb{R}^S$. If the utility is also *time-separable*, then $u(c_0, c_{1,s}) = v(c_0) + v(c_{1,s})$ and

$$\partial_0 u = \frac{\partial v(c_0)}{\partial c_0} \times (1, \ldots, 1) = v'(c_0) \times (1, \ldots, 1)$$

$$\partial_1 u = \left(\frac{\partial v(c_{1,1})}{\partial c_1}, \ldots, \frac{\partial v(c_{1,S})}{\partial c_S} \right) = \left(v'(c_{1,1}), \ldots, v'(c_{1,S}) \right)$$

and therefore $m_s = \frac{1}{\pi_s} \times \frac{\pi_s v'(c_{1,s})}{v'(c_0)} = \frac{v'(c_{1,s})}{v'(c_0)}$. For example, suppose $u(c_0, c_{1,s}) = \ln c_0 + \ln c_{1,s}$. Then

$$m = \left(\frac{c_0}{c_{1,1}}, \ldots, \frac{c_0}{c_{1,S}} \right)'$$

which shows that states where consumption is low are states in which m is high.

5.4 The Equity Premium Puzzle

Recall that for any asset j

$$\mathbb{E}\left[R^j\right] - R^f = -R^f \times Cov\left[m, R^j\right]$$

Using the assumptions from the last section we can write

$$\mathbb{E}\left[R^j\right] - R^f = -R^f \times \frac{Cov\left[\partial_1 u, R^j\right]}{E\left[\partial_0 u\right]}$$

so the Hansen-Jagannathan bound can be rewritten as

$$\sigma(m) \geq \frac{1}{R^f} \left| \frac{\mathbb{E}\left[R^j - R^f\right]}{\sigma(R^j)} \right|$$

that is,

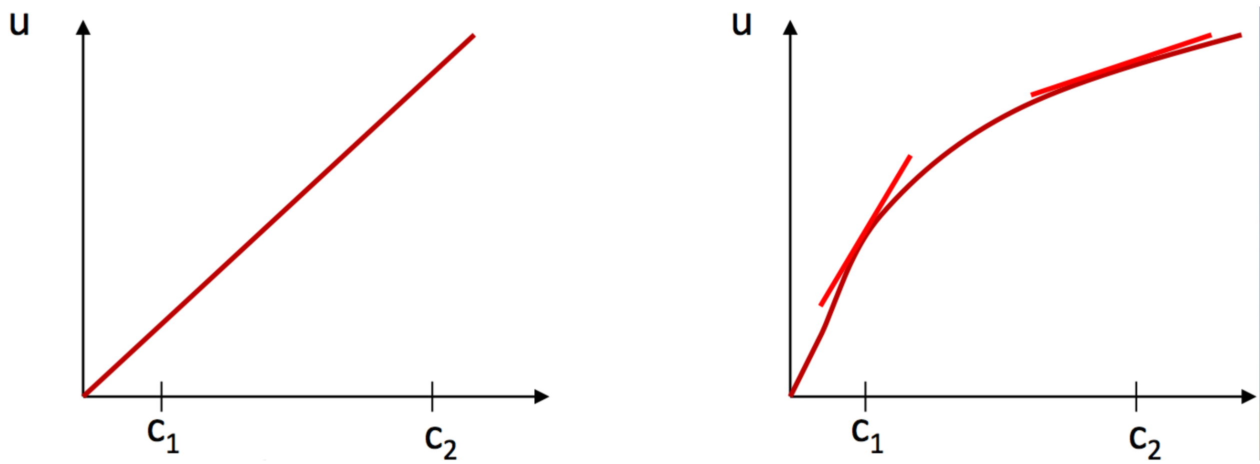
$$\sigma\left(\frac{\partial_1 u}{\mathbb{E}\left[\partial_0 u\right]} \right) \geq \frac{1}{R^f} \left| \frac{\mathbb{E}\left[R^j - R^f\right]}{\sigma(R^j)} \right|$$

However:

1. The right hand side of this inequality, the Sharpe Ratio of any asset, can be very high for some portfolios.

2. Consumption volatility is generally low.

Together, to be consistent with the Hansen-Jagannathan bound these two facts imply that the curvature of the utility function u must be very high. But curvature is synonimous for *risk aversion*, and to justify the empirical data we should accept that agents have an unrealistically high level of risk aversion. This inconsistency between theory and empirical facts is commonly referred to as the *Equity Premium Puzzle*.



Even if we allowed for an unrealistically high level of risk aversion, more problems emerge: due to the high degree of concavity of u there is also a low *elasticity of intertemporal substitution* (EIS) which would drive agents to strongly smooth consumption over time. But in the context of a von-Neumann Morgenstern utility function smoothing over time is equivalent to smoothing over states, and in the example of a CRRA utility function, the EIS parameter is just $\frac{1}{\gamma}$. Finally, the puzzle can be re-cast in terms of the risk-free interest rate, which the model predicts to be much higher than in reality.

Solutions to the puzzles have been proposed, which essentially depart from the von-Neumann Morgenstern expected utility paradigm and allow for a utility representation that separates risk-aversion from intertemporal elasticity of substitution.

5.5 Empirical Estimation: Generalized Method of Moments

Another great contribution by Lars Hansen (one of the 2013 Nobel Prize in Economics recipients) was the Generalized Method of Moments and its applications in the context of the canonical consumption-based asset pricing model. We know that under CRRA utility we have

$$\mathbb{E}_t \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(R_{t+1} - R_t^f \right) \right] = 0$$

That is, the expected discounted excess return should always be zero. How do we take this to data? How do we find parameters beta and gamma that best fit the data? How do we check this over many different times and returns, to see if those two parameters can actually explain empirical facts? What do we do about that conditional expectation $\mathbb{E}_t$, conditional on information in people's heads? How do we bring in all the variables that seem to forecast returns over time (e.g. by the dividend-price ratio) and across assets (value, size, etc.)? How do we handle the fact that return variance changes over time, and consumption growth may be autocorrelated?

When Hansen wrote his paper, this was a big headache. He suggested to just multiply by any variable z that we think forecasts returns or consumption, and take the unconditional average of this conditional average: the model predicts that the unconditional average obeys

$$\mathbb{E}\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R_{t+1}-R_t^f\right)\times z_t\right]=0$$

So we can just take this average in the data, and do this for lots of different assets R and lots of different instruments z. Finaly, we pick the β and γ that make some of the averages as close to zero as possible, and then look at the other averages and see how close they are to zero. (Hansen worked out the statistics of this procedure - how close should the other averages be to zero, and what is a good measure of the sample uncertainty in β and γ estimates - taking in to account a wide variety of statistical problems that may arise).

The results were not favorable to the consumption model: a huge γ is needed to fit the difference between stocks and bonds, questioning the validity of the utility function used, the measurements of consumption, and even of the whole consumption-based asset pricing framework.

Chapter 6

Mean-Variance Analysis and CAPM

In section 3.4 we discussed the State-Price Beta model, which essentially states that for an asset j

$$\mathbb{E}\left[R^j\right] - R^f = \beta_j \left(\mathbb{E}\left[R^*\right] - R^f\right)$$

Where $\beta_j = \frac{Cov\left(R^*, R^j\right)}{Var(R^*)}$ and R^* is the return of an asset whose payoff is equal to the discount factor m^* derived from the pricing kernel $q^* \in \langle X \rangle$. This is a very general statement about returns, but it does not help us identify what assets in reality have a return of R^*.

Suppose all agents have the same quadratic utility $u\left(c_0, c_1\right) = v\left(c_0\right) - \left(c_1 - \alpha\right)^2$. Recall that the expected utility is $\mathbb{E}\left[U\left(c\right)\right] = \sum_{s=0}^{S} \pi_s u\left(c_0, c_s\right)$ and therefore $m = \frac{\partial_1 u}{\mathbb{E}[\partial_0 u]}$: in the quadratic utility case we have

$$\partial_1 u = \begin{bmatrix} -2\left(c_1 - \alpha\right) & \cdots & -2\left(c_S - \alpha\right) \end{bmatrix}$$

So the excess return can be written as

$$\mathbb{E}\left[R^j\right] - R^f = -\frac{Cov\left(m, R^j\right)}{\mathbb{E}\left[m\right]} = -R^f \times \frac{Cov\left(\partial_1 u, R^j\right)}{\mathbb{E}\left[\partial_0 u\right]} = -R^f \times \frac{Cov\left(-2\left(c_1 - \alpha\right), R^j\right)}{\mathbb{E}\left[\partial_0 u\right]} =$$

$$= R^f \times \frac{2Cov\left(c_1, R^j\right)}{\mathbb{E}\left[\partial_0 u\right]}$$

This holds for any asset $x^j \in \langle X \rangle$, and can be generalized to any *portfolio* h:

$$\mathbb{E}\left[R^h\right] - R^f = R^f \times \frac{2Cov\left(c_1, R^h\right)}{\mathbb{E}\left[\partial_0 u\right]}$$

Now consider the *market portfolio* defined as $x^{mkt} = \sum_{j: x^j \in \langle X \rangle} w_j x^j$:[1] the above equation must hold

[1] Where $\{w_j\}$ are weights $w_j = \frac{h_j}{\sum_{k=1}^{J} h_k}$ for each (linearly independent) asset in the asset span.

for it too, and therefore dividing side by side we get

$$\frac{\mathbb{E}\left[R^h\right]-R^f}{\mathbb{E}\left[R^{mkt}\right]-R^f}=\frac{Cov\left(c_1,R^h\right)}{Cov\left(c_1,R^{mkt}\right)}$$

Where R^{mkt} is the return of the market portfolio. If moreover agents are homogeneous and live in an exchange economy, then we know that c_1 corresponds to the aggregate endowment and this is perfectly correlated with R^{mkt}:

$$\frac{\mathbb{E}\left[R^h\right]-R^f}{\mathbb{E}\left[R^{mkt}\right]-R^f}=\frac{Cov\left(R^{mkt},R^h\right)}{Var\left(R^{mkt}\right)}$$

But since $\beta_h=\frac{Cov\left(R^{mkt},R^h\right)}{Var\left(R^{mkt}\right)}$ we can write

$$\mathbb{E}\left[R^h\right]=R^f+\beta_h\left(\mathbb{E}\left[R^{mkt}\right]-R^f\right)$$

Which we call the *Market Security Line* (MSL). Note that in order for the above result to hold, R^* must be a linear function of R^{mkt} of the type

$$R^*=\frac{a+bR^{mkt}}{a+bR^f}$$

With $b<0$ since we know that the stochastic discount factor m is high (hence high R^*) in states in which the economy is doing poorly, so in which the market portfolio x^{mkt} is low (hence low R^{mkt}).

6.1 The Traditional Derivation of CAPM

We define the mean return of the portfolio h as

$$\mu_h\equiv\mathbb{E}\left[r_h\right]=\mathbb{E}\left[\sum_{j=1}^{J}w_jr_j\right]=\sum_{j=1}^{J}w_j\mu_j$$

Where $r_h=R^h-1$ is the net return[2] and h_k is the amount of asset k (as usual), so that the price of R^h is still 1. The variance of the portfolio is given by

$$\sigma_h^2\equiv Var\left(r_h\right)=w'Vw$$

And $\sigma_h\equiv\sqrt{Var\left(r_h\right)}$ is the standard deviation of portfolio h, where $w\in\mathbb{R}^J$ is the vector of weights introduced above and V is the covariance matrix of the assets in portfolio h.

[2]Note that we can always write the excess return in gross or net terms: $\mathbb{E}\left[R^j\right]-R^f=\mathbb{E}\left[r_j\right]-r_f$.

Definition: We say that portfolio A mean-variance dominates portfolio B if $\mu_A \geq \mu_B$ and $\sigma_A < \sigma_B$, or $\mu_A > \mu_B$ and $\sigma_A \leq \sigma_B$.

Note that in the (σ_h, μ_h) space, for a given value of μ_h and σ_h we can immediately find the subset of portfolios which are not dominated by the given μ_h and σ_h:

Definition: For given μ and σ, the *Efficient Frontier* is the locus of all non-dominated payoffs in the (σ_h, μ_h) space.

It follows directly from the definition of efficient frontier that no rational investor with mean-variance preferences would choose to hold a portfolio outside of the efficient frontier.

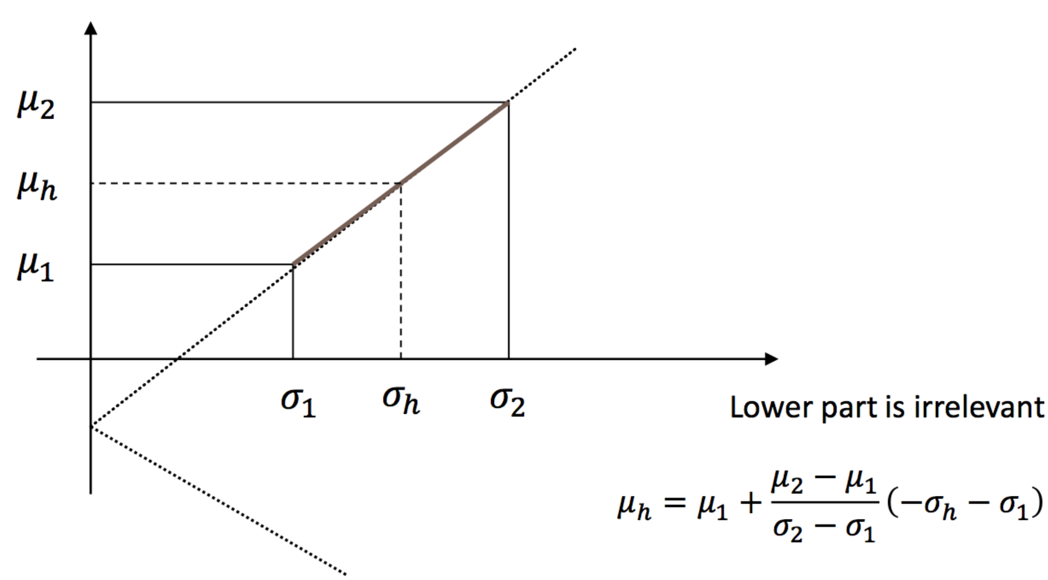
In the $J = 2$ case, for portfolio h we have $w_2 = 1 - w_1$, so the mean return is $\mu_h = w_1\mu_1 + (1 - w_2)\mu_2$ and the variance is $\sigma_h^2 = w_1^2\sigma_1^2 + (1 - w_1)^2 \sigma_2^2 + 2w_1(1 - w_1)\sigma_1\sigma_2\rho_{1,2}$. Note that in general, we can specify the vector μ of returns and covariance matrix V and back out the weights vector w: in the 2 assets example, for $\rho_{1,2} = 1$ we get

$$w_1 = \frac{\pm\sigma_h - \sigma_2}{\sigma_1 - \sigma_2}$$

So that plugging back in the expressions for μ_h we get

$$\mu_h = \mu_1 + \frac{\mu_2 - \mu_1}{\sigma_2 - \sigma_1}(\pm\sigma_h - \sigma_1)$$

Which looks like:



The Efficient Frontier: Two Perfectly Correlated Risky Assets

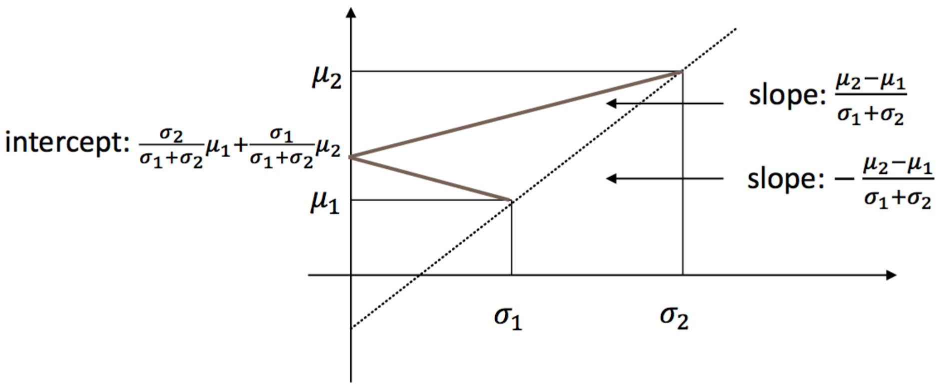
For $\rho_{1,2} = -1$ we get

$$w_1 = \frac{\pm\sigma_h + \sigma_2}{\sigma_1 + \sigma_2}$$

And therefore

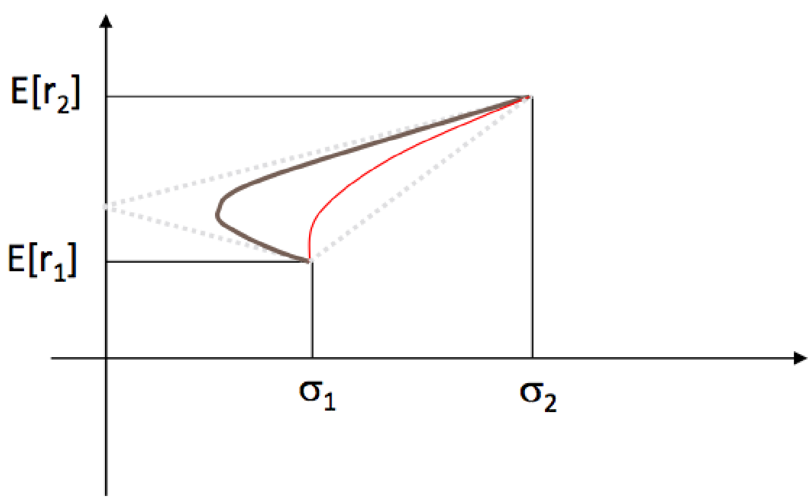
$$\mu_h = \frac{\sigma_2}{\sigma_1 + \sigma_2}\mu_1 + \frac{\sigma_1}{\sigma_1 + \sigma_2}\mu_2 \pm \frac{\mu_2 - \mu_1}{\sigma_1 + \sigma_2}\sigma_h$$

Which looks like:



The Efficient Frontier: Two Perfectly Negative Correlated Risky Assets

While for $\rho_{1,2} \in (-1, 1)$ we have



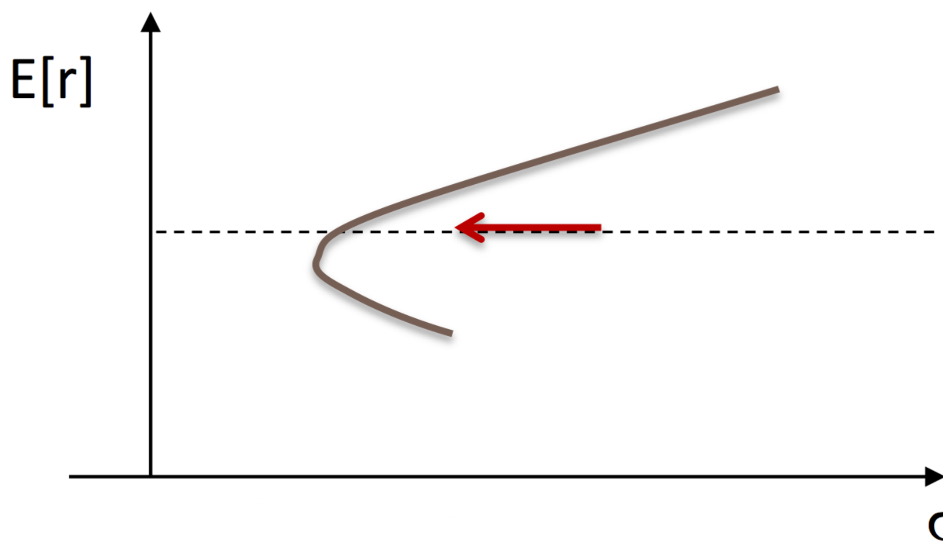
The Efficient Frontier: Two Imperfectly Correlated Risky Assets

The same concept generalizes to the J assets case: a frontier portfolio has minimum variance among all feasible portfolios with the same expected portfolio return, so the problem is

$$\max_{w} \frac{1}{2} w' V w$$

$$s.t.\ w'\mu = E,\ w'I = 1$$

Where $I = (1, \ldots, 1)$ and $E \in \mathbb{R}$ is a given (fixed) expected return. The idea is that we fix an expected return E and then look for the minimum variance possible given that expected return:



The FOCs are:

$$\frac{\partial \mathcal{L}}{\partial w} = Vw - \lambda\mu - \gamma I = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = E - w'\mu = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = I - w'I = 0$$

Where λ and γ are the Lagrange multipliers for the two constraints. Pre-multiplying the first FOC by $\mu' V^{-1}$ we get

$$\mu' w = \lambda \left(\mu' V^{-1} \mu\right) + \gamma \left(\mu' V^{-1} I\right) \equiv \lambda B + \gamma A$$

And since $\mu' w = w'\mu = E$ by the third FOC,

$$E = \lambda B + \gamma A$$

While pre-multiplying the same FOC by $I' V^{-1}$ we get

$$1 = I' w = \lambda \left(I' V^{-1} \mu\right) + \gamma \left(I' V^{-1} I\right) \equiv \lambda A + \gamma C$$

Since $\mu' V^{-1} I = I' V^{-1} \mu$. Solving for λ and γ we get

$$\lambda = \frac{E \times C - A}{D}$$

$$\gamma = \frac{B - E \times A}{D}$$

Where $D \equiv B \times C - A^2$. Therefore, pre-multiplying the first FOC by V^{-1} and plugging in λ and γ we have

$$w^* = \frac{E \times C - A}{D} V^{-1} \mu + \frac{B - E \times A}{D} V^{-1} I = \underbrace{\frac{B \times V^{-1} I - A \times V^{-1} \mu}{D}}_{\equiv g} + \underbrace{\frac{C \times V^{-1} \mu - A \times V^{-1} I}{D}}_{\equiv h} \times E = g + hE$$

Suppose now that for some portfolio with return r we set $E = \mathbb{E}[r]$. Similarly to the 2 assets problem, the solution portfolio weights are linear in the (expected) portfolio returns:

$$w^* = g + h\mathbb{E}[r]$$

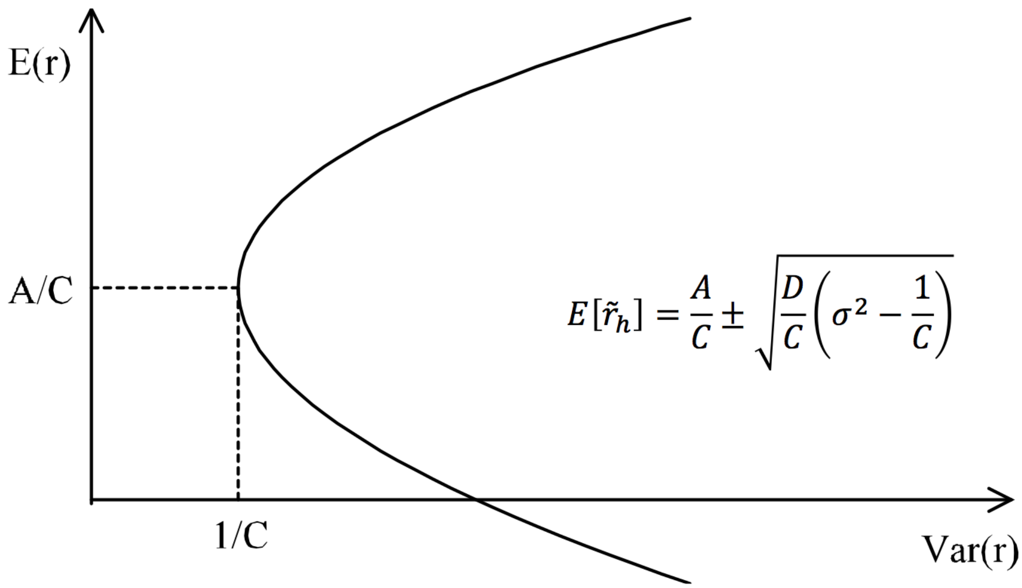
Note that for $\mathbb{E}[r] = 0$ and $\mathbb{E}[r] = 1$ we get g and $g + h$ respectively, which are therefore frontier portfolios as well. We established the following facts:

> **Proposition 1**: The entire set of frontier portfolios can be generated by g and $g+h$: any portfolio in the frontier is a linear combination of these two portfolios.
>
> **Proposition 2**: Any linear combination of frontier portfolios is also a frontier portfolio: the portfolio frontier can be described as linear combinations of any two frontier portfolios (not just g and $g + h$).

