NORTHWESTERN UNIVERSITY
J.L. KELLOGG GRADUATE SCHOOL OF MANAGEMENT

Tim Thompson
Finance D42
Winter, 1995

Cost of Capital Notes: Teaching Note

0. Introduction

The purpose of this teaching note is to discuss the estimation of the weighted average cost of capital for companies, divisions and projects. The cost of capital for an investment is an opportunity cost: it is the expected rate of return that investors in a project could earn in the capital market on other investments of similar risk.[1] Companies have many uses for such an opportunity cost of capital: 1. using a project's cost of capital as a *hurdle rate* to determine whether to make an investment in the project (i.e., using the cost of capital to calculate the NPV of a project); 2. using the cost of capital of a division or a subsidiary to determine whether the unit should be sold by the parent company; 3. using the cost of capital (for a project, division, or entire corporation) as a benchmark for performance measurement; 4. using the cost of capital for the entire corporation to calculate the value of the firm, which can be used to estimate a "true value" for the company's stock to compare with the "stock market" value of the company's stock, the purpose being to assess the attractiveness of repurchasing (or issuing) shares of the company's stock and to assess the vulnerability of the company to hostile attack by another company.

---

[1] Similar investments are investments with similar maturity (or duration, or if you are really into buzzwords, tenor), similar tax status to the holder, similar liquidity, and perhaps many other dimensions, not only similar risk. But we get the most mileage out of risk.

Consider an investment which generates a perpetuity of expected pretax operating cash flow of E(OI) per year. The investment's returns are taxed (at corporate level) at constant rate τ_c; interest expense is tax deductible (at the corporate level). The investment is assumed to be financed with debt and equity with market values D and E, respectively, and, by the market-value balance sheet constraint, we know that V, the market value of the investment, is equal to D + E. The capital structure separates the annual cash flow stream into a relatively low-risk stream of cash flows to debt holders and a relatively high-risk stream of cash flows to equity holders. The debt is perpetual debt with annual interest given by Int, so the cash flow to equity holders becomes $(E(OI)-Int)(1-\tau_c)$. The required rate of return on the debt from the holder's perspective is r_d; the required rate of return on the equity (cost of equity capital) is r_{eL}, where the L denotes the fact that this is equity in a levered firm.

We could define the project's hurdle rate as the rate of return (on the assets) at which the project merely returns enough cash flow (including the tax shield on debt interest) to pay both its debtholders and its equityholders their required dollar rates of return, respectively. If the project's expected return equals the required return on the project (hurdle rate), we have:

$$(E(OI)(1-\tau_c) + \tau_c Int = r_d D + r_{eL} E \qquad (1)$$

Dividing both sides of (1) by V, bringing the tax shield term to the right hand side, and recognizing that V = D+E, we define the project hurdle rate by r_a:

$$r_a = \frac{E(OI)(1-\tau_c)}{V} = r_d(1 - \tau_c)\left(\frac{D}{D+E}\right) + r_{eL}\left(\frac{E}{D+E}\right) \qquad \mathbf{(2)}$$

The far right hand side of (2) is the common definition of the weighted average cost of capital. Therefore, the weighted average cost of capital equals the hurdle rate on the project.

Each of the uses of the weighted average cost of capital discussed in the opening paragraph of this section involves *calculating the present value of a set of future expected cash flows at a discount rate given by the WACC.* Using the weighted average cost of capital as a discount rate to be applied to future cash flows depends on several assumptions. To put this issue in perspective, let's assume that markets are perfect: there are no taxes, corporate or personal, and there are no transactions costs, information costs, agency costs, etc. In this case the Modigliani-Miller (MM) theory holds and the cost of capital for an investment is not a function of how that asset is financed. In practice all you would need to know about the project is its unlevered cost of equity (its unlevered beta). The project cost of capital (assuming the Capital Asset Pricing Model (CAPM) is how we estimate the cost of equity) would be $r_a = r_{eu} = r_f + \beta_{eu}(r_m - r_f)$. Assuming an 8% risk free rate and an 8.5% risk premium on the market portfolio (in excess of risk frees) and an unlevered beta of 1.0 yields $r_{eu} = 16.5\%$.

This would be the project hurdle rate even if the project were financed with both debt and equity! Suppose the firm could finance up to 20% of the value of the project with risk-free debt and the rest with levered equity. Assuming the debt is riskless and the no-tax environment, the levered equity beta would be given by (See chapter 9 of Brealey and Myers) $\beta_{eL} = \beta_{eu}(1+(D/E)) = 1.0(1+(.2/.8)) = 1.25$, which yields (via CAPM equation) $r_{eL} = 18.725\%$. Combining this levered cost of equity with 8% (risk-free) cost of debt in a ratio of 80/20 yields a WACC (remember, no taxes in this example) of WACC = $.2(8\%) + .8(18.625\%) = 16.5\%$.[2] So in perfect markets, the cost of capital is independent of capital structure (as per MM).

---

[2] This example goes through if the debt is risky, but the numbers are harder to come up with, so I stuck with the easier case here.

When markets are not perfect, this is no longer true. In fact, the concept of valuation by DCF methods may not even be well defined given sufficient market imperfections. We will assume that the concept of DCF is applicable, but we incorporate that there are tax (both corporate and personal) benefits/costs (and perhaps other quantifiable benefits/costs associated with the financial structure of investments) in the real world. When there are tax benefits associated with the financing of assets, it is no longer the case that the investment and financing decisions are independent and the value of assets is a function not only of their operating value, but also of the net present value of the financing side-effects on value. There are two widely- used frameworks for incorporating financing side-benefits/costs into the valuation of assets: namely, the Adjusted Present Value (APV) and Adjusted Discount Rate (ADR) frameworks.[3]

The APV method asserts that the value of an investment (financed in a particular way) is the value of the investment as if it were unlevered *plus* the net present value of all financing side benefits:

$$APV = V_U + NPV(finacing\ \ side\ \ benefits)$$

where V_U is the present value of all the future cash flows of the investment discounted at the unlevered cost of equity, i.e.,[4]

---

[3] The reader is referred to Brealey and Myers, Chapter 19 for an introduction to the APV and ADR approaches. The interested reader is referred also to Thompson's teaching note, *Teaching Note: Valuation using the adjusted present value (APV) method vs. the adjusted discount rate (ADR) method (Theory)*. Also of interest may be HBS *Note on Adjusted Present Value*, HBS #9-293-092.

[4] UCF refers to the future cash flows being unlevered cash flows. To be more precise, I mean *operating free cash flow*. In the terminology I employ, "operating" is to be distinguished from "financing." Interest expenses, principal repayments, dividend payments are not included. Operating free cash flow is "free" in the sense that it is after all planned reinvestment in capital expenditures and required working capital investments. This terminology is consistent with capital budgeting treatment (see BM, Chapter 6). In other notes, I refer to this simply as free cash flow from operations and refer to it as FCF. In this context, the U helps the reader remember that there are no financing cash flows in this number.

$$V_U = \sum_{t=1}^{\infty} \frac{UCF_t}{(1 + r_{eu})^t} \tag{3}$$

The financing-side benefits include, but are not limited to, the tax shields afforded the asset by its ability to support debt in its capital structure: its debt capacity. The tax shields on the debt capacity of the asset in question can be forecast and discounted to the present and *added* to the unlevered value (along with other costs and benefits of financing).

The alternative approach to the APV method is to incorporate the tax benefits of debt financing (and other costs and benefits of financing, to whatever extent possible) into the discount rate (hence, the name adjusted discount rate method) to be applied to the unlevered free cash flow from operations *instead of adding them as in the APV method*. Brealey and Myers discuss two adjusted discount rates, the MM r* and the Miles-Ezzel r*. The adjusted discount rate which is of most practical importance, however, is the weighted average cost of capital. The value of a levered asset using the WACC of the asset to incorporate the tax benefits of leverage into the value of the project would be calculated as[5]

$$V_L = \sum_{t=1}^{\infty} \frac{UCF_t}{(1 + WACC)^t} \tag{4}$$

Including the tax shields of financing with debt into the discount rate is intuitively appealing because the fact that interest payments are tax deductible reduces the cost of using debt capital to the

---

[5] Equations (3) and (4) assume that the discount rate used to discount future cash flows is the same for each t, i.e., we are massaging the term structure of interest rates. We will continue to play this silly charade throughout this teaching note, but we will address the issue later in, hopefully, a less oblique fashion.

company. We have to be careful, however, about the assumptions behind using a formula like (4) to value levered assets: in particular, the WACC formula assumes that the debt to debt plus equity ratio (debt to capital ratio) will remain constant in the future --- that the company follows a target debt ratio policy.

