Real Estate Finance

Erwan Quintin*

Real Estate Department, University of Wisconsin

First draft: October 1, 2009

This version: Aug 20, 2010

*equintin@bus.wisc.edu, https://mywebspace.wisc.edu/quintin/web/.

Contents

Chapter 1

Course information

1.1 Instructor

Erwan Quintin, Assistant Professor, Real Estate Department

Office: 5257 Grainger Hall

E-mail : equintin@bus.wisc.edu

URL: https://mywebspace.wisc.edu/quintin/web/

1.2 Course Objective

My goal in this course is to discuss the theory and practice of real estate asset valuation. In the process, you will become familiar with the structure of the real estate industry, its key institutional features, as well as the mathematical and computational tools on which real estate professionals rely. We will begin by reviewing the foundations of financial economics, and will then discuss their applications to real estate questions.

1.3 Resources

The main sources of material for this course are my notes and the following textbook (hereafter referred to as GM): "Commercial Real Estate Analysis and Investments," D.M. Geltner, N.G. Miller, J. Clayton, P. Eichholtz, Cengage Learning, 2007.

As in all finance courses at this level, we are going to make heavy use of basic algebra and calculus, and I will expect you to become proficient with some advanced features of Excel. We will also learn how to use Argus, one of the most popular valuation tool in the commercial real estate industry.

As the course progresses, I will expect you to learn certain tools on your own. In all cases however, my teaching assistant and I are here to help you as the need arises. If our respective office hours do not suffice for this purpose, we can meet by appointment.

1.4 List of topics and GM readings

1. Preliminaries
   (a) The question (asset pricing) (GM 8, 9)
   (b) Real estate assets (GM 7, 20, 23)
   (c) Example: FRM yield vs. return (GM 16, 17)
   (d) Theory vs. practice
2. Asset pricing fundamentals
   (a) Notions of probability (GM 21)
   (b) Financial economies
   (c) Classical portfolio theory (GM 21)
   (d) CAPM (GM 22)
   (e) Option pricing (GM 27)
   (f) Modigliani-Miller (GM 13) ...
   (g) ... and beyond
   (h) WACC at work
3. Real estate investment analysis

   (a) Cash flow proformas (Argus) (GM 6, 11)
   (b) Discounted Cash Flow (DCF) approach (GM 10)
   (c) Ratio approach
   (d) Debt and Taxes (GM 13, 14, 15)

4. Mortgages
   (a) Legal framework (GM 16)
   (b) Basic mortgage algebra (GM 17)
   (c) Default and Prepayment risk
   (d) Refinancing
   (e) The underwriting process (GM 18)
   (f) Securitization (GM 20)
   (g) The foreclosure crisis

5. Advanced topics
   (a) Real options approach
   (b) Monte-Carlo simulations
   (c) Mortgage design puzzles

Chapter 2

Preliminaries

This chapter provides an overview of the basic question that concerns us in this course, and of the basic concepts we are going to employ to answer that question.

2.1 The basic question: asset pricing

Our goal in this course is to learn how to price real estate *assets.* In a broad sense, an asset is a store of value which may be owned by a person or a corporation. Practically, an asset is a string of payoffs, or more simply in most contexts, a string of cash flows. These flows are realized over time and typically involve some uncertainty.

This course can be decomposed into two basic tasks: describing the flows associated with an asset and their statistical distribution, and pricing them. The first task may seem trivial but turns out to be quite involved for a number of reasons. First, the nature of many modern assets is quite complicated. For instance, real estate markets abound with derivative instruments whose payoffs depend on some underlying assets. Second, doing a good job assigning a statistical distribution to future payoffs (in other words, *forecasting*) requires the right mix of solid mathematical tools, experience, and intuition. Forecasting, in fact, is an art that deserves a dedicated course. We will only discuss it in broad terms.

Our second task – pricing assets – involves the beautiful machinery that is modern finance. The theory and practice of asset pricing is founded on a compelling idea, known as *arbitrage*: similar assets should earn similar returns. Should similar or identical assets trade at two different prices, one could buy at the low price, then sell at the high price, enjoying a risk-free, potentially unbounded profit opportunity. In reality of course, transaction costs may partly hinder this correction mechanism, but it remains true that big price discrepancies are unlikely to survive for long, particularly in deep, well developed markets.

You may have noted that the previous paragraph used asset price and asset return almost interchangeably. This is not a coincidence: prices and returns are monotonically related. Consider for instance an asset – a zero-coupon, treasury bond, say – that pays \$1000 in one year. Assume the asset trades today for price $p = \$900$. The yearly return, r, on this asset, is

$$r = \frac{1000}{p} - 1 = \frac{1000}{900} \approx 11\%.$$

Since $p = 900(1+r)$, price and return convey the exact same information about the asset's value.

The last expression also suggests a very convenient way to think about returns. A payoff of \$1000 a year from now is the value in one year of \$900 invested at a rate of return of 11%. In finance, we say that \$900 is the *present value* of a payoff of \$1000 in one year at a *discount rate* of 11%.

A related way to think about r is that it is the discount rate that makes the present value of \$1000 in one year equal to the asset price today. That notion is called the *internal rate of return*, or IRR for short.

What if the payoff came in two years rather than one? What would be the yearly return in that case, assuming the same price? We are now looking for a rate of return r such that, \$900 invested for two consecutive years yields \$1000. After one year, \$900 invested at rate r yields $\$900(1+r)$, which we may reinvest for one year at, once again, return r, to produce:

$$(900(1+r))(1+r) = 900(1+r)^2$$

after two years. It follows that r must solve:

$$\begin{aligned} 900(1+r)^2 &= 1000 \\ \Longleftrightarrow 900 &= \frac{1000}{(1+r)^2} \\ \text{so that } r &\approx 5.4\% \end{aligned}$$

Once again, this says that the present value of \$1000 two years from now at yearly discount rate 5.4% is exactly the asset price today, namely \$900. Put another way, 5.4% is the IRR of an asset whose cost today is \$900 and yields \$1000 two years from now.

Note, of course, that delaying the date at which a payoff is received reduces the IRR, holding the current price constant. This is an illustration of the *time value of money*. Payoffs received earlier are more valuable than payoffs received late because they can be reinvested sooner. A dollar today is more valuable than a dollar a year from now because I can invest my dollar today for one year and start earning interest. Put yet another way, the present value of a dollar in one year is less than a dollar as long as positive return and essentially risk free investment opportunities exist.

Likewise, we can compute the return associated with strings of more than one cash-flows. Assume our asset continues to cost \$900, but now yields \$500 a year from now, and \$500 two years from now. The internal rate of return r must now solve:

$$\begin{aligned} 900 &= \frac{500}{(1+r)} + \frac{500}{(1+r)^2} \\ \text{so that } r &\approx 7\% \end{aligned} \tag{2.1}$$

Note that solving for r becomes a bit more complicated now. In your first homework, I ask you to solve this equation for r using algebra, then using Excel. In general,[1] as is typically the case in finance, one can only rely on computers or calculators to compute returns, which is why you need to become proficient with these tools.

Naturally, the size of future cash flows are not typically with certainty. In the previous example, assume for instance that the second year cash-flow is \$500 with probability q while it

[1]Niels Henrik Abel proved in 1824 that equations like this can be solved algebraically only provided the highest power involved is 4.

is zero otherwise. Then, we say that the expected cash-flow in year two is $500 \times q + 0 \times (1-q) = 500q$. An IRR may then be calculated as above using this expected value for year 2 and we may also calculate the present value of expected cash-flows at the appropriate discount rate.

With this language at hand, we are ready to start pricing assets. Say, for instance that your are a *securitizer* who is considering buying a thirty-year fixed-rate mortgage just originated by a bank. How much should you pay?

We are looking at up to 360 monthly cash flows over a thirty-year period. These cash flows, however, may be highly uncertain because, among other possibilities, the borrower could default, fail to make payments in time, or decide to pay off their mortgage early. While this sounds like a complicated problem, the principle to follow is trivial. This mortgage should yield the same internal return as similar mortgages out there. One would simply look at the price at which similar contracts trade, and compute the price at which the contract under review must trade to earn the same internal rate of return.

In practice, finding similar contracts may be difficult, but this is a difficulty that solid market research can overcome fairly convincingly. Once we have an appropriate rate of return, we can calculate the value of our expected 360 cash flows at that return to arrive at an estimated market value. This is known as the *discounted cash flow* approach.

A related approach to asset pricing involves focusing on key statistical properties of assets under study, constructing a *peer group* of similar assets and assuming that all assets in that group should trade at the same multiple of that particular statistical aspect. Assume for instance that we are trying to value a shopping mall in a given location. Two key statistics associated with this asset are the current levels of gross income (revenues) and net income (profits). The multiple or ratio valuation procedure consists of finding a set of similar properties, preferably in the same general location, which were recently sold, and computing the average ratio of the price to gross and/or net income. Assume for instance that similar properties market, on average, at 10 times yearly gross revenues. An estimate of the market value of the property under study can be found by applying that average multiple to its level of gross revenues.

In the real estate context, the two main multiple methods are called the capitalization rate approach and the gross rent multiplier (GRM) approach. The capitalization rate (or "cap rate" for short) for a given property is the ratio of its net operating income (NOI = gross rents minus operating expenses) to the property's value, while the GRM is the ratio of value to gross rents. To price a given property whose value is unknown, one simply finds a set of comparable properties (in terms of location, purpose, vintage . . .), computes the average cap rate and GRM for this peer group, and apply these ratios to the target property's NOI and gross rents. Simple as this sounds, this approach dominates practice in real estate markets, and most of the valuation conversation revolves around whether the properties that comprise the peer group are comparable to the target property.

2.2 Real estate assets

Real estate assets, like all assets, are entitlements to strings of cash flows to be realized in time over a certain horizon, and subject to some uncertainty. Real estate physical properties include land, and all structures affixed to it. These physical assets include undeveloped pieces of land, residences, office buildings, shopping malls, apartment complexes . . . . Commercial properties are assets held for a business purpose (for the purpose of generating revenues), while residential real properties are properties meant to be occupied by residents. Basically, any property other than single family homes and lots is a commercial property.

These physical properties are the bedrock of the real estate industry, but the industry also includes all assets whose payoffs depend – however remotely – from the cash flows generated by one or a set of physical properties. As a matter of basic accounting, all cash flows associated with a particular property are attached to a claim, or *liability*.

Some properties are owned directly by an investor. More generally, property cash flows are attached to a variety of *debt* and *equity* instruments. A debt contract defines a specifically defined claim (e.g. a fixed, uncontingent payment) to the asset's payoff, but carries no ownership right. An equity instrument (e.g. a public stock) gives its issuer ownership rights, but only a residual, or subordinated claim to the asset's payoffs. In the real estate world, the emblematic debt contract is the mortgage, while equity investments may be made by the asset's sole proprietor or a real estate investment trust (REIT).

These liabilities are themselves assets that may be owned directly by another single investor or financial intermediary (a bank), or owned via intermediate investment vehicles such as mortgage-backed securities (MBS). The resulting dicing and packaging of underlying property cash flows creates a myriad of real estate assets.

Another important division of real estate assets involved the market in which they are traded. Public assets are traded in markets with many buyers and sellers, observable transaction prices and volumes, and stringent disclosure rules. Private assets are traded in markets where transactions involve a limited number of buyers and sellers, and where transaction and financial information need not be disclosed (except, of course, to tax authorities.)

All these assets form the class of real estate assets, a set of assets whose fundamental value is governed by the value of real estate properties. Together with highly liquid holdings (cash and commercial paper), stocks, and bonds, they should be part of a well-balanced portfolio. Historically, real estate as a broad class has displayed roughly the same riskiness as investment-grade corporate bonds, and similar returns. They also provide a reasonable hedge against inflation risk since property values should in general rise and fall with the overall price level.

The class of real estate assets tends to grow in size with time as a result of financial innovation and deregulation. For instance, the 1960 real estate investment trust tax provision exempts REITs from federal taxation, sparing them the double taxation (corporate tax + individual income tax) imposed on US corporations. Further tax reforms implemented in the 1980s and 1990s further boosted the popularity of these investment vehicles.

On the debt side, the pooling of mortgages into single securities known as mortgage-backed securities was facilitated by legislation passed in the late 1980s to alleviate the Savings and Loans crisis. Bundling mortgages and selling the resulting security them in secondary markets – a process known as *securitization* – makes a lot of sense since it provides some diversification of contract-specific risk, and renders mortgages much more liquid (i.e. enhances the ease with which they can be converted into cash.)

As you probably know however, securitization has been blamed by some for much of the financial mess we find ourselves in. Reasons invoked for pointing the finger at securitizers vary. By enabling originators to pass them along quickly, securitization can reduce the incentives of lenders to do their homework. For instance, as long as no-documentation loans are accepted by securitizers, originators are willing to issue them since these contracts can immediately be sold, before – as they say – the peanut butter hits the fan. A related issue is that bundling may enable financial engineers to sneak bad assets into a large bundle of mostly decent assets. More generally, by setting guidelines on what they'll buy, securitizers limit the ability of originators to tailor contracts to their customers' needs.

What makes the role of securitizers even more controversial is that the biggest players in that game are so-called government-sponsored agencies (GSEs). What these agencies are willing to tolerate translates almost immediately into what contracts are originated. If, say, gentle political pressure is applied on GSEs to start buying contracts issued to high-risk borrowers (presumably in the benevolent hope to boost home-ownership), high-risk lending becomes the norm. Between 2005 and 2006, it is estimated that subprime loans (loans issued to high-risk borrowers) accounted for a quarter of all originations. When market conditions turned south, the housing-finance industry collapsed, taking down large financial institutions and, in turn according to many, the economy at large.

An important development in the real estate sector has been the growing popularity of *derivative* instruments. Derivatives allow investors to purchase some exposure to the returns of a particular asset without necessarily owning that asset. For instance, investors can now "swap" the returns of a particular real estate index for some fixed payment. Regardless of the complexity of the resulting instruments, valuing them continues to involve exactly two steps: describing the distribution of the associated payoff and, second, valuing that distribution via an arbitrage argument.

Another example of derivative that is receiving a lot of press during this default crisis are so-called credit-default-swaps (CDS). In this contract, one party swaps fixed premium payments for some insurance against the losses associated with credit events (default, say). CDS' are offered, for instance, on mortgage-backed securities. These contracts are said to have played a big role in the recent financial crisis. When housing markets crashed and default rates started rising a lot in mid-2006, protection sellers such as AIG had to swallow massive losses.

