Lecture Notes for the Course Investerings- og Finansieringsteori.

David Lando
Rolf Poulsen

January 2006

net present value, NPV

positivity of vectors

Chapter 3

Payment Streams under Certainty

3.1 Financial markets and arbitrage

In this section we consider a very simple setup with no uncertainty. There are three reasons that we do this:

First, the terminology of bond markets is conveniently introduced in this setting, for even if there were uncertainty in our model, bonds would be characterized by having payments whose size at any date are constant and known in advance.

Second, the classical net present value (NPV) rule of capital budgeting is easily understood in this framework.

And finally, the mathematics introduced in this section will be extremely useful in later chapters as well.

A note on notation: If $v \in \mathbb{R}^N$ is a vector the following conventions for "vector positivity" are used:

- $v \geq 0$ ("v is non-negative") means that all of v's coordinates are non-negative. ie. $\forall i$: $v_i \geq 0$.

- $v > 0$ ("v is positive") means that $v \geq 0$ and that at least one coordinate is strictly positive, ie. $\forall i$: $v_i \geq 0$ and $\exists i$: $v_i > 0$, or differently that $v \geq 0$ and $v \neq 0$.

- $v \gg 0$ ("v is strictly positive") means that every coordinate is strictly positive, $\forall i$: $v_i > 0$. This (when v is N-dimensional) we will sometimes write as $v \in \mathbb{R}^N_{++}$. (This saves a bit of space, when we want to indicate both strict positivity and the dimension of v.)

financial market
security price system
payment stream
portfolio
short position
long position
arbitrage opportunity

Throughout we use $v^\top$ to denote the transpose of the vector v. Vectors without the transpose sign are always thought of as column vectors.

We now consider a model for a financial market (sometimes also called a security market or price system; individual components are then referred to as securities) with $T+1$ dates: $0, 1, \ldots, T$ and no uncertainty.

Definition 1 *A* financial market *consists of a pair* (π, C) *where* $\pi \in \mathbb{R}^N$ *and* C *is an* $N \times T-$*matrix.*

The interpretation is as follows: By paying the price π_i at date 0 one is entitled to a stream of payments $(c_{i1}, \ldots, c_{iT})$ at dates $1, \ldots, T$. Negative components are interpreted as amounts that the owner of the security has to pay. There are N different payment streams trading. But these payment streams can be bought or sold in any quantity and they may be combined in *portfolios* to form new payment streams:

Definition 2 *A portfolio* θ *is an element of* $\mathbb{R}^N$. *The payment stream generated by* θ *is* $C^\top\theta \in \mathbb{R}^T$. *The price of the portfolio* θ *at date* 0 *is* $\pi \cdot \theta$ *(*$= \pi^\top\theta = \theta^\top\pi$*).*

Note that allowing portfolios to have negative coordinates means that we allow securities to be sold. We often refer to a negative position in a security as a *short* position and a positive position as a *long* position. Short positions are not just a convenient mathematical abstraction. For instance when you borrow money to buy a home, you take a short position in bonds.

Before we even think of adopting (π, C) as a model of a security market we want to check that the price system is sensible. If we think of the financial market as part of an equilibrium model in which the agents use the market to transfer wealth between periods, we clearly want a payment stream of $(1, \ldots, 1)$ to have a lower price than that of $(2, \ldots, 2)$. We also want payment streams that are non-negative at all times to have a non-negative price. More precisely, we want to rule out arbitrage opportunities in the security market model:

Definition 3 *A portfolio* θ *is an arbitrage opportunity (of type 1 or 2) if it satisfies one of the following conditions:*

1. $\pi \cdot \theta = 0$ *and* $C^\top\theta > 0$.

2. $\pi \cdot \theta < 0$ *and* $C^\top\theta \geq 0$.

Alternatively, we can express this as $(-\pi \cdot \theta, C^\top\theta) > 0$.

separating hyperplane
Stiemke's lemma
discount factors

The interpretation is that it should not be possible to form a portfolio at zero cost which delivers non-negative payments at all future dates and even gives a strictly positive payment at some date. And it should not be possible to form a portfolio at negative cost (i.e. a portfolio which gives the owner money now) which never has a negative cash flow in the future.
Usually type 1 arbitrages can be transformed into type 2. arbitrages, and vice versa. For instance, if the exists a $c_i > 0$, then we easily get from 2 to 1 But there is not mathematical equivalence (take $\pi = 0$ or $C = 0$ to see this).

Definition 4 *The security market is arbitrage-free if it contains no arbitrage opportunities.*

To give a simple characterization of arbitrage-free markets we need a lemma which is very similar to Farkas' theorem of alternatives proved in Matematik 2OK using separating hyperplanes:

Lemma 1 *(Stiemke's lemma) Let A be an $n \times m-$matrix: Then precisely one of the following two statements is true:*

1. There exists $x \in \mathbb{R}^m_{++}$ such that $Ax = 0$.
2. There exists $y \in \mathbb{R}^n$ such that $y^\top A > 0$.

We will not prove the lemma here (it is a very common exercise in convexity/linear programming courses, where the name Farkas is encountered). But it is the key to our next theorem:

Theorem 2 *The security market (π, C) is arbitrage-free if and only if there exists a strictly positive vector $d \in \mathbb{R}^T_{++}$ such that $\pi = Cd$.*

In the context of our security market the vector d will be referred to as a vector of discount factors. This use of language will be clear shortly.
Proof. Define the matrix

$$A = \begin{pmatrix} -\pi_1 & c_{11} & c_{12} & \cdots & c_{1T} \\ -\pi_2 & c_{21} & c_{22} & \cdots & c_{2T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\pi_N & c_{N1} & c_{N2} & \cdots & c_{NT} \end{pmatrix}$$

First, note that the existence of $x \in \mathbb{R}^{T+1}_{++}$ such that $Ax = 0$ is equivalent to the existence of a vector of discount factors since we may define

$$d_i = \frac{x_i}{x_0} \qquad i = 1, \dots, T.$$

complete market

Hence if the first condition of Stiemke's lemma is satisfied, a vector d exists such that $\pi = Cd$. The second condition corresponds to the existence of an arbitrage opportunity: If $y^\top A > 0$ then we have either

$$(y^\top A)_1 > 0 \text{ and } (y^\top A)_i \geq 0 \quad i = 1, \ldots, T+1$$

or

$$(y^\top A)_1 = 0 \text{ , } y^\top A \geq 0 \text{ and } (y^\top A)_i > 0 \quad \text{some } i \in \{2, \ldots, T+1\}$$

and this is precisely the condition for the existence of an arbitrage opportunity. Now use Stiemke's lemma. ■

Another important concept is market completeness (in Danish: *Komplethed* or *fuldstændighed*).

Definition 5 *The security market is complete if for every $y \in \mathbb{R}^T$ there exists a $\theta \in \mathbb{R}^N$ such that $C^\top \theta = y$.*

In linear algebra terms this means that the rows of C span $\mathbb{R}^T$, which can only happen if $N \geq T$, and in our interpretation it means that any desired payment stream can be generated by an appropriate choice of portfolio.