This teaching note deals with the *estimation* of the WACC for valuation of firms, divisions, projects, etc. Repeating the definition of WACC:

$$WACC = r_d(1 - \tau_c)\left(\frac{D}{D+E}\right) + r_{eL}\left(\frac{E}{D+E}\right) \tag{5}$$

To estimate the WACC of an investment, we must estimate its intended capital structure weights, D/(D+E) and E/(D+E), the required return on the equity and the debt of an investment in this project as a stand alone investment (or mini-firm) with the implied capital structure, and, lastly, the marginal corporate tax rate. The following sections take up each of these parameters in turn. Section 1 deals with issues involved in estimating the cost of equity via the capital asset pricing model and contains the bulk of the rest of this teaching note. Section 2 deals with estimating betas for divisional/project WACC estimation via the peer group method. Section 3 discusses the after-tax cost of debt. Section 4 discusses the capital structure weights to use in the weighted average cost of capital. Section 5 concludes.

1. Estimating the cost of equity capital with the CAPM

The capital asset pricing model is not the only available method for estimating the required rate of return on (or cost of) equity capital. A more intuitively appealing theory called the arbitrage pricing theory (APT) is starting to be used to a limited degree in corporate finance (and to a much larger degree in investments and portfolio management). One of the reasons we use the CAPM is because it is much more tractable, easy to use, framework, and the parameters are familiar to practitioners.

One drawback of using a model like the CAPM or APT is that we will have to use statistics estimated from historical data to make estimates of future expected risk premiums and future betas, etc. This is almost impossible to avoid in that the future betas and expected risk premiums are not observable. An alternative, however, is to use the current stock price and some kind of *market estimates of future dividends* and use a dividend discount model (such as the Gordon growth model, see Chapter 4 of BM) to calculate the market's expected return on the stock in question. If we have good estimates of the expected dividend over the next year, DIV_1, and level of permanent growth in dividends, g, we can estimate r by the following:

$$r_{eL} = (DIV_1/P_0) + g$$

Estimates of future dividends and growth in dividends are usually taken from analysts that follow the stock, frequently from Value Line. Great care must be taken in the use of a dividend discount model to estimate required return on common stocks. No one, in practice, really estimates perpetual growth rates on stocks, but rather the expected growth rate over next few years. The model is explicit that this growth in dividends should be the expected *permanent* growth rate. Small errors in the estimate of g greatly effect the r estimate. As was mentioned in BM Chapter 4, an analyst should always use a group of similar companies to estimate the parameters of such a model, and the model is likely to perform better for a company with much more predictable dividend stream, such as utility companies. For our purposes, the dividend discount model will not be acceptable in general.

The capital asset pricing model (CAPM) says that the required rate of return on investment i is given by

$$r_i = r_f + \beta_i \, (E(r_m) - r_f) \qquad \textbf{(6)}$$

where r_f is the risk-free rate of interest available to the investor, $E(r_m)$ is the required expected rate of return an investor would demand in order to invest in the "market" portfolio of all risky assets in the economy. $E(r_m)$-r_f, then, is the expected risk premium investors demand to earn to invest in the market portfolio rather than in risk free assets. The market portfolio is defined to have one unit of systematic (i.e., non-diversifiable, pervasive, general economic) risk, so the expected risk premium, $E(r_m)$-r_f, is the expected risk premium *per unit risk borne by the investor*. The beta of investment i, β_i, is the number of units of systematic risk inherent in investment i, so $\beta_i(E(r_m)$-$r_f)$ is the number of units of systematic risk of investment i times the expected risk premium per unit of risk, or, equivalently, the total risk premium that investors require to invest in investment i. Therefore, (6) says that the investor's required rate of return on an investment is given by the risk free rate, an amount the investor requires to part with funds today for later consumption even in a riskless investment, plus a risk premium based on the amount of beta risk of the investment.[6]

To estimate the cost of equity using the capital asset pricing model (i.e., making investment i the equity in question), we must estimate the risk free interest rate, r_f, the required rate of return on the market portfolio, $E(r_m)$ (and the proxy for the market portfolio, for that matter), and the beta risk of the equity. We consider these in turn.

1.a. The risk free rate

---

[6] I realize this is a review of stuff you already know, but if you need more of a review, see Chapters 7 and 8 of BM.

Keep in mind that we are trying to estimate the opportunity cost for investors *today* to invest in the project under consideration, rather than invest his/her funds in similar investments in the competitive capital market. The alternative expected rates of return that investors can earn *today* are prospective, or *forward-looking* expected returns. They have no necessary relation to what past returns on similar securities were. In general, we would prefer forward-looking expected returns to averages of past returns, *all else equal*.[7]

In general, expected future rates of return on securities are not observable. In the case of U.S. Treasury securities, however, the nominal yield to maturity is as close to a risk-free nominal rate of return as one can find. Under almost any possible future course of events, holders will receive the promised payments on U.S. Treasuries, so their ex post nominal return (if held to maturity) will equal their ex ante yield to maturity. Therefore, the yield to maturity on U.S. Treasury securities can be used as a proxy for a risk-free interest rate.

Why not use the long-run average of past Treasury rates rather than today's yield to maturity? People often ask this question. In particular, what if you believe that, in the future, interest rates could be higher than now or lower than now. A long historical average may be a better estimate of what future levels of interest rates will be over the life of an investment.

The last statement could, in fact, be true. But this wouldn't make the historical average of past interest rates the appropriate estimate for today's risk free rate. If current interest rates are higher than "average" or even than they will be in the future, it still reflects the interest rate that investors could earn if, instead of putting their money into your investment, they put their money in a risk-free vehicle. As such, the current Treasury rate correctly reflects the risk-free component of

---

[7] We have to be careful. All else is not always equal: see discussion below on the expected market risk premium.

their required rate of return, not what the average historical interest rate has been. Also, it should be pointed out that, if you are using a longer-term Treasury rate to generate your r_f estimate, that it incorporates *current market expectations* about future expected interest rates.

But doesn't that mean that my cost of capital will be changing with the level of interest rates?

Yeah, so what? The world is volatile; did you miss the news? Get used to it.

OK, tough guy, say I use a current Treasury yield to maturity, what maturity should I use?

This would seem to be a simple question, but it is not. That is why you will find different sources telling you different things about what to do. For example, in the 1994 Yearbook of *Stocks, Bonds, Bills and Inflation*, published by Ibbotson Associates in Chicago, they take a relatively agnostic approach, stating that the "two common choices for the nominal riskless rate are the U.S. Treasury bill yield and the yield on intermediate or long-term U.S. Treasury Bonds." Copeland, Koller and Murrin come down on the side of using the 10-year U.S. Treasury bond rate as their definition of the risk-free rate. Brealey and Myers, on the other hand, suggest that future expected *short-term* T-bill rates should be used as the definition of the risk-free interest rate.[8]

Let me make what I think is the important pronouncement here: whatever type of rate you choose to estimate the risk-free interest rate (the 10-year (or even longer term) U.S. Treasury rate, the future expected one-year T-bill embedded in the long-term Treasury rate, etc.) in the CAPM, you must estimate the expected premium on the market portfolio *in excess of the same type of risk free rate*. That is, you must be consistent in how you define the risk free interest rate and how you define

---

[8] See footnote 14 in Chapter 9 of the fourth edition of BM.

the estimated risk premium on the market portfolio.

That being said let's consider the following. None of the government securities under consideration (and no other securities I am aware of) are absolutely riskless: even under the assumption that the U.S. Treasury will absolutely not default on these obligations (reasonable enough), all have purchasing power risk --- the holder is guaranteed a certain number of dollars back, not a certain amount of consumption that said dollars will afford. Also, if we assume an investor puts his/her funds in U.S. Treasury securities of different maturities and assume that the investor has a particular holding period, then the nominal value of his/her investment at the end of this holding period is not riskless even if his/her money is all invested in short term T-bills. To amplify on this point, assume the investor's holding period is five years and consider the following strategies: 1. he/she can invest in one-year T-bills and roll them over each year until the end of five years; 2. he/she can invest in a ten-year T-Bond, reinvesting the annual coupons in other T-bonds, and at the end of five years, selling the entire portfolio for the market price of the portfolio at that date; 3. he/she can invest in five-year maturity zero-coupon Treasury strips. The first strategy is not riskless, because at the end of one year there is a new interest rate at which to reinvest; if interest rates drop from now until year five, then you earn a lower rate on reinvestments and will end up with less money at the end of year five. Thus the strategy has interest rate risk and the form of interest rate risk is called *reinvestment risk*. The second strategy has interest rate risk also: if interest rates rise, you will be reinvesting the coupons on the ten year bond at now-higher rates (this is the good news) but the price of the ten-year bond will go down (this is the bad news and is referred to as *price risk*). These two factors are partially offsetting and the outcome will depend on the movements of rates between now and year five. One of the worst possible outcomes may be if interest rates decline between now and the day before you sell, so that your reinvestments earn a lower rate of interest and then the day before you sell your bonds, interest rates go way above their level when you bought your ten-year bonds. Price risk and reinvestment risk add up to the total interest rate risk of a particular bond or bond portfolio and will, in general, be a function of the assumed investment horizon. The only nominally riskless investment for a given horizon will be strategy number three: a zero coupon bond maturing on the horizon date, and even then, there will be purchasing power risk.