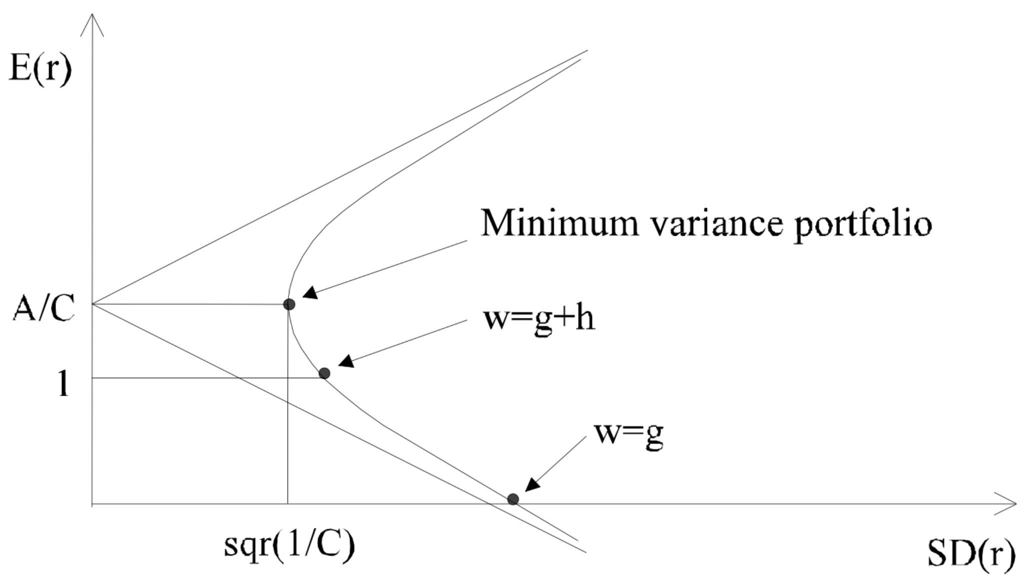
It is possible to show that

$$\sigma^2(r) = \frac{C}{D}\left[\mathbb{E}[r] - \frac{A}{C}\right]^2 + \frac{1}{C}$$

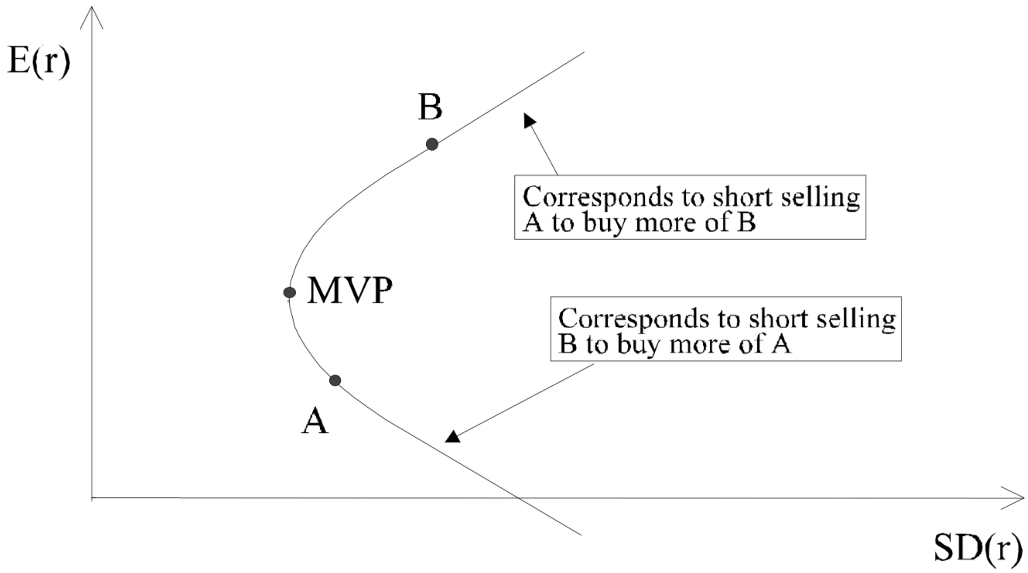
So that we know that 1) the expected return of the minimum variance portfolio is $\frac{A}{C}$, 2) the variance of the minimum variance portfolio is $\frac{1}{C}$, 3) is the equation of a parabola with vertex at $\left(\frac{1}{C}, \frac{A}{C}\right)$ in the expected return/variance space and of a hyperbola in the expected return/standard deviation space. That is, in the expected return-variance space $(Var(r), \mathbb{E}[r])$ space the set of frontier portfolios looks like:



Where A, B, C and D are constants. In the expected return-standard deviation space $(SD(r), \mathbb{E}[r])$ space the set of frontier portfolios is:



Given two portfolios A and B in the frontier:



We can see that points above B and below A correspond to portfolios in which we short-sell assets. How does all this change if we have a risk-free asset (i..e an asset whose variance is zero)? We can adapt the problem considered above as follows:

$$\max_{w} \frac{1}{2} w' V w$$

$$s.\,t.\; w'\mu + (1 - w' \cdot I) r_f = \mathbb{E}\left[r\right]$$

Where r_f is the net risk-free rate.[3] Note that in this problem weights automatically sum to one since $w'I + (1 - w'I) = 1$, so we don't have the second constraint in the problem without the risk-free asset. First order conditions yield

$$w^* = \lambda V^{-1} \left(\mu - I \cdot r_f\right)$$

So premultiplying both sides by $\left(\mu - I \cdot r_f\right)'$ yields

$$\lambda = \frac{\mathbb{E}\left[r\right] - r_f}{\left(\mu - I \cdot r_f\right)' V^{-1} \left(\mu - I \cdot r_f\right)}$$

Since $\left(\mu - I \cdot r_f\right)' w^* = \mu' w^* - I' w^* \cdot r_f = \mu \left(w^*\right)' - I \left(w^*\right)' \cdot r_f = \mathbb{E}\left[r\right] - r_f$ from the constraint in the problem above. Plugging this back in w^* we finally get

$$w^* = \frac{V^{-1} \left(\mu - I \cdot r_f\right)}{\left(\mu - I \cdot r_f\right)' V^{-1} \left(\mu - I \cdot r_f\right)} \times \left(\mathbb{E}\left[r\right] - r_f\right)$$

Where the denominator is the square of the *sharpe ratio* H. We have two important results at this

[3] That is, $r_f = R^f - 1$.

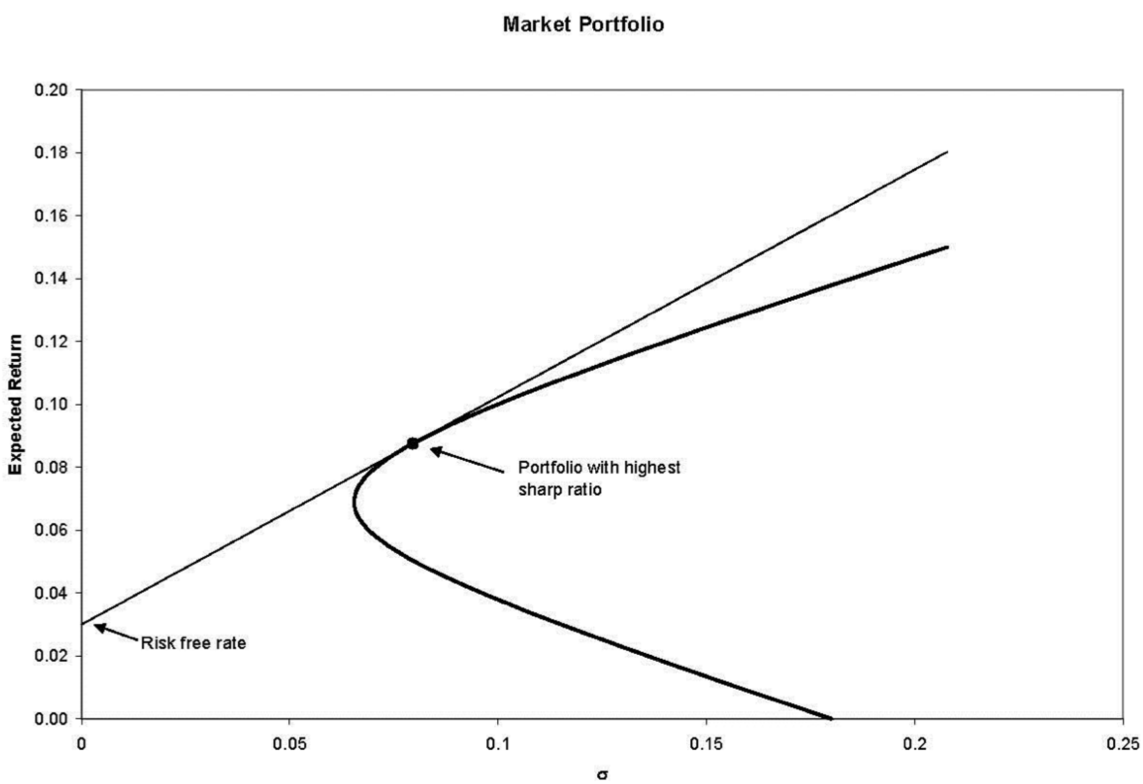
point:

Result 1: For any two frontier portfolios p and q, we have $\mathbb{E}[r_q] - r_f = \beta_{q,p}(\mathbb{E}[r_p] - r_f)$.

To see this, note that $Cov(r_q, r_p) = w_q' V w_p = w_q'(\mu - I \cdot r_f)\frac{\mathbb{E}[r_p]-r_f}{H^2} = \frac{(\mathbb{E}[r_q]-r_f)(\mathbb{E}[r_p]-r_f)}{H^2}$ and $Var(r_p) = \frac{(\mathbb{E}[r_p]-r_f)^2}{H^2}$, and dividing side by side yields the result. Importantly, this holds for any two frontier portfolio p, thus in particular it also holds for the market portfolio.

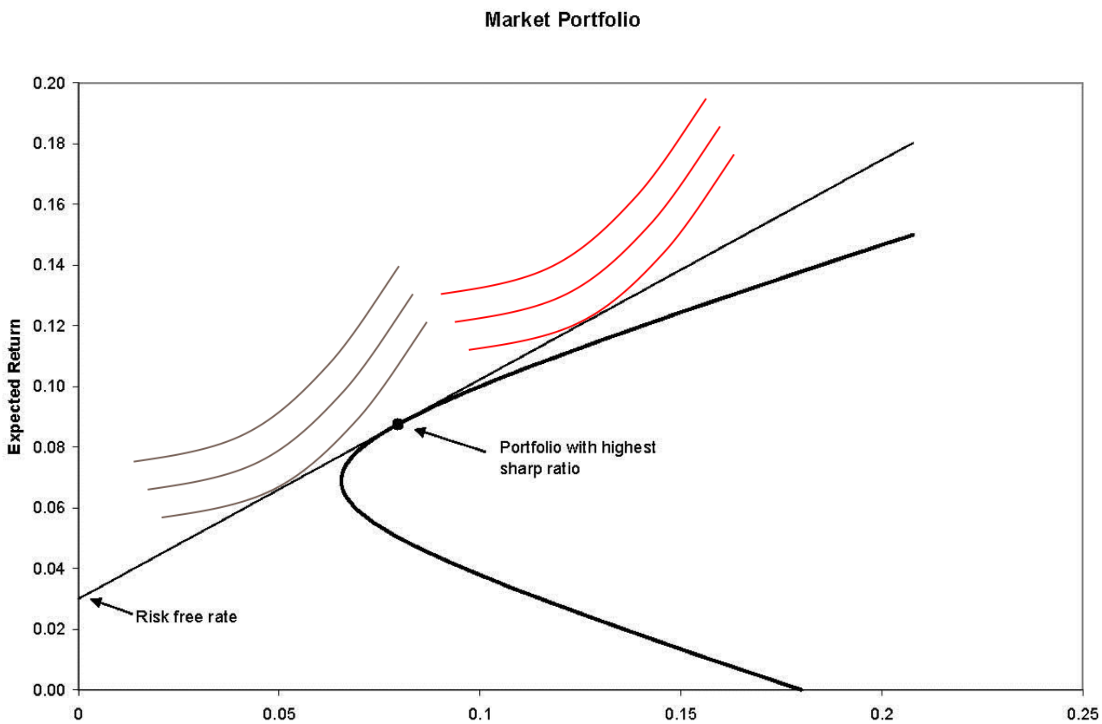
Result 2: The frontier is linear in the $(SD(r), \mathbb{E}[r])$ space.

This follows immediately from $Var(r_p) = \frac{(\mathbb{E}[r_p]-r_f)^2}{H^2}$ by taking the square root and rearranging: we get $\mathbb{E}[r_p] = r_f + H \times SD(r_p)$. Therefore the efficient frontier with a risk-free asset looks like the following:



6.1.1 Two Fund Separation

What is the result of individual optimization on the aggregate supply and demand of assets? That is, what can we say about the equilibrium state given what we have seen so far? We approach this problem in two steps: first we solve for the efficient frontier of the J risky assets, then we solve for the tangency point with the agents' indifference curves in the $(SD(r), \mathbb{E}[r])$ space. The advantage of this approach is that we have the same portfolio of J risky assets for different agents with differing risk aversion, and this makes it easier to apply equilibrium aguments. Represented in the graph below are the optimal portfolios of two investors with different degrees of risk aversion (the black and red indiffrerence curves).



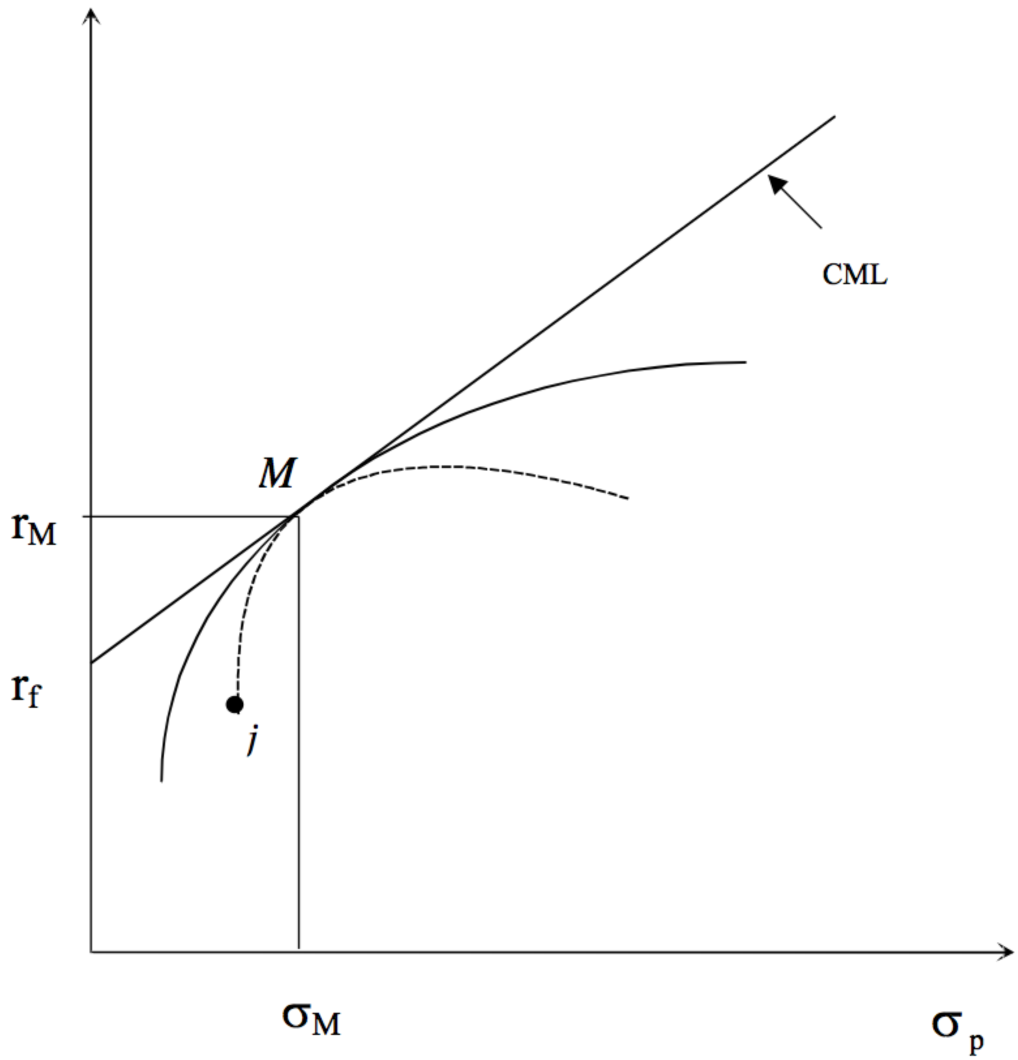
In this setup, mean-variance preferences represented by a utility function $U(\mu, \sigma^2)$ simply satisfy $\frac{\partial U}{\partial \mu} > 0$ and $\frac{\partial U}{\partial \sigma^2} < 0$: a simple example is $U(\mu, \sigma^2) = \mu - \frac{\rho}{2}\sigma^2$. As we have already seen, these preferences are equivalent to those of a Von Neumann-Morgenstern quadratic utility: if $u(X) = a + bX + cX^2$ then $\mathbb{E}[u(X)] = a + b\mu + c\sigma^2 + c\mu^2 = U(\mu, \sigma^2)$. Again as already seen, if asset returns are gaussian then also any portfolio is gaussian and therfore preferences can only depend on the first two moments. Moreover if agents have CARA utility function we know that the certainty equivalent for a lottery with mean μ and variance σ^2 is going to be $\mu - \frac{\rho}{2}\sigma^2$ where ρ is the absolute risk aversion of agents.

6.1.2 Equilibrium leads to CAPM

The theory examined so far only tells us about the demand of assets: prices are taken as given, and the composition of the optimal (risky) portfolio is the same for all investors. Setting the aggregate demand for assets equal to the supply of assets, that is, the market portfolio. Through CAPM we can find assets' equilibrium prices and agents' risk premium. As we have seen, the market portfolio is efficient (since it lies on the efficient frontier); moreover, all individual optimal portfolios are located on the half line originating at the point $(0, r_f)$. This half-line is called *capital market line* (CML) and can be written as

$$\mathbb{E}[r_h] = r_f + \frac{\mathbb{E}[r_{mkt}] - r_f}{\sigma(r_{mkt})}\sigma_h$$

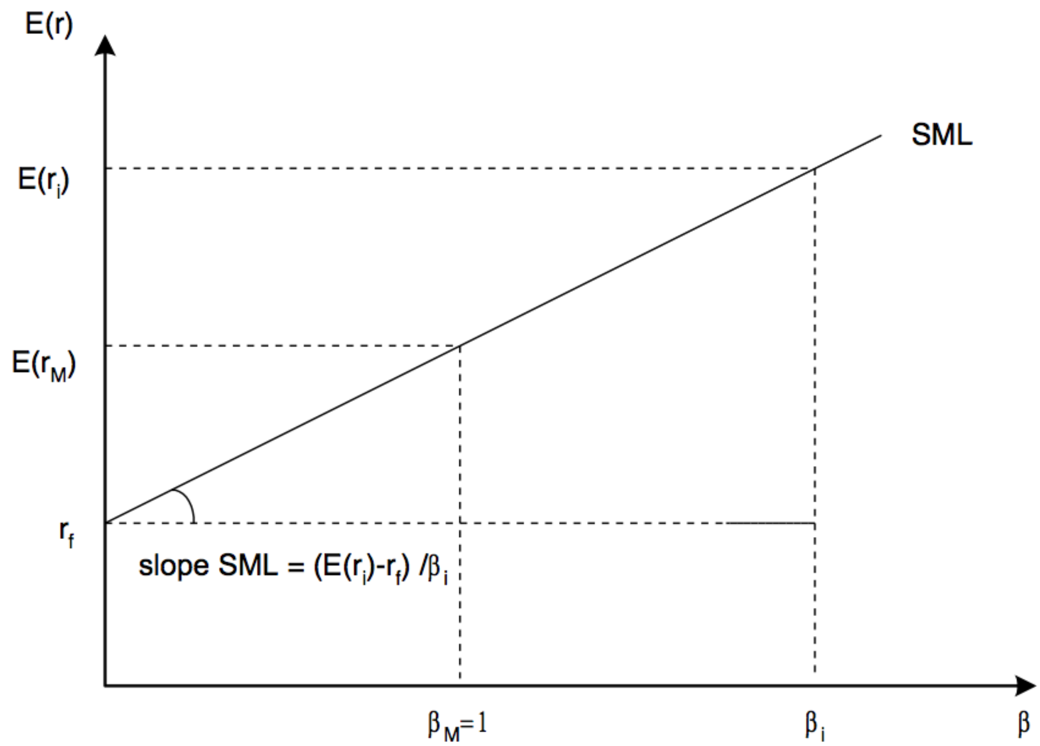
Note that the slope is the $\frac{\mathbb{E}[r_{mkt}]-r_f}{\sigma(r_{mkt})}$ sharpe ratio of the market portfolio.



Similarly, since $\frac{\sigma_h}{\sigma(r_{mkt})} = \beta_{h,mkt}$ we can write

$$\mathbb{E}[r_h] = r_f + \beta_{h,mkt} \left(\mathbb{E}[r_{mkt}] - r_f \right)$$

Which defines the *security market line* (SML)in the $(\beta, \mathbb{E}[r])$ space:



6.2 The Modern Approach

Given S states, each with probability $\pi_s > 0$, we define the *probability inner product* as

$$[x, y]_\pi \equiv \sum_{s=1}^{S} x_s y_s \pi_s = \sum_{s=1}^{S} \sqrt{\pi_s} x_s \sqrt{\pi_s} y_s = \mathbb{E}[xy]$$

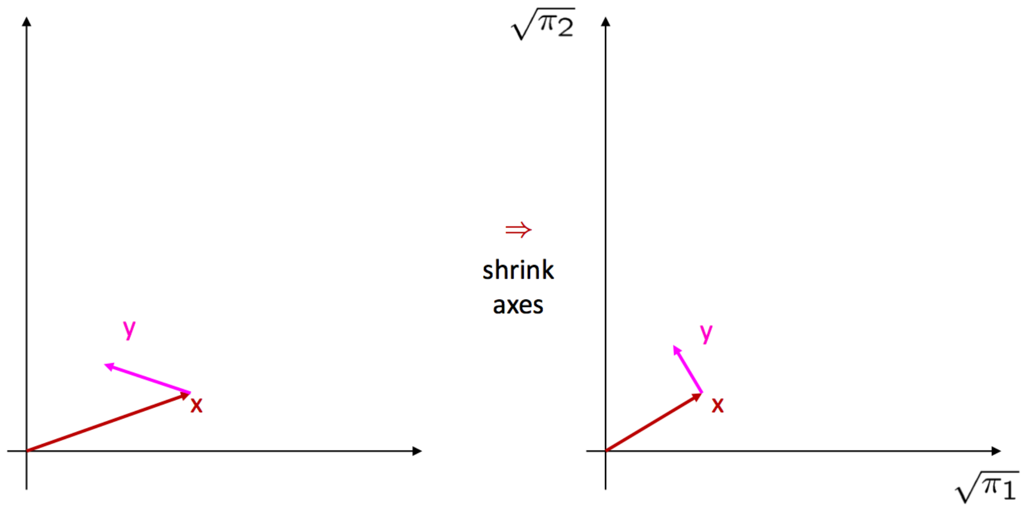
And the π-*norm* as

$$\|x\| = \sqrt{[x, x]_\pi}$$

Some of its properties are:

1. $\|x\| > 0$ for all $x \neq 0$ and $\|x\| = 0$ if $x = 0$.
2. $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$
3. $\forall x, y \in \mathbb{R}^S \ \ \|x + y\| \leq \|x\| + \|y\|$

To visualize what the probability inner product does, it is important to note that it is equivalent to an inner product in a space where the s^{th} axis has been shrinked by a factor of $\sqrt{\pi_s}$:



So that x and y are π-orthogonal if $[x, y]_\pi = 0$, that is, if $\mathbb{E}[xy] = 0$.