Theorem 3 *Assume that (π, C) is arbitrage-free. Then the market is complete if and only if there is a unique vector of discount factors.*

Proof. Since the market is arbitrage-free we know that there exists $d \gg 0$ such that $\pi = Cd$. Now if the model is complete then $\mathbb{R}^T$ is spanned by the columns of $C^\top$, ie. the rows of C of which there are N. This means that C has T linearly independent rows, and from basic linear algebra (look around where *rank* is defined) it also has T linearly independent columns, which is to say that all the columns are independent. They therefore form a basis for a T-dimensional linear subspace of $\mathbb{R}^N$ (remember we must have $N \geq T$ to have completeness), ie. any vector in this subspace has unique representation in terms of the basis-vectors. Put differently, the equation $Cx = y$ has at most one solution. And in case where $y = \pi$, we know there is one by absence of arbitrage. For the other direction assume that the model is incomplete. Then the columns of C are linearly dependent, and that means that there exists a vector $\widetilde{d} \neq 0$ such that $0 = C\widetilde{d}$. Since $d \gg 0$, we may choose $\epsilon > 0$ such that $d + \epsilon\widetilde{d} \gg 0$. Clearly, this produces a vector of discount factors different from d. ■

3.2 Zero coupon bonds and the term structure

zero coupon bond, ZCB
discount factors
forward rates
short rate

Assume throughout this section that the model (π, C) is complete and arbitrage-free and let $d^\top = (d_1, \ldots, d_T)$ be the unique vector of discount factors. Since there must be at least T securities to have a complete model, C must have at least T rows. On the other hand if C has exactly T linearly independent rows, then adding other securities to C will not add any more possibilities of wealth transfer to the market. Hence we can assume that C is am invertible $T \times T$ matrix.

Definition 6 *The payment stream of a zero coupon bond with maturity t is given by the t'th unit vector e_t of $\mathbb{R}^T$.*

Next we see why the words discount factors were chosen:

Proposition 4 *The price of a zero coupon bond with maturity t is d_t.*

Proof. Let θ_t be the portfolio such that $C^\top \theta_t = e_t$. Then

$$\pi^\top \theta_t = (Cd)^\top \theta_t = d^\top C^\top \theta_t = d^\top e_t = d_t. \quad \blacksquare$$

Note from the definition of d that we get the value of a stream of payments c by computing $\sum_{t=1}^T c_t d_t$. In other words, the value of a stream of payments is obtained by discounting back the individual components. There is nothing in our definition of d which prevents $d_s > d_t$ even when $s > t$, but in the models we will consider this will not be relevant: It is safe to think of d_t as decreasing in t corresponding to the idea that the longer the maturity of a zero coupon bond, the smaller is its value at time 0.

From the discount factors we may derive/define various types of interest rates which are essential in the study of bond markets.:

Definition 7 *(Short and forward rates.) The short rate at date* 0 *is given by*

$$r_0 = \frac{1}{d_1} - 1.$$

The (one-period) time t- forward rate at date 0, *is equal to*

$$f(0,t) = \frac{d_t}{d_{t+1}} - 1,$$

where $d_0 = 1$ by convention.

yield to maturity
term structure of interest rates

The interpretation of the short rate should be straightforward: Buying $\frac{1}{d_1}$ units of a maturity 1 zero coupon bond costs $\frac{1}{d_1}d_1 = 1$ at date 0 and gives a payment at date 1 of $\frac{1}{d_1} = 1 + r_0$. The forward rate tells us the rate at which we may agree at date 0 to borrow (or lend) between dates t and $t+1$. To see this, consider the following strategy at time 0 :

- Sell 1 zero coupon bond with maturity t.
- Buy $\frac{d_t}{d_{t+1}}$ zero coupon bonds with maturity $t+1$.

Note that the amount raised by selling precisely matches the amount used for buying and hence the cash flow from this strategy at time 0 is 0. Now consider what happens if the positions are held to the maturity date of the bonds: At date t the cash flow is then -1 and at date $t+1$ the cash flow is $\frac{d_t}{d_{t+1}} = 1 + f(0,t)$.

Definition 8 *The yield (or yield to maturity) at time* 0 *of a zero coupon bond with maturity* t *is given as*

$$y(0,t) = \left(\frac{1}{d_t}\right)^{\frac{1}{t}} - 1.$$

Note that

$$d_t(1 + y(0,t))^t = 1.$$

and that one may therefore think of the yield as an 'average interest rate' earned on a zero coupon bond. In fact, the yield is a geometric average of forward rates:

$$1 + y(0,t) = ((1 + f(0,0)) \cdots (1 + f(0,t-1)))^{\frac{1}{t}}$$

Definition 9 *The term structure of interest rates (or the yield curve) at date 0 is given by* $(y(0,1),\ldots,y(0,T))$.

Note that if we have any one of the vector of yields, the vector of forward rates and the vector of discount factors, we may determine the other two. Therefore we could equally well define a term structure of forward rates and a term structure of discount factors. In these notes unless otherwise stated, we think of *the term structure of interest rates* as the yields of zero coupon bonds as a function of time to maturity. It is important to note that the term structure of interest rate depicts yields of zero coupon bonds. We do however also speak of yields on securities with general positive payment steams:

yield to maturity
compounding periods
continuously compounded interest rate

Definition 10 *The yield (or yield to maturity) of a security* $c^\top = (c_1, \ldots, c_T)$ *with* $c > 0$ *and price* π *is the unique solution* $y > -1$ *of the equation*

$$\pi = \sum_{i=1}^{T} \frac{c_i}{(1+y)^i}.$$

Example 3 (Compounding Periods) In most of the analysis in this chapter the time is "stylized"; it is measured in some unit (which we think of and refer to as "years") and cash-flows occur at dates $\{0, 1, 2, \ldots, T\}$. But it is often convenient (and not hard) to work with dates that are not integer multiples of the fundamental time-unit. We quote interest rates in units of years^{-1} ("per year'), but to any interest rate there should be a number, m, associated stating how often the interest is compounded. By this we mean the following: If you invest 1 \$ for n years at the m-compounded rate r_m you end up with

$$\left(1 + \frac{r_m}{m}\right)^{mn}. \tag{3.1}$$

The standard example: If you borrow 1\$ in the bank, a 12% interest rate means they will add 1% to you debt each month (i.e. $m = 12$) and you will end up paying back 1.1268 \$ after a year, while if you make a deposit, they will add 12% after a year (i.e. $m = 1$) and you will of course get 1.12\$ back after one year. If we keep r_m and n fixed in (3.1) (and then drop the m-subscript) and and let m tend to infinity, it is well known that we get:

$$\lim_{m\to\infty} \left(1 + \frac{r}{m}\right)^{mn} = e^{nr},$$

and in this case we will call r the continuously compounded interest rate. In other words: If you invest 1 \$ and the continuously compounded rate r_c for a period of length t, you will get back e^{tr_c}. Note also that a continuously compounded rate r_c can be used to find (uniquely for any m) r_m such that 1 \$ invested at m-compounding corresponds to 1 \$ invested at continuous compounding, i.e.

$$\left(1 + \frac{r_m}{m}\right)^{m} = e^{r_c}.$$

This means that in order to avoid confusion – even in discrete models – there is much to be said in favor of quoting interest rates on a continuously compounded basis. But then again, in the highly stylized discrete models it would be pretty artificial, so we will not do it (rather it will always be $m = 1$).

annuity
serial loan
bullet bond
annuity

3.3 Annuities, serial loans and bullet bonds

Typically, zero-coupon bonds do not trade in financial markets and one therefore has to deduce prices of zero-coupon bonds from other types of bonds trading in the market. Three of the most common types of bonds which do trade in most bond markets are annuities, serial loans and bullet bonds. (In literature relating to the American market, "bond" is usually understood to mean "bullet bond with 2 yearly payments". Further, "bills" are term short bonds, annuities explicitly referred to as such, and serial loans rare.) We now show how knowing to which of these three types a bond belongs and knowing three characteristics, namely the maturity, the principal and the coupon rate, will enable us to determine the bond's cash flow completely.