The moral of this story is that none of these bonds are really riskless: they are at best proxies of the risk-free rate of interest. Maybe more importantly, that the definition of the risk free rate of interest *should depend on the timing of the cash flow being discounted*: i.e., maybe we should not be *massaging the term structure of interest rates* (see footnote 5 above) and we should allow for different r_f for each future cash flow. While this is probably the case, we will discuss the different choices under the heading of massaging the term structure of interest rates, because practitioners seem to be hell bent on being the best masseurs/masseusses on the block. A more subtle issue is that the "period length" in the above discussion which is defining "horizon" is not a function of the investment, but of the investor. The model assumes that investors have a natural horizon of investments which is important to them and the length of the horizon is not specified by the theory.

Method #1. Use the long-term Treasury Bond rate as proxy for r_f

What is "long term"? 10 years, 20 years, 30 years? Copeland, et al, make the following case for the 10 year Treasury rate.

> "First, it is a long-term rate that usually comes close to matching the *duration* of the cash flow being valued. Since the current Treasury bill rate is a short-term rate, it does not match *duration* properly. If we were to use short-term rates, the appropriate choice would be the

> short-term rates that are expected to apply in each future period, not today's short-term interest rate. In essence, the ten-year rate is a geometric weighted average estimate of the *expected* short-term Treasury bill rates (sic) over the evaluation horizon.
>
> Second, the ten year rate approximates the *duration* of the stock market index portfolio --- for example, the S&P 500 --- and its use is therefore consistent with the betas and market risk premiums estimated relative to these market portfolios.
>
> Finally, the ten-year rate is less susceptible to two problems involved in using a longer-term rate, such as the thirty-year Treasury-bond rate. Its price is less sensitive to unexpected changes in inflation, and therefore has a smaller beta than the thirty-year rate; and the liquidity premium built into ten-year rates may be slightly lower than that which is in thirty-year bonds. These are technical details, with a minor impact in normal circumstances. But they do argue in the direction of using a ten-year rate. [Italics added]"

The idea is that most valuation decisions faced by companies (e.g., capital budgeting decisions) involve discounting not only near-term cash flows, but also distant future cash flows. As such, we should define a "long term" discount rate. They argue that the ten-year bond matches the duration of cash flows of most investments better than the short term bill and its duration is closer to the duration of the stock market index. While a detailed discussion of duration models is beyond the scope of the note (some definitions can be found in an Appendix to this note), some issues about duration should be explained here.[9]

The most common definition of the duration of an asset with fixed nominal cash flows is a weighted average of the amount of time to each payment. Let a "bond" pay fixed nominal cash flows C_t at each date t=1,2,... T. Let the yield to maturity on this "bond" be Y. Let $PV(C_t)$ be the present value of cash flow C_t discounted at the yield to maturity of the "bond," Y: i.e., $PV(C_t) = C_t/(1+Y)^t$. The price of the bond, B_0, is the sum of the present values of each of the future cash flows discounted at the yield to maturity: i.e., $B_0 = \Sigma PV(C_t)$. The duration of the bond is given by:

---

[9] The interested reader is directed to BM, Chapter 25, section 2 for a brief discussion of bond duration. Any investments text will discuss this issue in more detail.

$$Duration = \left[\frac{PV(C_1)}{B_0} \times 1\right] + \left[\frac{PV(C_2)}{B_0} \times 2\right] + \ldots + \Bigg[\frac{PV(C_T)}{B_0} \times \qquad \textbf{(7)}$$

The definition of duration is the weighted average of the number of years to each cash flow, with weight on the ith year's cash flow equal to the percentage of the total present value of the "bond's" cash flows represented by the present value of the ith cash flow, where all the present values are calculated using the yield to maturity, Y, as the discount rate.

Duration has an equally important use in bond portfolio management, however: it measures the sensitivity of the price of a bond to movements in its yield to maturity. The duration in equation (7), under certain simplifying assumptions, can be shown to be equal to the percentage *decrease* in bond prices for a one percent *increase* in (1+Y). This implies that, for small decreases (increases) in the yield to maturity, the percentage price increase (decrease) of the "bond" is -D/(1+Y). So the price of a "bond" with a duration of 4 years and a yield to maturity of 11% would go up approximately 3.6% if the yield to maturity on the bond decreased to 10%.[10]

In the second paragraph in the above quote from Copeland, et. al., it seems as if they are referring to the second use of the term duration, i.e., that the duration of the stock index as measured by the sensitivity of the stock market index portfolio to changes in *interest rates* is similar in size to the *duration* of the ten-year Treasury bond. It is not clear that this is of any interest. The concept of duration as a price sensitivity to interest rate changes relies on the fact that the future cash flows are riskless, nominally fixed cash flows. To compare this value on bonds to the average stock price sensitivity to interest rates is not related, necessarily, to the length of time to cash flows on equity portfolios and, as such, the relation to the above argument about "matching" the length of life of the

---

[10] The duration of 4 divided by 1.11 is 3.6.

cash flows with that of the risk free rate chosen is unclear.[11]

The last paragraph from Copeland, et al, argues to use the ten-year bond rate rather than, say, the thirty-year bond rate. There are maturities in between, obviously. They argue that the ten-year bond price is less sensitive to unexpected changes in inflation and that the liquidity premium built into ten-year bond rates is smaller than that in thirty-year bond rates. Unexpected changes in inflation which affect ten- and thirty-year bond prices are those which affect long-term interest rates, so the authors are saying that the *duration* of the thirty year bond is longer than the ten-year bond. This is not really news and the graph on the following page shows that the difference in duration is not as large as you would imagine. The graph shows the duration on the y-axis as a function of the maturity on the x-axis of a coupon bond (no sinking fund payments, simply annual coupon payments with principal paid at maturity) for three coupon rates (4%, 8% and 12%) with yield to maturity of 8%. The duration of a ten year bond is about 7.2 years and the duration of a 30 year bond is about 12.7 years. There is a difference, but it is less than two times larger, not three times larger, as might be mistakenly assumed by the difference in maturity. The difference between the ten- and thirty-year bonds' durations is larger for the 4% coupon and smaller for the 12% coupon bonds.

It is also true that the liquidity premium for thirty year bonds is larger than for ten year bonds. But if that were a reason to use the 10 year bond, then you should use short term bonds rather than 10 year bonds, because there is a larger liquidity premium for ten year bonds than for short term

---

[11] The fact that stock prices are pervasively negatively related to changes in interest rates is, in itself, an interesting phenomenon. It argues that interest rate changes are a source of pervasive, general economic risk, which is probably not diversifiable. As such, *if this risk commands a risk premium*, it should be included in our estimate of the cost of capital of any investment based on the sensitivity of the investment to interest rate changes. Different measures closely related to interest rate changes such as the expected inflation rate, the slope of the yield curve, and others have shown up as important macroeconomic factors in some versions of the Arbitrage Pricing Theory.
See, for example, Chen, Roll and Ross (1986), and description of Alcar's APT! product in BM, Chapter 9. As we said earlier, treatment of APT is beyond the scope of this note.

bonds. Ibbotson Associates reports the yields to maturity on a 20-year constant maturity Treasury bond portfolio (Long Term Government Bond Yield) and on a 5-year constant maturity Treasury bond portfolio (Intermediate-term Government Bond Yield). The average spread between the Intermediate-term and the Long-term Yields (1926-1993, Arithmetic average) is .49% or 49 basis points.

To tie some of this argument together, we are trying to find a Treasury security to use as a proxy for a risk free rate. Not only are the bonds not truly riskless but, if investors have a specific planning horizon over which they invest, not everyone would agree on which bond is less or more risky. The required returns (the yields to maturity) on the bonds give an indication as to which is considered more/less risky. The difference between the yield to maturity on long-term T-bonds and short-term T-bills is, by convention, called the *liquidity premium*. But this is a risk premium just like any other. Longer term Treasury bonds have, on average, higher yields to maturity because of the risk of price changes associated with longer term bond holdings. This would indicate that, on average, investors view the short term T-bill as the least risky investment.

If the most important consideration is the *timing* of the cash flow, then the risk free rate should be defined as the rate of return on a strip Treasury security which pays its face value at the time of the cash flow and has no other cash flows. Therefore, we would have a different risk free rate for each cash flow's associated date, r_{ft}, t=1,...,T. If we define a different r_f for each cash flow date then we must have a different market risk premium for each date (i.e., the premium on stocks over one-year T-bills for t=1; the premium on stocks over 5 year T-bonds for t=5; etc.). But if we are going to massage the term structure of interest rates question, we must ask what the perspective of the investor is: do investors believe that a medium- to long- term T-bond reflects the *lowest risk* *alternative* to risky investments of similar length? If so, we can use the intermediate- or long-term T-bond rate for r_f as long as we remember to define the market risk premium as the return on stocks over the same maturity T-bond.