Derivatives markets have become very deep in the United States and many transactions involve counterparties who have no ties to the underlying assets. They enable agents to buy and sell exposure to real estate payoffs without actually dealing with real estate assets.

Another area of the real-estate industry where innovation has been fast and furious is the mortgage sector. A wave of deregulation culminating in 1982 with the Garn-Saint-Germain act empowered institutions to issue mortgages with variable payments. It also lifted legal restrictions on interest rates and fees an intermediary can charge, enhancing their potential ability to deal with high-default-risk customers. In addition, restrictions on loan-to-value ratios (the ratio of the loan balance to the asset's market value at origination), loan maturities, and the use of balloon payments (see next section), were all greatly relaxed in the process. Since then the gamut of mortgage contracts has expanded markedly. At the same time, the grand-daddy of them all, the conventional fixed rate mortgage (FRM) remains dominant.

As you may know the dominance of standard FRM contracts was challenged for a short period after the turn of the century. Between 2001 and 2006 especially, underwriting and lending standards were noticeably loosened and products that delay loan amortization (via, say, low down-payments and low principal payments early in the life of the loan) became very popular. These developments were favored by the fast development of securitization (mortgage-backed securities), with GSEs playing a big role. That phase, however, was followed by the so-called foreclosure crisis, which we will discuss at length in this class. The popularity of non-traditional mortgages, as we will discuss, has now greatly receded.

2.3 Example: mortgage yield vs. return

The emblematic debt contract in the real estate industry is the mortgage. We will devote a lot of time in this class to understanding the basic mechanics of this contract. This section introduces the basic features of mortgages, and briefly discusses how to compute their market value.

A mortgage is a debt contract used to finance the purchase of a real estate asset, and *secured* by this asset. In the event of default on the contract's stipulated obligations, the asset may be sold to cover the lender's losses.

A mortgage is characterized by a loan size at origination (the initial balance), a maturity (the maximum period of time over which the lender will receive payments from the borrower), the frequency of payments, an interest rate structure, and a payment structure. The contract rate in a given period is the rate of interest charged on the outstanding balance. The payment structure stipulates the payment to be made by the borrower to the lender in a given period of the contract.

Consider a mortgage of maturity T and with initial loan size b_0. In the case of thirty-year mortgage with monthly payments, $T = 30 \times 12 = 360$, and a typical balance at origination could be, say, $b_0 = \$200,000$. Denote the interest rate in period t by r_t. In the case of a standard fixed-rate mortgage, the rate does not vary with time, so that $r_t = r$ at all t. More generally, we could have rates that change on a fixed schedule (e.g., two years of low "teaser" rates, and then 28 years of higher rate), or depend on some market interest rate such as the rate banks charge each other on loans. In this introductory section, we will focus on fixed-rate contracts, which is by far the most popular choice in the United States.

All we need now is a payment structure. Here, the possibilities are endless. We could have payments rise or fall at fixed rate, or follow any pattern we wish. We will once again focus on the most common case where the payment m_t at date t is the same (m) for all t. An important decision to make is whether we want the loan to be fully paid-off (*a fully amortizing loan*), or leave a positive balance at the end of T periods to be paid in one *balloon* payment. Some loans, for instance, feature zero balance reduction, or *amortization*, until the very last period.

Here, we will focus on fully amortizing loans, i.e. loans whose balance is zero once payment number T is made. To make this formal, let b_t be the loan's balance after payment number t had been made. A fully amortizing loan of maturity T must be such that:

$$b_T = 0 \tag{2.2}$$

Now denote the fixed payment associated with the mortgage by m, while the fixed rate is r. In period t, interest payments are $b_{t-1}r$. Indeed, the balance before payment t is made (at the start of period t) is b_{t-1}. If $m > b_{t-1}r$, payments $m > b_{t-1}r$ go towards "principal", i.e.

reduce the balance, so that the remaining balance after payment t is:

$$\begin{aligned} b_t &= b_{t-1} - (m - b_{t-1}r) \\ \Longleftrightarrow b_t &= b_{t-1}(1+r) - m \end{aligned} \tag{2.3}$$

It is important to note that in general, mortgage payments can be such that $m < b_{t-1}r$, so that rather than decrease, debt actually increases from one period to the next, a possibility known as *negative amortization.* One popular contract where this often happens is the so-called option-ARM mortgage. Basically, that contract allows the borrower to choose the payment in a given period, within some limits. If the borrower chooses to make a payment lower than the interest they owe, the principal rises. Contracts that stipulate $m_t = b_{t-1}r$ over part of the contract or until maturity are called – for intuitively obvious reasons – interest-only mortgages (IOMs.) We will see in a minute that fully-amortizing fixed-rate loans have to be such that $m > b_{t-1}r$.

Given equation (2.3) that governs the evolution of the loan balance, the mortgage design question is: Given b_0 and r, what must m be so that $b_T = 0$? In other words, given the interest rate, what fixed payment is such that the initial balance is paid off after exactly T periods?

All we have to do to answer that question is solve things forward, using (2.3). We start with initial principal b_0. Then, we move after one payment to $b_1 = b_0(1+r) - m$. Then,

$$b_2 = b_1(1+r) - m = (b_0(1+r) - m)(1+r) - m = b_0(1+r)^2 - m[1 + (1+r)]$$

Continuing, we find that after t payments,

$$\begin{aligned} b_t &= b_0(1+r)^t - m[1 + (1+r) + (1+r)^2 + \ldots (1+r)^{t-1}] \\ \Longleftrightarrow b_t &= b_0(1+r)^t - m\sum_{i=0}^{t-1}(1+r)^i \end{aligned}$$

It follows that full amortization (i.e., equation 2.2) holds if and only if:

$$\begin{aligned}
& b_T = 0 \\
\Longleftrightarrow \quad & b_0(1+r)^T - m\sum_{i=0}^{T-1}(1+r)^i = 0 \\
\Longleftrightarrow \quad & b_0 - m\sum_{i=0}^{T-1}(1+r)^{i-T} = 0 \\
\Longleftrightarrow \quad & b_0 - m\sum_{i=0}^{T-1}\left(\frac{1}{1+r}\right)^{T-i} = 0 \\
\Longleftrightarrow \quad & b_0 - m\sum_{j=1}^{T}\left(\frac{1}{1+r}\right)^{j} = 0 \\
\Longleftrightarrow \quad & b_0 - m\frac{1-\left(\frac{1}{1+r}\right)^T}{r} = 0 \\
\Longleftrightarrow \quad & \boxed{m = b_0\frac{r}{1-\left(\frac{1}{1+r}\right)^T}} \qquad (2.4)
\end{aligned}$$

The algebra looks daunting, but it is easy. Just work it out. Going from the second line to the third involves dividing the whole expression by $(1+r)^T$. The fourth line is obtained by defining $j = T - i$, i.e. making a simple change of counting index. The next line follows from standard results on geometric sums (we will go into that in class), and the final step is trivial. Those of you who have taken finance before will recognize that this is the constant annuity formulae, one of the workhorse of the field which pops up everywhere.

In your first homework assignment, I ask you to design quirky mortgages by way of practicing these ideas. Here, however, we want to put this algebra to more fundamental work. Notice what the fourth step of the derivation above says:

$$\begin{aligned}
b_0 &= m\sum_{j=1}^{T}\left(\frac{1}{1+r}\right)^j \\
\Longleftrightarrow b_0 &= \frac{m}{1+r} + \frac{m}{(1+r)^2} + \ldots \frac{m}{(1+r)^T}
\end{aligned}$$

In words, the present value of all payments discounted at rate r is the initial balance. In fact, this is another way to define full amortization. Assuming no payments other than monthly and down payments (that is, assuming no "points" as we will discuss in the mortgage chapter), a mortgage is fully amortizing if and only if the present value of payments at the contract rate is the loan's initial balance.

This property/definition will come handy in many places in this class. For instance, assume that we want to design a mortgage where payments rise from month to month by $g\%$, where $g \geq 0$, and that we still want the loan to be fully amortizing at a contract rate of r. Such a contract is called a graduated payment mortgage, or GPM. Then, letting m_1 be the first payment, the payment in month t is $m_t = m_1(1+g)^{t-1}$ and m_1 must be such that:

$$b_0 = m_1 \sum_{t=1}^{T} \frac{(1+g)^{t-1}}{(1+r)^t} = \frac{m_1}{1+g} \sum_{t=1}^{T} \left(\frac{1+g}{1+r}\right)^t,$$

an expression which is easy (!) to solve using standard geometric sum manipulations.

Alternatively, there is an Excel-friendly way to solve something like this. Let $PV(1)$ be the value of the right-hand side of the equation above when m_1 is assumed to be 1. Then the correct value of m_1 is found simply by dividing b_0 by $PV(1)$.

Returning now to the FRMcase and looking at a mortgage from the point of view of the lender, the lender is paying b_0 at date 0 to buy a string of T cash flows m. From the point of view of the lender then, r, the contract rate, is the internal rate of return **if and only if** all payments are made as planned. This IRR conditional on all payments being made as planned is called the yield or yield-to-maturity or, after annualization, annual percentage rate (APR).

Now, all may go well, but in general, there is a non-negligible probability that the borrower will fail to make a payment, or at least fail to make it on time which, if nothing else, will cost the borrower valuable time and resources, as well as, more often than not, capital losses. It follows that the mortgage yield is an upper bound on its return. Assume for instance that there is only a probability one half that the final 10 payments will be made. Then the present value of the mortgage to the lender falls all else equal, so that its IRR must fall too. In your homework, I ask you to play with that idea.

Obviously, the lender will seek to charge a higher rate to people more prone to default to cover their expected losses and greater risk. This is exactly what we observe: subprime loans often carry exorbitant rates. But this is where things become subtle. Raising the rate causes payments to be higher and, therefore, makes default more likely so that raising the rate further may become necessary and so on ... This search process for a rate that makes the lender happy (what in math is called a *fixed point problem*) may have a solution but, then again, it may not. Some borrowers are simply too risky for leaders to be able to issue a profitable loan to them. (For more on this, see Corbae and Quintin, 2009.)

2.4 Theory vs. practice

In practice, most real estate professionals rely on a number of approaches, some inspired by finance theory, some not as much. We will try as much as possible to talk about what professionals do, whether or not we can motivate it from fundamentals. While we will emphasize formal tools in this class, experience and intuition are at least as determinant for success in this industry. Looking at this as Darwin would, real estate professionals who survive for a long time in the industry are doing something right, and one stands to learn a lot by observing how they operate.

Chapter 3

Foundations of Asset Pricing

This chapter develops the language one needs to ask asset-pricing questions, and describes models in which fundamental asset values can be calculated. As all models, models of financial economies make a number of assumptions that do not hold in practice. Financial economies are stylized worlds where exact answers to complicated questions can be given. These stylized worlds enable us, first, to emphasize and understand the fundamental determinants of asset values. They also provide us with approximate rules that can serve as useful benchmarks for the financial decisions practitioners need to make.

3.1 Notions of probability

As discussed in chapter 2, asset returns are subject to some uncertainty. They are, in other words, *random*, or *stochastic*. Entire books are devoted in many fields (including philosophy) to defining what words like "random" mean in a deep sense. For our purposes, we simply want to think about the situation where an asset's payoff depends in part on the outcome of a *random experiment*, an experiment whose outcome cannot be fully determined *a priori*.

The canonical random experiment is the flip of a coin. We can list all the possible outcomes of this experiment and assign probability to them quite easily. What we need is a framework that enables us to do this for all random experiments one can think of. For this we need some basic notions of probability theory. For those of you who want the gruesome details, a good place to start is chapter 5 of the macro theory notes available from my web page.

Outcomes of random experiments are draws from a set S (sometimes called the *universe*). For instance, if the experiment we have in mind is the roll of an ordinary dice, $S = \{1, 2, 3, 4, 5, 6\}$. Elements of S are called states of the world, or states, for short.

A set of states that is relevant for asset pricing is the set of possible states of the economy. For instance, we may write $S = \{\text{great depression, recession, stagnation, expansion, boom}\}$, where one presumably has a specific time period in mind. An *event* is a subset of the universe S. For instance, {expansion, boom} corresponds to an event we could describe as "the economy is growing." In our dice-rolling example, $A = \{2, 4, 6\}$ is the event that "the roll is even."

A *probability measure* or *probability distribution* on S is a function that assigns to every element of S a number between 0 and 1, and such that the sum of the probabilities of all elements is exactly one.

In our dice-casting example, $S = \{1, 2, 3, 4, 5, 6\}$, and, if the dice is fair, it is natural to set $p(s) = \frac{1}{6}$ for all elements s of S. With that *uniform* probability structure, the probability $p(A)$ of any event A is given by

$$p(A) = \frac{\#A}{\#S},$$

where $\#A$ is the number of elements event A contains. For instance, $P(\{2, 4, 6\}) = \frac{3}{6} = 0.5$.

The next notion we need is that of a *random variable*. A random variable X defined on S assigns to each draws s from S a value, or payoff. Here's an example. Consider a bet that pays one dollar if the roll of a dice turns out to be even and nothing otherwise. Letting X be the payoff associated with the bet, $X \in \{0, 1\}$, and we have that $P(X = 1) = P(s \in \{2, 4, 6\} = 0.5$ if the dice is fair. In this course, X will be some asset's payoff, and that payoff will depend on some random event, for instance on the likelihood that the economy at large will expand, or collapse.

Now we are ready to introduce some language with which you need to become intimately familiar. That language will give us formal ways to talk about an asset's average payoff, the volatility or risk associated with that payoff, and the extent to which some asset's payoff co-varies with the payoff of other assets. Once you understand these notions, you're ready to tackle quantitative work.