Let Z be the space of all linear combinations of vectors $z^1, \ldots z^n$. Given a vector $y \in \mathbb{R}^S$, the solution to the minimization problem

$$\min_{\alpha \in \mathbb{R}^n} \mathbb{E}\left[\left(y - \sum_{i=1}^{n} \alpha_i z^i\right)^2\right]$$

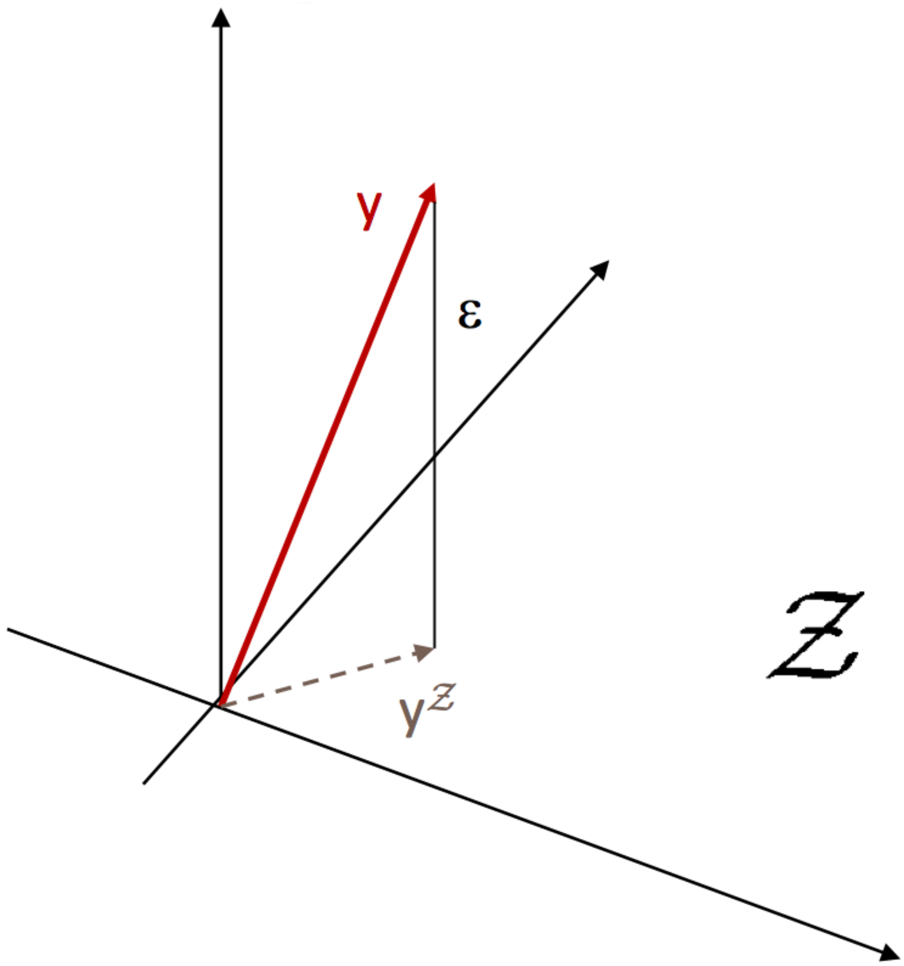
Is given by the FOC

$$\sum_{s=1}^{S} \pi_s \left(y_s - \sum_{i=1}^{n} \alpha_i z_s^i\right) \cdot z_s^i$$

for $i = 1, \ldots, n$. Let the solution be $\hat{\alpha}$, then we have $y^Z = \sum_{i=1}^{n} \hat{\alpha}_i z^i \in Z$. Define $\epsilon \equiv y - y^Z$, that is, the smallest distance between the vector y and the Z space. From the FOC we have

$$\sum_{s=1}^{S} \pi_s \left(y_s - y_s^Z\right) \cdot z_s^i = \mathbb{E}\left[z^i \epsilon\right] = 0$$

For each $i = 1, \ldots, n$: y^Z is the othogonal projection of y on Z. Moreover, y can be decomposed in two vectors, orthogonal to each other, one belonging to Z and one to the space orthogonal to Z: $y = y^Z + \epsilon$, $y^Z \in Z$ and $\epsilon \perp Z$.



Again, this language should sound familiar to that of linear regression: suppose we have S data points of the type $\{y_s, x_s^1, \ldots, x_s^K\}$ where

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_S \end{bmatrix}$$

Is the dependent variable,

$$X = \begin{bmatrix} x_1^1 & \cdots & x_1^K \\ \vdots & \ddots & \vdots \\ x_S^1 & \cdots & x_S^K \end{bmatrix}$$

Is the matrix of explanatory variables and

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$$

Is a vector of coefficients. We hypothesize the linear model

$$y = X\beta + \varepsilon$$

and therefore look for $\beta = \beta^*$ solution to the problem

$$\max_{\beta \in \mathbb{R}^k} \mathcal{L}(\beta) = \frac{1}{2}(y - X\beta)'(y - X\beta) = \frac{1}{2}\sum_{s=1}^{S}\left(y_s - \sum_{k=1}^{K}\beta_k x_s^k\right)^2$$

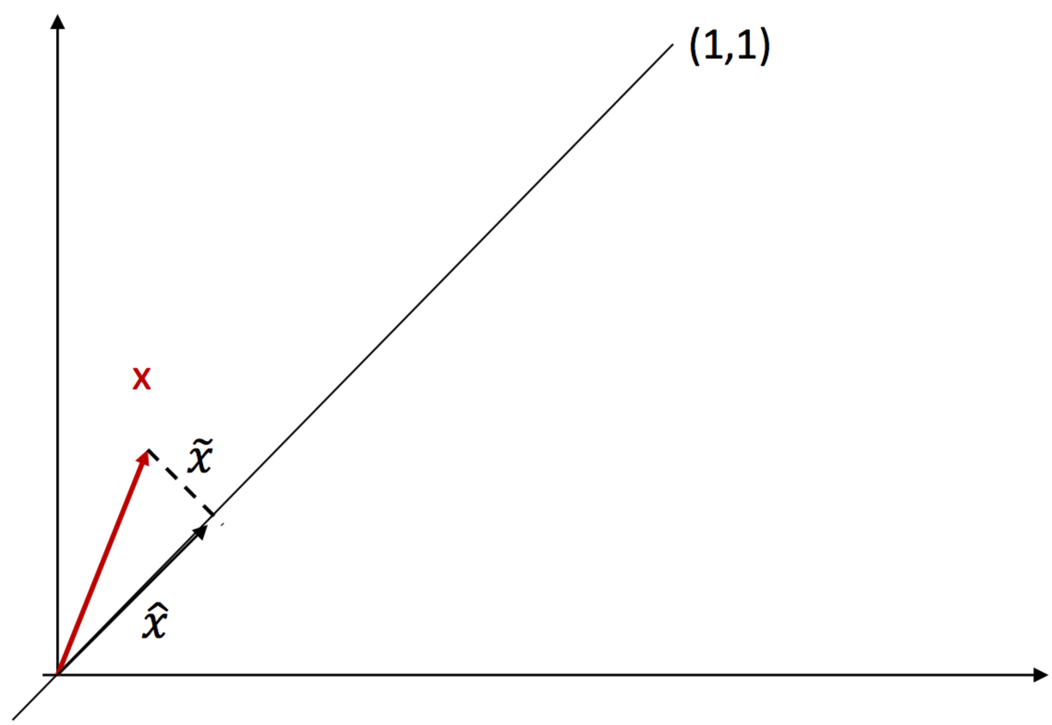
It turns out that $\beta^* = X\left([X', X]_\pi\right)^{-1}[X', y]_\pi$ and therefore $proj\,(y|X) = X\beta^* = X\left([X', X]_\pi\right)^{-1}[X', y]_\pi$, and

$$[y, y - X\beta^*]_\pi = [y, \varepsilon]_\pi = \mathbb{E}\left[\mathbb{E}\left[\varepsilon|y\right] \times y\right] = 0$$

Note that:

- $[x, y]_\pi = \mathbb{E}[xy] = Cov[x, y] + \mathbb{E}[x]\,\mathbb{E}[y]$
- $[x, x]_\pi = \mathbb{E}\left[x^2\right] = Var[x] + \mathbb{E}[x]^2$
- $\|x\| = \sqrt{\mathbb{E}[x^2]}$

Moreover, we can write x as $x = \hat{x} + \tilde{x}$, where $\hat{x}$ is a projection of x onto $\langle 1 \rangle$[4] and $\tilde{x}$ is a projection of x onto $\langle 1 \rangle^\perp$, the space orthogonal to $\langle 1 \rangle$.



[4]The space spanned by the vector $I = (1, \ldots, 1)$, so $\hat{x}$ can be written as $\hat{x} = \xi \cdot I$ where ξ is a scalar.

Moreover we have:

- $\mathbb{E}[x] = [x, I]_\pi = [\hat{x} + \tilde{x}, I]_\pi = [\hat{x}, I]_\pi = [\xi \cdot I, I]_\pi = \xi \cdot \underbrace{[I, I]_\pi}_{=1} = \xi$ constant.
- $Var[x] = \mathbb{E}\left[(x - E[x] \cdot I)^2\right] = \mathbb{E}\left[(\hat{x} + \tilde{x} - E[x] \cdot I)^2\right] = \mathbb{E}\left[(\xi \cdot I + \tilde{x} - \xi \cdot I)^2\right] = Var[\tilde{x}]$
- $\sigma_x \equiv \sqrt{Var[x]} = \|\tilde{x}\|_\pi$
- $Cov[x, y] = Cov[\tilde{x}, \tilde{y}] = [\tilde{x}, \tilde{y}]_\pi$

To see the last relation, just note that

$$Cov[x, y] = [x, y]_\pi - \mathbb{E}[x]\,\mathbb{E}[y] = [\hat{x}, \hat{y}]_\pi + [\tilde{x}, \tilde{y}]_\pi + \underbrace{[\hat{x}, \tilde{y}]_\pi}_{=0} + \underbrace{[\tilde{x}, \hat{y}]_\pi}_{=0} - \underbrace{\mathbb{E}[\hat{x}]\,\mathbb{E}[\hat{y}]}_{=[\hat{x}, \hat{y}]_\pi} = [\tilde{x}, \tilde{y}]_\pi$$

6.2.1 Pricing and Expectation Kernel

Let $\langle X \rangle$ be the space of available payoffs. We know that if there is no arbitrage and the physical probability $\{\pi_s\}$ is positive, then there exists a stochastic discount factor $m \in \mathbb{R}^S$ such that $m > 0$ and a valuation functional V such that for $z \in R^S$ $V(z) = \mathbb{E}[mz]$. Note that the stochastic discount factor m needs not be in the asset span. We defineed a pricing kernel as the state price vector q^* in the asset span $\langle X \rangle$: from now on we will equivalently refer to the pricing kernel as the stochastic discount factor m^*obtained from q^* (that is, $m^*_s = \frac{q^*_s}{\pi_s}$ for all $s = 1, \ldots, S$). Clearly, for m^* we have $V(z) = \mathbb{E}[m^* z]$ for all $z \in \mathbb{R}^S$, and if a state price vector exists, m^* is unique. Note that for any state price density $m \in \mathbb{R}^S$ and $z \in \mathbb{R}^S$ we have

$$\mathbb{E}[(m - m^*)\, z] = 0$$

And because $m = (m - m^*) + m^*$, it follows that m^* is the projection of m on $\langle X \rangle$. Clearly, if markets are complete $m = m^*$.

We define the *expectations kernel* as the vector $k^* \in \langle X \rangle$ such that for all $z \in \mathbb{R}^S$ $\mathbb{E}[z] = \mathbb{E}[k^* z]$. If the physical probability $\{\pi_s\}$ is positive the expectations kernel is unique. For any $z \in \mathbb{R}^S$ we have

$$\mathbb{E}[(I - k^*)\, z] = 0$$

So k^* is the projection of I on $\langle X \rangle$. Obviously, if a bond can be replicated then $k^* = I$. We are now ready to define the mean variance frontier in terms of projections.

Definition 1: We call the mean variance frontier the set $\mathcal{F} \equiv \{z \in \langle X \rangle \,|\, \nexists z' \in$: $\mathbb{E}[z'] = \mathbb{E}[z], V(z') = V(z$

Definition 2: We call ε the space generated by m^* and k^*.

We can decompose any vector $z \in \mathbb{R}^S$ as $z = z^\varepsilon + \epsilon$, where $z^\varepsilon \in \varepsilon$ and $\epsilon \perp \varepsilon$. Therefore $\mathbb{E}[\epsilon] = \mathbb{E}[\epsilon k^*] = 0$, $V(\epsilon) = \mathbb{E}[\epsilon m^*] = 0$ and $Cov(\epsilon, z^\varepsilon) = \mathbb{E}[\epsilon z^\varepsilon] = 0$ since $z^\varepsilon \in \varepsilon$ and $\epsilon \perp \varepsilon$. It follows that $Var[z] = Var[z^\varepsilon] + Var[\epsilon]$.

Claim: $\varepsilon = \mathcal{F}$.

We call *frontier returns* the returns of frontier payoffs with non-zero prices. Let

$$R^{m^*} = \frac{m^*}{V(m^*)} = \frac{m^*}{\mathbb{E}(m^* \cdot m^*)}$$

$$R^{k^*} = \frac{k^*}{V(k^*)} = \frac{k^*}{\mathbb{E}(k^*)}$$

If $z \in \mathcal{F}$ then we can write $z = \alpha m^* + \beta k^*$, so we can write frontie rreturns as

$$R^z = \underbrace{\frac{\alpha V(m^*)}{\alpha V(m^*) + \beta V(k^*)}}_{\equiv \lambda} R^{m^*} + \underbrace{\frac{\beta V(k^*)}{\alpha V(m^*) + \beta V(k^*)}}_{=1-\lambda} R^{k^*}$$

Graphically, frontier returns are payoffs with price equal to 1. Note that if $k^* = c \cdot m^*$ for some $c \in \mathbb{R}$, we have $R^z = R^{m^*} = R^{k^*}$. If instead $k^* \neq c \cdot m^*$ we can re-write the expression above as

$$R^z = R^{k^*} + \lambda \left(R^{m^*} - R^{k^*} \right)$$

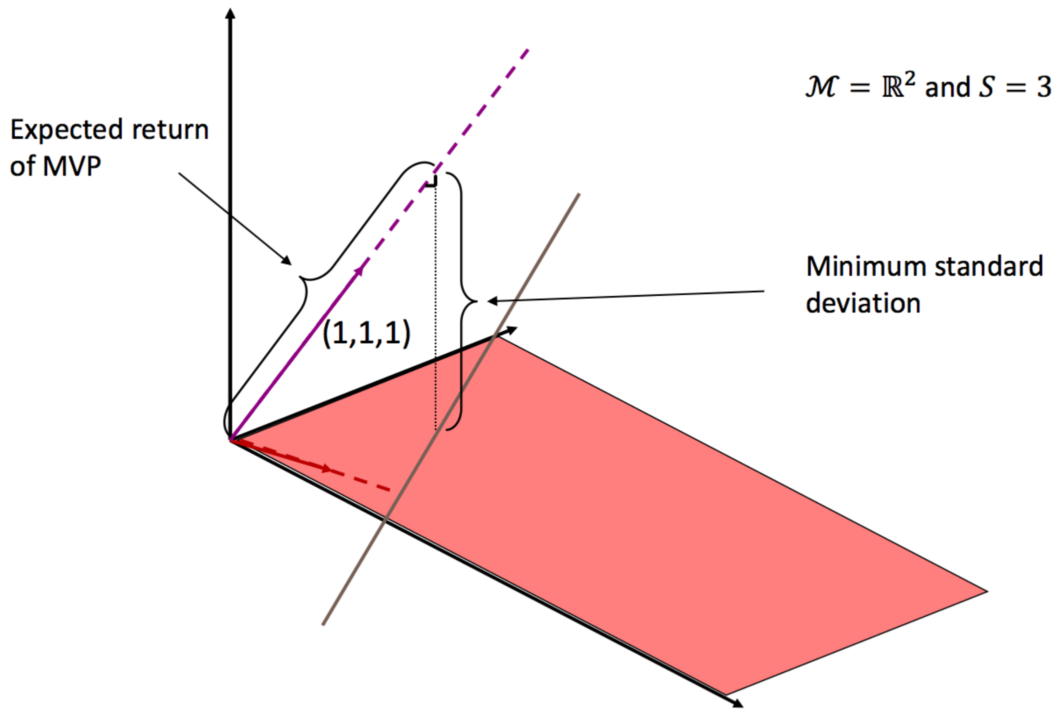
For which we have

$$\mathbb{E}[R^z] = \mathbb{E}\left[R^{k^*}\right] + \lambda \left(\mathbb{E}\left[R^{m^*}\right] - \mathbb{E}\left[R^{k^*}\right] \right)$$

and

$$Var[R^z] = Var\left[R^{k^*}\right] + 2\lambda Cov\left[R^{k^*}, R^{m^*} - R^{k^*}\right] + \lambda^2 Var\left[R^{m^*} - R^{k^*}\right]$$

The FOC with respect to λ of $Var[R^z]$ yields $\lambda_{\min} = -\frac{Cov\left[R^{k^*}, R^{m^*} - R^{k^*}\right]}{Var\left[R^{m^*} - R^{k^*}\right]}$, so the minimum varianc portfolio has a return of $R^{\min} = R^{k^*} + \lambda_{\min}\left(R^{m^*} - R^{k^*}\right)$. Graphically,



If a risk-free asset exists we have $k^* = I$ and variance and expectation of R^z simplify to:

$$Var\left[R^z\right] = \lambda^2 Var\left[R^{m^*}\right]$$

$$\mathbb{E}\left[R^z\right] = R^f + \lambda\left(\mathbb{E}\left[R^{m^*}\right] - R^f\right) = R^f \pm \sigma\left(R^z\right)\frac{\mathbb{E}\left[R^{m^*}\right] - R^f}{\sigma\left(R^{m^*}\right)}$$

> **Definition 3**: A return is *mean-variance efficient* if there is no other return with same variance but greater expectation.

Mean variance efficient returns are frontier returns with$\mathbb{E}\left[R^z\right] \geq \mathbb{E}\left[R^{\min}\right]$. If a risk-free asset can be replicated, we have:

1. Mean variance efficient returns correspond to returns with $\lambda = \lambda_{\min} = 0$
2. The portfolio with return equal to the pricing kernel is not mean variance efficient, since $\mathbb{E}\left[R^{m^*}\right] = \frac{\mathbb{E}[m^*]}{\mathbb{E}\left[(m^*)^2\right]} < \frac{1}{\mathbb{E}[m^*]} = R^f$, where the inequality follows from $Var\left[m^*\right] = \mathbb{E}\left[(m^*)^2\right] - \mathbb{E}\left[m^*\right]^2 > 0$.

Now take two frontier portfolios with returns $R^\lambda = R^{k^*} + \lambda\left(R^{m^*} - R^{k^*}\right)$ and $R^\mu = R^{k^*} + \mu\left(R^{m^*} - R^{k^*}\right)$: their covariance is given by

$$Cov\left[R^\lambda, R^\mu\right] = Var\left[R^{k^*}\right] + (\lambda + \mu)\, Cov\left[R^{k^*}, R^{m^*} - R^{k^*}\right] + \lambda\mu Var\left[R^{m^*} - R^{k^*}\right]$$

These two portfolios have zero covariance if

$$\mu = \mu_0 = -\frac{Var\left[R^{k^*}\right] + \lambda Cov\left[R^{k^*}, R^{m^*} - R^{k^*}\right]}{Cov\left[R^{k^*}, R^{m^*} - R^{k^*}\right] + \lambda Var\left[R^{m^*} - R^{k^*}\right]}$$

For all $\lambda \neq \lambda_{\min}$, μ_0 exists and clearly $\mu_0 = 0$ is a risk-free bond can be replicated.

6.2.2 Beta Pricing

Since R^λ and R^μ are frontier returns, we can write

$$R^\beta = R^\mu + \beta\left(R^\lambda - R^\mu\right) \tag{6.1}$$

Where R^β is again a frontier return. Consider any asset with payoff x^j: we know it can be decomposed in $x^j = \left(x^j\right)^\varepsilon + \epsilon^j$ with $V\left(x^j\right) = V\left(\left(x^j\right)^\varepsilon\right)$ and $\mathbb{E}\left[x^j\right] = \mathbb{E}\left[\left(x^j\right)^\varepsilon\right]$ since $\epsilon^j \perp \varepsilon$. The return of $\left(x^j\right)^\varepsilon$ is

$$R^j = \left(R^j\right)^\varepsilon + \frac{\epsilon^j}{V\left(x^j\right)}$$

Where $\left(R^j\right)^\varepsilon$ is a frontier return: using relation (1) above for $\left(R^j\right)^\varepsilon$ we can write

$$R^j = R^\mu + \beta_j\left(R^\lambda - R^\mu\right) + \frac{\epsilon^j}{V\left(x^j\right)}$$

Taking expectations we obtain the *Beta Pricing Equation*

$$\mathbb{E}\left[R^j\right] = \mathbb{E}\left[R^\mu\right] + \beta_j\left(\mathbb{E}\left[R^\lambda\right] - \mathbb{E}\left[R^\mu\right]\right)$$

Moreover, assuming $\mu = \mu_0$,

$$Cov\left[R^\lambda, R^j\right] = \beta_j Var\left[R^\lambda\right]$$

Since $R^\lambda \perp \frac{\epsilon^j}{V(x^j)}$, or $\beta_j = \frac{Cov\left[R^\lambda, R^j\right]}{Var\left[R^\lambda\right]}$.[5] If a risk-free asset exists, the beta pricing equation simplifies to

$$\mathbb{E}\left[R^j\right] = R^f + \beta_j\left(\mathbb{E}\left[R^\lambda\right] - R^f\right)$$

Note the similarity with the CAPM relation: here we have R^λ (a general frontier return) instead of the market return, which we need to identify to be able to price assets. CAPM is equivalent to this formulation with $R^\lambda = R^m$, that is, the market portfolio is a frontier return. We can derive the conditions under which the market portfolio is a frontier return: suppose there are two periods $t = 0, 1$. Agent i has individual endowmentw_1^i at time $t = 1$, and the aggregate endowment at $t = 1$

[5] Because $R^\lambda \in \mathcal{F}$ while $\epsilon^j \perp \mathcal{F}$.

is $\bar{w} = \bar{w}^{\langle X \rangle} + \bar{w}^{\perp}$ where $\bar{w}^{\langle X \rangle}$ is the orthogonal projection of $\bar{w}$ on $\langle X \rangle$ and $\bar{w}^{\perp}$ is orthogonal with respect to $\langle X \rangle$. The market payoff is $\bar{w}^{\langle X \rangle}$. Assume $V(\bar{w}^{\langle X \rangle}) \neq 0$, let $R^{mkt} = \frac{\bar{w}^{\langle X \rangle}}{V(\bar{w}^{\langle X \rangle})}$ and assume R^{mkt} is not the minimum variance return. If $R^{mkt\perp}$ is the frontier return that has zero covariance with R^{mkt}, then going back to the beta pricing equation for every security j we have

$$\mathbb{E}\left[R^j\right] = \mathbb{E}\left[R^{mkt^{\perp}}\right] + \beta_j \left(\mathbb{E}\left[R^{mkt}\right] - \mathbb{E}\left[R^{mkt^{\perp}}\right]\right)$$

With $\beta_j = \frac{Cov\left[R^{mkt}, R^j\right]}{Var\left[R^{mkt}\right]}$.[6] If a risk-free asset exists, the equation above becomes

$$\mathbb{E}\left[R^j\right] = R^f + \beta_j \left(\mathbb{E}\left[R^{mkt}\right] - R^f\right)$$

Which is the CAPM relation.

6.3 Testing CAPM

To test CAPM we simply test empirically its implications. There are two approaches: the *time-series approach* involves a regression of individual returns on the market returns, and tests for the null hypothesis that $\alpha_i = 0$:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,m}\left(R_{mkt,t} - R_{f,t}\right) + \varepsilon_{i,t}$$

On the other hand, the cross-sectional approach estimates the beta from a time series regression and then regresses the individual returns on the betas, always testing for $\alpha_i = 0$:

$$R_i = \alpha_i + \lambda \beta_{i,mkt} + \varepsilon_i$$

The empirical evidence shows that while excess returns for high-beta stocks are low, they are high for low-beta stocks (even though this effect has been weak since the early 1980s). Also, value stocks have high returns despite low betas and "momentum stocks" have high returns and low betas.

There are two main critiques to CAPM:

1. The *Roll critique* is that the CAPM is not testable because composition of true market portfolio is not observable in practice (even the S&P 500 is just an approximation).

2. The *Hansen-Richard critique* is that the CAPM could hold conditionally at each point in time, but certainly fails unconditionally.

[6]Note that the equation above always holds with two (linearly independent) assets.

Important counter-arguments include that the anomalies that make the CAPM fail empirically are the result of excessive data mining, or that these anomalies are concentrated in small and illiquid stocks, or yet that arkets are just inefficient and therefore it is impossible to test CAPM because it already implies market efficiency. Finally, there are several practical issues in estimating the regressions above: how do we estimate all the parameters we need for portfolio optimization? What is the market portfolio? Should we impose any trade restrictions? What about international assets and currency risk? Or even, should we assume a static market portfolio or one that changes over time?