Let the principal or face value of the bond be denoted F. Payments on the bond start at date 1 and continue to the time of the bond's maturity, which we denote τ. The payments are denoted c_t. We think of the principal of a bond with coupon rate R and payments $c_1, \dots, c_\tau$ as satisfying the following difference equation:

$$p_t = (1+R)p_{t-1} - c_t \qquad t = 1, \dots, \tau, \tag{3.2}$$

with the boundary conditions $p_0 = F$ and $p_\tau = 0$.

Think of p_t as the remaining principal right after a payment at date t has been made. For accounting and tax purposes and also as a helpful tool in designing particular types of bonds, it is useful to split payments into a part which serves as reduction of principal and one part which is seen as an interest payment. We define the reduction in principal at date t as

$$\delta_t = p_{t-1} - p_t$$

and the interest payment as

$$i_t = Rp_{t-1} = c_t - \delta_t.$$

Definition 11 *An annuity with maturity τ, principal F and coupon rate R is a bond whose payments are constant between dates 1 and τ, and whose principal evolves according to Equation (3.2).*

With constant payments we can use (3.2) repeatedly to write the remaining principal at time t as

$$p_t = (1+R)^t F - c\sum_{j=0}^{t-1}(1+R)^j \quad \text{for } t = 1, 2, \dots, \tau.$$

To satisfy the boundary condition $p_\tau = 0$ we must therefore have alfahage; $“alpha ˙n“rceil R $

$$F - c\sum_{j=0}^{\tau-1}(1+R)^{j-\tau} = 0,$$

so by using the well-known formula $\sum_{i=0}^{n-1} x^i = (x^n-1)/(x-1)$ for the summation of a geometric series, we get

$$c = F\left(\sum_{j=0}^{\tau-1}(1+R)^{j-\tau}\right)^{-1} = F\frac{R(1+R)^\tau}{(1+R)^\tau - 1} = F\frac{R}{1-(1+R)^{-\tau}}.$$

Note that the size of the payment is homogeneous (of degree 1) in the principal, so it's usually enough to look at the $F = 1$. (This rather trivial observation can in fact be extremely useful in a dynamic context.) It is common to use the shorthand notation

$$\alpha_{n\rceil R} = \text{ (“Alfahage”)} = \frac{(1+R)^n - 1}{R(1+R)^n}.$$

Having found what the size of the payment must be we may derive the interest and the deduction of principal as well: Let us calculate the size of the payments and see how they split into deduction of principal and interest payments. First, we derive an expression for the remaining principal:

$$\begin{aligned}
p_t &= (1+R)^t F - \frac{F}{\alpha_{\tau\rceil R}}\sum_{j=0}^{t-1}(1+R)^j \\
&= \frac{F}{\alpha_{\tau\rceil R}}\left((1+R)^t\alpha_{\tau\rceil R} - \frac{(1+R)^t - 1}{R}\right) \\
&= \frac{F}{\alpha_{\tau\rceil R}}\left(\frac{(1+R)^\tau - 1}{R(1+R)^{\tau-t}} - \frac{(1+R)^\tau - (1+R)^{\tau-t}}{R(1+R)^{\tau-t}}\right) \\
&= \frac{F}{\alpha_{\tau\rceil R}}\alpha_{\tau-t\rceil R}.
\end{aligned}$$

This gives us the interest payment and the deduction immediately for the annuity:

$$\begin{aligned}
i_t &= R\frac{F}{\alpha_{\tau\rceil R}}\alpha_{\tau-t+1\rceil R} \\
\delta_t &= \frac{F}{\alpha_{\tau\rceil R}}(1 - R\alpha_{\tau-t+1\rceil R}).
\end{aligned}$$

In the definition of an annuity, the size of the payments is implicitly defined. The definitions of bullets and serials are more direct.

bullet bond
serial loan

Definition 12 *A bullet bond*[1] *with maturity τ, principal F and coupon rate R is characterized by having $i_t = c_t$ for $t = 1, \ldots, \tau - 1$ and $c_\tau = (1 + R)F$.*

The fact that we have no reduction in principal before τ forces us to have $c_t = RF$ for all $t < \tau$.

Definition 13 *A serial loan or bond with maturity τ, principal F and coupon rate R is characterized by having δ_t, constant for all $t = 1, \ldots, \tau$.*

Since the deduction in principal is constant every period and we must have $p_\tau = 0$, it is clear that $\delta_t = \frac{F}{\tau}$ for $t = 1, \ldots, \tau$. From this it is straightforward to calculate the interest using $i_t = Rp_{t-1}$.

We summarize the characteristics of the three types of bonds in the table below:

	payment	interest	deduction of principal
Annuity	$F\alpha^{-1}_{\tau\rceil R}$	$R\frac{F}{\alpha_{\tau\rceil R}}\alpha_{\tau-t+1\rceil R}$	$\frac{F}{\alpha_{\tau\rceil R}}(1 - R\alpha_{\tau-t+1\rceil R})$
Bullet	RF for $t < \tau$ $(1 + R)F$ for $t = \tau$	RF	0 for $t < \tau$ F for $t = \tau$
Serial	$\frac{F}{\tau} + R\left(F - \frac{t-1}{\tau}F\right)$	$R\left(F - \frac{t-1}{\tau}F\right)$	$\frac{F}{\tau}$

Example 4 (A Simple Bond Market) Consider the following bond market where time is measured in years and where payments are made at dates $\{0, 1, \ldots, 4\}$:

Bond (i)	Coupon rate (R_i)	Price at time 0 ($\pi_i(0)$)
1 yr bullet	5	100.00
2 yr bullet	5	99.10
3 yr annuity	6	100.65
4 yr serial	7	102.38

We are interested in finding the zero-coupon prices/yields in this market. First we have to determine the payment streams of the bonds that are traded (the C-matrix). Since $\alpha_{3\rceil 6} = 2.6730$ we find that

$$C = \begin{bmatrix} 105 & 0 & & 0 \\ 5 & 105 & 0 & 0 \\ 37.41 & 37.41 & 37.41 & 0 \\ 32 & 30.25 & 28.5 & 26.75 \end{bmatrix}$$

[1] In Danish: Et stående lån

Clearly this matrix is invertible so $e_t = C^\top \theta_t$ has a unique solution for all $t \in \{1, \ldots, 4\}$ (namely $\theta_t = (C^\top)^{-1} e_t$). If the resulting t-zero-coupon bond prices, $d_t(0) = \pi(0) \cdot \theta_t$, are strictly positive then there is no arbitrage. Performing the inversion and the matrix multiplications we find that clean price

$$(d_1(0), d_2(0), d_3(0), d_4(0))^\top = (0.952381, 0.898458, 0.839618, 0.7774332),$$

or alternatively the following zero-coupon yields

$$100 * (y(0,1), y(0,2), y(0,3), y(0,4))^\top = (5.00, 5.50, 6.00, 6.50).$$

Now suppose that somebody introduces a 4 yr annuity with a coupon rate of 5 % . Since $\alpha_{4\rceil 5} = 3.5459$ this bond has a unique arbitrage-free price of

$$\pi_5(0) = \frac{100}{3.5459}(0.952381 + 0.898458 + 0.839618 + 0.7774332) = 97.80.$$

Notice that bond prices are always quoted per 100 *units* (e.g. $ or DKK) of principal. This means that if we assume the yield curve is the same at time 1 the price of the serial bond would be quoted as

$$\pi_4(1) = \frac{d_{1:3}(0) \cdot C_{4,2:4}}{0.75} = \frac{76.87536}{0.75} = 102.50$$

(where $d_{1:3}(0)$ means the first 3 entries of $d(0)$ and $C_{4,2:4}$ means the entries 2 to 4 in row 4 of C).