For the rest of this section, we assume that all random variables are defined on a given set of states S equipped with a probability distribution P. The expectation of random variable X is the weighted average of all possible outcomes, where the weights are the likelihood of states:

$$E(X) = \sum_{s \in S} X(s)P(s).$$

For example, if X is a variable that takes value one if the outcome of a fair dice roll is even, zero otherwise, then:

$$\begin{aligned} E(X) &= P(s=1) \times 0 + P(s=2) \times 1 + P(s=3) \times 0 \\ &+ P(s=4) \times 1 + P(s=5) \times 0 + P(s=6) \times 1 = 0.5. \end{aligned}$$

If the payoff is the same regardless of the state ($X(s) = x$ for all s), then we call X *risk-free* or *deterministic*. The closest thing we have in the United States to a risk-free asset is a Treasury bill (a short-term government bond). If, on the other hand, X is not risk-free, we call it risky. It is useful to have a measure of how risky a particular asset is. A natural way to think of risk is to look at deviations $X(s) - E(X)$ from the average in all possible states. For the purpose of measuring the unpredictability of a random variable however, whether the deviation is positive or negative is immaterial. A good way to measure risk, then, is to square deviations and weigh them by their respective likelihoods. The result is a measure called the variance of a random variable:

$$VAR(X) = E\left((X - E(X))^2\right) = \sum_{s \in S} (X(s) - E(X))^2 P(s).$$

For example, going back to the variable that takes value one if the outcome of a fair dice roll is even, then:

$$\begin{aligned} VAR(X) &= P(s=1) \times (0-0.5)^2 + P(s=2) \times (1-0.5)^2 + P(s=3) \times (0-0.5)^2 \\ &+ P(s=4) \times (1-0.5)^2 + P(s=5) \times (0-0.5)^2 + P(s=6) \times (1-0.5)^2 = 0.25. \end{aligned}$$

Note, importantly, that X is risk-free if and only if $VAR(X) = 0$.

A related measure is the square root of the variance, which is called the standard deviation. Taking the square root essentially offsets the fact that we squared deviations, so that risk and expectations are measured in similar units. We will denote this measure by σ_X, and it is defined by:

$$\sigma_X = \sqrt{VAR(X)}.$$

We also need a notion of how two random variables X and Y are related. Two random variables that tend to vary together will tend to take high values at the same time, and low-values at the same time. For instance, in general, we would expect GDP changes and the evolution of the stock market to move in the same direction. Likewise, two variables could be negatively related. We would expect average house prices and the unemployment rate to move in opposite directions. A key statistic that measures the relation between two variables X and Y is the covariance:

$$COV(X,Y) = \sum_{s \in S} (X(s) - E(X))(Y(s) - E(Y))P(s).$$

When the covariance is positive, high values of X and high values of Y tend to coincide.

Let's assume once again that X is a variable that takes value one if the outcome of a fair dice roll is even, and let's make Y the variable that takes value 1 if the roll is 4 or above, zero otherwise. Then $E(Y) = E(X) = 0.5$, and:

$$\begin{aligned}
COV(X,Y) &= P(s=1) \times (0-0.5)(0-0.5) P(s=2) \times (1-0.5)(0-0.5) \\
&+ P(s=3) \times (0-0.5)(0-0.5) + P(s=4) \times (1-0.5)(1-0.5) \\
&+ P(s=5) \times (0-0.5)(1-0.5) + P(s=6) \times (1-0.5)(1-0.5) \\
&= \frac{1}{12} \\
&\approx 0.083
\end{aligned}$$

The outcome of these two variables are, in other words, positively *correlated*

Because it is difficult to attach meaning to the size of covariances, one often "normalizes" them to generate a measure called the *coefficient of correlation* that varies from -1 to 1. It is defined by:

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}. \tag{3.1}$$

A coefficient of correlation of 1 means that X and Y are identical up to a change of unit and a level shift. In other words, for all s, $X(s) = aY(s) + b$, where $a > 0$ and b is some number. Conversely, a coefficient of correlation of -1 means that X and Y are essentially the opposite from one another. For all s that is, $X(s) = aY(s) + b$, where $a < 0$ and b is some number. A coefficient of zero means that X and Y are not linearly related at all. In our dice-casting example,

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{1}{12}}{\sqrt{0.25 \times 0.25}} = \frac{1}{3}.$$

Now we turn to the statistical facts that play a key role in finance. As we will see shortly, investment portfolios or strategies basically combine investments in a variety of assets. Assume then that X is the return on a dollar invested in a given asset. If we invest one dollar in asset X and one dollar in asset Y, the payoff in state s becomes $X(s) + Y(s)$. We have thus constructed a new random variable, which we can denote $X + Y$. Pushing the same idea further, assume that we invest fraction a of our wealth in asset X and fraction b in asset Y. In state s, that strategy returns $aX(s) + bY(s)$, and we have created yet another random variable which we may denote for short by $aX + bY$.

Here are the facts – that you should carefully check – that we will invoke over and over in this class. First, expectations are nice and linear:

$$E(aX + bY) = aE(X) + bE(Y)$$

Basically and conveniently, a portfolio return is the average of the returns. Next:

$$VAR(aX) = a^2 VAR(X) \Longleftrightarrow \sigma_{aX} = a\sigma_X,$$

so that standard deviations are nice and linear as well. Now, the fundamental idea behind risk diversification (we will work it out in class):

$$VAR(aX + bY) = a^2 VAR(X) + b^2 VAR(Y) + 2abCOV(X, Y)$$

If, for instance, $a = b = 0.5$ (split your wealth half in X and half in Y),

$$VAR(0.5X + 0.5Y) = 0.25VAR(X) + 0.25VAR(Y) + 0.5COV(X, Y).$$

Let's first take the case where the two variables are unrelated: $COV(X, Y) = 0$. Even in that case, the variance of the average is less that the average of the variance, because of the squaring of coefficients on the right hand side. Why is that? This is a baby version of the law of large numbers. If you flip a coin a 1000 times, you know that the fraction of heads will be pretty close to 50% if the coin is fair. If you flip it once only, the fraction will be either 0% or 100%, which implies a pretty high variance. In the portfolio context, if I put all my wealth in one asset, I am exposing myself to a lot of risk. If I spread my wealth around, the law of averages works in my favor.

If it so happens that $COV(X, Y) < 0$, I am reducing risk further. If I am at a casino and decide to put a chip on both black and red at the roulette, I have pretty much eliminated all risk. If I invest in two assets that covary negatively, low returns on one will tend to be mitigated by good returns on the other. If on the other hand $COV(X, Y) > 0$, things work in the opposite direction. I am compounding risk by investing in correlated assets. People who have all their savings in accounts attached to their employer are causing their labor income and their financial income to become correlated. That exposes them to a lot of risk, as evidenced by the hardship experienced by Enron employees, to pick one company.

Let's take two extreme examples to illustrate these ideas. Assume that $VAR(X) = VAR(Y)$ and that $\rho_{X,Y} = 1$. Then,

$$\begin{aligned}
VAR(0.5X + 0.5Y) &= 0.25VAR(X) + 0.25VAR(Y) + 0.5COV(X, Y) \\
&= 0.25VAR(X) + 0.25VAR(Y) + 0.5\sigma_X \sigma_Y \rho_{X,Y} \\
&= 0.5VAR(X) + 0.5VAR(Y)
\end{aligned}$$

so that there is no reduction in risk whatsoever. The derivation above uses equation (3.1), the fact that $\sigma_X = \sigma_Y$, and the fact that $\rho_{X,Y} = 1$. Conversely, if $VAR(X) = VAR(Y)$ and $\rho_{X,Y} = -1$, the same algebra shows that $VAR(0.5X + 0.5Y) = 0$ so that all risk is eliminated.

A key observation for what we are about to do is that if X is risk-free, then $COV(X,Y) = 0$ for any other random variable. Since X does not vary at all, it certainly does not co-vary with anything. It follows that if X is risk-free, $VAR(0.5X + 0.5Y) = 0.25VAR(Y)$ so that $\sigma_{0.5X+0.5Y} = 0.5\sigma_Y = 0.5\sigma_X + 0.5\sigma_Y$ since $\sigma_X = 0$. Combining a risk-free asset with a risky asset averages the risks, but provides no diversification boost.

We can go beyond two assets very easily. Specifically, if assets are labeled by $i \in \{1, 2, \ldots n\}$ and a_i is the fraction of the portfolio invested in asset i,

$$VAR\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j COV(X_i, X_j)$$

where the statement takes advantage of the fact that $COV(X,X) = VAR(X)$ for any given random variable (you should check.) You should convince yourself that things work out in the case $n = 2$.

In other words, the risk involved in a portfolio is a weighted average of the risk of each given asset, and of the covariances across assets. As a result of this, a very nice notational tool in finance given n assets is the so called variance-covariance matrix. It is a table which in row i and column j gives $COV(X_i, X_j)$. In particular, diagonal entries are variances (hence cannot be negative), and the matrix has to be symmetric about its diagonal since $COV(X_i, X_j) = COV(X_j, X_i)$. In a case with 3 assets, a typical variance covariance matrix could be:

$$\begin{bmatrix} 146 & 187 & 145 \\ 187 & 854 & 104 \\ 145 & 104 & 289 \end{bmatrix},$$

a matrix with which you will become very familiar on your homework.

A final fact of critical importance for what we are about to do is that the covariance operator is linear, in the following sense. If X, Y and Z are random variable and a and b

are two numbers, then

$$COV(aX + bY, Z) = aCOV(X, Z) + bCOV(Y, Z).$$

Now, we're in business.

3.2 Financial economies

Theoretical asset pricing relationships are derived in stylized, abstract environments called financial economies. Consider a world populated by a large number of investors. We will follow those investors between two dates: $t = 0$ and $t = 1$. At $t = 0$, investors have some wealth that they need to invest in assets (or securities, or properties, depending on the context) that pay some random return at $t = 1$. These assets include the whole gamut of assets available in modern economies, from cash to real estate or public stocks. In classical portfolio theory, the time between $t = 0$ and $t = 1$ is called the *holding period.*

There are N assets, and we denote the payoff of asset i by X_i. The price of asset i at date 0 is q_i, so that we can define the random return r_i on asset i as follows:

$$r_i(s) = \frac{X_i(s)}{q_i} - 1,$$

where s is the state of the world that prevails at date 1. Those returns are bona fide random variables. The expected return on security i is

$$\bar{r}_i \equiv E(r_i) = \frac{E(X_i)}{q_i} - 1.$$

Investors must decide how to allocate their wealth at date 0 across all available assets. Let α_i be the fraction of their wealth invested in asset i. We must have $\alpha_1 + \alpha_2 + \ldots \alpha_n = 1$ (these are fractions), but it may very well be the case that some α is negative. What does this mean to take a negative (or short) position in an asset? Instead of buying asset i, the investor is telling some other investor to give them some money now, and that he/she will pay asset i's payoff at date 1. To the other investor, this is exactly the same as purchasing the asset directly. To the short-seller, this means getting more money to invest in other assets now, at the cost of having to make payments at date 1. This is essentially how short-selling works (literally, short-selling means selling something you do not have yet, but promising to deliver the asset at a later date) in real markets.

For most of our purposes, it will turn out to be sufficient to assume that negative positions are only possible on the risk-free asset, when it happens to be present. Basically, there is a risk-free rate at which agents can lend and borrow at will. Almost all our asset pricing results require that possibility only, and carry through unchanged if we restrict positions on risky assets to be non-negative.

The return on the investor's portfolio is the random variable $\sum_{i=1}^{n} \alpha_i r_i$. From the previous section, we know how to calculate the expected return, and the risk associated with that portfolio return.

We will assume throughout that investors have mean-variance preferences. This means, in the language of economics, that the utility they derive from a specific portfolio depends only on its expected value (or mean), and its variance (or equivalently, standard deviation.) Investors, we assume, all prefer higher expected returns, holding variance the same. *Risk-averse* investors, at equal expected value, prefer portfolios with less variance. *Risk-loving* investors, at equal expected value, prefer portfolios with more variance. *Risk-neutral* investors are indifferent to risk.

A mean-variance investor is characterized by a utility function $U(\bar{r}, \sigma_r)$ that rises with the first argument (the mean return) and can rise or fall with the second (the standard deviation of the return.) This function can be represented on a two-dimensional chart with mean on the vertical axis and risk on the horizontal axis. Indifference curves are lines that connect all mean-risk combinations that give the investor the same utility. Indifference curves correspond to higher utility as we go north on the chart. They increase as we go east on the chart for risk-averse agents. (Why?) From now on, we will assume that agents are risk-averse.

On that same chart, we can plot the set of return-risk combination that are feasible given the set of existing assets. To get that set, we look at all possible portfolios $(\alpha_1, \alpha_2, \ldots \alpha_n)$ where, recall, $\alpha_1 + \alpha_2 + \ldots \alpha_n = 1$. Each possible portfolio gives a point in the feasible set. One "easily" shows that if there are at least 3 securities, the feasible set is a mass with no holes (it is *connected*, as they say in math.) In addition, if there is no risk-free asset, the north-west boundary bends outward (it is *concave*, as they say in math), as a result of risk-diversification. If there is a risk-free asset, the upper-boundary of the feasible set is a straight-line (as we will show in class.) Denoting the risk-free return by r_f, the feasible set starts a level r_f on the vertical axis and is tangent to the feasible set that would prevail if risk-free borrowing and lending was not present. We will work all of this out in class.

This north-west boundary of the feasible set is called the *efficient set*, and it plays a key role in classical portfolio theory. Portfolios on the efficient set are called *efficient portfolios.* A portfolio is efficient if it is not possible to find another portfolio with no more risk but a higher expected return, and it is not possible to find a portfolio with no less return but with less risk. Other portfolios are called inefficient: they would never be chosen by a risk-averse investor.

An *equilibrium* in this environment is a set of asset prices (or returns) and portfolios for each agent such that:

1. All agents choose the portfolio in the feasible set that maximizes their utility;

2. the total demand for each asset equals its supply.

The game now, basically, is to see what optimal behavior on the part of agents implies for agent returns. Before getting into more details, we should begin with the most important observations of them all, after quickly defining a couple terms:

Definition 1. *The* ***law of one price*** *holds if two portfolios that yield the exact same payoff in all states cost the same.*

Remark 1. *If there are no restrictions on short selling, the law of one price must hold in equilibrium.*

Proof. Assume the law of one price doesn't hold. Take the two portfolios that have identical payoffs but different prices. Buy the cheap one, sell the expensive one. The payoffs cancel each other at date $t = 1$. But the transaction has increased your wealth today, and that extra wealth can be invested at the risk-free rate to generate an additional risk-free payoff at date $t = 1$. In fact, investors will push that game as much as possible, seeking to take arbitrarily large positions in some asset or other, causing some excess demand or supply in at least one market, which is incompatible with equilibrium. $\square$

A deviation from the law of one price is an *arbitrage opportunity.*

Definition 2. *A **strong arbitrage portfolio** is a portfolio with a strictly negative price today and a non-negative payoff in all states.*

Remark 2. *(No arbitrage) If there are no restrictions on short selling, no strong arbitrage portfolio can exist in equilibrium.*

By (somewhat loose, but intuitively sound) extension, similar assets should earn similar returns, in equilibrium. This is the idea that governs most asset pricing exercises performed in practice.