6.4 Practical Issues

An investor looking to build a mean-variance portfolio has to estimate J means (one for each asset), J variances and $\frac{J(J-1)}{2}$ covariances.

6.4.1 Estimating Means

One way of eatimating means is the following: for any partition of $[0,T]$ with N intervals (so that $\Delta t = \frac{T}{N}$), $\mathbb{E}[R] = \frac{1}{\Delta t}\frac{1}{N}\sum_{i=1}^{N} R_{i\Delta t} = \frac{p_T - p_0}{N}$ (in log-prices), so knowing the first and last prices is sufficient.

A different approach is to take the log-return on the market X_k, with $k = 1, \ldots, n$ over a period of length h (that is, $\Delta = \frac{n}{h}$). We want to estimate the dynamics

$$X_k = \mu \cdot \Delta + \sigma \cdot \sqrt{\Delta}\varepsilon_k$$

Where $\varepsilon_k \overset{iid}{\sim} N(0,1)$. The standard estimator for μ is $\hat{\mu} = \frac{1}{h}\sum_{k=1}^{n} X_k$, which has mean μ and variance $\frac{\sigma^2}{h}$: remarkably, the accuracy of the estimator depends only upon the total length of the observation period h and not upon the number of observations n.

6.4.2 Estimating Variances

Using an "intuitive" estimator $\hat{\sigma^2} = \frac{1}{h}\sum_{k=1}^{n} X_k^2$ we get

$$\mathbb{E}\left[\hat{\sigma^2}\right] = \sigma^2 + \mu^2 \cdot \frac{h}{n}$$

$$Var\left[\hat{\sigma^2}\right] = \frac{2\sigma^2}{n} + \frac{4 \cdot \mu^2 \cdot h}{n^2}$$

Clearly, this estimator is biased since we did not subtract the expected return for each realization in the formula for $\hat{\sigma^2}$. The magnitude of the bias however is decreasing in n, so the accuracy can be improved my sampling the data more frequently.

Another approach is to do as we did for the means: for any partition of $[0,T]$ with N intervals (so that $\Delta t = \frac{T}{N}$), $\mathbb{E}[R] = \frac{1}{\Delta t}\frac{1}{N}\sum_{i=1}^{N} r_{i\Delta t} = \frac{p_T - p_0}{N}$. Then we take

$$Var[R] = \frac{1}{N}\sum_{i=1}^{N}\left(R_{i\Delta t} - \mathbb{E}[R]\right)^2 \to \sigma^2 \; as \; N \to \infty$$

In theory, observing the same time series at progressively higher frequencies increases the precision of the estimate. However in practice we run into a number of issues: for instance, over shorter time intervals increments are non-Gaussian; volatility is time-varying (so stochastic volatility, GARCH and other more complicated models are more appropriate) and there is always market microstructure "noise" that spoils the estimation.

6.4.3 Estimating Covariances

In theory, the estimation of covariances shares the features of variance estimation. In practice however, it is difficult to obtain synchronously observed time-series: this may require interpolation, which affects the covariance estimates. Moreover, the number of covariances to be estimated grows very quickly with the number of assets, sp that the resulting covariance matrices are oftenunstable. A possible solution is represented by "shrinkage estimators" (see Ledoitand Wolf, 2003, "Honey, I Shrunk the Covariance Matrix").

6.5 Unstable Portfolio Weights

A properly designed regression yields (positive) portfolio weights. There is a large literature on statistical tests for the significance of these weights: for instance, Britton-Jones (1999) test it for an international portfolio between 1977 and 1996 in 11 countries. The results are striking: weights vary significantly across time and in the cross section, and standard errors on coefficients tend to be very large.

	1977-1996		1977-1986		1987-1996	
	weights	t-stats	weights	t-stats	weights	t-stats
Australia	12.8	0.54	6.8	0.20	21.6	0.66
Austria	3.0	0.12	-9.7	-0.22	22.5	0.74
Belgium	29.0	0.83	7.1	0.15	66	1.21
Canada	-45.2	-1.16	-32.7	-0.64	-68.9	-1.10
Denmark	14.2	0.47	-29.6	-0.65	68.8	1.78
France	1.2	0.04	-0.7	-0.02	-22.8	-0.48
Germany	-18.2	-0.51	9.4	0.19	-58.6	-1.13
Italy	5.9	0.29	22.2	0.79	-15.3	-0.52
Japan	5.6	0.24	57.7	1.43	-24.5	-0.87
UK	32.5	1.01	42.5	0.99	3.5	0.07
US	59.3	1.26	27.0	0.41	107.9	1.53

6.6 The Black-Litterman Approach

So far, the approach to estimating portfolio weights focuses only on historical data. Since portfolio weights are very unstable, we need to somehow "discipline" our estimates. In particular, there are a few features that we would like to incorporate in our optimal weight estimation:

- Unusually high (or low) past return may not (on average) earn the same high (or low) return going forward
- Highly correlated sectors should have similar expected returns
- A "good deal" in the past (i.e. a good realized return relative to risk) should not persist if everyone is applying mean-variance optimization.

The Black-Litterman approach adds Bayesian statistics to the CAPM framework: it starts with a CAPM prior, dds "views" on the assets or the whole portfolio, and then updates weights using Bayes' rule.

6.6.1 The Black-Litterman Model

Priors: Suppose the returns of N risky assets (in matrix notation) are $r \sim N(\mu, \Sigma)$. We impose CAPM on the equilibrium risk premium of each asset, which is

$$\Pi = \gamma \cdot \Sigma \cdot w_{eq}$$

Where γ is the investors' corfficient of risk aversion and w_{eq} is the vector of equilibrium (market) portfolio weights. The investor is assumed to start with a Bayesian prior $\mu = \Pi + \epsilon^{eq}$, where $\epsilon^{eq} \sim N(0, \tau \cdot \Sigma)$: the variance of the equilibrium return estimates is assumed to be proportional to the variance of the returns (and τ is a scaling parameter).

Views: Investor views on a single asset affect many weights. Investor views regarding the performance of K portfolios (e.g. each portfolio can contain only a single asset). Let P be a $K \times N$ matrix of ortfolio weights, and let Q be a $K \times 1$ vector of views regarding the expected returns of these portfolios. Investor views are assumed to be imprecise:

$$P \cdot \mu = Q + \epsilon^{v}$$

Where $\epsilon^{v} \sim N(0, \Omega)$. Without loss of generality, Ω is assumed to be diagonal; ϵ^{eq} and ϵ^{v} are assumed to be independent. Recall that if X_1 and X_2 arenormally distributed as

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

Then the conditional distribution is given by

$$X_1|X_2 = x \sim N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

Posterior: The Black-Litterman posterior distribution of expected returns $R|Q$ is Gaussian with

$$\mathbb{E}[R|Q] = \left[(\tau \cdot \Sigma)^{-1} + P'\Omega^{-1}P\right]^{-1}\left[(\tau \cdot \Sigma)^{-1}\Pi + P'\Omega^{-1}Q\right]$$

$$Var[R|Q] = \left[(\tau \cdot \Sigma)^{-1} + P'\Omega^{-1}P\right]^{-1}$$

For example, consider the 2 assets case. You have a view on the (equally weighted) portfolio $\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 = q + \epsilon^{v}$. Then

$$\mathbb{E}[R|Q] = \left[(\tau \cdot \Sigma)^{-1} + \frac{1}{2\Omega}\right]^{-1}\left[(\tau \cdot \Sigma)^{-1}\Pi + \frac{q}{2\Omega}\right]$$

$$Var[R|Q] = \left[(\tau \cdot \Sigma)^{-1} + \frac{1}{2\Omega}\right]^{-1}$$

There are several advantages to this approach. First, returns are adjusted only partially toward the investor's views using Bayesian updating, thus recognizing that views may be due to estimation error and only highly precise or confident views are weighted heavily. Second, returns are modified in a way that is consistent with economic priors: highly correlated sectors have returns modified in the same direction. Finally, returns can be modified to reflect absolute or relative views, and the resulting weight have reasonable values and do not load up on the estimation error.

Exercises

1) Consider an economy in which there are three risky assets (A, B, and C) and one riskless asset. Asset A has an expected net return of 15% and the variance of its return is 0.20. Asset B has an expected net return of 20% and a variance of 0.40. Asset C has an expected net return of 55% and its variance is 0.60. The covariance between each pair of asset returns is zero. There are 3000 shares of A available in the economy and the current price of A is 20. There are 1500 shares of B available and its current price is 40. Finally, there are 10,000 shares of C outstanding and the current price of C shares is 18. Does there exists a riskless rate of return R^f so that given the current prices of the risky asset , the Sharpe-Lintner pricing equation $E\left[R^j\right] = R^f + \beta_{j,m}\left(E\left[R^{mkt}\right] - R^f\right)$? If so, what is that rate? If not, why is it impossible to find such a rate?

2) Show *as rigorously as possible* that $\varepsilon = \mathcal{F}$, that is, the space generated by the expectations and pricing kernel is the mean-variance frontier. Be careful: you need to show both $\mathcal{F} \subseteq \varepsilon$ and $\varepsilon \subseteq \mathcal{F}$.

3) A manufacturing firm has just been founded and needs to build a factory. It has a choice between two different types of factory (factory 1 and factory 2) and the cost of building each is $K_1 = 120$ and $K_2 = 50$, respectively. Only one factory can be built at this time. The table below lists the properties of the payoffs (X_1 and X_2) from the different projects, the gross market rate of return (R^{mkt}) and the gross risk free rate (R^f). You may assume that all the assumptions of the CAPM hold and that the management seeks to maximize the market value of the firm at all times. To simplify calculations, the net risk free rate of return is set to zero. You may also assume that any investments the firm makes have a negligible effect on the market portfolio.

$$\begin{array}{ccc} E\left[X_1\right] = 300 & E\left[X_2\right] = 150 & E\left[R^{mkt}\right] = 1.1 \\ Var\left[X_1\right] = 10,000 & Var\left[X_2\right] = 2,500 & Var\left[R^{mkt}\right] = 0.01 \\ Corr\left[X_1, X_2\right] = 0.4 & Corr\left[X_1, R^{mkt}\right] = 0.8 & Corr\left[X_2, R^{mkt}\right] = 0.2 \end{array}$$

a) In a meeting, an executive argues as follows: Since the expected (net) return on factory 2 is higher (200% vs. 150%), that one should be built. Another executive disagrees and argues that naturally one should always build factory 1, since the expected profit is larger than if factory 2 were built (180 vs. 100). Which (if any) of the two decision procedures suggested will maximize the firm value? Why?

b) In the end, the executives ask you to decide which factory to build. What is your decision? Compute the firm's total value V. If 100 shares are outstanding, what is the stock price P? (Assume that CAPM holds).

c) Suppose factory 2 has been built (do not assume this was the correct answer to part b)!) and a year has passed. It is time to expand, and the firm plans to build factory 1. However, it has no cash to fund the investment and needs to raise $K_1 = 120$ in the capital market. The firm decides to issue shares for a total value of 120. How many shares should it issue and at what price?

d) In a shareholder meeting, there is a heated argument about whether or not to go ahead with the building of the factory and the stock issue. An angry shareholder claims that although expected returns are impressive, factory 1 is an extremely risky investment, not least because most of its risk is not diversifiable. She argues that exposing the firm to this risk will reduce the value of her shares. A concerned executive responds by pointing out that since the values of the two factories are negatively correlated, the project does not really increase the riskiness of the firm's portfolio of assets that much and is therefore a good investment. What, in your opinion, is missing from the analysis by the shareholder and the executive? Are their arguments valid?

4) Assume that the market return is in the mean variance frontier. Suppose that there are three equally likely states. One security is risk free and has a gross return $R^f = 0.9$. A second security has a gross *return* vector of $(0, 1, 1)$. The market portfolio has a *payoff* of $(1, 2, 2)$.

a) Use the CAPM to price a security that pays off $(0, 0, 1)$.

b) Find the expectations kernel and pricing kernel.

c) Calculate the price of the same security (that pays off $(0, 0, 1)$) using the pricing kernel.

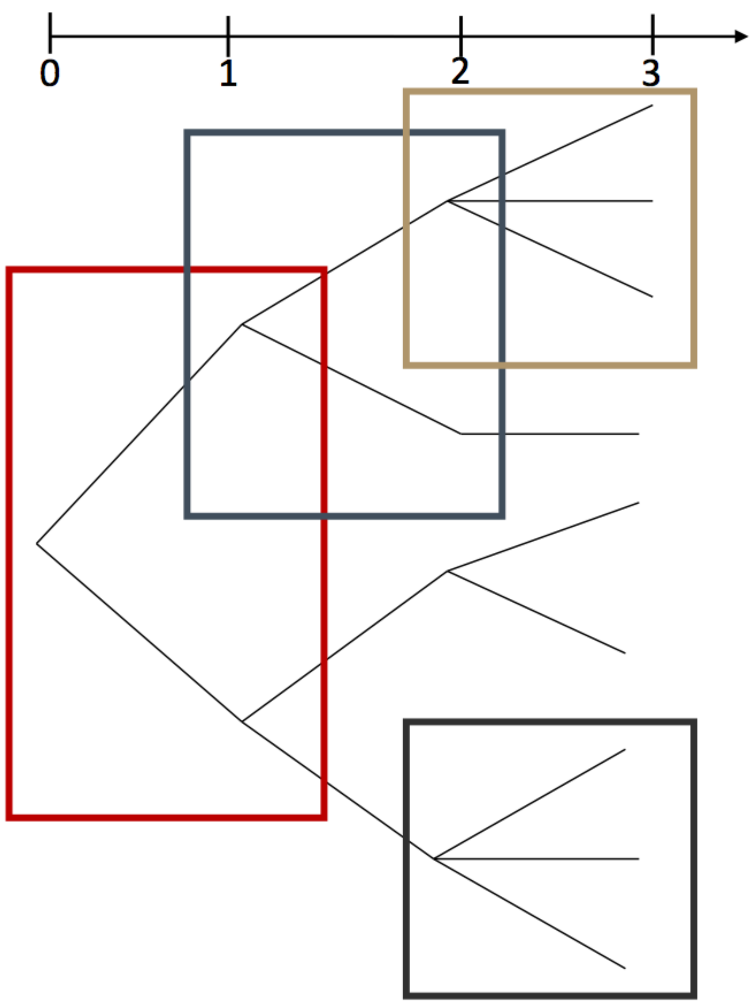
5) Show that $\mathbb{E}\left[\hat{\sigma^2}\right] = \sigma^2 + \mu^2 \cdot \frac{h}{n}$ and $Var\left[\hat{\sigma^2}\right] = \frac{2\sigma^2}{n} + \frac{4 \cdot \mu^2 \cdot h}{n^2}$.

6) Derive the expressions for the mean and variance of the Black-Litterman posterior distribution of expected returns $R|Q$.

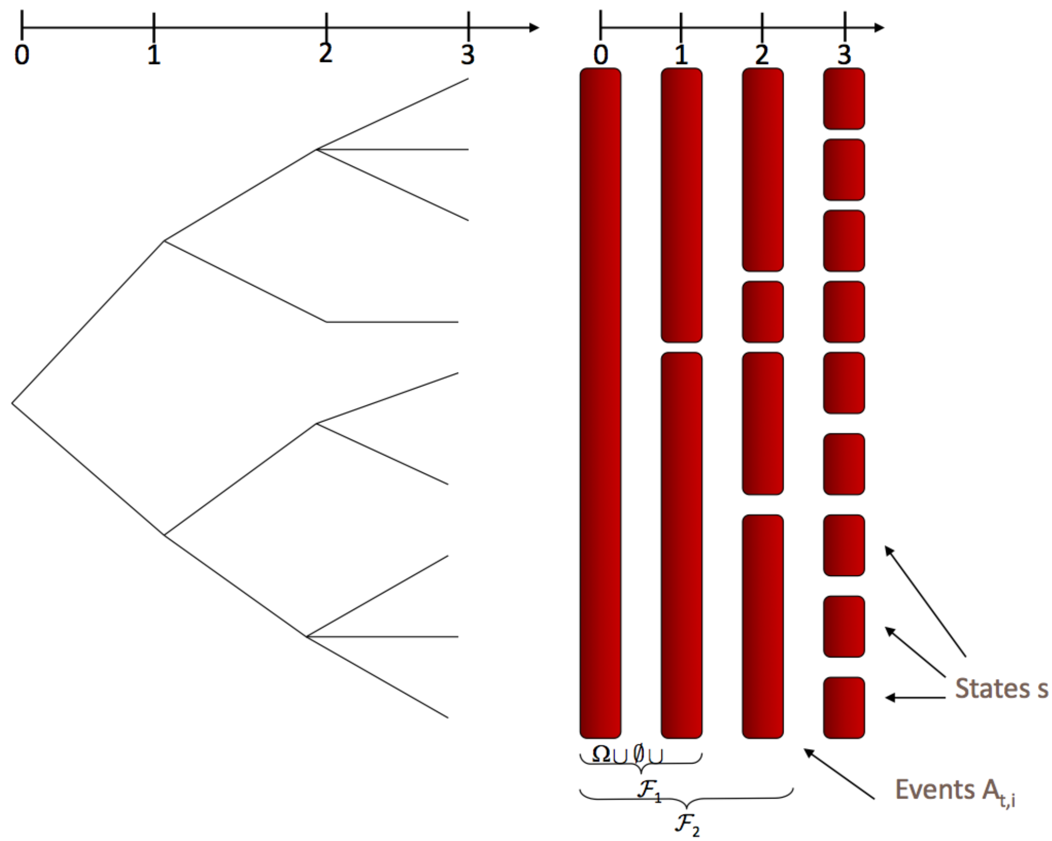
Chapter 7

The Multi-Period Model

How can we deal with a multi-period setup in which several outcomes are possible after each outcome realizes? One way to tackle this problem is to bring down the problem to a set of many “smaller” one-period model problems:



With this approach however, we need a more efficient way to define and work with the information available at each time step. A natural way to approach this is using the concept of *algebra*:



7.1 Model Setup

We begin by defining an algebra:

Definition 1: Given a state space Ω, an *algebra* $\mathcal{F}$ is a non-empty collection of subsets of Ω such that 1) $\emptyset \in \mathcal{F}$ 2) $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$ 3) $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$.

Definition 2: A random variable Y is *measurable* with respect to the algebra $\mathcal{F}$ if $\{\omega \in \Omega | Y(\omega) \leq y\} \in \mathcal{F}$ for all $y \in \mathbb{R}$.

Definition 3: A *stochastic process* is a collection of random variables $\{Y_t\}_{t=0}^T$.

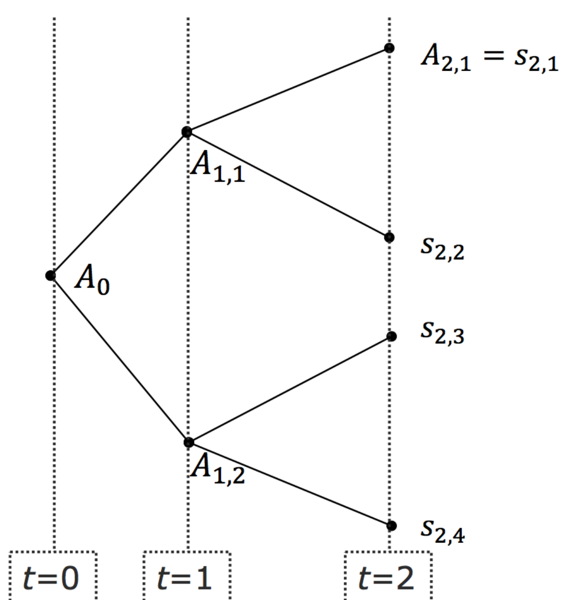
Definition 4: A filtration is a collection of algebras $\mathcal{F}_T = \{\mathcal{F}_s\}_{s=t}^T$ such that if $u \leq v$ then $\mathcal{F}_u \subseteq \mathcal{F}_v$.

Definition 5: A stochastic process $\{Y_t\}_{t=0}^T$ is *adapted* to the filtration $\mathcal{F}_T = \{\mathcal{F}_s\}_{s=t}^T$ if Y_s is measurable with respect to $\mathcal{F}_s$ for $s = t, \ldots, T$.

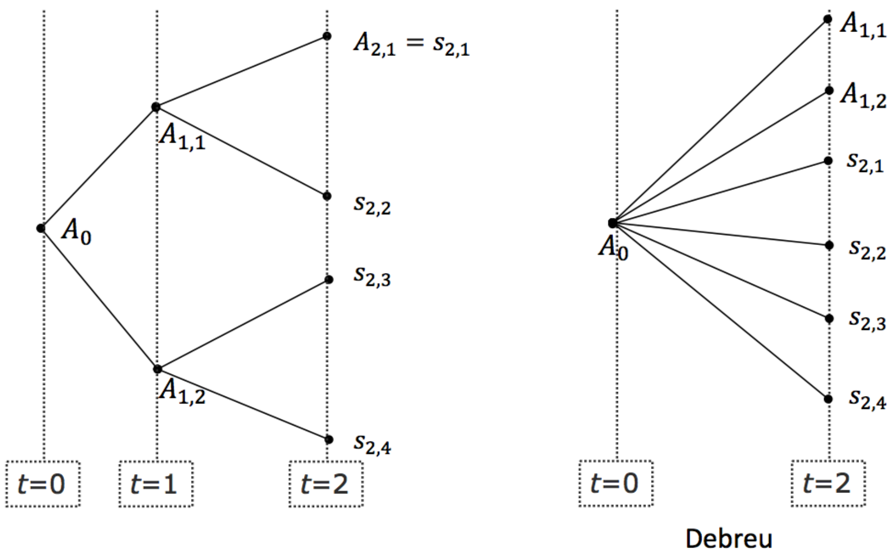
Adaptation is the requirement that people cannot see in the future. Moving from a static to a dynamic setting, we need to adapt our notation:

- Rather than asset holdings, we will talk about dynamic strategies
- Instead of asset payoff vectors x we will talk about the "next period payoff" $x_{t+1} + p_{t+1}$ of a strategy
- Instead of the asset span, we will talk about the subset of marketed dynamic strategies
- Market completeness can be interpreted in two ways: static completeness (Debreu completeness) and dynamic completeness (Arrow completeness)
- We defined no arbitrage with respect to asset holdings, but now it will be defined on dynamic strategies
- States $s \in \{s_1, s_2, \dots s_S\}$ generalize to *events* $A_{t,i}$ and final states $s_{t,i}$
- State prices q_s become *event* prices $q_{t,i}$
- The risk-free rate R_t^f varies over time
- the discount factor ρ_t is time-dependent and discounts from t to 0
- The risk-neutral probability is $\pi^{\mathbb{Q}}(A_{t,i}) = \frac{q_{t,i}}{\rho_t}$
- The pricing kernel now reads $M_t p_t^j = \mathbb{E}_t \left[M_{t+1} \left(p_{t+1}^j + x_{t+1}^j \right) \right]$, and $M_t = R_{t+1}^f \mathbb{E}_t \left[M_{t+1} \right]$

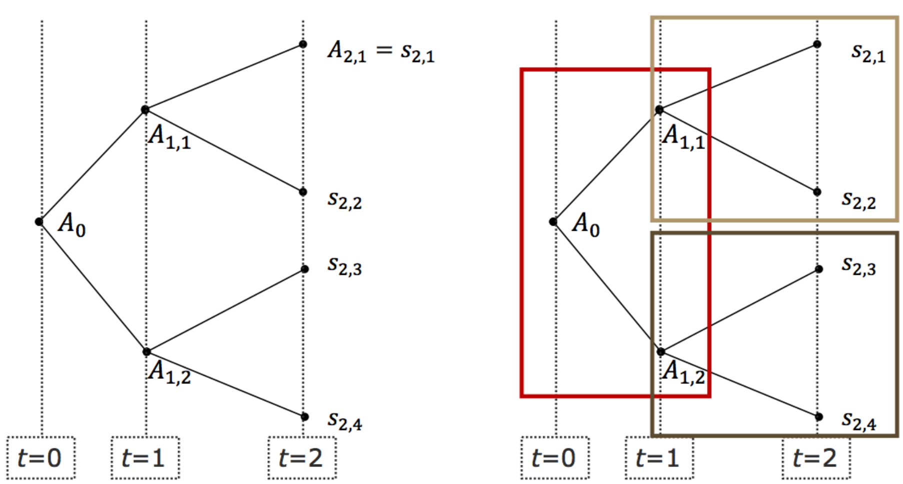
We will use trees to illustrate dynamics. In the picture below, the tree has 7 events. Every event has a positive probability attached to it, and we impose the consistency requirement that the probability of any event equals the sum of the events that follow from its node: for example, denoting $\pi_{t,s}$ the probability of $A_{t,s}$, we have $\pi_{1,1} = \pi_{2,1} + \pi_{2,2}$. Clearly, for each t we have $\sum_s \pi_{t,s} = 1$.