Method #2. Estimate the future expected short term T-bill rate embedded in the long-term T-bond rate.

On the other hand, if the market believes that the least risk alternative to investing (even for the long term) is to invest in short-term T-bills (and rolling them over until future dates) then the definition of the risk free alternative to investing in a risky project paying a cash flow at date t is investing in a series of one year T-bills and rolling them over until date t. The current short term (one year) T-bill rate is known, but future T-bill rates are uncertain. Copeland, et al, claim that (first paragraph in the above quote) that the long-term rate is a geometric weighted average estimate of the *expected* short term rates over the evaluation horizon. This is not true. The long term rate is the geometric average estimate of the *forward rates* over the evaluation horizon. If forward rates were equal to expected future spot rates, then there would be no liquidity premium. Therefore, what we should use for our estimate of the *expected future short term risk free rate* is today's long-term T-bond rate *minus* the expected liquidity premium built into the long-term T-bond rate:

$$\begin{aligned} \text{Expected future short-term } r_f &= \text{Long-term T-bond rate} \\ &\quad - \text{Expected liquidity premium} \end{aligned} \tag{8}$$

How do we estimate the liquidity premium that is built into long-term T-bonds over T-bills. Doing this systematically is beyond the scope of this course, but we could use a back-of-the-envelope measure by asking the question: on average in any given year, how much more do I earn on an investment in Long-term government bonds than in an investment in Short-term government bills? From 1926-1987 (Marriott case, FKMPR, pg. 450), long-term government bonds earned an average of 4.58% and one-year T-bills earned 3.54%; this *premium* of 1.04% could be viewed as the average premium that holders of long-term government bonds require over T-bills. Therefore, we could estimate the future expected short-term T-bill rates over the life of an investment to be given by:

$$\begin{aligned} Exp.\ future\ r_f &= Long-term\ T-bond\ rate \\ &\quad - Avg.\ premium\ on\ long-terms\ over\ short- \end{aligned} \tag{9}$$

For example, if the twenty-year T-bond rate is 7%, our estimate of the future expected short-term risk free rate is 7% - 1.04%, or about 6%.

The crucial thing to remember is that, if you are using method #2 to estimate the risk-free rate, the risk free rate you are estimating is a *short-term* risk free rate. Therefore, when you estimate the market risk premium, you must define it as the risk premium in excess of the short-term risk free rate.

1b. The expected risk premium on the market portfolio

The next parameter in equation (6) to estimate is $E(r_m)$, the expected return on the market portfolio.[12] Based on our earlier discussion, we would prefer a *forward-looking* estimate of $E(r_m)$. The true expected value of the future return on the market index, however, is unobservable. We could envision polling market participants and calculating some "consensus" estimate. Such polls are

---

[12] We could discuss what proxy for the market portfolio should be used and pros and cons of each, but we will simply assume that a broad-based market value-weighted index of common stocks such as the S&P 500 portfolio is a good proxy for the market portfolio to be used in the CAPM. In the following discussion, any reference to the market portfolio should be understood to be a reference to the proxy for the market portfolio.

made of economists' forecasts of many macroeconomic series and empirical tests are conducted on the statistical properties of the forecasts. Similarly, managers' and analysts' forecasts of earnings and other accounting variables are considered very valuable forecasts of these quantities. There are very few systematic forecasts concerning the future expected return on the market index and, if there were any, the forecasts would probably be too short-term for our purposes.

Some sources do calculate a "forecast" market risk premium of sorts. In the same flavor as using the yield to maturity on a Government bond to estimate the risk free rate, if we make a set of assumptions about the future dividends on the market index portfolio, we could use the current value of the index and the fact that the portfolio value is the present value of all the future dividends on the portfolio to calculate the *implied discount rate* which equates the present value of the future dividends to the current index value. For example, if we use the S&P 500 as our market index portfolio and we estimate the future dividends on the index portfolio as follows: a unit of the index would pay an expected dividend next year of DIV_1 and the dividends on the portfolio would grow into the indefinite future at an annual growth rate of g. The assumptions of the dividend growth model (Chapter 4, BM) would then hold and we could calculate the value of r which equates the present value of the infinite stream of dividends to the current value of the index, which we will call $Index_0$:

$$r = \frac{DIV_1}{Index_0} + g$$

Many investment banking companies estimate required returns for index portfolios using a variant of the above procedure. They may be estimating future growth in earnings; they may be using explicit forecasts for a few years and using a perpetuity (or growth perpetuity) method at the end of a few years.

However these methods are developed in practice, they have a large drawback: the assumption in the model that the market is pricing some forecasted dividend level in perpetuity (with growth usually). In practice, dividend amounts, dividend payout ratios, dividend yields and dividend growth rates are forecasted for many firms and for many portfolios, but the forecasts usually go out at most only few years (see Value Line forecasts for example on any Value Line tearsheet). In a perpetuity model, some level of dividends (at its attendant growth rate) has to be estimated to continue *forever*. This is quite important and whatever r is estimated for the portfolio is going to be subject to the assumptions of this perpetual dividend growth rate in any case. These perpetuity assumptions are particularly important and can make huge valuation differences. It is also the case that the user does not know the exact assumptions that the investment banking company is using in their model. It is not that the alternative methods available are without assumptions, but I am more comfortable with those assumptions and at least I know what those assumptions are.

Without a forward-looking estimate of $E(r_m)$, we must estimate it using historical data. If the return on the market portfolio in past time periods were considered *independent* drawings from the same probability distribution (or different probability distribution with the same mean return), we could simply use the time-series average of the return on the market portfolio: per the Marriott Cost of Capital case, the arithmetic average of the return on the S&P 500 was 12.01% per year from 1926-1987.

The most glaring problem with using this as our estimate of $E(r_m)$ is that the return on the market portfolio would not be *expected* to have a constant mean. When interest rates are very small (say 3%) should we expect $E(r_m)$ to be the same as when interest rates are very high (say 12%). Of course not. The level of interest rates (taken from the yield to maturity on a long-term Treasury bond) will incorporate the future inflation rate that investors expect will occur during the life of the bond in question. The expected return on the market portfolio will contain a "premium" for the level of expected inflation as well. We would expect higher $E(r_m)$ when interest rates are high and lower $E(r_m)$ when interest rates are low.

If we are willing to assume that, while the level of interest rates will change from period to period, investors require a **constant risk premium** over the risk free rate for the market index portfolio (i.e., for a beta risk of one), then we could calculate the historical average premium on the market portfolio in excess of the risk free rate, $r_{mt}-r_{ft}$, and use this time-series average as our estimate of $E(r_m)-r_f$. Then we could simply add this premium to our r_f estimate to obtain our estimate of $E(r_m)$.[13] We still have choices to make, however. What length of time series should we use over which to calculate the average market risk premium? Should we use an arithmetic average of past market risk premia or a geometric average? We take these issues up in turn.

One thing should be clear: if we are using a "short-term" definition of the risk free rate (e.g., Method #2 in section 1a above), then we must estimate the average return on the market index *in excess of the T-bill rate*; if we are, on the other hand, using a "long-term definition of the risk free rate (e.g., Method #1 in section 1a above), we must estimate the average return on the market index *in excess of the rate on a similar maturity T-bond*.

What length of time series should be used to estimate $E(r_m)-r_f$? This is a trade off. Ibbotson

---

[13] We are assuming here that the historical risk premia, $r_{mt}-r_{ft}$, are drawn from *independently* identically distributed probability distributions. There is evidence that stock returns are not iid; in particular that there is long-term autocorrelation in returns. We are abstracting from this issue here. We will mention this problem again below.

Associates provides data on the S&P500 and some Treasury series going back as far as 1926. Many people think that data from the more distant past is not relevant, because "things have changed." On the other hand, a quick look at the time series of annual returns on the S&P 500 (see the graph on the next page) gives you a feeling for how noisy this time series is. Trying to forecast the mean of a very noisy time series is like trying to tune in an AM radio station from a distant city: there is usually a lot of noise relative to the signal. This means that the more data we use, the better our forecast should be. On this tradeoff, we fall on the side of using more data and, to some extent believe that "the more things change, the more they stay the same."

What type of "averaging" process should be used: arithmetic or geometric? We prefer arithmetic averages to geometric averages; but, again, Copeland, Kollar and Murrin disagree. The simple argument purporting to use the geometric averaging goes as follows: you buy a non-dividend paying stock for $100 and hold it for two years. In the first year, the stock goes down to $50 (a -50% rate of return), the next year it returns to $100 (a +100% rate of return. The arithmetic average rate of return on the stock is 25% (i.e., (-50%+100%)/2). The geometric average rate of return gives you the actual holding period rate of return: $[(1-.5)(1+1.0)]^{1/2}-1 = 0\%$ rate of return. The geometric average gives you the correct measure of investment performance, but the incorrect measure of the **expected value** of future returns under some circumstances.