3.3 Classical portfolio theory

Classical portfolio theory derives asset-pricing implications in the environment we just described. Before discussing these, it is important to emphasize the big assumptions that we will maintain in the upcoming analysis:

1. Investors have mean-variance preferences;

2. Holding risk the same, investors prefer higher expected returns;

3. Holding expected returns the same, investors prefer portfolios with lower variances;

4. Investors can divide their wealth across assets however they wish;

5. No taxes, or transaction costs;

6. Investors have all the information they need about assets (they know the distribution of their payoffs.)

The fact that these assumptions clearly do not hold in reality means that our asset pricing formulas are approximate at best, or need to be amended when deviations from our stylized context become important. We will have to think long and hard, for instance, about the role of taxes.

From now on, we will also assume that there is a risk-free asset. That, too, is an approximation, since the most risk-free asset we can typically think of (a short-term government bond), is not literally risk-free. We also assume that agents can hold whatever position they want in that risk-free asset, that they can lend and borrow at will. Yet another strong assumption, obviously.

However, for the results we are about to develop, we don't need to assume that agents can short-sell risky assets. The ability to borrow at the risk-free rate actually render that option of no use. It would not change the feasible set in any way.

Assuming the presence of a risk-free asset pays off handsomely by yielding extremely strong implications. Since all investors face the same returns, they look at the same outer-boundary of the feasible set. That outer-boundary starts at the portfolio that puts all wealth in the risk-free asset, and is tangent to the feasible set one would get if that asset did not exist, at exactly one point. That tangency point is called the *market portfolio*. We will denote by r_m the random return of the market portfolio, while $\bar{r}_m = E(r_m)$ and $\sigma_m = \sigma_{r_m}$.

All investors try to end up with the indifference curve that is as far-north as possible while intersecting – eventually just touching – the flat feasible set. That means that all investors combine the same market portfolio with some lending or borrowing of the risk-free asset. This result is so nice that it ranks as a key theorem in finance:

Theorem 1. *(Two-portfolio theorem) In the set-up described above, in equilibrium, all agents hold a portfolio made of a positive investment in the market portfolio, and a positive or negative investment in the risk-free asset.*

Note that we don't even have to know details about the preferences of agents to conclude this. Note also that less risk-averse agents hold a smaller (possibly negative) position in the risk-free asset.

To make this theorem useful, we need to say more about the market portfolio, and try to think of an empirical counterpart for it. Note first that since all investors hold the market portfolio, all risky assets must have a strictly positive weight in it. Otherwise, the total demand for at least one asset would be zero, and that asset would therefore have to be in excess supply, which can't be in equilibrium. Second, since all investors hold all risky securities in the same proportions (namely the proportions of the market portfolio), the total holding of risky securities must also be in the same proportions as the market portfolio. Therefore, to compute the weights of the market portfolio, we only need to look at total holdings (the capitalization) of each risky security, and divide each of those holding by the sum of all risky holdings. A decent proxy for the market portfolio, then, is a capitalization-weighted market index such as the S&P500.[1]

In order to be really useful however, we need to go from this neat description of the efficient set to a way to price individual assets. We now turn to providing precisely one such method.

3.4 CAPM

Take asset i in our financial economy with risky payoff X_i. What should be its equilibrium price q_i? Equivalently, what should be its average return $\bar{r}_i = \frac{E(X_i)}{q_i} - 1$? Because investors are risk-averse the price should depend on the risk the security is adding to the market portfolio. The greater the risk it adds to that portfolio, the lower the price ought to be. The higher, in other words, the return investors will demand in equilibrium to be willing to hold that security. Now, how much risk the security adds to the market portfolio depends on its own variance (note: how it covaries with itself) and on how it covaries with other elements of the market portfolio, i.e. other risky assets.

The formalization of this idea yields one of the most famous and useful results in all of

[1]Decent only, since what is and what isn't included in the index is somewhat arbitrary, and many sectors, including real estate, are woefully underrepresented. The S&P500 offers many practical advantages however: data on it are widely available, and everybody has heard of it. If you're trying to pitch a valuation to a client, referring to a well-known index helps communication. Yet another reason why sticking to the S&P500 is that once go past the 500 largest capitalizations in the US, you are looking at assets that go through long periods without trades. The return information provided by, say, the Russell 2000 is often too stale to be of use.

finance – the Capital Asset Pricing Model (CAPM) – which can be stated, and demonstrated, as follows:

Theorem 2. (CAPM) *Under the assumptions of classical portfolio theory,*

$$\bar{r}_i = r_f + \beta_i(\bar{r}_m - r_f)$$

where

$$\beta_i = \frac{COV(r_i, r_m)}{VAR(r_m)}$$

is called security i's beta.

Proof. (A bit difficult, but any finance student needs to have seen this at least once.)

Consider the following thought experiment. Assume that we try to get the highest return we possibly can by combining some investment α_i in security i, some investment α_m in the market portfolio, and completing the portfolio by investing (possibly short-selling) $(1 - \alpha_i - \alpha_m)$ in the risk free asset, **all the while maintaining** the portfolio's risk at σ_m. In other words, we want to choose α_i and α_m to solve:

$$\max \alpha_i \bar{r}_i + \alpha_m \bar{r}_m + (1 - \alpha_i - \alpha_m) r_f$$

subject to:

$$\alpha_i^2 \sigma_i^2 + \alpha_m^2 \sigma_m^2 + 2\alpha_i \alpha_m \sigma_{im} = \sigma_m^2$$

where, for short, $\sigma_i = \sigma_{r_i}$ and $\sigma_{im} = COV(r_i, r_m)$. This is a constrained maximization problem the likes of which you encountered in intermediate microeconomics. To solve this, one writes a Lagrangian:

$$\mathcal{L} = \alpha_i \bar{r}_i + \alpha_m \bar{r}_m + (1 - \alpha_i - \alpha_m) r_f + \lambda \left(\sigma_m^2 - \alpha_i^2 \sigma_i^2 - \alpha_m^2 \sigma_m^2 - 2\alpha_i \alpha_m \sigma_{im} \right),$$

where λ is called a Lagrange multiplier. Then, we differentiate $\mathcal{L}$ with respect to the two choice variables and the Lagrange multiplier λ and set those derivatives to zero, yielding, in the first two cases:

$$\mathcal{L}_{\alpha_i} = \bar{r}_i - r_f - \lambda(2\alpha_i\sigma_i^2 + 2\alpha_m\sigma_{im}) = 0 \tag{3.2}$$

$$\mathcal{L}_{\alpha_m} = \bar{r}_m - r_f - \lambda(2\alpha_m\sigma_m^2 + 2\alpha_i\sigma_{im}) = 0 \tag{3.3}$$

Now here's the whole trick to this proof. Up to here, it was all brutal math, now we need to be clever. We know that the market portfolio is efficient. This means that setting $\alpha_m = 1$ and $\alpha_i = 0$ has to solve the problem above. If a solution exists that beats that proposal, it means that there exists a portfolio with the same risk as the market portfolio, but more return. That would contradict the fact that the market portfolio is efficient.

So $\alpha_m = 1$ and $\alpha_i = 0$ must solve both equations above. But note that equation (3.2) then implies that $2\lambda = \frac{\bar{r}_i - r_f}{\sigma_{im}}$. Plugging that fact into equation (3.3) and maintaining $\alpha_m = 1$ and $\alpha_i = 0$, we get:

$$\bar{r}_i - r_f = \frac{\sigma_{im}}{\sigma_m^2}(\bar{r}_m - r_f),$$

which is the CAPM equation. □

Having proved CAPM, let's put it to work. CAPM tells us what the expected return on an asset has to be given its covariance with the market portfolio. It tells us, in other words, what return investors are requiring from a given asset given its risk characteristic. It also tells us what aspect of the risk associated with an asset matters: its β. To get a bit more specific, CAPM says that the required return on an asset is the risk-free rate, plus a risk premium. The risk premium, in turn, is the product of two terms: the asset's β, and the risk premium $\bar{r}_m - r_f$ investors require from the market portfolio. In finance, it is customary to call β the *quantity of risk*, while $\bar{r}_m - r_f$ is called the *market price of risk.*[2]

Say then that you are looking at a property of a particular type. Under CAPM assumptions, all we need to know in principle is the property's beta to know what return we should

[2] I happen to think that this is unfortunate language. It would make much more sense to call $COV(r_i, r_m)$ the quantity of risk, while we should call $\frac{\bar{r}_m - r_f}{VAR(r_m)}$ the market price of risk. Indeed, $VAR(r_m) \equiv COV(r_m, r_m)$ is the covariance of the market portfolio with itself hence, in the CAPM sense, the risk of the market portfolio. Then, $\frac{\bar{r}_m - r_f}{VAR(r_m)}$ is the risk-premium on the market portfolio per unit of risk. That's the proper notion of market price of risk. Then you take the risk embodied in security i, which is $COV(r_i, r_m)$ and multiply that by the unit price of risk to get the premium. But finance chose to go in another direction, and we all have to live with it.

require from it, i.e. what discount rate we should apply to the cash flows associated with it. There are two practical problems. One is that we need a market portfolio. For that, it is best in most cases to follow standard practice and use the S&P500, as we have already discussed.[3] The second practical problem is more tricky. We may not have any historical information on the property's price or return, information that we would need to compute beta. There are several things we can do here. One is to try and find similar properties with publicly available info. The other, much easier, is to use the information contained in the returns of public REITs. REITs that specialize in all different sub-areas of real estate can be found, it is even possible for broad categories to find some that specialize in various regions. One can use the beta's of those REITs as a guide for the risk associated with a specific property purchase or development.

One concern with this approach could be that REITs are bundles of properties rather than single properties, and that therefore, the risk comparison is biased because REITs diversify away property-specific risk. That concern, however, is misplaced, since CAPM tells you precisely that specific risk does not affect the return on a given asset. This is actually a critical notion, and we need to make it more precise.

To that end, recall that $\beta_i = \frac{COV(r_i, r_m)}{VAR(r_m)}$. If you recall your statistics classes or look it up in a statistics textbook, you will see that this formulae is very well-known. It is the slope that you get if you regress r_i on r_m. If that does not speak to you (in that case, I'd recommend you read up a bit on linear regressions, this is the perfect opportunity to learn that very important tool), the main consequence of this observation combined with our CAPM theorem is that we can write r_i (recall, r_i the random return on asset i, not its expected return) as follows:

$$r_i = r_f + \beta_i(r_m - r_f) + \epsilon_i,$$

where, ϵ_i is a random variable such that $E(\epsilon_i) = 0$, and (this is the kicker) $COV(r_m, \epsilon_i) = 0$.

[3] Here's a good place to mention that common practice in the REIT industry is to use NAREIT, an index of returns on publicly traded REITs, as benchmark while invoking CAPM. From the point of view of theory, this common practice is deeply wrong. The benchmark in CAPM should be as broad as possible. If you are excluding all largest capitalizations from your benchmark, you cannot possibly be invoking CAPM. You are running a regression (in the sense we will discuss below) and engaging in a forecasting exercise, which is fine, but presenting this as a CAPM approach is betraying a complete misunderstanding of financial economics. If you are going to follow this standard practice – and one could make a case for it – leave CAPM out of it.

So now, using our nice formulas for variance, we get that:

$$VAR(r_i) = \beta_i^2 VAR(r_m) + VAR(\epsilon_i).$$

In other words, we can decompose the risk associated with security i into two independent, unrelated parts. The first, $\beta_i^2 VAR(r_m)$, is called the *systematic risk* of asset i. The second $VAR(\epsilon_i)$ is called the *specific risk*. CAPM tells us that, basically, we can ignore the specific component. The only part of risk that matter is the part that is correlated with the market portfolio.

A property with a lot of specific risk need not require a bigger return than a property with no specific risk. To put this another way, not only is it OK to use REITs data to get beta's for properties, theory is encouraging you to do so! REITs remove the part of risk that does not matter for betas and required returns. Intuitively, specific risk can be diversified away, market risk can't. Hence only the second form of risk matters.

There are legitimate concerns associated with the REIT-based approach however, but they have nothing to do with bundling per se (see GM, section 12.3, for a partial discussion.) REITs are much more liquid than direct holdings in specific properties, since they tend to be traded in very active market. That, of course, has value, and is reflected in returns. One should therefore add a liquidity premium to REIT-based estimates of returns to specific projects. There is a vast literature on liquidity premia on privately held assets which one can use for guidance in this respect. Next, and this applies to all CAPM applications that rely on historical data, past and future performance may have little to do with each other, particularly during turbulent periods. Third, REITs are exempt from corporate taxation (and thus provide no debt tax shield, see section 3.7 for a definition), while corporate investors get to deduct interest payments from taxable income. REIT returns are not helpful therefore for quantifying the effects of the financing mix one plans to employ in buying a specific property on the property's equity beta. More on ways to deal with that below. Fourth, specific property owners may manage specific properties in ways that affect the property's cash-flow in a way that should be reflected in the discount rate. Finally, public REITs have been popular since the IPO craze of the early 90s. After an initial boom during which most

REITs traded at a significant premium over the face value of their holdings, they sharply fell from grace around 1998 as investors moved on to the latest fad (dot-coms.) Since then REIT returns have been more stable but this gives us, at best, 10 years of clean return information on REITs. In summary, while useful, REIT returns are not a perfect proxy. More generally, finding publicly traded assets that approximate a given property's relevant features well is not an easy task.

One way to test whether CAPM "works" reasonably well in practice is to see what predictions it makes for various types of assets, and whether these predictions are borne out by the relevant evidence. As GM explain (see section 22.2), CAPM survives that basic test OK for broad classes of assets. As one looks at finer categories however, CAPM's empirical performance becomes much weaker. In real estate, this is particularly true for private real estate assets. While there are many reasons for why that may be the case that do not imply a full disqualification of CAPM as a useful tool for pricing these assets (see, again, section 22.2), this does suggest that one should use caution when interpreting the outcome of CAPM-based analysis in the context of specific investment projects.

Returning to the comfortable world of theory, a final way to illustrate the ideas contained in CAPM is to observe that it tells you, quite surprisingly perhaps, that a risky asset can in theory earn less than the risk-free rate. All we need to make this happen is to build a security that co-varies negatively with the market portfolio.