There are two ways to reduce this seemingly more complicated model to the one-period model. The first is to boil down every event to a final state, and apply what we know from the one-period model:



Note that if we simplify the tree in this way it follows that to achieve completeness we will need six independent assets, one for each final state. The second way is to reduce each branch of the tree to a one-period model:



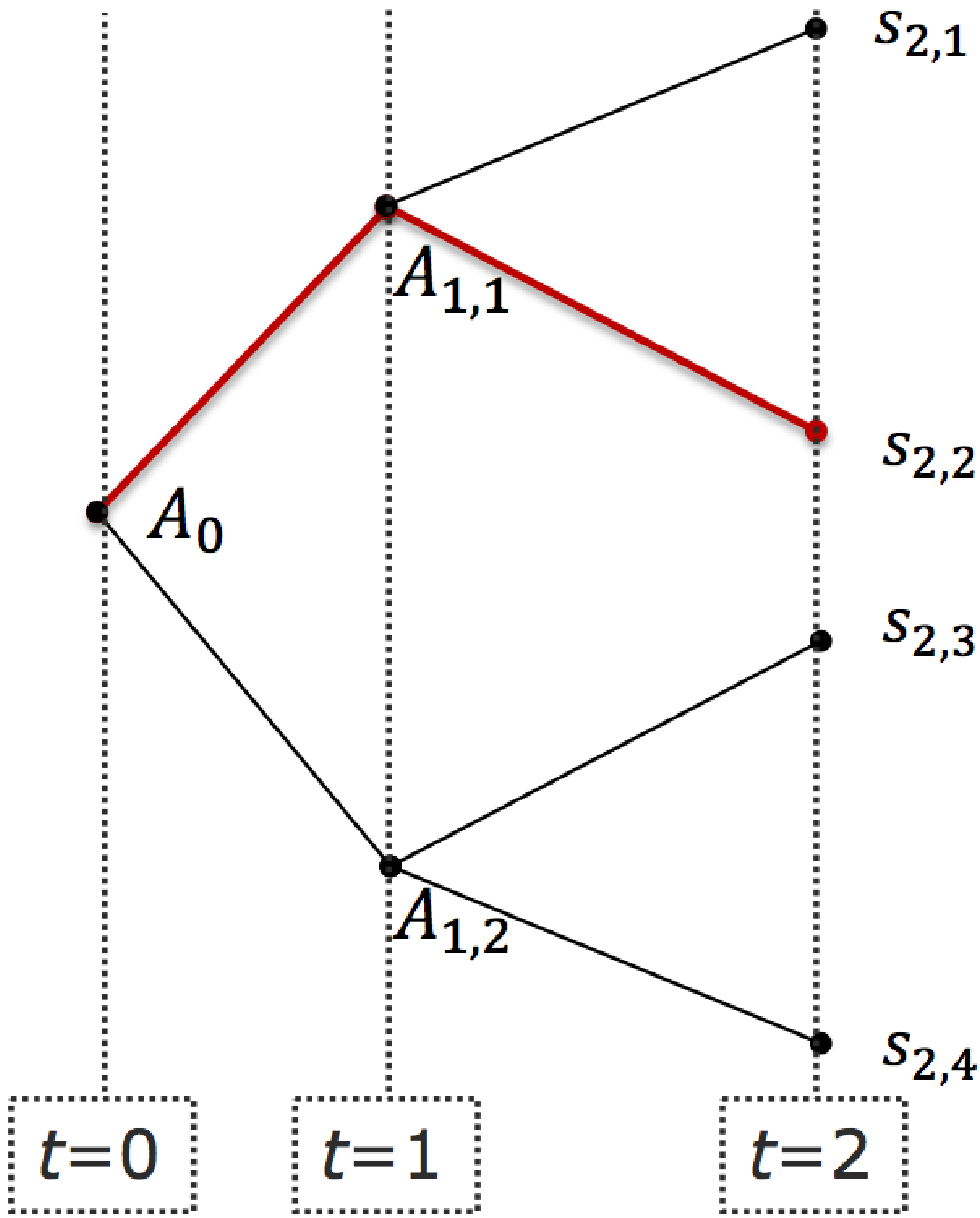
In this case, assets can be re-traded conditionally on the the occurrence of event $A_{1,1}$ or $A_{1,2}$. Completion can be achieved with two categories of assets: short-lived assets pay off only in the period following the one in which they are traded, while long-lived assets pay off over many periods.

7.2 Dynamic Trading and Market Completeness

Suppose now that asset can be traded in each period. This allows for dynamic completion, as well as for the existence of bubbles and ponzi schemes in an infinite horizon setting. We will study completeness with both short- and long-lived assets.

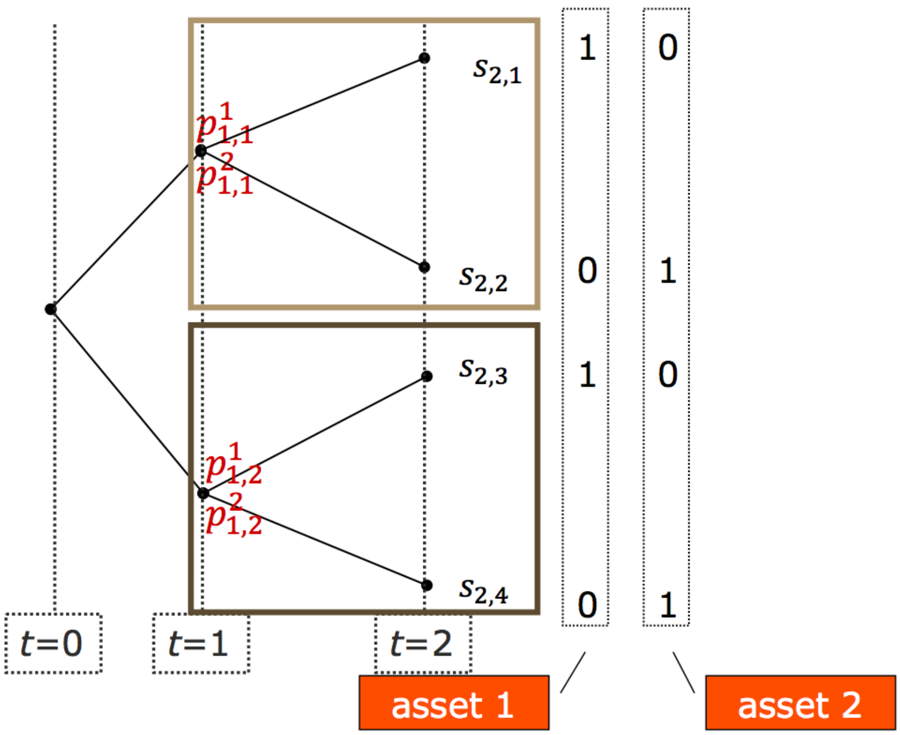
A short-lived asset is an asset that pays out only in the period immediately after the asset is issued. Without uncertainty and with T one-period assets (that is, T assets that can be bought in period $t-1$ at price p_t and pays off in period t for $t = 1, \ldots, T$), completeness requires that we are able to transfer wealth between ant two periods. For example, if zero coupon bonds are traded in each period and $T = 2$, at time $t = 0$ we can buy a quantity p_2 of bonds that pay off at time $t = 1$ at a total cost of $p_1 p_2$, and then in period $t = 1$ we can invest the proceeds from the bond we bought to buy one unit of the bond that pays at time $t = 2$: in this way we replicated a bond that is bought at $t = 0$ and pays \$1 at time $t = 2$, and therefore markets are complete. In general, for T periods the cost of rolling over this kind of short-lived bond strategy is just $p_t \cdot p_{t+1} \cdot \ldots \cdot p_{T-1}$, and completeness requires that we are able to transfer wealth between any two periods t and t' (not just consecutive).

With uncertainty the reasoning is very similar. Suppose the event tree is as follows:

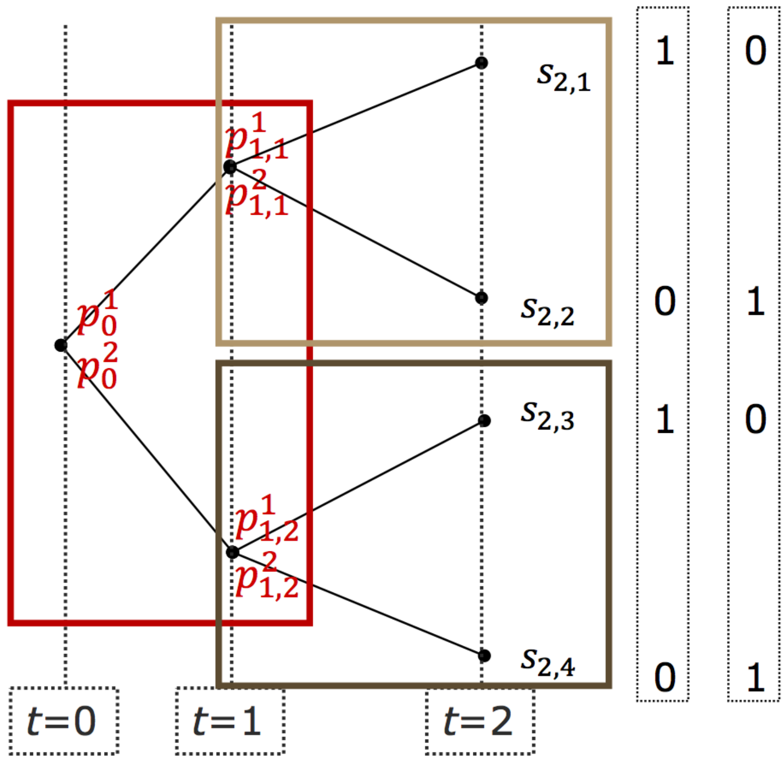


And we want to replicate the payoff of an asset that is bought at $t = 0$ and pays \$1 if event $A_{2,2} = s_{2,2}$ realizes. Let $q_{2,2}$ denote the state price corresponding to $A_{2,2}$ traded in $t-1$. Then, working backwards, in event $A_{1,1}$ we can buy a one-period asset that pays \$1 in event $A_{2,2}$ for a price $q_{2,2}$; to be able to dispose of an amount $q_{2,2}$ if event $A_{1,1}$ realizes, in event A_0 we need to buy a quantity $q_{2,2}$ of the asset that pays \$1 if event $A_{1,1}$ realizes, which costs $q_{1,1}$ per unit. Thus the price of the payoff of an asset that is bought at $t = 0$ and pays \$1 if event $A_{2,2}$ realizes is $q_{1,1} \cdot q_{2,2}$, and we can repeat the argument for all other events to see that this market achieves dynamic completeness.

With long-lived assets, consider a T-period model without uncertainty. There is a single asset paying \$1 in $t = T$, which is tradable in each period for a price p_t: this means that *sequentially*, there are T prices for this asset. The payoff can be transferred from period s to period $t > s$ by purchasing the asset at time s and selling it in period t. For example, suppose there are two assets that pay off like in the graph below.



The assets are long-lived, since they can be re-traded at $t = 1$ for the price $p^j_{t,s}$. Note that we have two assets, but each assets can be traded in three events; moreover, the price of each asset can be endogenously determined at time $t = 0$. Although one may be tempted to think that because we only have 2 assets markets cannot be complete, in fact dynamic trading provides a way to transfer wealth in all events:



Call *trading strategy* $[j, A_{t,i}]$ the cash flow of asset j that is purchased in event $A_{t,i}$ and sold one period later. There are six such trading strategies: $[1, A_0]$, $[1, A_{1,1}]$, $[1, A_{1,2}]$, $[2, A_0]$, $[2, A_{1,1}]$, $[2, A_{1,2}]$. For example, trading strategy $[1, A_{1,1}]$ costs $p^1_{1,1}$ and pays out \$1 in the first final state and zero in all other events; and trading strategy $[1, A_0]$ costs p^1_0 and pays out $p^1_{1,1}$ in event 1, $p^1_{1,2}$ in event 2, and zero in all the final states.

These six trading strategies give rise to a 6×6 payoff matrix:

Strategy	$[1, A_0]$	$[2, A_0]$	$[1, A_{1,1}]$	$[2, A_{1,1}]$	$[1, A_{1,2}]$	$[2, A_{1,2}]$
Event A_0	$-p^1_0$	$-p^2_0$	0	0	0	0
Event $A_{1,1}$	$p^1_{1,1}$	$p^2_{1,1}$	$-p^1_{1,1}$	$-p^2_{1,1}$	0	0
Event $A_{1,2}$	$p^1_{1,2}$	$p^2_{1,2}$	0	0	$-p^1_{1,2}$	$-p^2_{1,2}$
State $s_{2,1}$	0	0	1	0	0	0
State $s_{2,2}$	0	0	0	1	0	0
State $s_{2,3}$	0	0	0	0	1	0
State $s_{2,4}$	0	0	0	0	0	1

The payoff matrix is regular (and hence markets are complete) if and only if the submatrix

$$\begin{array}{|cc|} \hline p^1_{1,1} & p^2_{1,1} \\ p^1_{1,2} & p^2_{1,2} \\ \hline \end{array}$$

Is regular (that is, it has rank equal to 2). The components of this submatrix are the prices of the two assets conditional on period-1 events. There are cases in which these are collinear in equilibrium:

for instance when the per capita endowment is the same in event $A_{1,1}$ and $A_{1,2}$, in state $s_{2,1}$ and $s_{2,3}$, and in state $s_{2,2}$ and $s_{2,4}$, and if the probability of reaching state $s_{1,1}$ after event $A_{1,1}$ is the same as the probability of reaching state $s_{2,3}$ after event $A_{1,2}$ then the submatrix is singular (but, one might argue, then events $A_{1,1}$ and $A_{1,2}$ are effectively equivalent and could be collapsed into a single event). Moreover, a random square matrix is regular: this means that *generically,* the market is dynamically complete (that is, it is not complete only for information trees and asset structures which have zero probability of showing up in any application).

We call *branching number* the maximum number of branches that fan out from any event in the uncertainty tree: it turns out that this is also the number of assets necessary to achieve dynamic completeness. This generalizes to the continuum of events and continuous time case, in which a small number of assets is sufficient to achieve completeness because the need for many assets to achieve completeness that arises from a large number of possible events is offset by the ability to trade continuously, thus generating a large number of trading strategies. The classic example is the Black-Scholes formula: Cox, Ross and Rubinstein showed that is is possible to build a binomial tree model of Black-Scholes, with constant interest rates and in which the stock can only go up or down, in which the market is dynamically complete with just two assets: the stock and the risk-free bond. Thus an option on the stock can be replicated with dynamic delta-hedging.

7.3 The Multi-Period Stochastic Discount Factor

The idea of no arbitrage extends naturally to the multi-period setting: an arbitrage is a strategy that has either no cost today but some positive payoff along the tree or a negative cost today and no negative payoff along the tree. Moreover, it is equivalent to having no static arbitrage in each branch of the tree.

Recall that no arbitrage requires that there esxists a positive stochastic discount factor m_{t+1} for each subperiod t such that

$$p_t = \mathbb{E}_t \left[m_{t+1} \left(p_{t+1} + x_{t+1} \right) \right]$$

Define the multi-period stochastic discount factor as[1]

$$M_{t+1} = m_1 \cdot m_2 \cdot \ldots \cdot m_{t+1}$$

Multiplying each side by $m_1 \cdot m_2 \cdot \ldots \cdot m_t$ (which is measurable until the time t filtration) we get

$$M_t p_t = \mathbb{E}_t \left[M_{t+1} \left(p_{t+1} + x_{t+1} \right) \right]$$

Assuming for this asset there is a stream of cash flows $\{x_t\}_{t=1}^{\infty}$ where each x_t is a random payoff,

[1] We assume $m_0 = 1$ from now on.

we have for time $t = 0$

$$p_0 = \mathbb{E}_0 \left[M_1 \left(p_1 + x_1 \right) \right]$$

similarly

$$M_1 p_1 = \mathbb{E}_1 \left[M_2 \left(p_2 + x_2 \right) \right]$$

plugging in the equation for p_0 we have

$$p_0 = \mathbb{E}_0 \left[\mathbb{E}_1 \left[M_2 \left(p_2 + x_2 \right) \right] + M_1 x_1 \right] = \mathbb{E}_0 \left[\mathbb{E}_1 \left[M_2 p_2 + M_2 x_2 + M_1 x_1 \right] \right]$$

The Law of Iterated Expectations (LIE) states that, because of the property of filtrations that if $u \leq v$ then $\mathcal{F}_u \subseteq \mathcal{F}_v$, for any random variable X the time-u expectation of the random variable "time-v expectation of X" is equal to the time-u expectation of X. The intuition is as follows: consider your current guess of what the weather will be on sunday, and compare it to the guess that you will have on saturday about the weather on sunday. LIE states that your current guess on your guess for the sunday weather is the same as your current guess for the sunday weather. Therefore we get

$$p_0 = \mathbb{E}_0 \left[M_2 p_2 + M_2 x_2 + M_1 x_1 \right]$$

Iterating k steps forward we have

$$p_0 = \mathbb{E}_0 \left[M_k p_k \right] + \sum_{t=1}^{k} \mathbb{E}_0 \left[M_t x_t \right]$$

Taking the limit for $k \to \infty$ (and assuming $\lim_{k \to \infty} \mathbb{E}_0 \left[M_k p_k \right] = 0$) we have

$$p_0 = \sum_{t=1}^{\infty} \mathbb{E}_0 \left[M_t x_t \right]$$

In terms of projections, recall that $m^*_{t+1} = proj \left(m_{t+1} | \left\langle X_{t+1} \right\rangle \right)$, that is, there exists some h^*_t such that $m^*_{t+1} = X_{t+1} h^*_t$ and

$$p_t = \mathbb{E}_t \left[X'_{t+1} m^*_{t+1} \right] = \mathbb{E}_t \left[X'_{t+1} X_{t+1} \right] h^*_t$$

So that

$$h^*_t = \left(\mathbb{E}_t \left[X'_{t+1} X_{t+1} \right] \right)^{-1} p_t$$

And therefore

$$m^*_{t+1} = X_{t+1} \left(\mathbb{E}_t \left[X'_{t+1} X_{t+1} \right] \right)^{-1} p_t$$

Clearly we also have $M^*_t \equiv m^*_1 \cdot m^*_2 \cdot \ldots \cdot m^*_t \in \left\langle X_t \right\rangle$. Further, we can express m^*_{t+1} in terms of the covariance of returns Σ_t:

$$m^*_{t+1} = \mathbb{E}_t\left[m^*_{t+1}\right] + \left[p_t - \mathbb{E}_t\left[m^*_{t+1}\right]\mathbb{E}_t\left[X_{t+1}\right]\right]' \Sigma_t^{-1}\left[X_{t+1} - \mathbb{E}_t\left[X_{t+1}\right]\right]$$

Ans also in terms of covariance of excess returns $\Psi_t \equiv Cov_t\left[R^e_{t+1}\right]$:

$$m^*_{t+1} = \frac{1}{R^f_t} - \frac{1}{R^f_t}\mathbb{E}_t\left[R^e_{t+1}\right]' \Psi_t^{-1}\left[R^e_{t+1} - \mathbb{E}_t\left[R^e_{t+1}\right]\right]$$

Compare it to its continuous-time analogous[2]

$$\frac{dm^*_t}{m_t} = -r^f dt - \left(\mu + \frac{D}{p} - r^f\right)' \Psi_t^{-1} dZ_t$$

Recall that if the one-period stochastic discount factor m_t is not time-varying (that is, the distribution of m_t is i.i.d for each t), then the expectations hypothesis holds and the investment opportunity set does not vary. The corresponding R^*_t of a *single-factor* state-price beta model is then relatively easy to estimate, since over time we collect more and more realizations of R^*_t. However, if m_t (or the correspondingR^*_t) is time-varying, then we can assume that it depends on some *state variable* and need a *multi-factor* model to account for it. Indeed, suppose that $R^*_t = R^*(z_t)$, where z_t is some state variable. For example, suppose that z_t can be either 1 or 2 with equal probability. Then we could take all the periods in which $z_t = 1$ and back out $R^*(1)$, and find $R^*(2)$ in a similar way. Is it possible to do so in reality? The answer is no, because we have hedges *across* state variables.

7.4 Martingales

Let $\{X_t\}$ be a stochastic process and $\{x_t\}$ a sequence of its realizations. We say that $\{X_t\}$ is a *martingale* if $\mathbb{E}\left[X_{t+1}|X_t = x_t, X_{t-1} = x_{t-1}, \ldots, X_1 = x_1\right] = x_t$. Paul Samuelson argued in 1965 that asset prices have to be martingales in equilibrium:

$$p_t = \frac{1}{1 + r^f_t}\mathbb{E}^{\mathbb{Q}}_t\left[p_{t+1} + d_{t+1}\right]$$

this is trivial assuming assuming no dividends, no discounting and risk-neutral agents. Let's examine these cases one by one.

With discounting, in order for the price process to follow a martingale we simply need:

$$\mathbb{E}_t\left[\delta p_{t+1}\right] = p_t$$

With dividend payments, things are more complicated because p_t depends on the dividend of the

[2] Where Z_t is a Brownian Motion.

asset in period $t+1$ but p_{t+1} does not (it is the ex-dividend price). However, for example, the value of a fund that keeps reinvesting the dividends follows a martingale (LeRoy 1989): suppose the fund owns nothing but one unit of asset j. Assuming agents are risk-neutral, the value of the fund at time 0 is

$$f_0 = p_0 = \mathbb{E}\left[\sum_{t\geq 1} \delta^t x_t^j\right] = \delta \mathbb{E}\left[p_1 + x_1^j\right]$$

After receiving a (state contingent) dividend x_1^j, the fund buys more of asset j at price p_1, thus owning $1 + \frac{x_1^j}{p_1}$ units of the asset. The expected discounted value of the fund is then

$$\mathbb{E}\left[f_1\right] = \mathbb{E}\left[\delta p_1 \left(1 + \frac{x_1^j}{p_1}\right)\right] = \mathbb{E}\left[\delta\left(p_1 + x_1^j\right)\right] = p_0 = f_0$$

So the discounted expected value of the fund is a martingale.

A similar statement is true of the agent is risk-averse. The difference is that we have to discount with the risk-free interest rate instead of the agent's discount factor, and use risk-neutral probabilities (by now you should have an idea about why it is also called "equivalent martingale") instead of objective probabilities. Just like in the one-period model, we define the risk-neutral probability of event A as

$$\pi^{\mathbb{Q}}\left(A_{t,s}\right) = \frac{\pi_{t,s} M_{t,s}}{\rho_t}$$

Where $\rho_t = \left(\prod_{s=0}^{t} R_s^f\right)^{-1}$ is the risk-free discount factor between 0 and t. The initial value of the fund is

$$f_0 = p_0 = \mathbb{E}\left[\sum_{t\geq 1} M_t x_t^j\right] = \mathbb{E}\left[M_1 x_1^j + \sum_{t\geq 2} M_t x_t^j\right] = \mathbb{E}\left[m_1\left(x_1^j + \sum_{t\geq 2}\prod_{s=2}^{t} m_t x_t^j\right)\right] =$$

$$= \mathbb{E}\left[m_1\left(x_1^j + p_1\right)\right] = \mathbb{E}\left[M_1\left(x_1^j + p_1\right)\right]$$

Under the risk-neutral measure, this can be rewritten as

$$f_0 = \rho_1 \mathbb{E}^{\mathbb{Q}}\left[x_1^j + p_1\right] = \rho_1 \mathbb{E}^{\mathbb{Q}}\left[f_1\right]$$

Therefore, the properly discounted (ρ instead of δ) and properly expected ($\pi^{\mathbb{Q}}$ instead of π) value of the fund is indeed a martingale.