Example 5 (Reading the financial pages) This example gives concrete calculations for a specific Danish Government bond traded at the Copenhagen Stock Exchange(CSX): A bullet bond with a 4 % coupon rate and yearly coupon payments that matures on January 1 2010. Around February 1 2005 you could read the following on the CSX homepage or on the financial pages of decent newspapers

Bond type	Current date	Maturity date	Price	Yield
4% bullet	February 1 2005	January 1 2010	104.02	3.10 %

Let us see how the yield was calculated. First, we need to set up the cash-flow stream that results from buying the bond. The first cash-flow, π in the sense of Definition 8 would take place today. (Actually it wouldn't, even these days trades take a couple of day to be in effect; *valør* in Danish. We don't care here.) And how large is it? By convention, and reasonably so, the buyer has to pay the price (104.02; this is called the *clean price*) plus compensate the seller of the bond for the accrued interest over the period from January 1 to dirty price

February 1, ie. for 1 month, which we take to mean 1/12 of a year. (This is not as trivial as it seems. In practice there are a lot of finer - and extremely boring - points about how days are counted and fractions calculated. Suffice it to say that mostly actual days are used in Denmark.) By definition the buyer has to pay accrued interest of "coupon × year-fraction", ie. 4 × 1/12 = 0.333, so the total payment (called the *dirty price*) is $\pi = 104.35$. So now we can write down the cash-flows and verify the yield calculation:

Date	t_k	Cash-flow (c_k)	$d_k = (1 + 0.0310)^{-t_k}$	PV= $d_k * c_k$
Feb. 1 2005	0	- 104.35	1	
Jan. 1 2006	$\frac{11}{12}$	4	0.9724	3.890
Jan. 1 2007	$1\frac{11}{12}$	4	0.9432	3.772
Jan. 1 2008	$2\frac{11}{12}$	4	0.9148	3.660
Jan. 1 2009	$3\frac{11}{12}$	4	0.8873	3.549
Jan. 1 2010	$4\frac{11}{12}$	104	0.8606	89.505
				SUM = 104.38

(The match, 104.35 vs. 104.38 isn't perfect. But to 3 significant digits 0.0310 *is* the best solution, and anything else can be attributed to out rough approach to exact dates.)

Example 6 (Finding the yield curve) In early February you could find prices 4%-coupon rate bullet bonds with a range of different maturities (all maturities fall on January firsts):

Maturity year	2006	2007	2008	2009	2010
Clean price	101.46	102.69	103.43	103.88	104.02
Maturity year	2011	2012	2013	2014	2015
Clean price	103.80	103.50	103.12	102.45	102.08

These bonds (with names like `4%10DsINKx`) are used for the construction of private home-owners variable/floating rate loans such as "FlexLån". (Hey! How does the interest rate get floating? Well, it does if you (completely) refinance your 30-year loan every year or every 5 years with shorter maturity bonds.) In many practical contexts these are not the right bonds to use; yield curves "should" be inferred from government bonds. (Of course this statement makes no sense within our modelling framework.)
Dirty prices, these play the role of π, are found as in Example 5, and the (10 by 10) C-matrix has the form

$$C_{i,j} = \begin{cases} 4 & \text{if } j < i \\ 104 & \text{if } j = i \\ 0 & \text{if } j > i \end{cases}$$

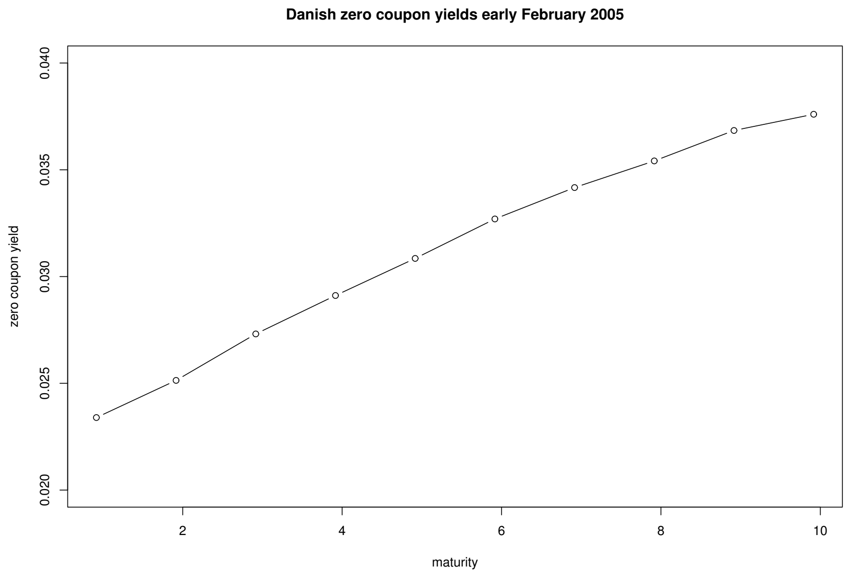

Figure 3.1: The term structure of interest rates in Denmark, February 2005. The o's are the points we have actually calculated, the rest is just linear interpolation.

The system $Cd = \pi$ has the positive ($\sim$ no arbitrage) unique ($\sim$ completeness) solution

$$d = (0.9788, 0.9530, 0.9234, 0.8922, 0.8593, 0.8241, 0.789, 50.7555, 0.7200, 0.6888)^\top.$$

and that corresponds to these (yearly compounded) zero coupon yields:

Maturity	0.92	1.92	2.92	3.92	4.92	5.92	6.92	7.92	8.92	9.92
ZC yield in %	2.37	2.55	2.77	2.95	3.13	3.32	3.48	3.61	3.75	3.83

as depicted in Figure 3.1.

Example 7 The following example is meant to illustrate the perils of relying too much on yields. Especially if they are used incorrectly! The numbers are taken from Jakobsen and Tanggaard.[2] Consider the following small bond

[2]Jakobsen, S. and C. Tanggard: Faldgruber i brugen af effektiv rente og varighed, finans/invest, 2/87.

market:

Bond (i)	100*Coupon rate (R_i)	Price at time 0 ($\pi_i(0)$)	100*Yield
1 yr bullet	10	100.00	10.00
2 yr bullet	10	98.4	10.93
3 yr bullet	10	95.5	11.87
4 yr bullet	10	91.8	12.74
5 yr bullet	10	87.6	13.58
5 yr serial	10	95.4	11.98

Now consider a portfolio manager with the following argument: "Let us sell 1 of each of the bullet bonds and use the money to buy the serial bond. The weighted yield on our liabilities (the bonds sold) is

$$\frac{100 * 10 + 98.4 * 10.93 + 95.5 * 11.87 + 91.8 * 12.74 + 87.6 * 13.58}{100 + 98.4 + 95.5 + 91.8 + 87.6} = 11.76\%,$$

while the yield on our assets (the bond we bought) is 11.98%. So we just sit back and take a yield gain of 0.22%." But let us look for a minute at the cash-flows from this arrangement (Note that one serial bond has payments $(30, 28, 26, 24, 22)$ and that we can buy $473.3/95.4 = 4.9612$ serial bonds for the money we raise.)