Something very nice about beta's is that they are nice and linear. Therefore, the beta of a portfolio is the weighted average of the beta's of the assets that comprise that portfolio. Consider for instance a portfolio made of share α of asset 1 with random return r_1 and share $(1-\alpha)$ of asset 2 with random return r_2. The resulting portfolio has random return $\alpha r_1 + (1-\alpha)r_2$ and, since COV is a linear operator, the following beta:

$$\beta_P = COV(\alpha r_1 + (1-\alpha)r_2, r_m) = \alpha COV(r_1, r_m) + (1-\alpha)COV(r_2, r_m) = \alpha\beta_1 + (1-\alpha)\beta_2.$$

This fact will come in very handy below.

3.5 Option Pricing

As we discussed in the previous chapters, derivatives are playing an ever larger role in real estate markets. Not only are financial assets whose payoffs derive from real properties more and more popular, we will argue below (and in the advanced topics chapter) that most real estate investments feature some aspects (called *real options*) that work exactly like derivative products. For all these reasons, we need to learn how to price derivatives. This section will derive formulas for pricing one of the most common type of derivatives: options. Discussing option pricing will also serve to illustrate yet again the power of arbitrage arguments.

An option is a contract where one party grants the other party the right (but not the obligation) to buy or sell a specific asset, at as specific price, within a specific time period. A *European option* is a contract where the buyer can exercise the option at and only at the expiration date. An *American option* can be exercised at any point before at the expiration date. Financial theory has provided us with many tools to price the first type of option. We will review those below. American options, on the other hand, are a much tougher nut to crack and theory has made very little progress towards providing general formulas to price them. If you want to become an instant superstar in finance, that is one area where a significant contribution would get you just that.

In a *call* option contract, the buyer purchases the right to buy an asset at given price called the *strike price* before some expiration date. Obviously, the buyer will only exercise the call option if the asset price rises above the strike price. We say in that case that the option is *in the money.* Of course the buyer may not exercise the option as soon as it is in the money, hoping that the price will rise further. The buyer has the option to wait for the perfect time, at the risk of course of losing the current value of the contract were he/or she to pull the trigger immediately. Conversely, a *put option* gives the buyer the right to sell the asset at a certain price.

Let us look at the simplest possible case. There are two dates as before ($t = 0$ and $t = 1$), and the same type of assets as in section 3.2. Let us further assume that there are only two states of the world, the “up state” and the “down state”, with probability p and $1 - p$, respectively. The underlying asset pays $u > 0$ if the up state occurs, and $d > 0$ otherwise, and we assume that $u > d$. It sells for price q today. A call option on this asset gives an investor the right to buy the asset at date 1 once we know what state has occured. If they buy in the first state, they get value u, otherwise they get value l. The strike price is s. The question is: given, this information, what must the market price of the call option be today?

The trick we are going to play to answer that question is the most common trick in finance: build a portfolio with exactly the same payoff as the option (it is called a *replicating portfolio*), and whose market value we know. Arbitrage then tells you that this value must be the option value.

The payoff of the call option is $\max\{u - s, 0\}$ in the up state and $\max\{d - s, 0\}$ in the down state. Indeed, the buyer will exercise the option if and only if it is in the money. Let us assume for simplicity and pretty reasonably in this case that $s > d$ so that $\max\{d-s, 0\} = 0$ hence the call option pays nothing in the down state. Let us assume also (just as reasonably) that $s < u$ making the payoff in the up state $u - s$.

To replicate the call's payoff using assets whose value we know. Let us try with a portfolio with two assets: quantity a of the underlying asset, and some investment b in the risk-free asset.

The payoff in the up state is $au + b(1 + r_f)$, while the payoff is the down state is $ad + b(1 + r_f)$. For this to replicate the desired payoff we must have

$$
\begin{aligned}
au + b(1 + r_f) &= u - s \\
ad + b(1 + r_f) &= 0
\end{aligned}
$$

Solving this system for a and b (you should do it) gives $a = \frac{u-s}{u-d}$ while $b = -\frac{(u-s)d}{(u-d)(1+r_f)}$. Note that this strategy involve taking a negative position in the risk-free asset, i.e borrowing. The cost of this portfolio is, at date zero:

$$
aq + b = \frac{u - s}{u - d} q - \frac{(u - s)d}{(u - d)(1 + r_f)}
$$

Because no arbitrage portfolio can exist in equilibrium, this must also be the value of the call option. This formulae is called the *binomial option pricing formulae.*

It may seem to you that it only applies to a very specific case: two dates, two states. However, we can multiply dates and use the formulae recursively to price contracts with many periods. What's more, one shows with some appeal to deep math that by making the distance between dates arbitrarily small, we can approximate asset with an infinity (a continuum, in fact) of states as well as we want. It has been shown, in fact, that the binomial option pricing formulae implies the famous, Nobel-prize winning Black -Sholes option pricing formulae.

In addition, a bit of work on the formulae shows two things: first, the higher the strike price, the lower the call option value (obviously.) Second, holding the expected return the same, more variance in the asset's return increases the option's value. Since the value of the down state is irrelevant, more distance between the two values implies a higher value in the high state, which is all that matters. This is true not only in the binomial case, but in full generality when it comes to options. All else equal, the more volatile the asset, the more valuable the option.

How about put options? I won't bore you with the algebra but it should be obvious that we can play a similar arbitrage trick for those as well. For details on all this, see the seminal paper of Cox et. al. (1979).

Option pricing is hot commodity in the real estate sector right now, and a skill that would greatly help you marketability in the industry if you take the time to master it. The reason for this is that options are part of essentially every real estate investment. For instance, buying a piece of land is buying, among other things, the option to build a structure on it. The strike price, in that case, is the building cost. The asset value is the property's value, once built. Likewise, when one builds a structure, one has the option to expand or upgrade it down the road. Again, the strike price is the building cost, the underlying asset's value is the market value of the upgrade, or the present value of the additional revenues associated with the upgrade. Recent research has shown that real options of this nature account for a very significant part of the value of most real estate assets.

3.6 Modigliani-Miller . . .

Besides asset pricing, the other fundamental branch of finance is called corporate finance. Corporate finance is concerned with optimal (profit maximizing) ways to finance asset purchases, i.e. the liability side of a business project or firm. The corner stone of that area of study is perhaps the best-known theorem in finance, which we owe to Modigliani and Miller, two Nobel-Prize winning economists (see Modigliani and Miller, 1958.)

As we discussed in the first chapter, the liability side of a commercial investment comprises two basic categories: debt and equity. The mix of debt and equity used to finance a particular asset holding (or the holding of a collection of asset, i.e., basically, a corporation) is called *capital structure*. The corporate finance question is: does capital structure affect an asset value? Does how an asset is financed affect, somehow, its value?

The Modigliani-Miller (MM) theorem says that absent taxes, information problems, transaction costs, limits on borrowing or short-selling, and default risk on debt contracts, an asset's value is independent of the capital structure. After our CAPM section, this should strike you as obvious: the value of an asset depends solely on its payoffs, specifically on their expected value and on how they covary with the market portfolio. How the asset is purchased does not enter this equation.

Obvious as it may seem now, the MM theorem shocked the profession 50 years ago. The profession had to be convinced once again with – you guessed it – an arbitrage argument.

To see how the argument works, we need to put ourselves in the model economy studied by MM. Instead of considering single holding periods, they consider assets that yield payoffs potentially "for ever." In such a world, consider two properties with the exact same random payoff X during all future periods. Note, for emphasis, that the probability distribution of payoffs remains unchanged over time. The first property is financed with a quantity D of debt and E of equity investment. Letting V^L be the market value of the asset today, we must have, as a matter of accounting:

$$V^L = E + D$$

as along as E and D are understood to be the market value of the debt and equity attached to the property. Here, L stands for *levered*, more on that terminology soon. In addition to assuming that the asset yields the same random payoff in each period for ever, MM also assume the asset's capital structure remains the same for ever.[4]

The second property is financed through equity alone, and has a market value at date 0 of V^U (U for *unlevered*.) Assume, by way of contradiction, that $V^U < V^L$ (for concreteness, we could work with the opposite inequality.) We will show that this situation creates an arbitrage opportunity for investors, which, as we know, can't be in equilibrium.

Let X be the random payoff associated the property. Note that the payoff is what it is by assumption, it is a fundamental characteristic of the asset. Assume an investor purchases fraction α of the levered firm's equity, at cost αE. The random payoff associated with this strategy at each date will be $\alpha(X - r_f D)$.

The investor, instead, could sell fraction α of the levered firm equity, and purchase an amount αV^U of equity in the unlevered firm, by borrowing αD at the risk-free rate, making the initial outflow

$$\alpha V^U - \alpha D = \alpha(V^U - D) < \alpha(V^L - D) = \alpha E.$$

Basically, the investor is using his ability to borrow to lever its purchase of the unlevered asset. If, somehow, the market value of the unlevered asset is less than that of the levered asset, acquiring a given share of the ownership of the unlevered asset is cheaper. Note that this gives the investor a random payoff of $\alpha X - \alpha r_f D$ at each date.

Now, we have the contradiction we wanted. First, the two investment strategies yield exactly the same payoff at all dates. Second, the second investment strategy is strictly cheaper. That's an arbitrage opportunity and we know that this cannot happen in equilibrium.

Exactly what do we learn from the MM theorem? Not so much that capital structure does not matter in practice. We learn, instead, that if it matters, it must be because of the frictions which MM assume away. Taxes, in particular, make capital structure matter. So do the fact that more debt makes financial distress more likely, or, on the other hand, may serve to discipline managers by reducing the free cash flows they get to play with. MM tells

[4]Whenever one applies the techniques we are about to describe, one should keep these assumptions in mind. Big deviations from them should lead to adjustment in techniques. In practice of course, hardly anyone bothers making those adjustments.

us where not to look for optimal financing criteria. More on this in the final section of this chapter.

As if that weren't enough, MM give us even more in that extraordinary 1958 paper of theirs. Define

$$r^U = \frac{X}{V^U} \left(\equiv \frac{X}{V^L} = \frac{X}{E+D} \right),$$

to be the return on equity in the unlevered firm. Now note that the return r^E on equity in a levered firm is

$$r^E = \frac{X - r_f D}{E} = \frac{X}{E+D} \frac{E+D}{E} - r_f \frac{D}{E} = r^U + \frac{D}{E}(r^U - r^f).$$

This algebra may not look like it, but it contains tons of information. First, as long as $r^U > r^f$ (a safe bet since equity is risky while debt isn't), more debt raises the return on equity, a phenomenon known as leverage. Financing a high expected return by borrowing low expected return funds (debt) raises the returns to the residual claimant.

Why is that? That's because equity becomes riskier when more debt is involved. To see this trivially, note that

$$VAR(r^E) = VAR(r^U) \left(1 + \frac{D}{E}\right)^2.$$

But now that we know more finance, we can be even more sophisticated. Let β^U be the beta of the unlevered firm, while β^L is the beta of equity in the levered firm. Then:

$$\beta^L = \beta \left(r^U + \frac{D}{E}(r^U - r^f) \right) = \beta \left(r^U + \frac{D}{E} r^U \right) = \left(1 + \frac{D}{E}\right) \beta^U.$$

This derivation uses two facts. First, since r^f is risk-free it does add to the asset's risk and can be dropped. Second, as we have shown, the beta of a combination of assets is the average of the betas. This says, yet again, that equity becomes riskier as leverage $\left(\frac{D}{E}\right)$ rises.

Finally, these properties tell us that the average return required from the asset by all stake-holders (again, to emphasize, as long as there are no frictions), the so-called weighted-average cost of capital, or WACC, is the same regardless of the capital structure. Indeed:

$$WACC = \frac{E}{E+D} E(r^E) + \frac{D}{E+D} r^f = \frac{E}{E+D} \left(E(r^U) + \frac{D}{E}(E(r^U) - r^f) \right) + \frac{D}{E+D} r^f = E(r^U),$$

an expression which does not depend on D, or E. This is called MM's proposition II. (What a paper!)

One easily shows (you will show it in your homework) that, under the MM assumptions, the value of the property is the present value of all its cash flows discounted at the WACC. Loosely speaking, $V^U = V^L > 0$ means that the property produces enough cash-flows to compensate all its stakeholders. Furthermore, under MM assumptions, this is true regardless of the capital structure used to finance the asset. This motivates the single most important method used in practice to value assets: compute the WACC associated with the purchasing of the asset, and discount all future flows associated with the asset using the WACC. More on this after we introduce frictions that make capital structure relevant.

3.7 . . . and beyond

Now we need to go beyond MM, because 1) reality is not frictionless, 2) cash-flows do not have the same distribution over time, 3) the debt level is not constant over time either. Let's begin by introducing frictions. Here, since this is not a corporate finance class, we will deal only with the role of taxes. Taxes matter for capital structure because interest on debt is tax deductible at the corporate level in the US, as in most countries. Payments to equity are not. Debt thus represent a "tax-shield" for a US corporation that faces taxation. If it faces a tax rate of τ, debt D reduces its tax liability in a given period by $Dr^f\tau$ (see below for the mechanics.) Effectively, this raises the asset's payoff.

Now the first question that ought to come to your mind is that if debt offers such a clear advantage, why not push it all the way and forget about equity altogether? Answering this question well is beyond the scope of this class. Suffice it to say that as leverage rises, problems may arise. For instance, the risk of financial distress rises. There are, therefore, decreasing returns to milking the tax advantage of debt, and, typically, the optimal level of the debt to value ratio is well below 100%.

Having said that, what becomes of the Miller-Modigliani results when the property owner is subject to taxation, and when debt payments are tax deductible, as is the case for instance when the owner is a corporation? We need to make several corrections. First, note that taxable cash flows from the property are now given by

$$X - Dr^f,$$

so that after tax cash-flows are

$$X - \tau(X - Dr^f) = (1-\tau)X + \tau Dr^f.$$

The very last part of the expression (τDr^f) is called the debt tax shield. Given these calculations, the property can be viewed as the sum of two assets (two strings of cash flows): property income net of taxes, and the present value of the tax shield. Furthermore, since (by assumption, in reality this often needs to be amended) the tax shield is risk-free, it can be treated as a risk-free perpetuity. Its present value is (check that you get the last equality):

$$\sum_{t=1}^{+\infty} \frac{\tau Dr^f}{1+r^f} = \tau D.$$

The value of the risky, after-tax cash flows, for its part, is the value of the asset if it is purchased without any debt. It follows that, with taxes:

$$V^L = V^U + \tau D.$$

The exact nature of this expression is highly MM-specific, but fundamental idea holds in near-complete generality (as long as markets function well enough.) The value of a deal is the fundamental value of the cash flows – what the cash flows would be worth if the project were all equity financed – plus the net present value of the financing. This is the adjusted-present value principle. On a given deal:

$$APV = NPV(property) + NPV(financing),$$

i.e., the adjusted present value is the net present value of the property plus the net present value of the financing used to finance its purchase.