We can now introduce a *fifth* asset pricing formula. Let $P_s(t,T)$ be the time-s price of a (zero-coupon) bond to be purchased at time t and with maturity T. We know the pricing relation is

$$P_t(t,T) = \mathbb{E}_t\left[m_{t+1} P_{t+1}(t+1,T)\right]$$

Dividing side by side the pricing relation for a generic asset j and rearranging we get

$$\frac{p_t^j}{P_t(t,T)} = \mathbb{E}_t\left[\frac{P_{t+1}(t+1,T)\,m_{t+1}\left(x_{t+1}^j + p_{t+1}^j\right)}{P_{t+1}(t+1,T)\,\mathbb{E}_t\left[m_{t+1}P_{t+1}(t+1,T)\right]}\right] = \mathbb{E}_t^{\mathbb{F}_T}\left[\frac{x_{t+1}^j + p_{t+1}^j}{P_{t+1}(t+1,T)}\right]$$

Where the expectation is taken with respect to the risk-forward measure $\mathbb{F}_T$ for which $\pi_s^{\mathbb{F}_T} = \frac{\pi_s P_{t+1}(t+1,T)m_{t+1}}{\mathbb{E}_t[m_{t+1}P_{t+1}(t+1,T)]}$ (note that $\pi_s^{\mathbb{F}_T} \geq 0$ and $\sum_{s=1}^{S}\pi_s^{\mathbb{F}_T} = 1$ guarantee that $\pi^{\mathbb{F}_T}$ is a probability distribution). This asset pricing formula was speficically devised to price bond options, and boils down to the risk-neutral measure if $t+1 = T$ like in the one-period model case.

7.5 Ponzi Schemes and Rational Bubbles

Allowing for an infinite horizon allows agents to borrow an arbitrarily large amount and roll over debt forever, without ever repaying the debt. This is referred to as a *Ponzi scheme*: it allows infinite consumption. Consider a model with infinite horizon, no uncertainty and a complete set of short-lived bonds. Let z_t be the amount of bonds in the portfolio maturing in period t, and let p_t be the price of one such bond in period $t-1$ (just like before). The consumption path with $c_t = y_t + 1$ is feasible: note that the agent consumes more than his endowment in each period, forever! This can be financed with increasing debt:

$$z_{t+1} = \frac{z_t - 1}{p_{t+1}}$$

For all $t \geq 0$ and $z_0 = 0$. However, Ponzi schemes can never be part of an equilibrium because they remove the existence of a utility maximum, since the choice set of the agent is unbounded. Therefore we need an additional constraint, called *transversality condition*:

$$\lim_{t\to\infty} p_t z_t \geq 0$$

Which implies that the value of debt cannot diverge to infinity, or equivalently, that all debt must eventually be redeemed. There are also additional solutions to this problem that incorporate a "bubble" component to the asset price: consider again a model with no uncertainty and a consol delivering \$1 in each period, forever. Our formula states that

$$p_t M_t = \mathbb{E}_t\left[M_{t+1}\left(p_{t+1} + x_{t+1}\right)\right]$$

In our case $x_t \equiv 1$ and we have no uncertainty, so we can write

$$M_t p_t = \mathbb{E}_t\left[M_{t+1}p_{t+1}\right] + \mathbb{E}_t\left[M_{t+1}x_{t+1}\right]$$

Solving forward we obtain

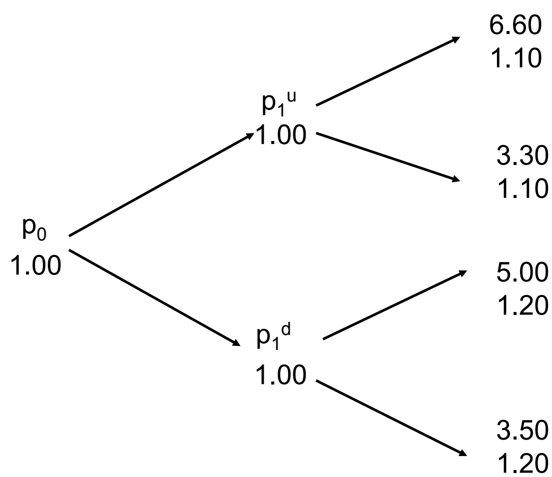
$$p_0 = \underbrace{\sum_{t=1}^{\infty} \mathbb{E}_0 \left[M_t x_t \right]}_{Fundamental\ Value} + \underbrace{\lim_{k \to \infty} \mathbb{E}_0 \left[M_k p_k \right]}_{Bubble\ Component}$$

The fundamental value is the price given by the static-dynamic model, however repeated trading gives rise to the possibility of a rational bubble. For example, fiat money can be understood as an asset with no dividends. In the static-dynamic model, such an asset would have no value (because the present value of zero is zero). But if there is a bubble on the price of fiat money, then it can have a positive value (see Bewley, 1980). In asset pricing theory, we often rule out bubbles simply by imposing

$$\lim_{k \to \infty} \mathbb{E}_0 \left[M_k p_k \right] = 0$$

Exercises

1) Consider the following stochastic processes for a risk free bond

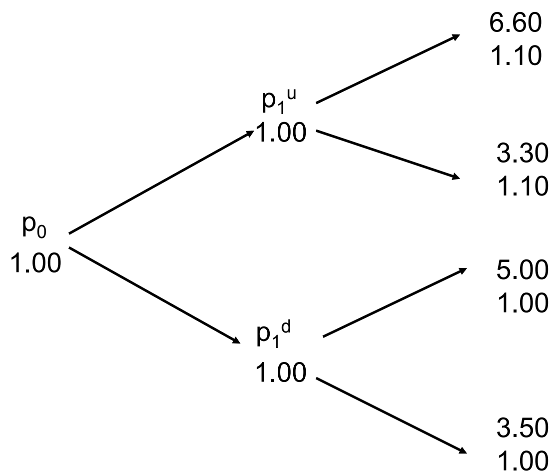


Note that **no asset** pays dividends. The price of the one-period risk-free bond is normalized to 1 in each period. Suppose that the **risk-neutral probability** of an "up" tick is 1/3 and that of a "down" is 2/3, for every point in time.

a) Is the market statically complete? Is the market dynamically complete?

b) What is the event price of the event {up, up} and of the event of only one up, ie, {up, up}∪{up, down}?

Now suppose we change the processes in the following manner:



c) Is the market statically complete? Is the market dynamically complete?

2) Consider a security market model with three dates, $t = 0, 1, 2$ and five states of the world, $s = 1, ..., 5$. Investors have no information at time 0 and full information at time 2. At the intermediate date, their information partition consists of the two sets $A = \{1, 2, 3\}$ and $B = \{4, 5\}$. There are two assets with the following price and dividend process:

- $p_0^1 = 3$, $p_1^1(A) = 3/2$, $p_1^1(B) = 3$;
- $d_1^1(A) = 3/2$, $d_1^1(B) = 2$, $d_2^1(s) = s$;
- $p_0^2 = 9/4$, $p_1^2(A) = 3$, $p_1^2(B) = 6/5$;
- $d_1^2(A) = 1$, $d_1^2(B) = 4/5$, $d_2^2(s) = 6 - s$.

a) Is this price-dividend system arbitrage-free?

b) Is the contingent claim with the safe payoff stream $y_t = 1$ $(t = 1, 2)$ attainable? If it is, calculate its arbitrage price and a replicating portfolio strategy.

3) Consider a security market with three dates, $t = 0, 1, 2$, and five states of the world, $s = 1, \ldots, 5$. The information structure is described by the following sequence of partitions of the state space:

$$\begin{aligned}\mathcal{P}_0 &= \{1, \ldots, 5\} \\ \mathcal{P}_1 &= \{A, B, C\} \\ \mathcal{P}_2 &= \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}\end{aligned}$$

where $A = \{1,2\}$, $B = \{3,4\}$ and $C = \{5\}$. There are four securities with terminal payoffs given by the matrix

$$D = \begin{pmatrix} 2 & 4 & 5 & 1 \\ 1 & 1 & 2 & 1 \\ 3 & 2 & 4 & 1 \\ 4 & 5 & 7 & 1 \\ 5 & 6 & 10 & 1 \end{pmatrix}$$

with the j-th column corresponding to the j-th security. The prices of these securities at date $t = 0, 1$ are given by

$$\begin{pmatrix} 3.6 & 4 & 6.6 & 1 \\ 1.5 & 2.5 & 3.5 & 1 \\ 3.25 & 2.75 & 4.75 & 1 \\ 5 & 6 & 10 & 1 \end{pmatrix}$$

where the j-th column lists the price of the j-th security at date 0, event A, event B and event C, respectively. The securities do not pay dividends prior to the terminal date

1. Verify that this securities market permits no arbitrage.

2. Is the market dynamically complete?

3. Compute the initial no-arbitrage price of the following three securities:

   (a) A call option on security 1 with an exercise price of 3.5.

   (b) A down-and-under call option on security 1 with an exercise price of 3.5. This is a call option with an extra provision - if the price of security 1 ever drops below the exercise price, then the option becomes worthless.

   (c) A convertible security that in each state pays at the terminal date the largest payout of the securities 1-4 in that state.

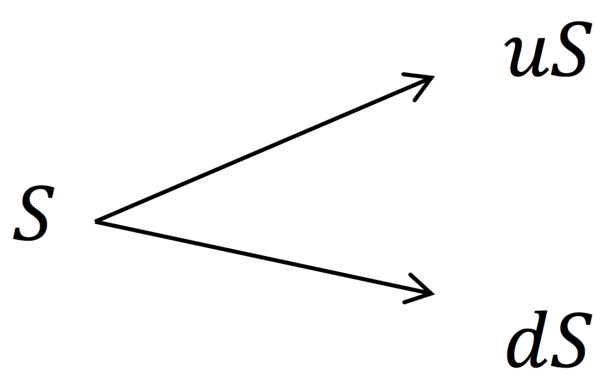
Chapter 8

Multi-Period Model: Options

Consider a European call option with strike K and maturity T. Its payoff is $(S_T - K)^+$, with no cash flows between $t = 0$ and $t = T$. Unfortunately we are unable to statically replicate this payoff using just the stock a risk-free bond: we need to dynamically hedge: that is, we need to engage in a strategy where the required stock (and bond) position changes for each period, until maturity. Since the replication strategy depends on specified random process of stock price, we need to impose a specific model for the evolution of the stock. In a discrete time setting, the canonical model is the Cox-Ross-Rubinstein binomial model. We assume that:

- The stock pays no dividends.
- The length of a period is h and over each period we assume that the stock can either go up to $S_{t+h} = u \times S_t$, or down to $S_{t+h} = d \times S_t$, so for each period t the distribution is binomial.
- The gross risk-free rate between periods is $R^f = e^{r^f \cdot h}$, and by no arbitrage we require $d < R^f < u$.

We start with a one-period binomial tree (for $h = 1$):



Suppose we buy ("go long" in trading jargon) Δ stocks and B bonds. The payoff from portfolio $C_0 = \Delta S_0 + B$ is

$$C_1 = \Delta S_1 + B \cdot R^f = \begin{cases} \Delta u S_0 + B \cdot R^f & if\ the\ stock\ goes\ up \\ \Delta d S_0 + B \cdot R^f & if\ the\ stock\ goes\ down \end{cases}$$

Call C_u the option payoff in the "up" state and C_d the option payoff in the "down" state. The replicating strategy must satisfy

$$C_u = \Delta u S_0 + B \cdot R^f$$

$$C_d = \Delta d S_0 + B \cdot R^f$$

Solving for Δ and B we get

$$\Delta = \frac{C_u - C_d}{S_0 (u - d)}$$

$$B = \frac{u C_d - d C_u}{R^f (u - d)}$$

Δ can be interpreted as the sensitivity of call price to a change in the stock price, or equivalently, how much of the stock we should hold to hedge the option: for instance, to hedge a long call position we need to sell Δ units of the stock. Substituting Δ and B in C_0 we get

$$C_0 = \frac{C_u - C_d}{S_0 (u - d)} S_0 + \frac{u C_d - d C_u}{R^f (u - d)} = \frac{1}{R^f} \left[\frac{R - d}{u - d} C_u + \frac{u - R}{u - d} C_d \right]$$

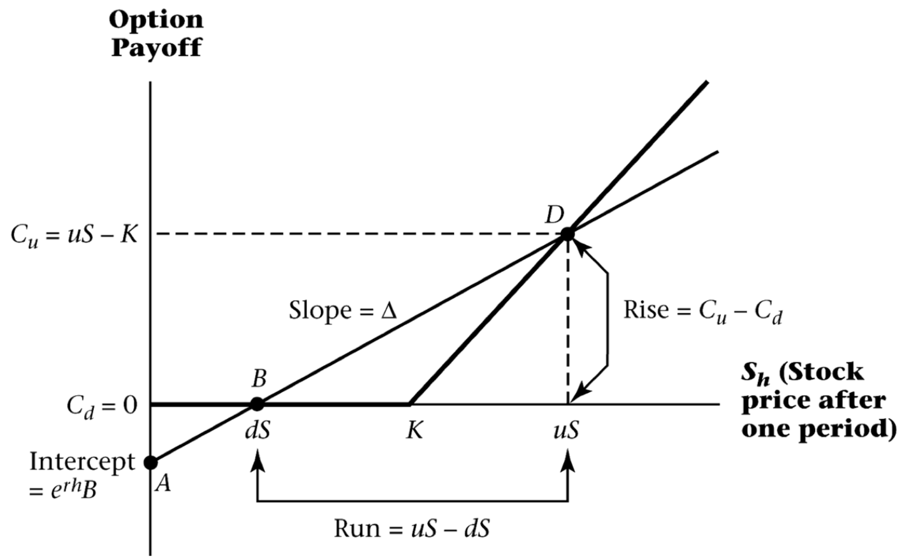
Define $\pi^{\mathbb{Q}} \equiv \frac{R^f - d}{u - d}$ and note that $\frac{u - R^f}{u - d} = 1 - \pi^{\mathbb{Q}}$:

$$C_0 = \frac{1}{R^f} \left[\pi^{\mathbb{Q}} C_u + \left(1 - \pi^{\mathbb{Q}} \right) C_d \right] = \frac{1}{R^f} \mathbb{E}^{\mathbb{Q}} \left[C_1 \right]$$

Therefore the option price is the discounted payoff of the option under the equivalent martingale (risk-neutral) measure $\mathbb{Q}$. Note that $\mathbb{Q}$ is also the probability distribution that would justify the current stock price in a risk-neutral world:

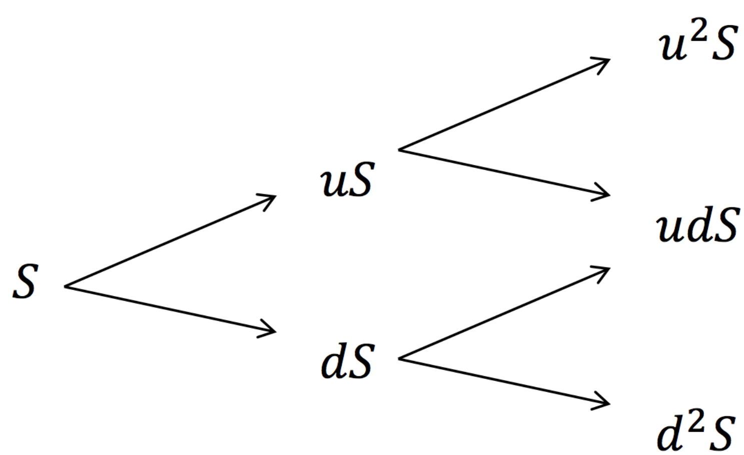
$$S_0 = \frac{1}{R^f} \left[\pi^{\mathbb{Q}} S_0 u + \left(1 - \pi^{\mathbb{Q}} \right) S_0 d \right]$$

Note how we never even mentioned the physical probability measure π, since we are working with relative asset pricing.



8.1 Two-Period Binomial Tree

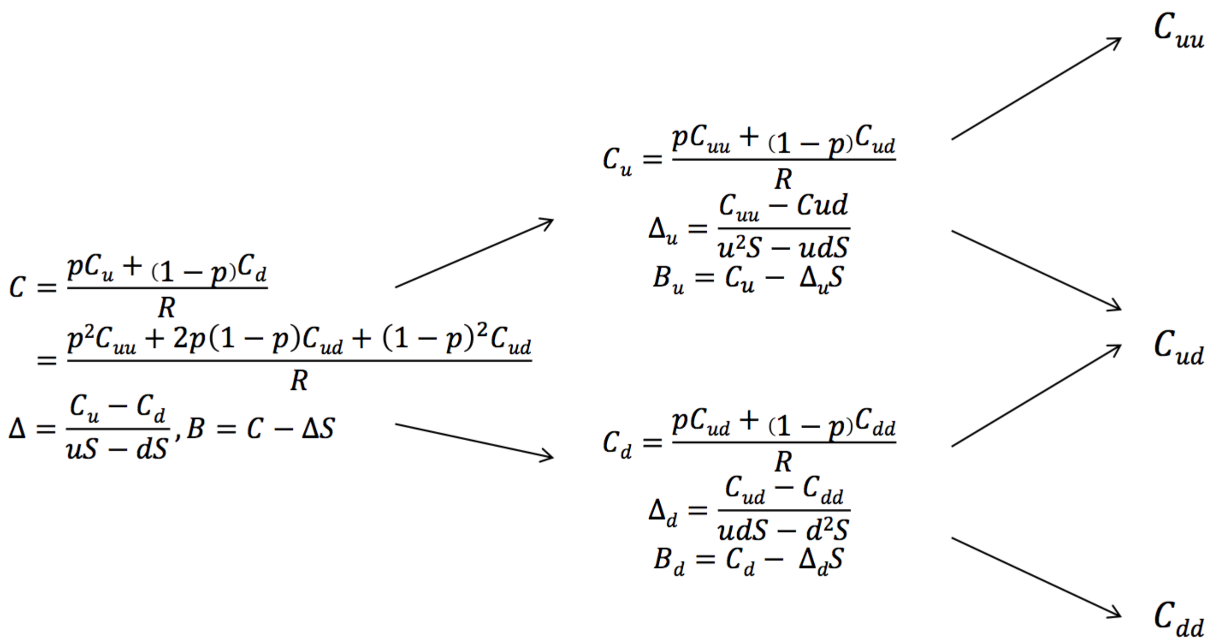
A Two-Period Binomial Tree is just a concatenation of single-period binomial trees:



To price the option at time $t = 0$ we just apply the same procedure backwards, starting from the two sub-trees in the final period. To summarize,

1. We compute the risk-neutral probability $\pi^{\mathbb{Q}}$ from the stock price
2. We plug these probabilities in the formula for C at the final two nodes, and once we find the option prices for each of the two possible states in time $t = 1$ we repeat the procedure for the initial node
3. We find Δ and B at each time to find the replicating strategy.

The general procedure for a two-period tree is illustrated below (from now on, let $p = \pi^{\mathbb{Q}}$):



Suppose we observe an option price on the market of \$36, while the cost of the dynamic replication strategy is \$34.08: then e can take advantage of this arbitrage opportunity by selling the option and buying the synthetic portfolio. Today we pocket \$1.92, and at maturity our cash outflows equal our inflows.

8.2 The Relation with the Black-Scholes Model

The Black-Scholes option pricing model introduced in chapter 1 can be viewed as the limit of a binomial tree, in which the mumber of periods n (and therefore of states) goes to infinity. To see this, take $u = e^{\sigma\sqrt{\frac{T}{n}}}$, $d = \frac{1}{u} = e^{-\sigma\sqrt{\frac{T}{n}}}$ and $R^f = e^{r^f \cdot \frac{T}{n}}$ where T is the time to expiration and σ is the annualized standard deviation of the stock log-returns. The general binomial formula for a European call on non-dividend paying stock n periods from expiration is:

$$C_0 = \frac{1}{R^f}\left[\sum_{j=0}^{n} \frac{n!}{j!\,(n-j)!} p^j \,(1-p)^{n-j} \left(u^j d^{n-j} S - K\right)^+\right]$$

Plugging in u, d and R^f and taking the limit as $n \to \infty$ we get

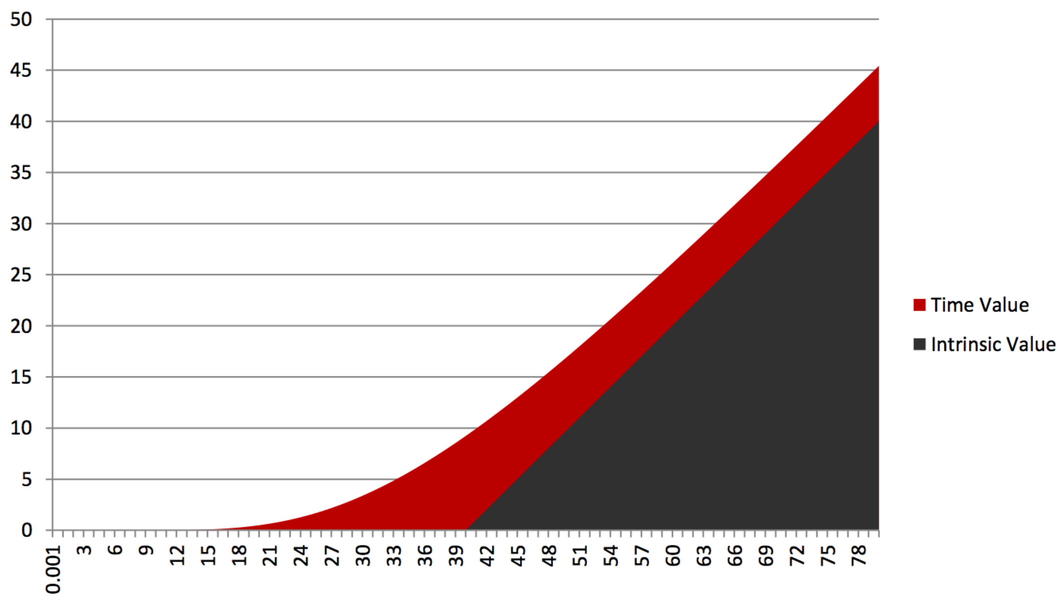
$$S_0 N\left(\frac{\ln\frac{S_0}{K} + \left(r^f + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) - Ke^{-rT} N\left(\frac{\ln\frac{S_0}{K} + \left(r^f - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)$$

Let $d_1 = \frac{\ln \frac{S_0}{K} + \left(r^f + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \frac{\sigma}{2}\sqrt{T}$, we can interpret the trading strategy under the Black-Scholes formula as $\Delta = N(d_1) \in [0,1]$ and $B = -Ke^{-rT}N(d_2)$. We can also find the price of a put option: by put-call parity, we have

$$P_0 = C_0 - S_0 - Ke^{-r^f T} = Ke^{-r^f T}N(-d_2) - S_0 N(-d_1)$$

Note that the put option has $\Delta = -N(-d_1) \in [-1,0]$

We call *intrinsic value* of the option the quantity $(S_t - K)^+$ and time value of the option the quantity $C_t - (S_t - K)^+$:



8.3 Delta-Hedging

A stock has a delta equal to one, while a portfolio of J assets - where each asset j has notional amount N_j - has a delta $\triangle_p = \sum_{j=1}^{J} \Delta_j N_j$. We say that the portfolio is *delta neutral* if $\Delta_p = 0$. Delta-hedging a portfolio is only a "perfect" hedge (yielding the same final payoff as the corresponding option) with continuous trading: computing the delta is a linear approximation to the option value, but the option price is convex and therefore the second- ahd higher-order derivatives matter. Delta-hedging when trading is not continuous is effective only for smal price changes. Delta-Gamma hedging reduces this convexity risk, and involves trading with other options.

8.4 The Volatility Smile

The Black-Scholes model assumes that the volatility σ is constant over time and through strikes. This turns out to be a bad assumption: suppose we take option prices as they are observed in the market, and then back out the volatility implied by these prices for many values of the time to maturity T and strike K. It turns out that the volatility implied by options on the market exhibits a volatility "smile", while the Black-Scholes model implies a constant parameter σ in the (T, K) space. Interestingly, this effect was almost absent before the 1987 "Black Monday" stock market crash, and became very evident afterwards. This means that the Black-Scholes mdoel underprices out of the money puts (and thus in the money calls) and overprices out of the money calls (and thus in the money puts).

One way to get around this problem is to use a different model for the volatility of the stock called "stochastic volatility". Other issues in the Black-Scholes include stochastic interest rates, bid-ask spreads, other transaction costs, etc.

8.5 Collateralized Debt Obligations

Collateralized Debt Obligations (CDOs) are derivatives which repackage cash-flows from a set of assets. Payoffs are divided in tranches: in the event of underperformance of the underlying assets, the Senior tranche is paid out first, the Mezzanine second, and the Junior tranche last. Option theory is very useful in pricing CDOs: the tranches can be priced using analogues from option pricing formulas, and it is possible to estimate the "implied default correlations" between the underlying assets that correctly price the tranches.