	Time 0	1	2	3	4	5
Liabilities						
1 yr bullet	100	-110	0	0	0	0
2 yr bullet	98.4	-10	-110	0	0	0
3 yr bullet	95.5	-10	-10	-110	0	0
4 yr bullet	91.8	-10	-10	-10	-110	0
5 yr bullet	87.6	-10	-10	-10	-10	-110
Assets						
5 yr serial	-473.3	148.84	138.91	128.99	119.07	109.15
Net position						
	0	-1.26	-1.19	-1.01	-0.93	-0.75

So we see that what have in fact found is a sure-fire way of throwing money away. So what went wrong? The yield on the liability side is not 11.76%. The yield of a portfolio is a non-linear function of all payments of the portfolio, and it is not a simple function (such as a weighted average) of the yields of the individual components of the portfolio. The correct calculation gives that the yield on the liabilities is 12.29%. This suggests that we should perform the exact opposite transactions. And we should, since from the table of cash-flow we see that this is an arbitrage-opportunity ("a free lunch"). But how can we be sure to find such arbitrages? By performing an analysis similar to that in Example 4, i.e. pick out a sufficient number of bonds to construct zero-coupon bonds and check if all other bonds are priced correctly. If not it is easy to see how the arbitrage-opportunities are exploited. If we pick out the 5 bullets and do this, we find that the correct price of the serial is 94.7, which is confirmation that arbitrage-opportunities exists in the market. Note that we do not have to worry if it is the serial that is overpriced or the bullets that are underpriced.

capital budgeting
internal rate of return
net present value, NPV

Of course things are not a simple in practice as in this example. Market imperfections (such as bid-ask spreads) and the fact that there are more payments dates the bonds make it a challenging empirical task to estimate the zero-coupon yield curve. Nonetheless the idea of finding the zero-coupon yield curve and using it to find over- and underpriced bonds did work wonders in the Danish bond market in the '80ies (the *19*80'ies, that is).

3.4 IRR, NPV and capital budgeting under certainty.

The definition of *internal rate of return* (IRR) is the same as that of yield, but we use it on arbitrary cash flows, i.e. on securities which may have negative cash flows as well:

Definition 14 *An internal rate of return of a security* $(c_1, \ldots, c_T)$ *with price* $\pi \neq 0$ *is a solution* $y > -1$ *of the equation*

$$\pi = \sum_{i=1}^{T} \frac{c_i}{(1+y)^i}.$$

Hence the definitions of yield and internal rate of return are identical for positive cash flows. It is easy to see that for securities whose future payments are both positive and negative we may have several IRRs. This is one reason that one should be very careful interpreting and using this measure at all when comparing cash flows. We will see below that there are even more serious reasons. When judging whether a certain cash flow is 'attractive' the correct measure to use is net present value:

Definition 15 *The PV and NPV of security* $(c_1, \ldots, c_T)$ *with price* c_0 *given a term structure* $(y(0,1), \ldots, y(0,T))$ *are defined as*

$$PV(c) = \sum_{i=1}^{T} \frac{c_i}{(1+y(0,i))^i}$$

NPV criterion

$$NPV(c) = \sum_{i=1}^{T} \frac{c_i}{(1+y(0,i))^i} - c_0$$

Next, we will see how these concepts are used in deciding how to invest under certainty.

Assume throughout this section that we have a complete security market as defined in the previous section. Hence a unique discount function d is given as well as the associated concepts of interest rates and yields. We let y denote the term structure of interest rates and use the short hand notation y_i for $y(0,i)$.

In capital budgeting we analyze how firms should invest in projects whose payoffs are represented by cash flows. Whereas we assumed in the security market model that a given security could be bought or sold in any quantity desired, we will use the term *project* more restrictively: We will say that the project is scalable by a factor $\lambda \neq 1$ if it is possible to start a project which produces the cash flow λc by paying λc_0 initially. A project is not scalable unless we state this explicitly and we will not consider any negative scaling.

In a complete financial market an investor who needs to decide on only one project faces a very simple decision: Accept the project if and only if it has positive NPV. We will see why this is shortly. Accepting this fact we will see examples of some other criteria which are generally inconsistent with the NPV criterion. We will also note that when a collection of projects are available capital budgeting becomes a problem of maximizing NPV over the range of available projects. The complexity of the problem arises from the constraints that we impose on the projects. The available projects may be non-scalable or scalable up to a certain point, they may be mutually exclusive (i.e. starting one project excludes starting another), we may impose restrictions on the initial outlay that we will allow the investor to make (representing limited access to borrowing in the financial market), we may assume that a project may be repeated once it is finished and so on. In all cases our objective is simple: Maximize NPV.

First, let us note why looking at NPV is a sensible thing to do:

Proposition 5 *Given a cash flow $c = (c_1, \ldots, c_T)$ and given c_0 such that $NPV(c_0; c) < 0$. Then there exists a portfolio θ of securities whose price is c_0 and whose payoff satisfies*

$$C^\top \theta > \begin{pmatrix} c_1 \\ \vdots \\ c_T \end{pmatrix}.$$

Conversely, if $NPV(c_0; c) > 0$, *then every* θ *with* $C^\top \theta = c$ *satisfies* $\pi^\top \theta > c_0$.

Proof. Since the security market is complete, there exists a portfolio θ^c such that $C^\top \theta^c = c$. Now $\pi^\top \theta^c < c_0$ (why?), hence we may form a new portfolio by investing the amount $c_0 - \pi^\top \theta^c$ in some zero coupon bond (e_1, say) and also invest in θ^c. This generates a stream of payments equal to $C^\top \theta^c + \frac{(c_0 - \pi^\top \theta^c)}{d_1} e_1 > c$ and the cost is c_0 by construction.
The second part is left as an exercise. ■

The interpretation of this lemma is the following: One should never accept a project with negative NPV since a strictly larger cash flow can be obtained at the same initial cost by trading in the capital market. On the other hand, a positive NPV project generates a cash flow at a lower cost than the cost of generating the same cash flow in the capital market. It might seem that this generates an arbitrage opportunity since we could buy the project and sell the corresponding future cash flow in the capital market generating a profit at time 0. However, we insist on relating the term *arbitrage* to the capital market only. Projects should be thought of as 'endowments': Firms have an available range of projects. By choosing the right projects the firms maximize the value of these 'endowments'.