To see that the arithmetic average is the correct measure of expected returns in a discrete-time case, assume that there are two periods and that the two observed returns are the only two possible returns in each period and that they are equally likely in each period. Given this, the possible terminal stock prices at the end of two periods can be shown in a binomial diagram as follows:

Time 0	Time 1	Time 2
		400
	200	
100		100
	50	
		25

The probability distribution of the stock price at Time 2 is:

Stock price at Time two, S_2	Probability that stock price=S_2	Weighted stock price, S_2xprob
400	.25	100.00
100	.50	50.00
25	.25	6.25
	Expected stock price, Time 2	156.25

The expected value of the two-year rate of return is, then $(156.25/100)^{1/2}$ - 1, or 25% per year. So, if the stock price ends up at $100 after two years, your holding period rate of return was 0%, but from an ex ante perspective, the expected value of the holding period rate of return was 25% (per year). So, if each year is an independent, identically distributed rate of return on the stock with the above probability distribution, the best estimate of the future expected rate of return is the sample mean of the rates of return, or 25%.

To CKM's credit, they do recognize this argument:

> "The difference between the arithmetic and geometric averages is that the former infers expected returns by assuming independence, and the latter treats the observed historical path as the single best estimate of the future...

> The arithmetic average is biased by the measurement period...[because it] depends on the length of interval chosen. For example, an average of monthly returns will be higher than an average of annual returns. The geometric average, being a single estimate for the entire interval, is independent to the choice of interval...
>
> Finally, empirical research by Fama and French (1988), Lo and MacKinlay (1988) and Poterba and Summers (1988) indicates that a significant long-term negative autocorrelation exists in stock returns. Hence, historical observations are not independent draws from a stationary distribution."

To clarify some of these points, remember that we are using the past history of stock returns as a sample to *estimate* the expected value of a probability distribution. Second, remember that when we calculate the "returns" for each historical period to use in the arithmetic average method, each of these returns is a geometric rate of return. Next, the reason that the arithmetic averaging method is "biased" by the length of interval chosen (this is referring to using monthly versus daily versus annual returns, not to what years are included in the sample) is that the returns are usually calculated as effective per period rates of return.[14] If the periodioc rates of return were calculated as logarithmic rates of return (assuming instead continuously compounded returns), the arithmetic average would be the same as the geometric average.[15]

These points make it clear that the first and third paragraphs of CKM's support for the geometric average method are false. The geometric average *is the arithmetic average* of the

---

[14] For example, if a stock begins the year at $100, pays no dividends during the year and ends up at $150, this would be a 50% rate of return. (150/100)-1 = .5.

[15] To see this, suppose a stock starts out at $100 and earns the following effective rates of return in years one, two and three: 8%, 10%, then 6%. The arithmetic average annual rate of return is clearly 8%. The terminal value of the stock is $100(1.08)(1.10)(1.06) = 125.928. The annualized rate of return (geometric average) = $(1.25928)^{1/3}$ - 1 = 7.9876529% per year. Suppose, instead, we had calculated the continuously compounded rate of return on the stock in each year (rather than the 8%, 10% and 6% which were annually compounded rates): ln(1.08) = .076961041, ln(1.1) = .095310180, ln(1.06) = .058268908. The arithmetic average of the continuously compounded rates of return (sum them up and divide by 3) is .07684710. The effective annual rate associated with this continuously compounded rate is $e^{.07684710}$ - 1 = 7.9876529%, exactly equal to the geometric average rate of return.

continuously compounded rates of return, restated to be an effective annual rate (see the prior footnote). So the geometric average method assumes that the sample arithmetic mean of past continuously compounded rates of return is the best estimate of the future expected value. Hence the geometric average method is assuming independent identically distributed past continuously compounded returns, which invalidates the first and third paragraphs of CKM's critique.[16]

Seemingly the most damning comment that CKM make about the arithmetic average method is that it is biased by the measurement period. They should say that it *depends* on the measurement period used to estimate the expected value. This is true. But the fact is that the method which is correct depends on whether you assume that future cash flows occur at discrete time periods or continuously through time. The common applications of DCF that are used in practice, such as equations (3) or (4) are discrete time representations of future cash flows: we assume that the "annual" cash flow occurs at the end of period t, for t = 1, ..., N. We understand that this is an approximation to the fact that the annual cash flow comes in more or less continuously throughout the course of each year. So it seems to be a matter of choice whether you want to stick to the discrete model you are using or whether you believe that the continuous-time nature of the underlying cash flow process indicates that a continuous time model should supplant this model. In our opinion, we are saying that our discrete time approximation is close enough to estimate the expected value of

---

[16] In particular, the evidence from Fama and French (1988), Lo and MacKinlay (1988) and Poterba and Summers (1988) all indicate that long run returns on the market portfolio are negatively autocorrelated. For example, that if the return on the market over the last few years has been lower than average, we would expect the return on the market to be above its long term average over the next few years. In this case, the long run average would be incorrect as the expected return on the market over the next few years, irrespective of whether we are using the arithmetic or geometric average as our "long run average." What we need is a *conditional* estimate of the future expected return on the market (or the market risk premium), based on the autocorrelation in the process and where the most recent observed returns on the market have been (i.e., above or below average). Further discussions of conditional expected returns on the market are beyond the scope of this teaching note.

the annual cash flows of the project, so you should use a discount rate that is consistent with the discrete time model you are using when you forecast the cash flows. This discount rate would be the the arithmetic average of annual rates of return if you are discounting annual cash flows.

To see that the CAPM is assuming the use of an effective per period rate of return (and hence arguing for the best estimate of the expected effective per period rate of return, consider the security market line:

$$E(R_i) \;=\; r_f \;+\; \beta_i \;[E(r_m) \;-\; r_f] \tag{11}$$

Notice that E(r_i) depends on the period length chosen and is the effective rate of return per period. Let the beginning of the period be time 0 and let the end of the period be time 1, and if we assume that security i pays no dividends in the period, we can substitute in for E(r_i) the definition of the per period rate of return on security i:

$$\frac{E(P_{i1}) - P_{i0}}{P_{i0}} \;=\; \frac{E(P_{i1})}{P_{i0}} \;-\; 1 \;=\; r_f \;+\; \beta_i[E(r_m) \;-\; r_f] \tag{12}$$

We want to convert equation (12) to a "valuation equation," i.e., isolating P_{i0} on the left hand side. So, using expressions in the middle and the far right hand side of (12), add one to both sides, multiply both sides by P_{i0} to yield:

$$E(P_{i1}) \;=\; P_{i0}\left[1 \;+\; r_f \;+\; \beta_i \;[E(r_m) - r_f]\;\right] \tag{13}$$

Substituting in from equation (11) and dividing both sides of (13) by the bracketed term on the right hand side of (13) and isolate P_{i0} on the left hand side:

$$P_{i0} = \frac{E(P_{i1})}{[1 + E(r_i)]} \tag{14}$$

Equation (14) says to discount the expected value of a cash flow to occur at the end of a "period" using the CAPM, you should divide it by one plus the *expected value of the per period rate of return*. This is consistent with the arithmetic averaging method as long as the interval of the data is the same as the periodicity as that of your expected future cash flows. So the arithmetic average of historical annual market risk premia is consistent with calculating the present value of annual expected future cash flows, even though we know that the discreteness of the future cash flows is an approximation to the actual stream of future cash flows.

This is of considerable importance in practice: per 1994 Stocks, Bonds, Bills and Inflation published by Ibbotson Associates, the average return on the S&P500 in excess of Treasury returns from 1926-1993 can be summarized by the following table:

	Arithmetic Average	Geometric Average
Average return on S&P500 in excess of T-bill rate (1926-1993)	8.6%	7.6%
Average return on S&P500 in excess of T-bond yield (1926-1993)	7.2%	5.2%

Using the geometric rather than the arithmetic average risk premium results in, on average, a 1% to 2% lower cost of equity estimate and results in commensurately higher valuations. If we were cynical (no need for the subjunctive here, really) we would argue that this is the reason that many consulting firms suggest to use the geometric average risk premium.

The table below summarizes the methods which should be used to estimate the risk free rate and the expected risk premium on the market portfolio (and the values these methods would give in December of 1993):

Method	General/Specific	Risk free rate	Expected market risk premium
1.	General	Current long term Treasury bond rate	Arithmetic average of returns on stocks in excess of T-bond yields
	December 1993 value	20-year bond rate = 6.5%	7.2%
2.	General	Current long term Treasury bond rate less historical average premium on long term bonds over short term bills	Arithmetic average of returns on stocks in excess of T-bill rates
	December 1993 value	6.5% - 1.4% = 5.1%	8.6%
3.	General	Estimate a different r_f for each future cash flow by the current zero coupon yield on a Treasury instrument maturing at the date of the cash flow	For each t, arithmetic average of returns on stocks in excess of same maturity Treasuries

Method 3 on the above table is simply included to indicate a possible method if we were to model explicitly the term structure of interest rates.