Exercises

1) Consider the binomial model with $u = 1.2$, $d = 0.9$, $R = 1.1$ and initial asset price $S = 100$.

1. Calculate the price of a call option with exercise price $K = 100$ and $n = 4$ periods left until expiry.
2. Calculate the prices of the European and American put options with the same exercise price and time to expiry as in (a).
3. Calculate the price of an option which, at the end of the third period, gives its holder the right to purchase the underlying asset at the minimum price realized over the life of the option.

2) A stock price is currently \$30. During each 2-month period for the next 6 month it will increase by 8% or reduce by 10%. The annual risk-free interest rate is 5% (continuous compounding). Use a three-step tree to calculate the value of a derivative (called turbo option) that pays off $S_T \times (30 - S_T)^+$, where S_T is the stock price in 6 months.

Chapter 9

Multi-Period Model: Fixed Income

In this chapter we will cover the basic fixed income tools in the multi-period model introduced in the last chapter. Let us begin with the U.S. Treasuries, these include:

- T-Bills, with maturity less than one year, usually with no coupon payments, and are usually sold "at discount", that is, the price is lower than the face value.
- T-Notes (1-10 year maturity) and T-Bonds (10-30 year maturity) pay a semiannual coupon and sell "at par", that is, when they are issued the price is the same as the face value.

Let $r_t(t_1, t_2)$ denote the (annualized) interest rate from t_1 to t_2 prevailing at time t, and let $P_t(t_1, t_2)$ be the price of a bond quoted at t to be purchased at time t_1 and maturing at t_2. We call *yield to maturity* the percentage increase in dollars earned from the bond (annualized). It follows that at time $t = 0$ the price of a zero-coupon bond (ZCB) purchased at the same time and with maturity t is

$$P_0(0, T) = \frac{1}{(1 + r_0(0, t))^t}$$

The *yield curve* is a plot of yields of zero coupon bonds as a function of their maturity. Usually, long-term bond yields are higher than short-term bond yields – in which case we say that the yield curve is *upward sloping* – but sometimes short-term bond yields are higher than long-term bond yields – which we would call a *downward sloping* or *inverted* yield curve. The yield curve sometimes has humps or other shapes as well. Formally we can represent the yield curve as the annualized bond yield $r_0(0, t) \equiv y_t$ as a function of time. We can determine the *forward rate* $r_0(t_1, t_2)$ as follows: suppose you buy a ZCB with maturity t_1, and at maturity you use all the proceeds from this purchase to buy a bond with maturity t_2. By no arbitrage, this strategy must have the same price as buying a ZCB at time $t = 0$ with maturity $t = t_2$:

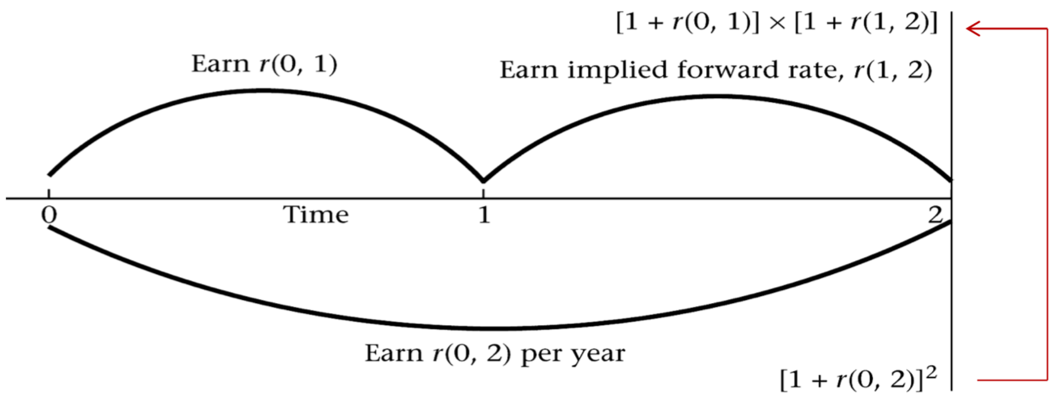
$$P_0(0, t_2) = P_0(0, t_1) \times P_0(0, t_2)$$

Which can be rewritten as

$$(1 + r_0(0, t_1))^{t_1} (1 + r_0(t_1, t_2))^{t_2 - t_1} = (1 + r_0(0, t_2))^{t_2}$$

Solving for the forwatd rate $r_0(t_1, t_2)$,

$$r_0(t_1, t_2) = \left[\frac{(1 + r_0(0, t_2))^{t_2}}{(1 + r_0(0, t_1))^{t_1}}\right]^{\frac{1}{t_2 - t_1}} - 1 = \left(\frac{P_0(0, t_1)}{P_0(0, t_2)}\right)^{\frac{1}{t_2 - t_1}} - 1$$



A coupon bond can be similarly decomposed in the sum of $n + 1$ ZCBs: suppose the (annualized) coupon payment is c and the bond pays a coupon on each date $\{T_1, \ldots, T_n\}$ (with $T_i - T_{i-1} = \frac{1}{2}$ since coupon bonds usually pay semi-annually):

$$B_t(t, T_n) = \sum_{i=1}^{n} \frac{c}{2} \times P_t(t, T_i) + P_t(t, T_n)$$

Note that in order for the bond to sell at par, i.e. $B_t(t, T_n) = 1$, it must be that $c = 2 \times \frac{1 - P_t(t, T_n)}{\sum_{i=1}^{n} P_t(t, T_i)}$.

Finally, we can also write

$$B_t(t, T_n) = \sum_{i=1}^{n} \frac{c}{2} \times \frac{1}{(1 + r_t(t, T_i))^{T_i - t}} + \frac{1}{(1 + r_t(t, T_n))^{T_n - t}}$$

We define the yield to maturity of a coupon bond as the $y \in \mathbb{R}_+$ that solves

$$B_t(t, T_n) = \sum_{i=1}^{n} \frac{c}{2} \times \frac{1}{(1 + y)^{T_i - t}} + \frac{1}{(1 + y)^{T_n - t}}$$

9.1 Duration

Duration is the sensitivity of a bond's price to changes in interest rates. It is sometimes interpreted as the "average time" it takes to get the money back (although this is not exactly correct). It is defined as

$$D(y) = -\frac{dB(y)}{dy}$$

Where y is the bond's yield to maturity. Let X_t be the amount the coupon bond pays at date t, then we can write

$$B_t\left(t, T_n\right) = \sum_{i=1}^{n} \frac{X_{T_i}}{\left(1+y\right)^{T_i - t}}$$

And therefore the duration is equal to

$$D(y) = -\frac{1}{(1+y)} \times \sum_{i=1}^{n} (T_i - t) \frac{X_{T_i}}{\left(1+y\right)^{T_i - t}}$$

An alternative form of the duration is the Macaulay duration, defined as

$$MD\left(y\right) = -\frac{dB(y)/B\left(y\right)}{dy/\left(1+y\right)} = \frac{\left(1+y\right)}{B\left(y\right)} D\left(y\right)$$

Which for a coupon bond is equal to

$$MD\left(y\right) = \frac{1}{B\left(y\right)} \sum_{i=1}^{n} (T_i - t) \frac{X_{T_i}}{\left(1+y\right)^{T_i - t}}$$

Duration is useful in the context of portfolio immunization: suppose we own a bond with maturity t_1 and we wish to trade a quantity N of a bond ith maturity t_2 so that the total duration of the portfolio is zero. We set

$$D_1(y_1) + N \times D_2(y_2) = 0$$

That is,

$$N = -\frac{D_1(y_1)}{D_2(y_2)} = -\frac{B_1\left(y_1\right) MD_1\left(y_1\right)}{B_2\left(y_2\right) MD_2\left(y_2\right)} \frac{1+y_2}{1+y_1}$$

When interest rates move by a small amount, the total value of the portfolio made by 1 unit of the t_1-maturity bond and N units of the t_2-maturity bond will not change. This will not be true for larger changes in the interest rates however, as we are only using a first order Taylor (linear) approximation; moreover, while the duration is computed with respect to the yield to maturity which is constant and moves in parallel shifts, the yield curve changes in a non-parallel way, so the change in value of the portfolio may not be exactly zero even with a small change in interest rates.

9.2 The Term Structure of Interest Rates

Bond prices carry all the information on intertemporal rates of substitution, which is primarily affected by expectations and indirectly by risk considerations. A collection of (real) interest rates for different times to maturity is a meaningful predictor of the market expectations for future economic developments: for instance, more optimistic expectations will produce an upward sloping term structure of interest rates. To see this, consider a risk-free ZCB;

$$P_0(0,t) = \mathbb{E}[M_t] = \frac{1}{(1+y_t)^t}$$

Thus the yield is

$$y_t = (P_0(0,t))^{-\frac{1}{t}} - 1 = \frac{1}{\delta}\left(\frac{\mathbb{E}[u'(c_t)]}{u'(c_0)}\right)^{-\frac{1}{t}} - 1$$

Define the growth rate g_t as the solution to

$$(1+g_t)^t = \frac{c_t}{c_0}$$

For each t. If the representative agent has CRRA utility with parameter γ, a first order Taylor approximation yields

$$y_t \approx \gamma\mathbb{E}[g_t] - \ln[\delta]$$

so the (real) yield curve measures expected growth over different time horizons. Note that a second order Taylor approximation would include u''' terms, so uncertainty on g_t would also matter. If the agent is prudent, then uncertainty about g_t lowers the yield. The term structure of (real) interest rates is upward sloping when the expected growth increases over time, and long-term uncertainty on the growth rate is generally smaller than the short-term uncertainty (this makes intuitive sense, for instance, if the growth rate is viewed as mean reverting).

9.3 The Expectations Hypothesis

The expectations hypothesis comprises three equivalent statements about the pattern of riskless bond yields across maturity:

1. The T period yield is the average of the expected future one-period yields
2. The current forward rate equals the expected future spot rate.
3. The expected holding period returns are equal on bonds of all maturities.

The term structure of interest rates is derived from the cross-section of bond prices at a particular point in time; but how does the term structure evolve with time? This question is especially relevant for an investor who is trying to decide what kind of bonds to invest into, or what kind of loan to take. The expectations hypothesis is the classic theory for understanding the shape of the yield curve: it is the traditional benchmark for thinking about the expected value of future yields. We can state the expectations hypothesis in three mathematically equivalent forms:

1. The T-period yield is the average of expected future one-period yields: $r_0(0,T) = \frac{1}{T}\mathbb{E}_0\left[r_0(0,1) + r_1(1,2) + ..\right.$ (plus a risk premium)

2. The forward rate equals the expected future spot rate: $r_0(T-1,T) = \mathbb{E}_0\left[r_{T-1}(T-1,T)\right]$ (plus a risk premium)

3. The expected holding period returns are equal on bonds of all maturities: $\mathbb{E}_0\left[\frac{P_n(n,T)}{P_0(0,T)}\right] = \mathbb{E}_0\left[\frac{1+r_0(0,T)}{1+r_n(n,T)}\right] = 1 + r_0(0,n)$ (plus a risk premium)

The first form reflects a choice between two ways of getting money from 0 to T. You can buy a T period bond, or roll-over T one-period bonds. Risk neutral investors will choose one over the other strategy until the expected T-period return is the same. The three forms are mathematically equivalent: if every way of getting money from t to $t+1$ gives the same expected return, then so must every way of getting money from $t+1$ to $t+2$, and, chaining these together, every way of getting money from t to $t+2$.

The expectations hypothesis explains the shape of the yield curve: if the yield curve is upward sloping, according to the expectations hypothesis this is because short term rates are expected to rise in the future. The expectations hypothesis can be seen as a response to a classic misconception: if long term yields are 10% but short term yields are 5%, an unsophisticated investor might think that long-term bonds are a better investment. The expectations hypothesis shows how this may not be true: according to it, future short rates are expected to rise, and this means that rolling over the short-term bonds at a really high rate, say 20%, would give the same long-term return. When the short term interest rates rise in the future, long-term bond prices decline. Thus, the long-term bonds will only give a 5% rate of return for the first year.

Consider another example. Suppose we want to invest our money for two years, we have three choices:

1. Buy a 2-year ZCB with gross yield $1 + r_0(0,2) = 1 + y_2$

2. Buy a 1-year ZCB and roll it over when it matures. The expected gross yield is $(1+y_1)\,\mathbb{E}\left[(1+r_1(1,2))\right]$

3. Buy a 3-year ZCB and sell it after 2 years. The expected gross yield is $\mathbb{E}\left[\frac{1+y_3}{1+r_2(2,3)}\right] = (1+y_3)\,\mathbb{E}\left[\frac{1}{1+r_2(2,3)}\right]$.

Note that only strategy 1 is fully risk-free. It is possible that the risky strategies have a premium over the riskless strategy, called *term premia* (a particular form of risk premium). It is common to add a constant risk premium and still refer to the resulting model as the expectations hypothesis. One end of each of the three statements does imply more risk than the other. A forward rate is known, while the future spot rate is not. Long-term bond returns are more volatile than short term bond returns. Rolling over short term real bonds is a riskier long-term investment than buying a long term real bond. These risks will generate expected return premia if they covary with the discount factor, and our theory should reflect this fact.

The price of a t-period ZCB price is

$$P_0(0,t) = \mathbb{E}[M_t] = \mathbb{E}[m_1 \cdot \ldots \cdot m_t]$$

So if we buy t one-period ZCBs rolled over $t-1$ times we get that the expected price for this strategy is

$$P_0(0,1) \cdot \mathbb{E}[P_1(1,2)] \cdot \ldots \cdot \mathbb{E}[P_t(t-1,t)] = \mathbb{E}[m_1] \cdot \mathbb{E}[m_2] \cdot \ldots \cdot \mathbb{E}[m_t]$$

These two strategies yield the same expected return if

$$\mathbb{E}[m_1 \cdot \ldots \cdot m_t] = \mathbb{E}[m_1] \cdot \mathbb{E}[m_2] \cdot \ldots \cdot \mathbb{E}[m_t]$$

Which holds, for instance, if m_t is not serially correlated, or in a world with full certainty or risk-neutral agents, or yet in an "iid" world: in this case, there are no term premia. However, if the m_t's are serially correlated (for instance because the growth process is serially correlated) then the expectation hypothesis fails. Allowing an arbitrary time-varying risk premium, the model is a tautology, of course. Thus, the entire content of the "expectations hypothesis" augmented with risk premia is in the restrictions that the risk premium is constant over time. However, the constant risk premium model does not do well empirically.

9.4 Futures

Recall that futures are exchange-traded forward contracts. Typical features include:

- There are standardized features, including the delivery date, the location and the procedures
- A clearinghouse matches the by or sell orders, keeps track of the customers obligations and payments and is the effective counterparty of its customers

Unlike forward contracts, they are settled daily through mark-to-market accounting, which lowers credit risk for the counterparties. Moreover, they are highly liquid, so it is much easier to offset an existing position. For instance, futures on the S&P 500 index have a fixed notional value unit equal to $250 times the value of the index. It is a cash-settled contract for which the open interest (the total number of buy/sell pairs) is known by the public at any time. The margin and mark-to-market work as follows: there is an initial margin to be posted when the contract is traded, plus a maintenance margin that is about 70-80% of the initial margin. As the underlying moves, counterparties may be subject to requirements to increase the maintenance margin (a practice called "margin call") and the value of each contract is marked to the value implied by the market each day. For short-dated contracts the difference between forwards and futures is rather small, while it can be large for long-dated contracts or whenever interest rates are highly correlated with the underlying asset.

The time t price of a future contract is always zero at initiation, so the price of the forward using the equivalent martingale measure is:

$$0 = \mathbb{E}_t^{\mathbb{Q}}\left[\rho_T\left(F_{0,T} - S_T\right)\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\rho_T\right]\mathbb{E}_t^{\mathbb{Q}}\left[F_{0,T} - S_T\right] - Cov_t^{\mathbb{Q}}\left[\rho_t, S_T\right]$$

Assuming a constant continuously compounded interest rate r,

$$F_{0,T} = \mathbb{E}_t^{\mathbb{Q}}\left[S_T\right]$$

So the future price of a stock $F_{0,T}$ equals the risk-neutral discounted value of the stocl at time T.

Other than in equities, forwards and futures are also used in currencies and commodities. In currencies, suppose r_y is the interest rate prevailing in Japan and x_0 is the current $/¥ exchange rate. The prepaid forward will be equal to $F_{0,T}^p = x_0 e^{-r_y T}$ (since by deferring delivery of the currency we lose interest income from bonds denominated in that currency) and therefore the currency forward is $F_{0,T} = x_0 e^{(r-r_y)T}$ where r is the US interest rate. Note that $F_{0,T} > x_0$ whenever the domestic interest rate exceeds the foreign interest rate. Commodity forward prices can be described by the same formula as that for financial forward prices: $F_{0,T} = S_0 e^{(r-\delta)T}$, but for commodities δ is the commodity lease rate (the return an investor would make by buying and lending the commodity).

The *forward curve* is the set of prices for different expiration dates for a given commodity. When it is upward sloping, we say that the market is in *contango*; when downward sloping, we say the market is in *backwardation* (the same market can be partly in contango and partly in backwardstion).

Forward rate agreements (FRAs) are over-the-counter contracts that guarantee a borrowing or lending interest rate on a given notional amount. They can settle either in arrears, in which case the amount exchanged is $(r_{quarterly} - r_{FRA}) \times Notional$, or at the time of borrowing, in which case the amount exchanged is $\frac{(r_{quarterly} - r_{FRA})}{1 + r_{quarterly}} \times Notional$. FRAs can be replicated using ZCBs. An example of standardized FRAs is the Eurodollar futures, where the Eurodollar is the interest rate prevailing on dollar-denominated accounts abroad (there is one key difference with FRAs however: the settlement structure of Eurodollar contracts favors borrowers, therefore the rate implicit in

Eurodollar futures is greater than it would be for an FRA rate). The payoff at expiration is $(Futures\ Price - (100 - r_{LIBOR})) \times \$2500$. Recently, Eurodollar futures took over T-bill futures as the preferred contract to manage interest rate risk: this is partly because LIBOR tracks the corporate borrowing rates better than the T-bill rate.

A similar type of contract is a *repurchase agreement*, or *repo*. This entails selling a security with an agreement to buy it back at a fixed price: the underlying security is held as collateral by the counterparty. A repo is therefore a form of collateralized borrowing often used by securities dealers to finance inventory. Normally a "haircut" is charged by the counterparty to account for credit risk.

9.5 Swaps

A *swap* is a contract in which two parties agree to exchange a floating against a fixed stream of cash flows, commonly used by companies to hedge a stream of risky payments. A single-payment swap is equivalent to a forward contract: the price today is zero and the amount exchanged at expiration is $Notional \times (r_{floating} - r_{fixed})$. Swaps can settle either in cash (financial settlement), with no exchange of notional amount, or in phisical settlement with exchange of notional (or delivery of the asset). The market value of a swap is zero at interception; once the swap is struck, its market value will generally no longer be zero because forward prices for oil and interest rates will change over time, and even if prices did not change, the market value of swaps will change over time due to the changes in the implicit borrowing and lending rates. It is possible to exit the swap contract by entering into an offsetting swap transaction (with or without the same counterparty). The market value of the swap is the difference in the present value of payments between the original and new swap rates. The notional of the swap is the amount on which the interest payments are based, while the life of the swap is called "swap term" or "swap tenor". The market-maker is a counterparty to the swap wishing earns fees for offering this service, not to take on interest rate risk. Therefore, the market-maker will hedge the floating rate payments by using, for example, forward rate agreements. Finally, we call swap rate *payer* the counterpart which is receiving the floating rate (and paying the fixed rate), and swap rate *receiver* the opposite counterparty.

Suppose there are n swap settlement dates occurring on dates $\{t_i\}_{i=1}^{n}$. As before, the (implied) forward rate between t_i and t_{i-1} is $r_0(t_{i-1}, t_i)$ and the price of a ZCB with maturity t_i is $P_0(0, t_i)$. Call the fixed swap rate R, then we can write

$$0 = \sum_{i=1}^{n} P_0(0, t_i)\left[R - r_0(t_{i-1}, t_i)\right]$$

Which can be solved for R to yield:

$$R = \frac{\sum_{i=1}^{n} P_0(0,t_i)\, r_0(t_{i-1},t_i)}{\sum_{i=1}^{n} P_0(0,t_i)} = \sum_{i=1}^{n} r_0(t_{i-1},t_i) \left(\frac{P_0(0,t_i)}{\sum_{i=1}^{n} P_0(0,t_i)} \right)$$

Thus, the fixed swap rate is as a weighted average of the implied forward rates, where zero-coupon bond prices are used to determine the weights. Using the fact that $r_0(t_{i-1},t_i) = \frac{P_0(0,t_{i-1})}{P_0(0,t_i)} - 1$ we can also write:

$$R = \frac{\sum_{i=1}^{n} P_0(0,t_i)\, r_0(t_{i-1},t_i)}{\sum_{i=1}^{n} P_0(0,t_i)} = \frac{\sum_{i=1}^{n} \left[P_0(0,t_{i-1}) - P_0(0,t_i)\right]}{\sum_{i=1}^{n} P_0(0,t_i)} = \frac{P_0(0,t_0) - P_0(0,t_n)}{\sum_{i=1}^{n} P_0(0,t_i)} = \frac{1 - P_0(0,t_n)}{\sum_{i=1}^{n} P_0(0,t_i)}$$

Which is equivalent to the formula for the coupon of bond trading at par: the swap rate is the coupon rate on a par coupon bond (and a firm that swaps floating for fixed rates ends up with the economic equivalent of a fixed- rate bond).

The set of swap rates at different maturities is called *swap curve*. The swap curve should be consistent with the interest rate curve implied by the Eurodollar futures contract, which is used to hedge swaps: recall that the Eurodollar futures contract provides a set of 3-month forward LIBOR rates. In turn, zero-coupon bond prices can be constructed from implied forward rates. Therefore, we can use this information to compute swap rates. The *swap spread* is the difference between swap rates and Treasury-bond yields for comparable maturities.

A deferred swap is a swap that begins at some date in the future, but its swap rate is agreed upon today. The fixed rate on a deferred swap beginning in k periods is computed as

$$R = \frac{\sum_{i=k}^{n} P_0(0,t_i)\, r_0(t_{i-1},t_i)}{\sum_{i=k}^{n} P_0(0,t_i)}$$

Note that for $k = 1$ we are back to our previous case. An *amortizing* swap is a swap whose notional value is declining over time (e.g. a floating rate mortgage); an *accreting* swap is a swap where the notional value grows over time. In both cases, the fixed swap rate is still a weighted average of implied forward rates, but now the weights also involves the changing notional Q_t:

$$R = \frac{\sum_{i=1}^{n} Q_{t_i} P_0(0,t_i)\, r_0(t_{i-1},t_i)}{\sum_{i=1}^{n} Q_{t_i} P_0(0,t_i)}$$

The advantage of swaps is that they allow firms to separate credit and interest rate risk: by swapping its interest rate exposure, a firm can pay the short-term interest rate it desires, while the long-term bondholders will continue to bear the credit risk.

Exercises

1) Suppose that the return of a zero coupon bond of maturity i satisfies:

$$r_i = \gamma_i f + \epsilon_i$$

where $Cov(\epsilon_i, \epsilon_j) = 0$ whenever $i \neq j$, and $Cov(\epsilon_i, f) = 0$.

Currently the yield of all zero coupon bonds is 5% annually compounded, that is the price of a bond that matures in i years equals $(1 + .05)^{-i}$. You are holding a coupon bond that pays a coupon of 6 in one, two and three years and pays 106 in 4 years.

a) What is the price of this coupon bond if there is no arbitrage?

b) Calculate the sensitivity of this coupon bond with respect to f, as a function of the γ_i's.

c) Suppose now that $\gamma_i = i$, and that $\sigma^2(\epsilon_i) = \sigma^2$ for each $i = 1, 2, 3, 4$. You want to hedge the factor sensitivity of the coupon bond. You are told that you can short *a single bond* among the zero-coupons with maturities 1,2,3 years. Explain how you would choose the "best" maturity and how much would you short of the zero-coupon bond to hedge the common factor exposure. Give some intuition for your result.