What becomes of MM's other results? Recall that the required return on equity goes up when debt is introduced, essentially to compensate shareholders for the additional risk they must assume. With taxes however, there are now two offsetting factors. More debt increases risk as before. But now we've added a risk-free component to cash flows (the tax shield). Let's compute the net result of these conflicting forces. Denote the unlevered return by r^U, as before. Then:

$$r^U = \frac{X(1-\tau)}{V^U} = \frac{X(1-\tau)}{E+D-\tau D} = \frac{X(1-\tau)}{E+(1-\tau)D}$$

On the other hand,

$$r^E = \frac{X(1-\tau) - r^f(1-\tau)D}{E}.$$

These expressions should look very familiar from the previous section, and the same argument as in that section shows that:

$$r^E = r^U + \frac{(1-\tau)D}{E}(r^U - r^f),$$

and, that,

$$\beta^L = \left(1 + \frac{(1-\tau)D}{E}\right)\beta^U$$

Risk associated with leverage, and the compensation for it, both rise. Finally, the weighted-average cost of capital, in this case, is:

$$WACC = \frac{E}{E+D}E(r^E) + \frac{D}{E+D}(1-\tau)r^f.$$

3.8 WACC at work

How do people use all those ideas in practice? By far the most common approach to valuing a property (outside at least of theory-free ratio methods, or payback period criteria) consists of the following steps:

1. Produce an expected path of cash flows associated with the asset (using, typically, standard market analysis, more on that in the next chapter)

2. Compute an expected return on equity which involves the following substeps:

(a) Find a set of comparable assets with publicly available information, or find the appropriate REITs data

(b) Estimate the beta of these comparable assets

(c) Unlever their beta (see below for a discussion)

(d) Compute an average unlevered beta

(e) Re-lever this average beta given the specific finance mix used the finance the purchase of the asset, and the appropriate tax rate

(f) Calculate the required return on equity using CAPM

3. Compute the WACC using this return on equity, the cost of debt, and the appropriate tax rate

4. Discount cash flows using this WACC

How well does this standard approach mesh with what theory says we should do? The brutal truth is that the fit is poor, and we are now going to get into why that is the case. Before we do so, however, one should keep in mind that it is not clear that a vastly superior alternative with comparable ease of use is available out there. What's more, the method is the industry standard, and there is something to be said for using a criterion with a long track record and which everybody understands.

Another good aspect of the WACC approach is that even if it gives an approximate answer at best, a positive NPV under WACC means loosely speaking that the property is expected to generate sufficient flows to compensate all its stake-holders. This is an intuition that is easy to market to clients. Even there though you have to be careful: expected paths are just that, an average of all possible paths. It may well be the case that a positive NPV investment under WACC actually comes with a very high probability of failing to meet expected returns, not to mention downright failure to meet even obligatory payments.

Returning to deeper problems, under the MM assumptions (particularly constant expected cash flows and constant debt), one can show that discounted after-tax flows at the WACC gives the right answer. The problem is that those assumptions are heroic, and don't come close to being met in most cases. Miles and Ezzel (ME, 1980) generalize the MM assumptions and show that the WACC approach works in theory even when expected cash flows are not constant, as long as the leverage ratio (the debt to equity ratio) remains constant over time at market values. Still unlikely to be met exactly in most cases, but, at least, it gives us a bit more flexibility.

Second comes the issue of how one should unlever betas. A big problem is that different models tell us different things. By far the most common way to unlever beta is to use the expression implied by the MM model:

$$\beta^U = \frac{1}{\left(1 + \frac{(1-\tau)D}{E}\right)} \beta^L$$

Unfortunately, this is inconsistent with the more flexible ME environment, and strictly speaking, only works as long as the MM assumptions are met. In this class, we will use this unlevering method nonetheless, because it is standard practice. Relevering the beta is done simply by using the same MM expression in reverse. Note importantly that if one uses REIT returns as a proxy for the equity beta, the result should be delevered assuming that $\tau = 0$ since REITs are exempt from corporate taxation (see pp600-601 in GM for more on this.) The relevering step should be done using the investor's effective tax rate in all cases.

Third comes a set of CAPM-related issues we have already discussed: CAPM only holds strictly under CAPM assumption, it is not clear that we have a solid counterpart for the market return (although, as I argued above, using the S&P500 is probably best on practical grounds), and we don't even have a proper counterpart for the risk-free rate (although, as I argued above, using T-bill returns is probably best on practical grounds.)

Finally, it gets worse. This method is typically applied inside investment firms that manage a variety of projects. Theory tells us in no ambiguous terms that WACCs should be project-specific. The peer group of assets used to estimate the required return on equity should be picked based on the project's type. In addition, the weights used in the WACC calculation should reflect the financing used for this particular investment, not the company's average mix. Instead, in practice, investment companies often compute a once-and-for-all, one-size-fits-all WACC to be applied to all investment projects. There is only one way to put this: that's not right.

Like all of us, you will get to sweep all those issues under the rug in practice. But you should have heard them at least once.

Chapter 4

Real estate investment analysis

Now it is time to apply the tools we have learned to actual real estate valuation problems. In doing so, we are going to discuss a lot of industry-specific details. But you should not lose sight of the fact that the basic idea behind investment analysis is trivial. First, we need to describe the cash flows associated with a given property. Second, we need to estimate the market value of those flows. One way to do this is to discount expected cash flows at the appropriate rate, which requires estimating a discount rate. Another way, much simpler, uses basic information about comparable properties to infer the market value of the property under study.

4.1 Cash flow pro-formas

We begin with the part of real estate investment analysis which is common to all real estate valuation approaches: describing the current and future cash flows expected to be associated with a given property. It is standard to present revenue and cash flow projections in a "proforma table", a table that presents projections of the key elements of cash-flow calculations for a revenue-generating property. Typically, projections are made 5 to 10 years into the future on a eminem yearly basis, but both the horizon and the frequency may have to be adjusted in the context of a specific project.

In this section, it is probably best for concreteness to think of a property that amounts to a collection of rentable units, such as an office building, and apartment complex, a shopping center, a self-storage business ... but it should be clear that pro-formas can be written for any sort of asset.

The bottom line of each column of a pro-forma is the property-before-tax cash flow (PBTCF) which corresponds to the property's expected revenue in a given period minus the expected cost. To compute the expected revenue we begin by calculating the potential gross income (PGI, also called rent roll) which is the product of the rentable space and the expected rent per square foot. It is, in other words, the revenue the property is expected to generate if it operates at full capacity, under an assumed market value of rent. Here it is helpful to carry out calculations rentable unit by rentable unit, hopefully for added precision.

Physical capacity (rentable space) is the trivial element of this calculation. Expected rent per square foot, on the other hand, is not easy to project with certainty. If we are looking at an existing property, the natural thing to do is to look at past values of this item, and to project them into the future using statistical and market analysis. This basically involves describing current and future supply and demand conditions in the market of interest, and, based on this description, inferring a market clearing rent rate. This is much, much easier said that done (read chapter 6 in GM, for more details) and requires stating a full-blown econometric model to be done properly. In practice, at least in preliminary calculations, it is sometimes a good idea to outsource those calculations to institutions that specialize in this sort of exercise. If one goes that route, marketwide numbers need to be adjusted for specific features of the property under study that make it different from the average property in a given market. These specific factors include for instance recent upgrades or recent damage, on a unit-by-unit basis.

Much as we might wish otherwise, properties seldom operate at full capacity, and the next important item to project, therefore, is the expected vacancy rate for each unit in each period over the horizon of the projection. Again, there are two sources information for this: past vacancy rates for the property if available, and market analysis. The vacancy allowance in a given period is the part of PGI lost to vacancy. The property's net operating income (NOI) then, is given by:

$$\text{NOI} = \text{PGI - Vacancy Allowance} + \text{Other Income (if any) - Operating Expenses.}$$

Other (non-rent) income includes the expected income from any optional services provided to tenants. Operating expenses, for their part, are all the costs associated with operating the property in a given period. Since we are doing a cash-flow analysis, these expenses should include all costs that imply a cash outlay. Pure accounting expenses, such as depreciation, should not enter this part of the calculation.[1]

Operating expenses come in two basic forms: *fixed expenses* do not vary with the occupancy rates, while *variable expenses* do depend on occupancy.[2] Utility bills are the most obvious form of variable cost. Property taxes, for instance, are fixed, at least if one has a reasonably short period of analysis in mind. Labor costs can be of either type. Management expenses – at least for small properties – tend to be fixed, contractual labor expenses (cleaning and maintenance, say) tend to vary with occupancy.

Lease contracts sometimes include quirky expense agreements, such as *expense stops*, under which tenants pay expenses if they exceed a specified amount. This gives tenants the flexibility to use units more intensively than normal when needed, at their own expenses. We will cover the full menagerie of lease covenants in class.

Operating expenses occur basically at the same frequency as cash flows. More infrequent and often larger expenses are called *capital expenses.* They include building improvements or long-term maintenance (replacing a roof, say), among other things.

The line between operating and capital expenses can be fuzzy. Abstracting from tax considerations, the only thing that matters is that all expenses be taken into account. With taxes however, what is capital and what is operating matters a lot. Operating expenses are expensed immediately. The tax man, however, requires in most countries that capital expenses be expensed over several years via depreciation deductions. In that context, for instance, building betterment is completely different from building maintenance. More on this when we bring up taxes later in this chapter.

[1] One important exception to this cash-only principle is the value of the owner's time when they are managing the property directly. Even if they do not write a paycheck to themselves, this is tacitly a labor expense.

[2] What is fixed and what is variable depends of course on the length of the period analysis. In the long-run, all expenses are variable.

Having introduced capital expenses, we may now define:

$$\text{PBTCF=NOI- Capital expenditures,}$$

which is the bottom line in a world without debt or taxes.

Now we are only missing one piece, which is the property's terminal value at the end of the period of analysis. If we were to own the property literally for 10 years, we would sell it at the end and receive its market value. Even if we plan to hold the property longer, we need to include the present value of all operating cash-flows that do not appear in our arbitrarily truncated pro-forma. GM and much of the industry call this terminal value the final reversion cash flow. A reversion cash flow is received whenever all or part of the property is sold. Reversion cash flows sometimes appear within the period of analysis if a unit or two or divested or sold before the last period. The terminal cash-flow always appears in the last period.

How are terminal values calculated? There are two basic things we can do. One is back-of-the-envelope calculations, based on the multiple approach. This works as follows, guess the PGI, NOI, or PBTCF one period after the last period of analysis, estimate the average ratio of property values to these characteristics, and apply those ratios to your financial estimates. More on that in the peer-asset section below. Another approach is to guess PBTCF in the year after the last period, guess a permanent growth rate of this cash-flow, and discount the resulting infinite perpetuity at the appropriate discount rate.

After arriving at this last number, our pro-forma is complete. You should read the numerical example in GM's section 11.1.8 in detail. In your homework, I ask you to replicate a 5-year version of exhibit 11.3 using the trial version of Argus that came with your textbook.

4.2 Discounted Cash Flow approach

The DCF approach comprises three basic steps:

1. Produce an expected path of cash flows associated with the asset (we now know how to do that)

2. Find the required return on the asset

3. Discount cash flows using this discount rate

As we discussed in the past two chapters, the appropriate d before-tax discount rate in this context is the before-tax return investors require on assets similar to the property under consideration.

A good place to start is to look at historical evidence on returns in the class of assets in which the property belongs. As GM discuss, the total return on large, institutional quality commercial properties as a broad class of asset has been of the order of 10% a year since the mid-70s. Not a bad place to start for a first set of NPV calculations. It is important to recognize however that institutional investors tend to buy large, premium, relatively safe properties. Average returns on non-institutional investments are higher by 100 to 200 basis points.

We can refine this by looking at specific sub-classes of real estate assets (shopping malls, apartment complexes, office buildings ...) using, for instance, evidence from publicly traded REITs. In using this evidence, one should by way of sensitivity analysis consider different sample periods, and in particular, produce numbers based on more recent time periods.

Another source of data on returns come from survey of investors, which basically ask investors what returns they require from various types of investment (see GM, section 11.2.)

Finally, one can use widely available estimates of *capitalization rates* or cap rate for short, the ratio of NOI or PBTCF to the value of the property. Assume that the current (PBTCF) cap rate is y (for yield) and that the PBTCF is expected to grow at a rate g for ever. Letting r be the total return, i.e. the appropriate discount rate for a particular property and letting V be the property's current value, we must have, by definition of r:

$$V = \sum_{t=1}^{+\infty} \frac{Vy(1+g)^t}{(1+r)^t}.$$

Indeed, the first cash flow is $Vy(1+g)$ by definition of the cap rate and it then grows at an exponential rate g. For this equation to hold, we must have, assuming that $r > g$:[3]

$$1 = y \sum_{t=1}^{+\infty} \frac{(1+g)^t}{(1+r)^t} = y \frac{\frac{1+g}{1+r}}{1 - \frac{1+g}{1+r}} = \frac{y(1+g)}{r-g},$$

which implies:

$$r = y + g + yg.$$

Since yg is a very small number (the product of two small percentages is a very small number), we get as a good approximation, the holy trinity© of real estate:

$$r \approx y + g.$$

Given survey evidence on y, a guess about g, we have yet another guess for the total return.

An important warning here (see GM) is that most cap rates are NOI cap rates. Since we are working with pro-formas where PBTCF is the bottom line, we need to use PBTCF cap rates. Most surveys, unfortunately, inquires about NOI cap rates, and the numbers that come from those surveys, accordingly, should be adjusted.

Using all these sources of information, a discount rate decision can be made. All we need then is to discount cash flows at that rates. This gives the present value (PV) of the property. Given a proposed price, we can then calculated the NPV of the property as PV - Price. A positive NPV means that the property yields a return in excess of the return one would normally require from an asset of this type and is, in that sense, a profitable use of funds. Alternatively, given a price and a set of cash flows, one can calculate the property's IRR and simply ask if this IRR exceeds the required return.

[3] If $r < g$ we're looking at an asset with an infinite market value, which can't be in any financial economy equilibrium.