1c. Beta of equity

We could devote at least another 30 pages to different methods of estimating beta. Some of this discussion was covered in Finance I and a more detailed discussion of beta estimation will be left to the Investments class. The perspective we take here is to assume that the reader is a user of financial statistics as opposed to a producer of financial statistics. For example, we assume that you are using a beta book or an financial information database such as Alcar's APT! or BARRA's database as your source of beta estimates.[17] Beta estimates provided by investment banking firms and financial consulting firms are statistical estimates calculated from regressing the returns on the stock in question against returns on some market index (or indices in the case of APT betas). There are many different methods that the firms use to make these estimates and a detailed discussion of all these methods is beyond the scope of this note. We will try to hit some high points.

There are several things to remember about betas: 1. they are statistics, measured with error; 2. they are innately *levered betas*, because the stock returns used in the regression are the returns on a stock levered at the current leverage level of the company; 3. betas are probably biased also; 4. betas estimated by regressions using daily returns (and often even weekly returns) are questionable because of two problems with these data --- the non-synchronous trading/non trading problem and the bid-ask bounce problem.

Betas are measured with error

The following page is Table 9-2 from BM and is a page out of a Merrill Lynch beta book. If we were interested in estimating the cost of equity for any of the firms on this page we could simply look up its beta in this book. Take, for example, Digital Optronics Corp. Its beta is .52. If we were to assume that Digital Optronics had no debt in its capital structure, then, its WACC would be equal to its cost of equity which, if we use method 1 from section 1 of this note, we would get (assuming the beta as of December 1993 is the same as on Table 9-2 of BM)

---

[17] The beginning of Chapter 9 of Brealey and Myers discusses the use of a beta book to arrive at beta estimates of traded stocks.

WACC = r_e = 6.5% + .52(7.2%) = 10.24%.

How confident are we in this estimate of the cost of equity? Not very. The R^2 of the market model regression is only 3% --- the standard error of the beta estimate is 1.33! This means that a 95% confidence interval on the beta estimate would be (-2.14,3.18), which is huge. This is an extreme case, but underscores the fact that beta estimates are statistical estimates, not God-given parameter values.

Measured betas are innately levered betas

Suppose we knew that Digital Optronics had 20% market value of debt relative to its market value of debt plus equity, i.e., D/(D+E) = .2 (rather than being unlevered as in the above example). Suppose we also knew that Digital Optronics marginal tax rate was 35% and that their cost of debt, r_d, was 9%. What would its WACC be? Assuming that DO's capital structure was 20% debt during the time period over which the beta was estimated and is expected to be in the future also. The definition of WACC (equation 5) calls for the *levered cost of equity*, r_{eL}, to be plugged into the formula. Is 10.24% the levered required rate of return on equity? Yes, because the estimated beta, .52, is for a stock which has 20% debt in the capital structure; therefore, the required return (from CAPM) is the levered equity return associated with 20% debt in the capital structure. Given that we expect DO's capital structure in the future to contain 20% debt to capital also, 10.24% is the correct value of r_{eL} to plug in equation (5):

$$\text{WACC} = (.20)(9\%)(1-.35) + (.8)(10.24\%) = 9.36\%$$

What if we knew that DO's debt to capital ratio was 20% during the time period that the beta was estimated, and that its underlying asset risk (its unlevered beta risk) has not changed, but that, for the indefinite future, they intend to lever the company up to 30% debt to capital? Is our r_{eL} correct for this WACC? No. We would have to recalculate the cost of equity for the new degree of financial leverage. It is also possible, if not likely, that the cost of debt will be higher at the more highly levered capital structure. We will simply assume that the new cost of debt is 9.5% and that the marginal tax rate of the company will not have changed. How do we unlever and relever the cost of equity/beta?

Calculating the unlevered beta of the firm from its estimated levered beta depends on many assumptions, including the assumed regime of taxation, the assumed time pattern and riskiness of the debt tax shields, and the assumed time pattern and riskiness of future cash flows to the firm. We will maintain an assumption that there are corporate and personal taxes, that the corporate tax code offers a tax incentive for firms to issue debt (the corporate tax deductibility of interest payments) and that the personal tax code (and the nature of debt income vs. equity income) offers a tax incentive for firms to issue equity (the lower tax rate on realized capital gains than on ordinary income, the greater ability to defer recognition of income on equity relative to debt, and the greater ability to generate offsetting tax losses to offset realized gains with equity securities relative to debt securities). The latter incentive for firms to use equity arises out of the fact that, holding risk constant, an equity instrument offerring lower personal tax incidence will require a lower pre-personal tax yield than a higher taxed debt security. The net effect of these two partially offsetting influences is summarized by T*, a term referred to as the net tax gain per dollar of debt financing:[18]

---

[18] The derivation of T* is given in Thompson's teaching note, *Financial Leverage and the Cost of Equity Capital*.

$$T^{*} = \left[1 - \frac{(1-\tau_c)(1-\tau_{pe})}{(1-\tau_{pd})}\right] \qquad \textbf{(15)}$$

where τ_{pd} and τ_{pe} are the marginal tax rates of debt and equity investors *implied by the required rates on debt and equity in the market*. An exact value for T* is not readily available: the fact that there is an implicit personal tax rate being paid by debt issuers is clear, because taxable debt yields are always in excess of (holding risk, maturity and liquidity constant, if this is possible) municipal rates. Getting a good estimate of τ_{pe}, however, is a quixotic task. Prior to the Tax Reform Act of 1986, T* probably was in the range of 18% to 25%. Between 1986 and 1991, T* was probably between 25% and 30%. Since then, T* has probably been in the area of 20%-28%.

Under the above assumptions of T*, and assuming the levered equity and debt betas are β_{eL} and β_d, respectively (i.e., assuming that the debt of the company is not riskless, but has some beta risk), the unlevering formula for betas is given below:[19]

$$\beta_{eU} = \left[\frac{(1-T^{*})\ L}{1-T^{*}L}\right]\beta_d + \left[\frac{1-L}{1-T^{*}L}\right]\beta_{eL} \qquad \textbf{(16)}$$

where L is the debt to debt plus equity ratio (i.e., debt to value ratio) of the firm consitent with the leverage of the firm implied by its estimated beta (i.e., its historical debt ratio, not the future target

---

[19] Pages 191-192 of BM go through unlevering and relevering betas in the no-taxes case. Page 469 discusses unlevering betas in the corporate taxes case. Equation (16) is the same as the formula given for β_{assets} in the middle of page 469 of BM, generalized for corporate and personal taxes. The Salomon Bros. reading on cost of capital (in the case packet) assumes the debt beta is zero in the corporate taxes model. A detailed derivation of these formulas is given in Thompson's teaching note, *Financial leverage and the Cost of Equity Capital*, op cit.

debt ratio).[20]

Once you have the unlevered beta, you have to relever it to the future target capital structure weights. To do this, you must have an estimate of what the debt beta will be at the new target capital structure, $\beta_{d,target.}$ Then you rearrange equation (16) using the target D/E ratio:[21]

$$\beta_{eL,target} = \beta_{eU} + (1-T^{*})\left(\frac{D}{E}\right)_{target}\left(\beta_{eU} - \beta_{d,target}\right) \qquad \textbf{(17)}$$

To apply (16) and (17), we need the current debt beta and the target debt beta of the company. Debt betas are much more difficult to estimate in practice than are equity betas: debt securities are often not publicly traded at all, and when they are, they are often traded very

---

[20] While equation (16) looks pretty complicated, it has a simple interpretation. See that the two bracketed terms sum to one and can be characterized as weights, so that the unlevered equity beta is a weighted average of the levered equity beta and the debt beta. The only difference is that the weights are not L and 1-L, because the denominator is not the value of the levered firm, but the value of the unlevered firm, which is less than the value of the levered firm by the value of the tax shields of debt in the levered firm's capital structure. This interpretation is the same as that on page 469 of BM.

[21] Notice that, if the debt betas in equations (16) and (17) are zero, equations (16) and (17) collapse to the more familiar formulas below (16') and (17'):

$$\beta_{eU} = \frac{\beta_{eL}}{\left[1+(1-T^{*})\frac{D}{E}\right]} \qquad \textbf{(16')}$$

$$\beta_{eL,target} = \beta_{eU}\left[1+(1-T^{*})\left(\frac{D}{E}\right)_{target}\right] \qquad \textbf{(17')}$$

infrequently. Fama and French (1993) give the following beta estimates for debt securities ranging from 1-5 year government bonds to low-grade corporate bonds, measured over the time period July, 1963 to December, 1991:

Fama-French Debt Betas by Credit Quality (July, 1963 - December, 1991)

	1-5 yr Govt.	**6-10 yr Govt.**	**Aaa Corp.**	**Aa Corp.**	**A Corp.**	**Baa Corp.**	**Below Baa**
OLS Beta	0.08	0.13	0.19	0.20	0.21	0.22	0.30
t-stat	5.24	5.57	7.53	8.14	8.42	8.73	11.90

If we have a feel for the bond rating of the company's current debt level and for its target debt level, we may use the Fama French estimates as "ballpark" bond betas to use in equations (16) and (17).