Some times when performing NPV-calculations, we assume that 'the term structure is flat' . What this means is that the discount function has the particularly simple form

$$d_t = \frac{1}{(1+r)^t}$$

for some constant r, which we will usually assume to be non-negative, although our model only guarantees that $r > -1$ in an arbitrage-free market. A flat term structure is very rarely observed in practice - a typical real world term structure will be upward sloping: Yields on long maturity zero coupon bonds will be greater than yields on short bonds. Reasons for this will be discussed once we model the term structure and its evolution over time - a task which requires the introduction of uncertainty to be of any interest. When the term structure is flat then evaluating the NPV of a project having a constant cash flow is easily done by summing the geometric series. The present value of n payments starting at date 1, ending at date n each of size c, is

$$\sum_{i=1}^{n} cd^i = cd \sum_{i=0}^{n-1} d^i = cd \frac{1-d^n}{1-d}, \qquad d \neq 1$$

Gordon's growth formula
capital budgeting

Another classical formula concerns the present value of a geometrically growing payment stream $(c, c(1+g), \ldots, c(1+g)^{n-1})$ as

$$\begin{aligned} & \sum_{i=1}^{n} c\frac{(1+g)^{i-1}}{(1+r)^i} \\ = \;& \frac{c}{1+r}\sum_{i=0}^{n-1}\frac{(1+g)^i}{(1+r)^i} \\ = \;& \frac{c}{r-g}\left(1-\left(\frac{1+g}{1+r}\right)^n\right). \end{aligned}$$

Although we have not taken into account the possibility of infinite payment streams, we note for future reference, that for $0 \leq g < r$ we have what is known as *Gordon's growth formula:*

$$\sum_{i=1}^{\infty}\frac{c(1+g)^{i-1}}{(1+r)^i} = \frac{c}{r-g}.$$

3.4.1 Some rules that are inconsistent with the NPV rule.

Corresponding to our definition of internal rate of return in Chapter 3, we define an internal rate of return on a project c with initial cost $c_0 > 0$, denoted $IRR(c_0; c)$, as a solution to the equation

$$c_0 = \sum_{i=1}^{T}\frac{c_i}{(1+x)^i}, \qquad x > -1$$

As we have noted earlier such a solution need not be unique unless $c > 0$ and $c_0 > 0$.

Note that an internal rate of return is defined without referring to the underlying term structure. The internal rate of return describes the level of a flat term structure at which the NPV of the project is 0. The idea behind its use in capital budgeting would t hen be to say that the higher the level of the interest rate, the better the project (and some sort of comparison with the existing term structure would then be appropriate when deciding whether to accept the project at all). But as we will see in the following example, IRR and NPV may disagree on which project is better: Consider the projects shown in the table below (whose last column shows a discount function d):

date	proj 1	proj 2	d
0	-100	-100	1
1	50	50	0.95
2	5	80	0.85
3	90	4	0.75
IRR	0.184	0.197	-
NPV	19.3	18.5	-

Project 2 has a higher IRR than project 1, but 1 has a larger NPV than 2. Using the same argument as in the previous section it is easy to check, that even if a cash flow similar to that of project 2 is desired by an investor, he would be better off investing in project 1 and then reforming the flow of payments using the capital market.

Another problem with trying to use IRR as a decision variable arises when the IRR is not uniquely defined - something which typically happens when the cash flows exhibit sign changes. Which IRR should we then choose?

One might also contemplate using *the payback method* and count the number of years it takes to recover the initial cash outlay - possibly after discounting appropriately the future cash flows. Project 2 in the table has a payback of 2 years whereas project 1 has a payback of three years. The example above therefore also shows that choosing projects with the shortest payback time may be inconsistent with the NPV method.

3.4.2 Several projects

Consider someone with $c_0 > 0$ available at date 0 who wishes to allocate this capital over the $T+1$ dates, and who considers a project c with initial cost c_0. We have seen that precisely when $NPV(c_0; c) > 0$ this person will be able to obtain better cash flows by adopting c and trading in the capital market than by trading in the capital market alone.

When there are several projects available the situation really does not change much: Think of the *i'th* project (p_0^i, p) as an element of a set $P_i \subset \mathbb{R}^{T+1}$. Assume that $0 \in P_i$ all i representing the choice of not starting the i'th project. For a non-scalable project this set will consist of one point in addition to 0.

Given a collection of projects represented by $(P_i)_{i \in I}$. Situations where there is a limited amount of money to invest at the beginning (and borrowing is not permitted), where projects are mutually exclusive etc. may then be described abstractly by the requirement that the collection of selected projects $(p_0^i, p^i)_{i \in I}$ are chosen from a feasible subset P of the Cartesian product $\times_{i \in I} P_i$. The NPV of the chosen collection of projects is then just the sum of the NPVs of the individual projects and this in turn may be written as the NPV of the sum of the projects:

$$\sum_{i\in I} NPV(p_0^i; p^i) = NPV\left(\sum_{i\in I}(p_0^i, p^i)\right).$$

Hence we may think of the chosen collection of projects as producing one project and we can use the result of the previous section to note that clearly an investor should choose a project giving the highest NPV. In practice, the maximization over feasible "artificial" may not be easy at all.

Let us look at an example from Copeland and Weston (1988): .

Example 8 Consider the following 4 projects

project	NPV	initial cost
1	30.000	200.000
2	16.250	125.000
3	19.250	175.000
4	12.000	150.000

Assume that all projects are non-scalable, and assume that we can only invest up to an amount of 300.000. This capital constraint forces us to choose, i.e. projects become mutually exclusive to some extent. Clearly, with no constraints all projects would be adopted since the NPVs are positive in all cases. Note that project 1 generates the largest NPV but it also uses a large portion of the budget: If we adopt 1, there is no room for additional projects. The only way to deal with this problem is to stick to the NPV-rule and go through the set of feasible combinations of projects and compute the NPV. It is not hard to see that combining projects 2 and 3 produces the maximal NPV given the capital constraint. If the projects were assumed scalable, the situation would be different: Then project 1 adopted at a scale of 1.5 would clearly be optimal. This is simply because the amount of NPV generated per dollar invested is larger for project 1 than for the other projects. Exercises will illustrate other examples of NPV-maximization.

The moral of this section is simple: Given a perfect capital market, investors who are offered projects should simply maximize NPV. This is merely an equivalent way of saying that profit maximization with respect to the existing price system (as represented by the term structure) is the appropriate strategy when a perfect capital market exists. The technical difficulties arise from the constraints that we impose on the projects and these constraints easily lead to linear programming problems, integer programming problems or even non-linear optimization problems.

However, real world projects typically do not generate cash flows which are known in advance. Real world projects involve risk and uncertainty and therefore capital budgeting under certainty is really not sophisticated enough for a manager deciding which projects to undertake. A key objective of this course is to try and model uncertainty and to construct models of how risky cash flows are priced. This will give us definitions of NPV which work for uncertain cash flows as well.

3.5 Duration, convexity and immunization.

3.5.1 Duration with a flat term structure.

In this chapter we introduce the notions of duration and convexity which are often used in practical bond risk management and asset/liability management. It is worth stressing that when we introduce dynamic models of the term structure of interest rates in a world with uncertainty, we obtain much more sophisticated methods for measuring and controlling interest rate risk than the ones presented in this section.

Consider an arbitrage-free and complete financial market where the discount function $d = (d_1, \dots d_T)$ satisfies

$$d_i = \frac{1}{(1+r)^i} \text{ for } i = 1, \dots, T.$$

This corresponds to the assumption of a flat term structure. We stress that this assumption is rarely satisfied in practice but we will see how to relax this assumption.

What we are about to investigate are changes in present values as a function of changes in r.. We will speak freely of 'interest changes' occurring even though strictly speaking, we still do not have uncertainty in our model.