4.3 Ratio valuation procedure

The DCF approach is nice enough and founded on solid financial economics principles but, in practice, investors rely more frequently on a simpler approach: the ratio valuation procedure. The idea behind this approach is simple and intuitive. Build a set of properties similar to the one your are considering, with known current operating income or gross income as it is often called, NOI and PBTCF. In practice, a set of 5 to 10 similar properties is common, with known market values. Known, presumably, because there was a recent transaction or at least a recent, solid appraisal performed on the property. Next divide the property's market value by each of the three financial measures to obtain an average Gross income multiple, a NOI multiple, and a PBTCF multiple. Finally, apply this average multiple to the financials of the property under consideration and voila, as the French say.

The peer group should be made of properties that are as similar to the property under consideration as possible in terms of size, location, purpose, vintage, capital structure and, generally, all things that obviously matter for value. In practice, perfect matches are impossible to find. So fudge factors, otherwise known as lack of comparability discounts, are often applied to the number that comes out of ratio methods. Corrections are often made also for any advantage the buyer may have in terms of managing the property. A sole proprietor, for instance, may estimate that not having to deal with the dissension that characterizes some partnerships, or not having to delegate management, has value. This is often referred to as a control premium.

An alternative to going out and putting together your own peer property data is use survey or historical evidence on cap rates. A rule of thumb here is that commercial properties typically trade at roughly 10 times their yearly NOI. There is some variation across property types and quality (see exhibit 11-7a) of course, and one should take that into account in their choice of multiple.

When does this approach give the right answer, namely the same answer as correctly done DCF with the correct discount rate? Barring a very lucky coincidence – generically, as economists say – this approach gives the right answer if and only if the peer properties are not simply similar to the property under consideration, but are in fact identical to it up to scaling up or down. That is, the ratio of financials between the peer group and the target property has to be the same now, and for ever. That assumption is never met exactly, obviously. The only hope is that it is met approximately.

Still, the exercise is always very useful. If a buyer is asking for a multiple that is way off the average multiple, they had better provide a justification for it if the deal is to make sense. Defending a surprisingly high multiple requires solid evidence that the property's cash flows are going to grow faster than those attached to the properties in the peer group, or, for that matter, the industry at large.

In addition, the exercise complements the DCF approach nicely. Once again, if what comes up from the DCF analysis differs from what the multiple approach yields, there had better be a reason. Comparing the two sets of numbers provides a nice framework for investment project discussions around a meeting table. If, on the other hand, the two approaches give similar answers, it builds confidence in your general analysis.

As outlined here, the method produces an estimate of the total value of the property. The value of equity can then be estimated by deducting the value of any debt on books and adding any cash on books attached to the investment.

4.4 Debt and Taxes

As we discussed in the last chapter, many investors are subject to taxation, and the investment analysis should therefore reflect the effects of taxes. We know for instance that when there are taxes, the finance mix used in making an investment can affect its profitability.

There are two ways to do things properly in this context: discount after-tax cash flows at the WACC (see the last section of the previous chapter), or discount cash-flows after taxes and debt service payments at the required return on equity. In the second case, the value one gets is the value of the equity stake in the property, i.e. $E \equiv V^L - D$ in the notation we used in the previous chapter. As we discussed there, the first method is the most common in general capital budgeting, at least after the ubiquitous ratio calculations. In real estate however, there seems to be a preference for the second method. We will discuss both methods.

First, we need to amend our pro-forma tables to produce a couple of new objects. One object we need to calculate in each period is taxable income. In most cases, depreciation and interest payments on debt are deductible, so that:

$$\text{Taxable Income} = \text{NOI - Depreciation Expense - Interest Payments.}$$

As I mentioned above, this is one place where the distinction between capital expenditure and operating expense is critical. Operating expenses can be deducted immediately (recall that it is part of the NOI calculation), capital expenditures can only be deducted according to a schedule set by the federal government for each capital asset type. And the law in this respect changes often. For instance, one common trick the federal government uses to give investment a boost when it needs it is to temporarily allow business to expense capital purchases faster. Then,

$$\text{Income Tax} = \text{Taxable Income} \times \text{Investor's tax rate,}$$

and,

$$\text{PATCF} = \text{PBTCF - Taxable Income,}$$

where PATCF stands for property-after-tax cash flow. Of course, equity holders only get cash flows net of whatever goes to debt-holders, so that:

$$\text{EATCF} = \text{PATCF - Debt service payments,}$$

where EATCF stands for equity-after-tax cash flow.

Reversion cash flows, likewise, should now be reported net of capital gains, if any. On a basic level, capital gains due on the sale of a property are what we get at the time of the sale - what we paid for the property. To the tax man, what we "paid" for the property is:

$$\text{Adjusted basis} = \text{Book Value} + (\text{Total Capital Expenditures - Total Depreciation}).$$

The idea here is that the initial outlay (the book value) plus all capital additions are what we effectively paid for the property. But tax authorities already let us deduct part of our capital expenditures as depreciation expenses. Since the tax man is not one to gift you a double deduction of one item (they will tax you twice on the same item often, but Lord forbid they would let you deduct one item twice), total depreciation over the life of the investment has to be deducted from the cost basis. Now, we get:

$$\text{Capital Gains} = \text{Net Sale Proceeds - Adjusted Basis.}$$

Now, you'd think that the calculation of taxes owed would simply involve a multiplication of capital gains by a uniform rate. That'd be too simple. Instead, tax authorities in their infinite wisdom have decreed that distinction needs to be made between too pieces of the overall accounting gain. Note that:

$$\begin{aligned}\text{Net Sale Proceeds - Adjusted Basis} &= \\ \text{Net Sale Proceeds} &- (\text{Book Value} + \text{Total Capital Expenditures}) \\ &+ \text{Total Depreciation}\end{aligned}$$

The first piece of the calculation is taxed at the capital gains tax rate while the second is taxed at the typically higher *depreciation recapture tax rate.*

Thank God for Turbo Tax and for the tortured souls who choose for whatever reason to dedicate their life to tax accounting. We can then deduct the overall amount due in taxes at reversion from the net sale proceed to get the reversion cash flow net of taxes. For equity-level calculations, part of this cash flow will often go to making a final debt principal payment which is reflected as a final debt service payment that equals in most cases any remaining loan balance.

All we need to do now is to discount total PATCF flows at the WACC to get an estimate of the total value of the property, or discount EATCF at the required return on equity to get an estimate of the equity value of the property. In the last chapter, we discussed how one should compute the return on equity, and we know that even that "best-practice approach" is fraught with heroic approximations. In the real estate context, one should be even more humble since more often than not, we are trying to put a price tag on a private asset with little historical evidence to rely upon. Section 14.3.1 in GM is appropriately humble in that respect. My personal taste is to report an overall IRR before taxes, an overall IRR after taxes, and an equity level IRR after taxes. Then one should present the best available evidence they can find on what average returns of these three sorts are on comparable assets out there and let the client decide for themselves whether there is enough of a gap between these returns to justify investing in a given property.

4.5 Sensitivity analysis and simulations

All valuation exercise rely on strong assumptions, and it is important to be upfront about the consequences of big deviations from these assumptions. It is very important however to understand that sensitivity analysis is different from risk analysis. We are not asking and one should never ask: what is the project's NPV if the worst case scenario materializes? Such a question is simply meaningless and betrays a complete misunderstanding of basic finance principles. The one and only NPV is obtained by discounting average expected cash flows at the appropriate discount rate. The place to capture pure risk (the fact that cash flows are uncertain) is the discount rate. In fact, if one computes the NPV by discounting all possible cash flow paths and weighting them by their likelihood, one gets the same answer as simply discounting expected cash flows. Simulations can play no role in assessing pure risk.

Instead, the point of running simulations or scenario analysis is to ask: what if I am wrong in my estimates of expected cash flows? More specifically, take all elements we used in arriving at baseline forecast. The sensitivity analysis question is: what would the NPV be given specific changes in a particular subset of these elements? For instance, "What if the expected resale value is overstated?" is one type of question clients often ask (see GM p248.)

Simulations are also useful as a way to check that we did not make a subtle mistake in computing average cash flows. Pure mathematical errors in producing financial averages are very common. People often employ short-cuts that lead to catastrophic mistakes. Simulations, in principle, minimize that risk. One plots all possible future paths under a given set of assumptions. A probability-weighted average over all these realizations produces the correct expected cash flow path, without resorting to algebraic shortcuts that may or may not be OK.

Here's an illustration of a common mistake. Consider a project where the vacancy rate is 10% on average. At the expected 90% occupancy, net revenues are positive and the investor can keep operating the building without encountering any financial issue. However, the 10% vacancy average hides quite a bit of variance: the rate is 0% in each period with 80% probability, while it is 50% the rest of the time. When the vacancy rate is 50%, net revenues are significantly negative, and the investor has to dip into his own funds to keep the project running. Because his own funds are limited, three consecutive bad vacancy draws would force him to shut down the project, at a significant loss. Assume that the analyst computes average NOI as the average PGI minus the average vacancy allowance minus the average operating expenses. Since at this average NOI financial distress is not an issue, he/she does not factor that possibility into his calculations. The result is an overstated project value.[4] Correctly simulating every possible path instead of pre-averaging each sub-element of the calculation greatly minimizes the risk of error. Of course, the potential for coding error now enters the picture, and triple-checking simulations is necessary.

[4]This situation is analogous to what has become a textbook personal finance mistake. Assume a retiree calculates the withdrawal rate from their saving that will enable them to maintain a certain income level for a certain number of years using average savings returns. This calculation ignores the fact that savings returns vary a lot. In fact, as long as there is enough variance in this return and the planning horizon is long enough, one can show that these calculations will cause the retiree the run out of money with near certainty before the desired date.

Chapter 5

Mortgages

Mortgages are part of just about every real estate investment project, whether residential or commercial. They are a debt contract secured at least in part by the underlying property. This chapters discusses the key features of this ubiquitous contract, and the various risks lenders have to consider when they issue it.

5.1 Legal framework

Legally, a mortgage comprises two distinct part: a promissory note and a mortgage deed (a.k.a. security deed or deed of trust.) The first part stipulates the contractual obligation by the borrower, or mortgagor, to make a specified list of payments. The deed states the legal claim of the lender (or mortgagee) to the property as collateral in the event of default. In lien theory states[1] the deed of trust states the right the lender has to force a foreclosure[2] when the borrower fails to meet the obligations laid out in the promissory note. In title theory states instead, the lender retains ownership control of the property until the loan has been paid off.

Typically, several claims or liens are attached to a particular property. Even when only one mortgage is attached to the property, local authorities have a so-called tax lien over the property and can, as such, initiate the foreclosure process. In all cases, tax liens come first.

[1] Lien means link in French.

[2] "Early close": act of terminating the borrower's current or potential ownership right over the asset.

When there are several mortgages, they are organized in order of seniority. The first lien or first mortgage has first claim to post-tax foreclosure proceeds, while subsequent liens get the residual proceeds. Obviously, this makes junior liens more risky than senior liens, which is reflected in yields at origination. More on that in the default section.

Any lien holder can initiate a foreclosure suit. In fact, a junior lien holder may have the greatest incentive to start the foreclosure process. When signals that the borrower may be heading towards default begin to mount, time is of the essence from the point of view of residual claimants. Delaying the settlement process carries cost (more missed payments, property value erosion for lack of care or market depreciation) which are eventually borne by the marginal claim holder.

Mortgages usually carry oodles of quirky covenants. GM pp394-96 goes into details, here we will only discuss the prominent ones. Most mortgages contain acceleration clauses which stipulates conditions under which the entire loan balance becomes due. There are two big events that cause acceleration: default and sale. The due-on-sale clause, in particular, gives the lender the right to make the entire balance of the loan due when the property is sold by the borrower. Similar clauses, which come under various names, enable the borrower to accelerate the loan if default occurs.

A critical aspect of mortgages is whether or not they come with a prepayment clause. That clause, when present, gives the borrower the right to make payments beyond those stated in the promissory note, thereby accelerating loan amortization. Almost all residential mortgages in the United States allow prepayment. From the point of view of the lender, this adds uncertainty to the distribution of mortgage payments, as we will discuss in the prepayment risk section. Commercial mortgages, on the other hand, typically do not feature prepayment clauses. Instead, they usually stipulate penalties associated with prepayment. Often in fact, these penalties are so high that prepayment risk is all but eliminated.

Another important aspect of the mortgage contract are the stipulations of the extent of recourse in the event of default. In some states, the liability of borrowers can in principle extend to assets beyond the property that collateralizes the loan.[3] Nonrecourse loans are loans that carry an exculpatory clause that basically limits the liability of the borrower to

[3] The law differs a lot across states when it comes to recourse. See Ghent and Kudlyak, 2009, for details.

the property. Recourse loans, in states where they are permitted, are subject to deficiency judgments initiated by lien holders. Whether deficiency matters much for default rates hence mortgage pricing is the subject of a very active literature. Conventional wisdom is that deficiency judgments are of limited value in practice because they are costly for the lender, and in many default situations, borrowers have little left to go after. The recent foreclosure crisis is making testing the validity of this view a critical research question.

These legal technicalities are of limited importance when all goes according to the plan and payments are made on time as they should be until the loan matures. They matter greatly, on the other hand, in the event of default, i.e. when the borrower fails to meet part of their contractual obligations. In the event of default, lenders can do a number of things. First, they can attempt to avoid the unpleasantries and costs associated with the foreclosure process by negotiating a solution with the borrower, in a process called a workout. Specifically, the loan can be modified for instance by making payments lower for a while in exchange for higher payments later, or even by adjusting the loan balance down. In practice, lenders are loath to engage in loan modification, for a variety of reasons. For one thing, this could simply delay the inevitable (see Adelino et. al., 2009), which is costly. In addition, making loan modification a distinct possibility would cause incentive problems associated with asymetric information. It would give borrowers an incentive to misrepresent their financial difficulties or, worse, to take action to put them in a situation where a balance reduction becomes necessary.

When default cannot be avoided, another nonlitigious way to avoid legal costs is a procedure called *deed in lieu of foreclosure*. Under that procedure, the borrower turns over the property to the lender who then releases them of any remaining loan obligation.

The lender also has the option to *sue for specific performance* rather than invoke the deed of trust, a step that is often meant to scare the borrower straight. Initiating the foreclosure process also carries this fear-factor feature. In the residential context, it is estimated that most foreclosures initiated do not end up in liquidations. Before the end of the process from start to liquidation, the borrower may find a way to make good on missed payments and arrears (this is called curing the loan), or a short sale (a regular sale of the house by the borrower where the proceeds falls short of the outstanding balance) can take place if both parties agree to it. At the end of the day, this is a lot like a deed in lieu.