As an example, suppose that DO's debt rating at 20% debt would be Aa rated, but their rating would drop to A if they were levered up to 30%. As such, $\beta_{d,current} = .2$, $\beta_{d,target} = .21$. Applying (16) and (17) to DO's equity beta (assuming T*=20%) we would estimate DO's unlevered beta

$$\beta_{eU} = [(1-.20)(.20)/(1-.20(.20))](.2) + [(1-.20)/(1-.20(.20))](.52) = 0.467$$

and then relever it to the target debt to capital of 30%:[22]

$$\beta_{eL,target} = .467+(1-.2)(.3/.7)[.467 - .21] = 0.555$$

---

[22] Had we used equations (16') and (17'), i.e., had we assumed debt betas were zero for these calculations, we would have estimated β_{eU} to be .43 and β_{eL},target to be .58. The relevered beta is certainly within rounding error of the one calculated in the text. This will usually be the case if you unlever betas then relever them to leverage levels close to where you started from. The larger the difference between the initial debt level and the target debt level, the larger the effect will be of using equations (16) and (17) rather than (16') and (17'). This will be particularly important in the APV method, where the unlevered cost of equity is used as the discount rate. Because betas will be estimated initially as levered betas and subsequently unlevered, using equation (16) rather than (16') will result in potentially very different betas. In particular, the unlevered equity beta will be, in general, larger than if we used (16') and will be more appropriate.

The relevered cost of equity becomes 6.5% + .555(7.2%) = 10.49%, and the weighted average cost of capital is given by

WACC = (.3)(9.5%)(1-.35) + .7(10.49%) = 9.20%.

Notice that, even though T* is used in the unlevering/relevering formulas, the correct tax rate to apply to the cost of debt in the WACC formula is τ_c.

Measured betas may be biased

The second last column on Table 9-2 from BM lists an "adjusted" beta. BM point out (page 188-189) that adjusted betas give better predicted betas, that they are closer to one than raw betas and that the Bayesian statistics needed to understand them are beyond the scope of their book. The reason that raw betas should be "shrunk" back towards one (raw betas less than one should be increased, raw betas greater than one should be decreased) is simply the "regression to the mean" phenomenon.[23,24]

While a discussion of Bayesian statistics is also beyond the scope of this note, the actual adjustment made to the betas on Table 9-2 is not. Take any "Beta" (raw beta) on the table. Calculate the absolute value of the difference between the raw beta and one. Call this absolute value the "distance to the goal line," and move the beta one third of the distance to the goal line towards one.

---

[23] Think of the estimated beta as being equal to the true beta plus some random mean zero estimation error. The error in beta estimation could be positive or negative, but when we are looking at a very large (small) beta, chances are we are seeing not only a high (low) true beta, but also a positive (negative) error term. In a predictive sense, the true beta will still be the same value, but the expected value of the future error term is zero, so we would expect that the future beta will be smaller (larger) than our estimate.

[24] We like to think of this as the Cy Young effect, where the winner of the Cy Young award (awarded to the top pitcher in each of the Major Leagues in each year) on average follows the year of the award with a year where his performance is below the performance of the year that yielded him the award. This is often blamed on getting older, or fat and lazy, or that the league is "gunning" for you, but is also completely consistent with regression to the mean.

I know this sounds stupid, but it is referred to as the "one third of the distance to the goal line" approach. E.g., Diasonics's beta is 1.48; the absolute value of the difference between the raw beta and one is .48; one third of this amount is .16, subtract .16 from 1.48 and the adjusted beta is 1.32. Similarly, Digital Optronics's beta is .52, or .48 below one, .52 + 1/3(.48) = .70. In each of the cases this simple adjustment is within rounding error of the Bayesian adjusted beta. This method is used by Value Line also to adjust betas for forecasting purposes.

It is good to know that services use such an adjustment procedure and to know there are the following problems with using a simple one-third of the distance to the goal line approach: 1. the procedure assumes that the prior belief about the beta of the firm is that it is one, which is the average beta of all stocks, but is not our prior belief about the average beta in the airline industry (very much larger than one) or the food industry (less than one), so we are shrinking betas to one indiscriminately with this method; 2. each beta is shrunk towards one by the same proportional factor, whereas some betas are measured with very little error (standard error of beta) and others (like Digital Optronics) are measure with very much error. Logically, for betas measured with a lot of error, our faith in the estimate is less than for betas measured with less error. It seems reasonable that betas measured with more error should be shrunk proportionately *more* towards our prior beliefs about their true mean than betas measured with very little error.

A final note on adjusted betas is that you should apply the adjustment (if you are going to use one) prior to any relevering/unlevering adjustments. The error is in the statistical estimation method coming up with the original beta.

Beware of betas estimated by regressions using daily returns

What is the frequency of the data used in the regressions run by Merrill Lynch underlying Table 9-2 beta estimates? Monthly. It doesn't say so on the Table, but it does in the accompanying BM text. You could have known this simply by seeing the 60's. It is very common (for no particularly good reason) for betas to be estimated using five years of monthly returns on the stock and the market index, which is 60 months. For the stocks with less than 60 months, it is usually because fewer than 60 monthly returns were available for the stock.

All betas are measured with some error, but you have to be a little wary of daily betas (and even sometimes weekly). Stocks do not all trade continuously, many are thinly traded, trading only a few times per day, sometimes only a few times per week. If the market index is the S&P500, it is composed of more actively traded stocks. The market may have an up day on day t and we would expect, conditional on this information, that stock i would go up also (according to its beta). But it may not trade on day t, so the observed price will not have gone up from the prior day. The next time it trades, however, it is likely to be up (based on the earlier market movement). In measured betas, this will show up as the stock's return being correlated with yesterday's market return. The covariance between the stock and the market return on the same day will be lower than it should be and the covariance between the stock's return and yesterday's market return (which should be about zero) will be higher than it should be. Since the concurrent (same day) covariance goes into the beta estimate and the non-concurrent covariance doesn't, the beta will be underestimated. An adjustment for this problem is, in fact, pretty apparent: if the stock is correlated with yesterday's market return too, add that in as part of the beta estimate. In fact, this is approximately how financial economists adjust for some of the non-trading problem in beta estimation.[25]

Because the non-trading problem is worse the greater is the possibility that stocks have not traded during the interval of the data being used to estimate the beta, it is far less of a problem for monthly data.

There is also a bid-ask bounce bias in estimating betas, which is also exacerbated by the use of daily returns. In general, using monthly data to estimate stock betas is advisable and if daily or weekly data is used, you should know enough to ask if some adjustment has been made for the non-synchronous trading problem.

2. Estimating divisional/project WACC via the peer group method

While the protracted discussion in Section 1 discusses how to estimate r_e for companies, we are ultimately interested in generating divisional or project specific hurdle rates as well. The cost of capital for an investment is an opportunity cost --- the expected rate of return that investors could earn on alternative investments in the capital market with the same risk as the investment. In general, the company wide cost of capital will not be the correct cost of capital to use for a specific investment: only if the riskiness of the project is the same as the riskiness of the company.

What we would like to have, and hold as the ultimate goal, is a project specific cost of capital based on the beta risk of the project. As a practical matter, however, it is almost impossible to estimate directly the beta risk of a project. Remember we estimate betas by regressing market returns on the investment against the return on the market portfolio. Individual projects are not publicly

---

[25] There are several methods in the literature for adjusting daily and weekly (even monthly, in some instances) for the non-synchronous trading problem. Two you may hear about are the Scholes-Williams and the Dimson beta adjustments.

traded, only publicly traded securities are. So, in practice we must estimate the beta risk of the project by estimating the betas of publicly traded peer group companies, each of which has been selected to be "like" the project in question.[26] What we really want is that the underlying business beta risk (unlevered beta) of the peer companies be close to the underlying business beta risk of the project. Here is the rub: the more finely we define the project in question, the more difficult it is to find publicly traded companies which simply are a mirror image of the project. This is why divisional costs of capital are often estimated: a division is usually an entity which is in a well-defined industry group and/or has assets with similar systematic risk. It is usually feasbile to find a set of publicly traded companies that are similar to a division of a multidivision company. This is why the beta estimation procedure often ends with estimates of divisional betas rather than project specific betas.

How does the analyst select the group of companies to be in the peer group? This is where the practice of finance becomes as much art as science. The idea is to find a group of firms with the same underlying business risk as the division. Logically, companies could be in different industries and have the same underlying business beta, but the method that is generally used is to find firms in the same industry as the division. You try to pick companies as much like the division as possible: similar products, similar customers, similar size, similar production/distribution technologies and methods, etc. By making the comparison as close as possible, you are increasing the likelihood that the companies face the same systematic risks. However, if you force the comparison to be too identical, you will end up with too few peer group companies (in fact, in the extreme case, you will

---

[26] An alternative to the peer group method of estimating beta is to estimate what are referred to as accounting betas, which use accounting information (and the covariance of functions of these accounting data with similar accounting aggregates for a market index) instead of market price data to estimate betas. A discussion of accounting betas is beyond the scope of this note, but this is a valuable topic which may be brought up in the Security Analysis class.

have none). Which brings us to ...