With a flat term structure, the present value of a payment stream $c = (c_1, \dots, c_T)$ is given by

$$PV(c;r) = \sum_{t=1}^{T} \frac{c_t}{(1+r)^t}$$

We have now included the dependence on r explicitly in our notation since what we are about to model are essentially derivatives of $PV(c;r)$ with respect to r.

duration, Macaulay
convexity

Definition 16 *Let c be a non-negative payment stream. The Macaulay duration $D(c;r)$ of c is given by*

$$D(c;r) = \left(-\frac{\partial}{\partial r} PV(c;r)\right) \frac{1+r}{PV(c;r)} \tag{3.3}$$
$$= \frac{1}{PV(c;r)} \sum_{t=1}^{T} t \frac{c_t}{(1+r)^t}$$

The Macaulay duration and is the classical one (many more advanced durations have been proposed in the literature). Note that rather than saying it is based on a flat term structure, we could refer to it as being based on the yield of the bond (or portfolio).

If we define

$$w_t = \frac{c_t}{(1+r)^t} \frac{1}{PV(c;r)}, \tag{3.4}$$

then we have $\sum_{t=1}^{T} w_t = 1$, hence

$$D(c;r) = \sum_{t=1}^{T} t\, w_t.$$

Definition 17 *The convexity of c is given by*

$$K(c;r) = \sum_{t=1}^{T} t^2\, w_t. \tag{3.5}$$

where w_t is given by (3.4).

Let us try to interpret D and K by computing the first and second derivatives[3] of $PV(c;r)$ with respect to r.

$$PV'(c;r) = -\sum_{t=1}^{T} t\, c_t \frac{1}{(1+r)^{t+1}}$$
$$= -\frac{1}{1+r} \sum_{t=1}^{T} t\, c_t \frac{1}{(1+r)^t}$$
$$PV''(c;r) = \sum_{t=1}^{T} t\,(t+1) \frac{c_t}{(1+r)^{t+2}}$$
$$= \frac{1}{(1+r)^2} \left[\sum_{t=1}^{T} t^2 c_t \frac{1}{(1+r)^t} + \sum_{t=1}^{T} t c_t \frac{1}{(1+r)^t}\right]$$

[3]From now on we write $PV'(c;r)$ and $PV''(c;r)$ instead of $\frac{\partial}{\partial r} PV(c;r)$ resp. $\frac{\partial^2}{\partial r^2} PV(c;r)$

Now consider the relative change in $PV(c;r)$ when r changes to $r + \Delta r$, i.e. duration, modified

$$\frac{PV(c;r+\Delta r) - PV(c;r)}{PV(c;r)}$$

By considering a second order Taylor expansion of the numerator, we obtain

$$\begin{aligned}\frac{PV(c;r+\Delta r) - PV(c;r)}{PV(c;r)} &\approx \frac{PV'(c;r)\Delta r + \frac{1}{2}PV''(c;r)(\Delta r)^2}{PV(c;r)} \\ &= -D\frac{\Delta r}{(1+r)} + \frac{1}{2}(K+D)\left(\frac{\Delta r}{1+r}\right)^2\end{aligned}$$

Hence D and K can be used to approximate the relative change in $PV(c;r)$ as a function of the relative change in r (or more precisely, relative changes in $1+r$, since $\frac{\Delta(1+r)}{1+r} = \frac{\Delta r}{1+r}$).

Sometimes one finds the expression *modified duration* defined by

$$MD(c;r) = \frac{D}{1+r}$$

and using this in a first order approximation, we get the relative change in $PV(c;r)$ expressed by $-MD(c;r)\Delta r$, which is a function of Δr itself. The interpretation of D as a price elasticity gives us no reasonable explanation of the word 'duration', which certainly leads one to think of quantity measured in units of time. If we use the definition of w_t we have the following simple expression for the duration:

$$D(c;r) = \sum_{t=1}^{T} t\, w_t.$$

Notice that w_t expresses the present value of c_t divided by the total present value, i.e. w_t expresses the weight by which c_t is contributing to the total present value. Since $\sum_{t=1}^{T} w_t = 1$ we see that $D(c;r)$ may be interpreted as a 'mean waiting time'. The payment which occurs at time t is weighted by w_t.

Example 9 For the bullet bond in Example 5 the present value of the payment stream is 104.35 and $y = 0.0310$, so therefore the Macaulay duration is

$$\frac{\sum_{k=1}^{4} t_k c_k (1+y)^{-t_k}}{PV} = \frac{475.43}{104.35} = 4.556$$

while the convexity is

$$\frac{\sum_{k=1}^{4} t_k^2 c_k (1+y)^{-t_k}}{PV} = \frac{2266.35}{104.35} = 21.72,$$

and the following table shows the the exact and approximated relative chances in present value when the yield changes:

Yield	△yield	Exact rel. (%) PV-change	First order approximation	Second order approximation
0.021	-0.010	4.57	4.42	4.54
0.026	-0.005	2.27	2.21	2.24
0.031	0	0	0	0
0.036	0.005	-2.15	-2.21	-2.18
0.041	0.010	-4.27	-4.42	- 4.30

Notice that since PV is a decreasing, convex function of y we know that the first order approximation will underestimate the effect of decreasing y (and overestimate the effect of increasing it).

Notice that for a zero coupon bond with time to maturity t the duration is t. For other kinds of bonds with time to maturity t, the duration is less than t. Furthermore, note that investing in a zero coupon bond with yield to maturity r and holding the bond to expiration guarantees the owner an annual return of r between time 0 and time t. This is not true of a bond with maturity t which pays coupons before t. For such a bond the duration has an interpretation as the length of time for which the bond can ensure an annual return of r :

Let $FV(c; r, H)$ denote the (future) value of the payment stream c at time H if the interest rate is fixed at level r. Then

$$\begin{aligned} FV(c; r, H) &= (1+r)^H PV(c; r) \\ &= \sum_{t=1}^{H-1} c_t (1+r)^{H-t} + c_H + \sum_{t=H+1}^{T} c_t \frac{1}{(1+r)^{t-H}} \end{aligned}$$

Consider a change in r which occurs an instant after time 0. How would such a change affect $FV(c; r, H)$? There are two effects with opposite directions which influence the future value: Assume that r decreases. Then the first sum in the expression for $FV(c; r, H)$ will decrease. This decrease can be seen as caused by *reinvestment risk:* The coupons received up to time H will have to be reinvested at a lower level of interest rates. The last sum will increase when r decreases. This is due to *price risk* : As interest rates fall the value of the remaining payments after H will be higher since they have to be discounted by a smaller factor. Only c_H is unchanged.