5.2 Basic mortgage algebra

The point of view we adopt in this chapter is that of the lender. What sort of mortgage should they be willing to issue to a given borrower? In other words, what return should they require from a particular loan type? To answer those questions, we need to first understand basic mortgage payment algebra. We already covered the main concepts in section 2.3, this section provides a few more details.

As we have discussed, the first two features of a mortgage are its maturity (the maximum period of time over which the lender will receive payments from the borrower) and the frequency of payments. Typical maturities range from 1 (for construction loans, mainly) to 30 years, and the typical frequency of payments is monthly. Both elements give us the maximum number of payments the lender may receive over the life of the contract, which we will denote by T. The other three objects of importance are the initial loan balance b_0, the interest rate schedule $\{r_t\}_{t=1}^T$ which for each period $t = \{1, 2, 3, \ldots T\}$ of the contract gives the rate of interest charged to the borrower, and finally, a payment schedule $\{m_t\}_{t=1}^T$ which lists cash transfers from the borrower to the lender in each period in which the contract is active. That's it, we have a mortgage.

Mortgage algebra defines for each period the interest owed and the loan balance at the start and at the end of the period. Let b_t be the loan balance after payment number t has been made. As we discussed section 2.3, the interest payment at date t is $b_{t-1}r_t$ while the balance at the end of date t, or, equivalently, at the start of date $t+1$, is:

$$b_t = b_{t-1}(1 + r_t) - m_t.$$

A mortgage is called fully amortizing if it is paid in full when the last payment has been made, i.e. $b_T = 0$. Any outstanding balance at the end of date T is called a balloon payment and is due immediately. Fully amortizing vs. balloon is a bit of an arbitrary delineation, obviously. We could change the last payment to

$$\tilde{m}_T = m_T + b_T$$

and turn a balloon mortgage into a fully amortizing mortgage. That language makes sense however in the context of specific mortgages that feature fairly smooth payments over the life of the contract. A balloon payment, basically, refers to an unusually high payment.

Equipped with all this language, we can start designing mortgages. Section 2.3 already described FRMs and IOMs, so let's simply discuss or revisit here a few common mortgages for the sake of practice. A constant amortization mortgage is a mortgage such that

$$b_t = b_0 - \frac{t}{T} b_0,$$

i.e. such the same proportion of the principal is paid in each period. Given a rate schedule $\{r_t\}_{t=1}^T$, the payment at date t has to be

$$m_t = b_{t-1} r_t + \frac{b_0}{T}.$$

The idea in mortgage design is always the same: name the feature you want, then reverse-engineer the payments that deliver that feature.

A constant payment mortgage is such that $m_t = m$ for all t. FRMs are one type of CPM, but CPMs can be constructed even when rates vary over time. Assuming for simplicity a fixed yield $r_t = r$ for all t and full amortization leads to the standard FRM algebra we already went through once. Specifically, for all t,

$$b_{t+1} = b_t(1 + r) - m,$$

which, solving forward from b_0 (do it) implies:

$$b_T = b_0(1 + r)^T - m(1 + (1 + r) + \dots (1 + r)^{T-1}) = b_0(1 + r)^T - m\frac{(1 + r)^T - 1}{r}.$$

Since we must have $b_T = 0$, the payment in each period has to be:

$$m = b_0 \frac{(1+r)^T r}{(1+r)^T - 1}.$$

A CPM with varying contract rates can be built similarly, at the cost of making notation messy.

For more fun algebra along those lines, see GM 17. But we have more important things to do. As we discussed in section 2.3, the yield to maturity (YTM) is the mortgage's internal rate of return when and only when all payments are made as scheduled. As we also discuss there, the yield is not the contract's IRR. First, the borrower may choose to make early payments, or default on part of the contract they owe. These possibilities are in fact so frequent and important that they deserve a section of they own. We will conclude this section by discussing reasons why a mortgage's contract rate may not be its YTM.

We've already shown that for plain-vanilla FRMs, the contract rate is in fact the YTM. Some FRMs, however, feature cash transfers from the lender to the borrower at origination, usually quoted as a percentage of the initial balance, and called *points*. A one point fee means an payment at origination worth 1% of b_0. This adds a new cash flow to the computation of the loan's YTM, hence makes it higher than the contract rate. Points serve a number of purposes as we will see. Basically, imposing a one time fixed cost on the borrower reduces any gains associated with shortening the life of the loan.

As for mortgages with adjustable rates (ARMs), the YTM coincides with the initial contract rate if the initial contract rate is also the expected contract rate in all future periods. Garden-variety ARMs typically feature contract rates that are a some money market interest rate plus a premium or margin. Sometimes, the initial premium is low to entice borrowers with low early payments. If for that reason or any other we expect the rate to rise over time, the YTM should be computed under these expected rates rather than the initial rate. In practice, the YTM for ARMs where the margin does not rise with time is computed just like it would be for an FRM assuming that the initial rate will remain the same over the life of the loan. More careful calculations are made when and only when initial indexation is partial.

5.3 Default and Prepayment risk

The main risk facing the lender is that payments on the mortgage will not be made as planned. The borrower may default on one payment or more, leading in some cases to foreclosure, i.e. the seizure of the asset by the lender. In that case, the lender may not recover all of the outstanding principal. Furthermore, even in the rare case when they do somehow recover their money, they may not be able to reinvest the funds at the mortgage contract rate, causing flows to be lower than planned. The borrower may also choose to prepay the loan. Even when pre-payment penalties are steep as they tend to be on commercial loans, market rates may fall so much that the borrower still chooses to close the loan early, among other scenarios. Since prepayment is particularly attractive when market interest rates are low, these causes payments once again to be lower than planned.

Note that from the point of view of the borrower, prepayment and default are options (see Elul, 2006, for a very insightful article on this idea.) Take prepayment first. If market rates fall below the contract rate, the present value of the mortgage (its market value) rises above the loan balance. The borrower can essentially buy back its debt obligation at the loan face value. The prepayment option, therefore is a call option with strike price the loan balance and where the asset price is the market value of all future payments associated with the mortgage. The borrower is initially long in the underlying asset, full prepayment enables them to go short in that same asset and offset their long position. Default, for its part, is obviously a put option. The borrower sells the asset to the lender for the mortgage balance. If the market value of the asset has fallen below the loan balance (if the borrower has *negative equity*) the borrower gains the difference, at the expense of the lender.

Pricing these options is difficult for many reasons, but in particular because they are American options rather than European options. As we discussed in chapter 3, the former are much tougher to price than the latter, and require advanced mathematical techniques. Even though going into those details is well beyond the scope of this class, thinking of these objects as options is useful even on a conceptual level. As we have discussed, options are particularly valuable when the underlying asset's price is volatile. Anything that makes the property's market value volatile (turbulent economic times, say) makes default a valuable option, and should imply higher contract rates. When interest rates are volatile (following the 1970's oil shocks for instance or when the monetary policy regime becomes more difficult to assess), the market value of the mortgage is likely to vary a lot, making the prepayment option highly valuable which, once again should be reflected in higher contract rates.

We also know that American options are not usually exercised as soon as they are in the money. Exercising the option means killing the option to exercise it later, when it could be even more valuable. In the default context, we should not expect all borrowers with negative equity to default. This prediction is borne out by the available evidence.[4] In the prepayment context, we should not expect borrowers to prepay or refinance (see next section) the minute market rates fall below contract rates. Of course, especially in the residential context, borrowers with negative equity have other incentives not to default: keeping their home, preserving their credit history ... Likewise, in the prepayment context, the interest rate gain must be sufficiently high to cover transaction costs. The option aspect of these decisions adds another reason for borrowers not to be too quick on the trigger.

While pricing all these options is of independent interest, all that lenders need in practice is a forecast of the likelihood of default and/or prepayment at origination, as well as the likely cost associate with these events. Lenders need to know for instance the likelihood of foreclosure at each possible age of the contract, and the cost associate with default, namely the fraction of principal they expect to collect upon foreclosures. Fortunately, empirical estimates of these objects based on historical data are readily available.

5.4 Refinancing

One of the main reason to prepay a mortgage is to take advantage of lower rates. Consider a mortgagor whose mortgage $\{b_0, T, \{r_t^{OLD}, m_t^{OLD}\}_{t=1}^T\}$ is in period k of its life, so that k payments have been made, while $T-k$ payments remain to be made. The superscript "OLD" refers to the fact that the mortgage was originated at some point in the past. Assume that mortgages with maturity $T-k$, initial balance b_k, and terms $\{r_i^{NEW}, m_i^{NEW}\}_{i=1}^{T-k}$ are currently available. Switching to this new mortgage costs C in transaction costs.

[4]See Gerardi et. al. 2009 for a discussion of the evidence in this respect.

Note that typically C will depend on the current size of the loan, although not perfectly since it captures fixed costs such as administrative expenses associated with drafting a new mortgage. This fixed component of C is a big reason why professionals tell you that you should not refinance on a loan that you expect to kill soon because you are planning to move or because your are close to maturity, or a loan with a small balance. Short-life and/or small loans give you less scope to recoup the fixed portion of the refinancing cost.

To make our life simpler, let us assume that $r_i^{NEW} = r^{NEW}$ for all i. Dealing with the more general case is not difficult conceptually, but it makes notation messier. Refinancing amounts to replacing the remaining part of the old string of payments $\{m_t^{OLD}\}_{t=k+1}^{T}$ with the new one available on the market, namely $\{m_i^{NEW}\}_{i=1}^{T-k}$. The net cash flows associated with this swap are $\{m_{k+i}^{OLD} - m_i^{NEW}\}_{i=1}^{T-k}$. If the present value of this string of cash flows is greater than transaction costs, refinancing is *potentially* a positive operation in net present value terms.

There are two complications to deal with. First, what discount rate should we apply to $\{m_{k+i}^{OLD} - m_i^{NEW}\}_{i=1}^{T-k}$? Common practice is to use r^{NEW} the current yield available on the market for loans of maturity $T-k$ and initial balance b_k. The idea is that r^{NEW} reflects what investments of maximum maturity $T-k$ are yielding, hence is the relevant opportunity costs of capital. There are a few issues with this logic, but it'll do for our purposes.

Now, observe that:

$$\begin{aligned} PV(\{m_{k+i}^{OLD} - m_i^{NEW}\}_{i=1}^{T-k}) &= \sum_{i=1}^{T-k} \frac{m_{k+i}^{OLD} - m_i^{NEW}}{(1+r^{NEW})^i} \\ &= \sum_{i=1}^{T-k} \frac{m_{k+i}^{OLD}}{(1+r^{NEW})^i} - \sum_{i=1}^{T-k} \frac{m_i^{NEW}}{(1+r^{NEW})^i} \\ &= \sum_{i=1}^{T-k} \frac{m_{k+i}^{OLD}}{(1+r^{NEW})^i} - b_k. \end{aligned}$$

The last part of the algebra follows from the fact that r^{NEW} is the new loan's IRR when all payments are made as planned.

So the rule now becomes simple: refinancing is *potentially* a positive operation in net present value terms if the present value of old payments at currently offered rates exceeds the current loan balance plus transaction costs.

What is up with this nagging "*potentially*" we've been dragging around this entire section? The problem is that the calculations we have carried out so far ignore the option value associated with refinancing. In a nutshell, what if rates drop a bit further in the next period? One could say that we could just refinance again if that is the case, but that would imply bearing the transaction costs yet again, and in fact, may no longer be valuable at all if the incremental gain is too small to cover C. Obviously, the option of waiting one more period to refinance has value. Therefore, killing that option comes at cost, so that $PV(\{m_{k+i}^{OLD} - m_i^{NEW}\}_{i=1}^{T-k}) - C$ overstates the NPV of refinancing.

Calculating the true NPV requires calculating the value of that American option, and we know that to be quite difficult. Without getting into details, what we have learned already tells us a few important things. First, obviously, one should not refinance the minute $PV(\{m_{k+i}^{OLD} - m_i^{NEW}\}_{i=1}^{T-k}) > C$. One should wait until there is a positive gap between net benefits and costs. How large a gap? It depends on the value of the option to wait. We know that the option will be most valuable when the value of refinancing in the future is uncertain, which is the case for instance in volatile environments.

Things become even more complicated when you recall that mortgages also incorporate a default put option, and recognize the fact that the value of this other option may be affected by the refinancing decisions. Lower payments reduce incentives to default, for one thing. The resulting problem becomes complicated enough that it is best left to our friends the *quants* to figure out. Professionals typically use the basic NPV calculations we have outlined in this section as a first pass, and then add some rule-of-thumb threshold to make a final decision. More sophisticated institutions rely on computer algorithms which enable one to formulate hypothesis on the behavior of future rates, and refine the final decision.

5.5 The underwriting process

See power point notes.

5.6 Securitization

See power point notes.

5.7 The foreclosure crisis

See power point notes, chapter 1.

Chapter 6

Advanced topics

6.1 Real options approach

6.2 Monte Carlo simulations

6.3 Mortgage design puzzles

Chapter 7

Bibliography

Adelino, M., Gerardi, K., and Willen, P. S., 2009, “Why Don’t Lenders Renegotiate More Home Mortgages? Redefaults, Self-Cures, and Securitization,” Federal Reserve Bank of Atlanta Working Paper 2009-17.

Corbae, D. and Quintin, E., 2009, “Mortgage Innovation and the Foreclosure Boom”, mimeo.

Cox, J. C., S. A. Ross, and Rubinstein, M., 1979, “Option Pricing: A Simplified Approach,” *Journal of Financial Economics* 7: 229-263.

Elul, R., 2006, “Residential Mortgage Default,” Federal Reserve Bank of Philadelphia Business Review, Third Quarter.

Gerardi, K., Lehnert, A., Sherland, S. M., and Willen, P. S.“Making Sense of the Subprime Crisis,” Federal Reserve Bank of Atlanta Working Paper 2009-2.

Ghent, A. C. and Kudlyak, M., 2009, “Recourse and residential mortgage default: theory and evidence from U.S. states,” Federal Reserve Bank of Richmond working paper 09-10.

Miles, J. A. and Ezzel, J. R., 1980, “The Weighted Average Cost of Capital, Perfect Capital Markets, and Project Life: A Clarification,” *Journal of Financial and Quantitative Analysis*, 15: 719-730.

Modigliani, F. and Miller, M., 1958, “The Cost of Capital, Corporation Finance and the Theory of Investment,” *American Economic Review* 48: 261297.