Chapter 10

The Multi-Period Equilibrium Model

Going back to the CAPM model discussed in chapter 6, recall that if the one-period stochastic discount factor m_t is not time-varying (that is, the distribution of m_t is i.i.d for each t), then the expectations hypothesis holds and the investment opportunity set does not vary. The corresponding R_t^* of a *single-factor* state-price beta model is then relatively easy to estimate, since over time we collect more and more realizations of R_t^*. However, if m_t (or the correspondingR_t^*) is time-varying, then we can assume that it depends on some *state variable* and need a *multi-factor* model to account for it. Indeed, suppose that $R_t^* = R^*(z_t)$, where z_t is some state variable. For example, suppose that z_t can be either 1 or 2 with equal probability. Then we could take all the periods in which $z_t = 1$ and back out $R^*(1)$, and find $R^*(2)$ in a similar way. Is it possible to do so in reality? The answer is no, because we have hedges *across* state variables.

10.1 Dynamic Hedging Demand INCOMPLETE

Trade-off: Low return realization in next period. Good since opportunity going forward is high, so you invest more, but bad since marginal utility is high, so you consume and invest less. High return realization in next period. In terms of CRRA utility: if γ is higher (lower) than 1, then first (second) effect dominates. If $\gamma = 1$ (log-utility), then the two effects perfectly offset each other.

Illustration with noise trader risk: suppose the fundamental value v is constant but the price is noisy (due to noise traders). If the asset is underpriced, say $p = 0.9$, it might be even more underpriced in the next period. Myopic risk-averse investor: buy some of the asset and push price towards 1, but not fully. Forward-looking risk-averse investor: there can be intermediate losses if the price declines in next period, but then the investment opportunity set improves even more i.e. if returns are bad, then there are greater opportunities (dynamic hedge).

10.2 Intertemporal CAPM

The static problem is equivalent to the dynamic problem in a few special cases:

- In a general ICAPM setting, if agents have CRRA utility with $\gamma \neq 1$ and changing investment opportunity sets.
- In a CRRA utility setting, i.i.d returns and constant risk-free rate. In particular, short- and long-run investors have the same portfolio weights and the fraction of wealth invested in each asset is time-invariant (Merton 1971).
- Agents have log utility with non-i.i.d. returns.

If the βs of each subperiod CAPM are time-independent, then conditional and unconditional CAPM are equivalent. If βs are time-varying they may co-vary with R_m and hence CAPM equation does not hold for unconditional expectations: an additional co-variance terms has to be considered, and we move from a single- to multi-factor setting.

The following model is due to Merton (1973). The Bellman equation is given by

$$V\left(W_t, z_t\right) = \max_{\{c_t\}} \left\{u\left(c_t\right) + \delta \mathbb{E}_t\left[V\left(W_{t+1}, z_{t+1}\right)\right]\right\}$$

Where $W_{t+1} = R^W_{t+1}\left(W_t - c_t\right)$ and R^W_{t+1} is the optimal portfolio. The FOC yields:[1]

$$0 = u'\left(c_t\right) - \delta \mathbb{E}_t\left[V_W\left(W_{t+1}, z_{t+1}\right) R^W_{t+1}\right]$$

Since $V_W\left(W_{t+1}, z_{t+1}\right) = \delta \mathbb{E}_t\left[V_W\left(W_{t+1}, z_{t+1}\right) R^W_{t+1}\right]$ by the envelope theorem, we get that for all t

$$u'\left(c_t\right) = V_W\left(W_t, z_t\right)$$

So the one-period pricing eqaution reads

$$\mathbb{E}\left[R^j_{t+1}\right] - R^f_{t+1} = -Cov_t\left[\frac{u'\left(c_{t+1}\right)}{\mathbb{E}\left[u'\left(c_{t+1}\right)\right]}, R^j_{t+1}\right] =$$

$$= -Cov_t\left[\frac{V_W\left(W_{t+1}, z_{t+1}\right)}{\mathbb{E}\left[V_W\left(W_{t+1}, z_{t+1}\right)\right]}, R^j_{t+1}\right]$$

Taylor expanding V_W around $\left(W_t, z_t\right)$ we get

$$V_W\left(W_{t+1}, z_{t+1}\right) \approx V_W\left(W_t, z_t\right) + V_{WW}\left(W_t, z_t\right) \Delta W_{t+1} + V_{Wz}\left(W_t, z_t\right) \Delta z_{t+1}$$

[1] Here V_W denotes the derivative of V with respect to its first argument.

Plugging this in the expression above we obtain

$$\mathbb{E}_t\left[R^j_{t+1}\right]-R^f_{t+1}=-Cov_t\left[\frac{V_W\left(W_t,z_t\right)+V_{WW}\left(W_t,z_t\right)\Delta W_{t+1}+V_{Wz}\left(W_t,z_t\right)\Delta z_{t+1}}{\mathbb{E}\left[V_W\left(W_{t+1},z_{t+1}\right)\right]},R^j_{t+1}\right]=$$

$$=\underbrace{-\frac{V_{WW}\left(W_t,z_t\right)}{\mathbb{E}_t\left[V_W\left(W_{t+1},z_{t+1}\right)\right]}}_{R_R\ Coefficient}Cov_t\left[\Delta W_{t+1},R^j_{t+1}\right]\underbrace{-\frac{V_{Wz}\left(W_t,z_t\right)}{\mathbb{E}_t\left[V_W\left(W_{t+1},z_{t+1}\right)\right]}}_{Additional\ "Risk\ Factor"}Cov_t\left[\Delta z_{t+1},R^j_{t+1}\right]$$

Now assume the representative agent has CRRA utilityfunction, then

$$1=\mathbb{E}_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}R^j_{t+1}\right]$$

With a second-order Taylor approximation we get

$$0=\ln\delta-\gamma\mathbb{E}_t\left[\triangle c_{t+1}\right]+\mathbb{E}_t\left[r^j_{t+1}\right]+\frac{1}{2}\left[\gamma^2 Var_{cc}+Var_{jj}-2\gamma Cov_{cj}\right]$$

Where $c_t\equiv\ln C_t$, $r^j_t\equiv\ln R^j_t$, $Var_{cc}\equiv Var_t\left[\Delta c_{t+1}\right]$, $Var_{jj}\equiv Var_t\left[\Delta r^j_{t+1}\right]$ and $Cov_{cj}\equiv Cov_t\left[\Delta c_{t+1},\Delta r^j_{t+1}\right]$. For the market portfolio $j=m$, this can be rewritten as

$$\mathbb{E}_t\left[\triangle c_{t+1}\right]=\mu_m+\frac{1}{\gamma}\mathbb{E}_t\left[r^m_{t+1}\right]$$

Where

$$\mu_m=\frac{1}{\gamma}\ln\delta+\frac{1}{2}\left[\gamma Var_{cc}+\frac{Var_{mm}}{\gamma}-2Cov_{cm}\right]$$

The budget constraint is

$$W_{t+1}=R^m_{t+1}\left(W_t-C_t\right)$$

Or equivalently

$$\frac{W_{t+1}}{W_t}=R^m_{t+1}\left(1-\frac{C_t}{W_t}\right)$$

Or, in logs,

$$w_{t+1}=r^m_{t+1}+\ln\left(1-e^{c_t-w_t}\right)$$

Taylor expanding $\ln\left(1-e^{c_t-w_t}\right)$ around $(\bar{c},\bar{w})$ we get

$$\ln\left(1-e^{c_t-w_t}\right)\approx\ln\left(1-e^{\bar{c}-\bar{w}}\right)-\frac{e^{\bar{c}-\bar{w}}}{1-e^{\bar{c}-\bar{w}}}\left(c_t-w_t-\left(\bar{c}-\bar{w}\right)\right)$$

Letting $\rho = (1 - e^{\bar{c}-\bar{w}})$ we can write

$$w_{t+1} = r^m_{t+1} + k + \left(1 - \frac{1}{\rho}\right)(c_t - w_t)$$

Where $k = \ln(1 - e^{\bar{c}-\bar{w}}) + \frac{e^{\bar{c}-\bar{w}}}{1-e^{\bar{c}-\bar{w}}}(\bar{c} - \bar{w})$ is a constant. Using the identity $\Delta w_{t+1} = \Delta c_{t+1} + (c_t - w_t) - (c_{t+1} - w_{t+1})$ we have

$$c_t - w_t = \sum_{k=1}^{\infty} \rho^k \left(r^m_{t+1} - \Delta c_{t+k}\right) + \frac{\rho k}{1-\rho}$$

Taking expectations,

$$c_t - w_t = \sum_{k=1}^{\infty} \rho^k \mathbb{E}_t \left[r^m_{t+k} - \Delta c_{t+k}\right] + \frac{\rho k}{1-\rho}$$

Combining this with the log-linearized budget constraint we get

$$c_{t+1} - \mathbb{E}_t c_{t+1} = \sum_{k=0}^{\infty} \rho^k \left(r^m_{t+k+1} - \mathbb{E}_t \left[r^m_{t+k+1}\right]\right) - \sum_{k=0}^{\infty} \rho^k \left(\Delta c_{t+k+1} - \mathbb{E}_t \left[\Delta c_{t+k+1}\right]\right)$$

Combining this with $\mathbb{E}_t [\triangle c_{t+1}] = \mu_m + \frac{1}{\gamma}\mathbb{E}_t \left[r^m_{t+1}\right]$ we get

$$c_{t+1} - \mathbb{E}_t c_{t+1} = r^m_{t+1} - \mathbb{E}_t r^m_{t+1} + \left(1 - \frac{1}{\gamma}\right) \sum_{k=0}^{\infty} \rho^k \left(r^m_{t+k+1} - \mathbb{E}_t \left[r^m_{t+k+1}\right]\right)$$

Finally, this implies that

$$Var_{jc} = Var_{jm} + \left(1 - \frac{1}{\gamma}\right) Var_{jh}$$

Where $Var_{jh} = Cov_t \left[r^j_{t+1}, \sum_{k=0}^{\infty} \rho^k \left(r^m_{t+k+1} - \mathbb{E}_t \left[r^m_{t+k+1}\right]\right)\right]$, which is the covariance of asset j with a hedge portfolio h. For a risk-free asset the log-Euler equation simplifies to

$$0 = \ln\delta - \gamma \mathbb{E}_t [\Delta c_{t+1}] + r^f_{t+1} + \frac{1}{2}\gamma^2 Var_{cc}$$

Then we can write the Consumption CAPM as

$$\mathbb{E}_t \left[r^j_{t+1}\right] - r^f_{t+1} = -\frac{Var_{jj}}{2} + \gamma Var_{jc}$$

And finally, the Intertemporal CAPM as

$$\mathbb{E}_t \left[r^j_{t+1}\right] - r^f_{t+1} = -\frac{Var_{jj}}{2} + \gamma Var_{jm} + (\gamma - 1) Var_{jh}$$

Appendix: The Bellman Equation

Under certain conditions, we can show that a Bellman equation will have a well- behaved solution. There are a handful of ways to solve such an equation. Here, we will consider two in the context of a simple problem where $u(c) = \log(c)$ and $g(w, c) = w - c$.

1) Iteration: in this method, we begin with some guess for $V_0(w)$ (for example, $V_0(w) = 0$) and solve the first order condition to find $V_1(w) = \max_c \{u(c) + \beta V_0(g(w, c))\}$. We iterate again to obtain V_2, V_3, etc. and the true value function is given by the limit $\lim_{k \to \infty} V_k$.

1. Solve the first order condition to obtain $c_0 = h_0(w)$.
2. Plug $h_0(w)$ into $V_1(w) = u(h_0(w)) + \beta V_0(g(w, h_0(w)))$.
3. Repeat to obtain $V_2(w)$, $V_3(w)$, and so on. Can you guess what is $V_k(w)$?
4. Take the limit $\lim_{k \to \infty} V_k$ (for the specific example $u(c) = \log(c)$ and $g(w, c) = w - c$, use your knowledge of geometric series).

2) Guess and Verify: in this case, we guess a form of the value function $V(w) = A + B\log(w)$ and solve for the corresponding coefficients A and B that make this guess correct. In particular, we leave the coefficients A and B as variables in the problem, plug in the consumption c implied by the FOC and then solve for A and B that make the Bellman Equation consistent.

1. Begin with $V(w) = A + B\log(w)$ and plug into the Bellman equation to obtain $V(w) = \max_c \{\log(c) + \beta \log(w - c)\}$.
2. Take the FOC with respect to c and plug the solution back into the equation above to solve for the right-hand side.
3. Note that the left-hand side of the Bellman Equation must also be equal to $V(w) = A + B\log(w)$. Thus, the coefficient on the right-hand side on $\log(w)$ must be equal to B. Also, the constant on the right-hand side must be equal to A. Compute A and B and verify that this result yields the same V obtained with the iteration method.

Extension on Chapter 7: Multi-Period Utility

Ponzi Schemes and Rational Bubbles Allowing for an infinite horizon allows agents to borrow an arbitrarily large amount and roll over debt forever, without ever repaying the debt. This is referred to as a *Ponzi scheme*: it allows infinite consumption. Consider a model with infinite horizon, no uncertainty and a complete set of short-lived bonds. Let z_t be the amount of bonds in the portfolio maturing in period t, and let p_t be the price of one such bond in period $t-1$ (just like before). The problem the agent solves is

$$\max_{\{c_t\}} \sum_{t=0}^{\infty} \delta^t u\left(c_t\right)$$

$$s.t.\ c_0 - w_0 \leq -p_1 z_1,\ c_t - w_t \leq z_t - p_{t+1} z_{t+1}$$

For all $t > 0$. The consumption path $\{c_t\}$ such that $c_t = w_t + 1$ is feasible: note that the agent consumes more than his endowment in each period, forever! This can be financed with increasing debt:

$$z_{t+1} = \frac{z_t - 1}{p_{t+1}}$$

For all $t \geq 0$ and $z_0 = 0$. However, Ponzi schemes can never be part of an equilibrium because they remove the existence of a utility maximum, since the choice set of the agent is unbounded. Therefore we need an additional constraint, called *transversality condition*:

$$\lim_{t \to \infty} p_t z_t \geq 0$$

Which implies that the value of debt cannot diverge to infinity, or equivalently, that all debt must eventually be redeemed. There are also additional solutions to this problem that incorporate a "bubble" component to the asset price: consider again a model with no uncertainty and a consol delivering $1 in each period, forever. Our formula states that

$$p_t M_t = \mathbb{E}_t\left[M_{t+1}\left(p_{t+1} + x_{t+1}\right)\right]$$

In our case $x_t \equiv 1$ and we have no uncertainty, so we can write

$$M_t p_t = M_{t+1} p_{t+1} + M_{t+1}$$

Solving forward we obtain

$$p_0 = \underbrace{\sum_{t=1}^{\infty} M_t}_{Fundamental\ Value} + \underbrace{\lim_{T\to\infty} M_T p_T}_{Bubble\ Component}$$

Recall that according to our "static" dynamic model we should have $p_0 = \sum_{t=1}^{\infty} M_t$, with $M_t = m_1 \times m_2 \times \ldots \times m_t$. Recall that $M_t = \delta^t \frac{u'(c_t)}{u'(c_0)}$, so

$$p_0 = \sum_{t=1}^{\infty} \delta^t \frac{u'(c_t)}{u'(c_0)}$$

At time t this becomes

$$p_t = \sum_{s=t+1}^{\infty} \delta^{s-t} \frac{u'(c_s)}{u'(c_t)}$$

So at $t+1$ we have (since $m_t = \delta \frac{u'(c_t)}{u'(c_{t-1})}$)

$$p_{t+1} = \sum_{s=t+2}^{\infty} \delta^{s-t-1} \frac{u'(c_s)}{u'(c_{t+1})} = \frac{1}{\delta} \frac{u'(c_t)}{u'(c_{t+1})} \sum_{s=t+2}^{\infty} \delta^{s-t} \frac{u'(c_s)}{u'(c_t)} = \frac{1}{\delta} \frac{u'(c_t)}{u'(c_{t+1})} \left(\sum_{s=t+1}^{\infty} \delta^{s-t} \frac{u'(c_s)}{u'(c_t)} - \delta \frac{u'(c_{t+1})}{u'(c_t)} \right) =$$

$$= \frac{1}{\delta} \frac{u'(c_t)}{u'(c_{t+1})} \left(p_t - \delta \frac{u'(c_{t+1})}{u'(c_t)} \right) = \frac{1}{m_{t+1}} (p_t - m_{t+1})$$

Which can be rewritten as

$$p_t = p_{t+1} m_{t+1} + m_{t+1}$$

Solving this forward we get again

$$p_0 = m_1 + m_1 m_2 + \ldots = \sum_{t=1}^{\infty} M_t + \lim_{T\to\infty} M_T p_T$$

The fundamental value is the price given by the static-dynamic model, however repeated trading gives rise to the possibility of a rational bubble. For example, fiat money can be understood as an asset with no dividends. In the static-dynamic model, such an asset would have no value (because the present value of zero is zero). But if there is a bubble on the price of fiat money, then it can have a positive value (see Bewley, 1980). In asset pricing theory, we often rule out bubbles simply by imposing

$$\lim_{T\to\infty} M_T p_T = 0$$

Time Preferences We will assume a Von Neumann-Morgenstern time-separable utility function

$$u(c_0) + \sum_{t>0} \delta(t) \mathbb{E}[u(c_t)]$$

Where $\delta(t) \in (0, 1]$ is the discount factor and $\delta(t) > \delta(t+1)$. $\delta(t)$ also represents the intertemporal rate of substitution. The agent solves for the whole comsumption plan $(c_0, c_1, \ldots)$ (for each event on the tree) at $t = 0$. The two main functional forms are (1) exponential discounting: $\delta(t) = \delta^t$ and

$$\mathbb{E}\left[\sum_{t\geq 0} \delta^t u(c_t)\right]$$

In this functional form the agent prefers early resolution of uncrtainty if this affects his actions, and is indifferent otherwise. The other main functional form for $\delta(t)$ is the hyperbolic discounting formulation:[2]

$$\mathbb{E}\left[u(c_0) + \beta \sum_{t>0} \delta^t u(c_t)\right]$$

Which can be expressed in recursive form and is connected to the preference for the timing of uncertainty resolution.

We can extend our consumption-savings problem encountered in chapter 4 to a multi-period setting: the problem becomes

$$\max_{\{s_t, a_t\}_{t=1}^{T-1}} \mathbb{E}\left[\sum_{t=0}^{T} \delta^t U(c_t)\right]$$

$$s.t.\ c_T = s_{T-1}(1 + r_f) + a_{T-1}(r_T - r_f)$$

$$c_t + s_t \leq s_{t-1}(1 + r_f) + a_{t-1}(r_T - r_f)$$

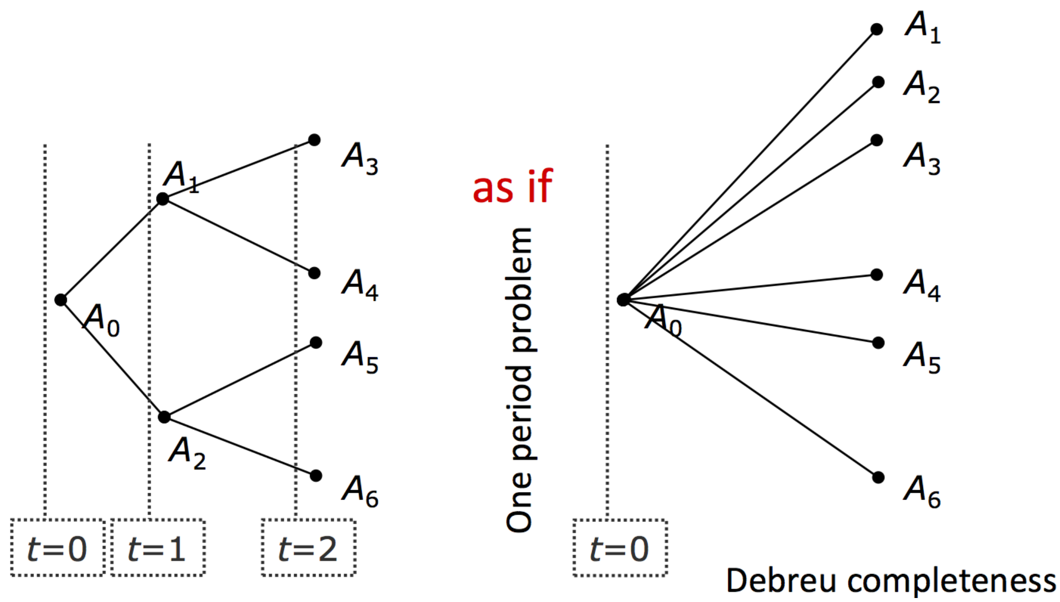
$$c_0 + s_0 \leq Y_0$$

A result similar to that found in chapter 4 holds as well:

> **Theorem** (Merton, 1971): For CRRA utility, constant r_f and i.i.d. $\{r_t\}_{t=1}^{T}$ the ratio $\frac{a_t}{s_t}$ is time-invariant

A "Static" Dynamic model The Debreu completeness refers to a setting in which information is gradually revealed over many periods, but all decisions on asset trading are made at time $t = 0$: decision making is static, even if the model is dynamic.

[2]More precisely, this is a special case of hyperbolic discounting.



If Arrow-Debreu securities conditional on each event A are tradable, the representative agent solves

$$\max_{\{c_t\}\in B(y)} \sum_{t\geq 0} \delta^t E\left[u\left(c_t\right)\right]$$

Where $B(y)$ is the intertemporal budget constraint (as a function of the income stream $\{y_A\}$):

$$q_A\left(y_A - c_A\right) \leq w_A$$

Where q_A is the state price of the event A. The FOC with respect to $c_{t(A_0)} = c_0$ yields

$$u'\left(c_0\right) = \lambda$$

Since $\pi_0 = 1$, $t\left(A_0\right) = 0$ and $q_{A_0} = 1$. Assuming the constraint is binding, for a generic event A instead we get

$$\delta^{t(A)} \pi_A u'\left(c_A\right) = \lambda q_A$$

Therefore we obtain the equilibrium stochastic discount factor:

$$\frac{q_A}{\pi_A} = \delta^{t(A)} \frac{u'\left(c_A\right)}{u'\left(c_0\right)} \equiv M_{t(A)}\left(A\right)$$

We call the vector M_t the *multi-period stochastic discount factor*. Note that this can be written as

$$M_{t(A)}\left(A\right) = \delta^{t(A)} \frac{u'\left(c_A\right)}{u'\left(c_0\right)} = \underbrace{\left(\delta \frac{u'\left(c_{\psi_1(A)}\right)}{u'\left(c_0\right)}\right)}_{m_{\psi_1(A)}} \cdot \underbrace{\left(\delta \frac{u'\left(c_{\psi_2(A)}\right)}{u'\left(c_{\psi_1(A)}\right)}\right)}_{m_{\psi_2(A)}} \cdots \underbrace{\left(\delta \frac{u'\left(c_{\psi_{t(A)}(A)}\right)}{u'\left(c_{\psi_{t(A)-1}(A)}\right)}\right)}_{m_{\psi_{t(A)}(A)}} = \prod_{t=1}^{t(A)} m_{\psi_t(A)}$$

Where $m_{\psi_t(A)}$ is the usual "one-period ahead" stochastic discount factor. The fundamental pricing formula for an asset x paying a cash flow stream $\left\{x_t^j\right\}_{t=1}^{\infty}$ is just

$$p^j = \sum_{t=1}^{\infty} \mathbb{E}\left[M_t x_t^j\right]$$

If the representative agent is risk-neutral, $M_t = \delta^t$ and

$$p^j = \sum_{t=1}^{\infty} \delta^t \mathbb{E}\left[x_t^j\right]$$

Bibliography

[1] Markus K. Brunnermeier, Lecture Notes for the course FIN 501: Asset Pricing, 2013.

[2] Jakub Jurek, Lecture Notes for the course FIN 521: Fixed Income Models and Applications, 2012.

[3] John H. Cochrane, Asset Pricing revised edition, Princeton University Press, 2002.

[4] John Y. Campbell, Andrew W. Lo, A. Craig McKinay, The Econometrics of Financial Markets, Princeton University Press, 1996.

[5] Robert L. McDonald, Derivatives Markets, Pearson series in finance, 2009.

[6] Kerry E. Back, Asset Pricing and Portfolio Choice Theory, Oxford University Press, 2010.

[7] Testing Merton's Model for Credit Spreads on Zero-Coupon Bonds, Gordon Gemill, 2003.

[8] Stephen Morris, Lecture Notes for the course ECO 317: The Economics of Uncertainty, 2011.

[9] Edwin J. Elton, Martin J. Gruber, Deepak Agrawal, and Christopher Mann, Explaining the Rate Spread on Corporate Bonds, Journal of Finance, 2001.

[10] John Y. Campbell and Robert J. Shiller, Valuation Ratios and the Long-Run Stock Market Outlook, The Journal of Portfolio Management, 1998.