The natural question to ask then is for which H these two effects cancel each other. At such a time point we must have $\frac{\partial}{\partial r} FV(c; r, H) = 0$ since an infinitesimal change in r should have no effect on the future value. Now, immunization

$$\begin{aligned}\frac{\partial}{\partial r}FV(c;r,H) &= \frac{\partial}{\partial r}\left[(1+r)^H PV(c;r)\right] \\ &= H(1+r)^{H-1}PV(c;r) + (1+r)^H PV'(c;r)\end{aligned}$$

Setting this expression equal to 0 gives us

$$\begin{aligned} H &= \frac{-PV'(c;r)}{PV(c;r)}(1+r) \\ \text{i.e. } H &= D(c;r)\end{aligned}$$

Furthermore, at $H = D(c;r)$, we have $\frac{\partial^2}{\partial r^2}FV(c;r,H) > 0$. This you can check by computing$\frac{\partial^2}{\partial r^2}\left((1+r)^H PV(c;r)\right)$,reexpressing in terms D and K, and using the fact that $K > D^2$. Hence, at $H = D(c;r)$, $FV(c;r,H)$ will have a minimum in r. We say that $FV(c;r,H)$ is *immunized* towards changes in r, but we have to interpret this expression with caution: The only way a bond really can be immunized towards changes in the interest rate r between time 0 and the investment horizon t is by buying zero coupon bonds with maturity t. Whenever we buy a coupon bond at time 0 with duration t, then to a first order approximation, an interest change immediately after time 0, will leave the future value at time t unchanged. However, as date 1 is reached (say) it will *not* be the case that the duration of the coupon bond has decreased to $t-1$. As time passes, it is generally necessary to adjust bond portfolios to maintain a fixed investment horizon, even if r is unchanged. This is true even in the case of certainty.

Later when we introduce dynamic hedging strategies we will see how a portfolio of bonds can be dynamically managed so as to truly immunize the return.

3.5.2 Relaxing the assumption of a flat term structure.

What we have considered above were parallel changes in a flat term structure. Since we rarely observe this in practice, it is natural to try and generalize the analysis to different shapes of the term structure. Consider a family of structures given by a function r of two variables, t and x. Holding x fixed gives a term structure $r(\cdot, x)$.

For example, given a current term structure $(y_1, \ldots, y_T)$ we could have $r(t,x) = y_t + x$ in which case changes in x correspond to additive changes in the current term structure (the one corresponding to $x = 0$). Or we could have $1 + r(t,x) = (1 + y_t)x$, in which case changes in x would produce multiplicative changes in the current (obtained by letting $x = 1$) term structure.

duration

Now let us compute changes in present values as x changes:

$$\frac{\partial PV}{\partial x} = -\sum_{t=1}^{T} t c_t \frac{1}{(1+r(t,x))^{t+1}} \frac{\partial r(t,x)}{\partial x}$$

which gives us

$$\frac{\partial PV}{\partial x}\frac{1}{PV} = -\sum_{t=1}^{T} \frac{t w_t}{1+r(t,x)} \frac{\partial r(t,x)}{\partial x}$$

where

$$w_t = \frac{c_t}{(1+r(t,x))^t}\frac{1}{PV}$$

We want to try and generalize the 'investment horizon' interpretation of duration, and hence calculate the future value of the payment stream at time H and differentiate with respect to x. Assume that the current term structure is $r(\cdot, x_0)$.

$$FV(c; r(H,x_0), H) = (1+r(H,x_0))^H PV(c; r(t,x_0))$$

Differentiating we get

$$\begin{aligned}\frac{\partial}{\partial x} FV(c; r(H,x), H) &= (1+r(H,x))^H \frac{\partial PV}{\partial x} \\ &\quad + H(1+r(H,x))^{H-1}\frac{\partial r(H,x)}{\partial x} PV(c; r(t,x))\end{aligned}$$

Evaluate this derivative at $x = x_0$ and set it equal to 0:

$$\left.\frac{\partial PV}{\partial x}\right|_{x=x_0} \frac{1}{PV} = -H \left.\frac{\partial r(H,x)}{\partial x}\right|_{x=x_0} (1+r(H,x_0))^{-1}$$

and hence we could define the duration corresponding to the given parametrization as the value D for which

$$\left.\frac{\partial PV}{\partial x}\right|_{x=x_0} \frac{1}{PV} = -D \left.\frac{\partial r(D,x)}{\partial x}\right|_{x=x_0} (1+r(D,x_0))^{-1}.$$

The additive case would correspond to

$$\left.\frac{\partial r(D,x)}{\partial x}\right|_{x=0} = 1,$$

and the multiplicative case to

$$\left.\frac{\partial r(D,x)}{\partial x}\right|_{x=1} = 1 + y_D.$$

Note that the multiplicative case gives us the duration measure (called the Fisher-Weil duration)

duration, Fisher-Weil
barbell strategy

$$D_{mult} = -\frac{\partial PV}{\partial x}\frac{1}{PV} = \sum_{t=1}^{T} t w_t$$

which is just like the original measure although the weights of course reflect the structure $y(0,t)$.

Example 10 (Macaulay vs. Fisher-Weil) Consider again the small bond market from Example 4. We have already found the zero-coupon yields in the market, and find that the Fisher-Weil duration of the 4 yr serial bond is

$$\frac{1}{102.38}\left(\frac{32}{1.0500} + \frac{2*30.25}{1.0550^2} + \frac{3*28.5}{1.0600^3} + \frac{4*26.75}{1.0650^4}\right) = 2.342,$$

and the following table gives the yields, Macaulay durations based on yields and Fisher-Weil durations for all the coupon bonds:

Bond	Yield ( )	M-duration	FW-duration
1 yr bullet	5	1	1
2 yr bullet	5.49	1.952	1.952
3 yr annuity	5.65	1.963	1.958
4 yr serial	5.93	2.354	2.342

So not much difference.
Similarly, the Fisher-Weil duration of the bullet bond from Examples 5, 6 and 9 is 4.552, whereas its Macaulay duration was 4.556.

3.5.3 The Barbell: Messing with your head

We finish this chapter with an example (with something usually referred to as a barbell strategy) which is intended to cause some concern. Some of the claims are for you to check!

A financial institution issues 100 million $ worth of 10 year bullet bonds with time to maturity 10 years and a coupon rate of 7 percent. Assume that the term structure is flat at $r = 7$ percent. The revenue (of 100 million $) is used to purchase 10-and 20 year annuities also with coupon rates of 7%. The numbers of the 10 and 20 year annuities purchased are chosen in such a way that the duration of the issued bullet bond matches that of the portfolio of annuities. Now there are three facts you need to know at this stage. Letting T denote time to maturity, r the level of the term structure and γ the coupon rate, we have that the duration of an annuity is given by

$$D_{ann} = \frac{1+r}{r} - \frac{T}{(1+r)^T - 1}.$$

Note that since payments on an annuity are equal in all periods we need not know the size of the payments to calculate the duration.

The duration of a bullet bond is

$$D_{bullet} = \frac{1+r}{r} - \frac{1+r-T(r-R)}{R\left((1+r)^T - 1\right) + r}$$

which of course simplifies when $r = R$.

The third fact you need to check is that if a portfolio consists of two securities whose values are P_1 and P_2 respectively, then the duration of the portfolio $P_1 + P_2$ is given as

$$D(P_1 + P_2) = \frac{P_1}{P_1 + P_2} D(P_1) + \frac{P_2}{P_1 + P_2} D(P_2).$$

Using these three facts you will note that a portfolio consisting of 23.77 million dollars worth of the 10-year annuity and 76.23 million dollars worth of the 20-year annuity will produce a portfolio whose duration exactly matches that of the issued bullet bond. By construction the present value of the two annuities equals that of the bullet bond. The present value of the whole transaction in other words is 0 at an interest level of 7 percent. However, for all other levels of the interest rate, the present value is strictly positive! In other words, any change away from 7 percent will produce a profit to the financial institution. We will have more to say about this phenomenon in the exercises and we will return to it when discussing the term structure of interest rates in models with uncertainty. As you will see then, the reason that we can construct the example above is that we have set up an economy in which there are arbitrage opportunities.