Using a pure-play company vs. a peer group of companies

There is a school of thought that says that the analyst should use the one company which most closely matches the company being modelled: the pure-play company. This is in contrast to using a set of companies that are more-or-less similar to the company. The peer group method is to be preferred to the pure-play method for estimating the beta of a division for a reason we mentioned in Section 1c.: betas are statistical estimates measured with error. The benefit of estimating betas for several peer companies and then averaging the estimates is that the estimation error of each of the betas will not be highly correlated. The averaging will tend to reduce our total estimation error (same argument as the argument for diversifying a portfolio). The situation is clearly a tradeoff: if you include companies in the peer group that are less like the division, then the beta estimate may be biased, but if you include too few peers, then the beta estimate may have a very large standard error. Judgement must certainly play a role here.

To demonstrate the use of peer group companies to estimate a divisional beta, consider the information on the table below:

Company	Market value Debt/(Debt+Equity)	Company size (Mkt val Equity) ($ millions)	Est. Beta	Std. Err. Beta Estimate
A	0.25	200	1.20	.35
B	0.00	1,150	0.80	.20
C	0.14	850	0.85	.25
D	0.36	500	0.75	.46

We will assume that the peer group companies are sufficiently "like" the division to be useful as a peer group. The table gives the market value debt to total capital ratio (D/D+E), the market value of the equity, the estimated beta and the standard error of each peer company. The betas are estimated by regressing the stock returns of the peer company against returns on the market index; as such they are *levered betas* at the degree of leverage indicated by the market value debt ratio. Our hypothesis is that the peer companies have the same *business risk* as the division, so we want to estimate the peer companies' unlevered betas. We can use equation (16') (and its attendant assumptions) to estimate the unlevered betas of each of the peer groups. Assuming T*=.20, we get

$$\beta^{A}_{eU} = 1.20/[1+(1-.2)(.25/.75)] = .947$$

$$\beta^{B}_{eU} = .80/[1+(1-.2)(0)] = .80$$

$$\beta^{C}_{eU} = .85/[1+(1-.2)(.14/.86)] = .752$$

$$\beta^{D}_{eU} = .36/[1+(1-.2)(.36/.64)] = .517$$

So we have four estimates of the unlevered risk of the division ranging from .517 to .947. What should our estimate be? We average these estimates. The simple average of the above unlevered betas is (.947 + .8 + .752 + .517)/4, or .754. Is this our best estimate, however? Often it is argued that you should weight the betas by some measure of the size of the firm, such as the market value of equity. This doesn't necessarily make sense for estimation of the division's beta, however: what we want to weight each of the betas by is by a measure of how much we think the peer is "like" the division or how "correct" the beta is. Unless the degree of similarity of the peer to the division is directly proportional to the size of the peer group company, weighting by size is not sensible. If we believe that each of the peers is just as good a peer as the others, then equal weighting is fine.

On the other hand, we have a measure of how "correct" the betas are, at least a measure of how precisely they are measured: the standard errors. The higher the standard error, all else equal, the lower is the precision of the beta estimate. We should weight the more precisely estimated betas more heavily than the less precisely estimated betas. A measure of the precision of the beta estimate has to be inversely related to the standard error. One common measurement of precision is the inverse of the squared standard error: Precision (β) = $1/[\text{S.E.}(\beta)]^2$. A precision-weighted average peer group unlevered beta is calculated in the table below:

Company	Unlevered Beta	Precision of beta	Precis. Weight	Weight x β_{eU}
	.947	8.1632653	.151	.143
A	.80	25	.464	.371
B	.752	16	.297	.223
C	.517	4.7258929	.088	.045
D		53.889	1.00	.782

The number at the bottom of the precision column is the sum of the precision numbers and is used as the denominator to calculate the precision weights (which clearly should sum to one). Notice that the weight on company D's beta is much smaller than in the equally-weighted case because its beta is measured with more error. The number at the bottom of the the weighted unlevered betas is the weighted average of the unlevered betas of the peers using the precision weights. Therefore our estimate of the unlevered beta of the division is .782.

Next, we need to determine the target debt ratio for the division, relever the beta to the target debt ratio of the division, plug in the CAPM to yield the relevered cost of equity, r_{eL} (target), estimate the cost of debt if the division were a stand alone company with the target capital structure and plug the estimates into equation (5) to arrive at the WACC of the division.

Determining the target debt ratio for the division

The capital structure weights to use in the WACC are the intended future market value weights D/D+E and E/D+E. When we say "how the assets are financed," we are referring to the capital structure that the company intends to maintain for the foreseeable future. In particular, we are not referring to the specific source of financing that may be used to purchase some new investment. For example, if a company is considering an investment in its corrugated packaging division and the new investment will be paid for by taking a loan out from a bank, the bank loan rate does not represent the cost of capital for the investment. If this investment doesn't generate enough cash flow to repay its debt, the debt will be paid off out of the remainder of the corporations cash flow: the debt rate reflects not the riskiness of the project, but the riskiness of having a debt-priority position against the entire corporation's income and cash flow. In addition, should the investment pay more than enough cash flow to repay the bank debt, the remainder of the cash flow will belong to the corporation and can go to pay other obligations and ultimately can go to the equityholders. Therefore, the specific funds used to purchase an asset are irrelevant. The way the asset is *financed* refers to the capital structure that represents the claims to the cash flows generated by the assets.

Should we use the "optimal capital structure" to define the weights for divisional WACC's or should we simply use the corporation's intended capital structure in total. What do we mean by an optimal capital structure?

Different assets support different amounts of debt in their capital structure. Assets differ in the volatility of the cash flows and income they generate; assets differ in their ability to be recognized as valuable collateral (so that you can borrow at fairer rates); some assets are commodities (often easily leveraged) others are very specialized assets (not as easily leveraged); some intangible assets are very hard to lever against (management quality, good consultants), others are easier to lever up (brand name recognition).[27]

We will assume for the purposes of this note that there is such a thing as an optimal capital structure that we would expect the company to be levered to, all else equal.[28] A company could own divisions in industries with different debt capacities and if we used its corporate capital structure weights in the divisional WACC's we would be understating the WACC of low debt capacity division (and overstating the WACC of the high debt capacity divisions). Let's use some numbers to be more concrete: suppose a company has two divisions, Division A which makes a commodity oil-based product with a debt capacity (optimal D/D+E) of 45% and Division B which makes building products, a highly cyclical business with debt capacity of 15%. Suppose the two divisions have the same market value $V_A = V_B = 100$. As a stand alone business, Division A could support .45(100) or 45 of debt; Division B could support .15(100) or 15 of debt. Next, suppose that the corporation has 60 in debt (a D/D+E ratio at the corporate level of 30%). Then we can use D/D+E = 45% as the target capital structure for Division A's WACC and D/D+E = .15 as the target capital structure for Division B's WACC, because these ratios are *consistent with the corporation's overall use of debt in its capital structure*.

What if the corporation uses a different amount of debt at the corporate level than the weighted average of the debt capacities of the divisions? In our example, what if the corporation intended to have 20% debt in its capital structure (rather than the 30% weighted average of the debt

---

[27] For a discussion of debt capacity, see BM, Chapter 19, page 459.

[28] Be careful not to assume that the industry average debt level is the optimal level of debt for a corporation in that industry. This leads to the lemmings theory of capital structure and I don't want you to think like lemmings.

capacities of the divisions)? We can still "allocate" different debt amounts to the different divisions that reflect their *relative* debt capacities and which is still consistent with the overall corporate use of debt in its capital structure. All we have to do is "scale down (or up)" the debt capacities by a factor that makes the total debt equivalent to the target corporate debt: If our example corporation has a debt to capital ratio of .2, rather than .3 (the weighted average of 45% and 15%, with equal weights), we can multiply the divisional debt capacities by .2/.3 (or .6667), yielding (45%)(.6667) = 30% for Division A and (15%)(.6667) = 10% for Division B. Notice the weighted average of our adjusted debt capacities (.5)(30%) + (.5)(10%) = 20%, the total corporate debt to capital ratio.

The hard part of the above method is to figure out the weights, the respective value of each of the divisions. In practice you will not know these quantities and will probably use some available measure as a "proxy" for value: divisional sales, divisional operating income, divisional identifiable assets, etc. None is perfect and you might as well do it a couple of different ways; divisional operating income would be best for most firms, but not all.

A much easier alternative is to simply use the corporation's target debt to capital ratio for *each* of the divisions. This number is definitionally consistent with the corporate use of debt in its capital structure but does not reflect the differential debt capacities of different divisions. If the divisions' debt capacities are reasonably close to each other, then the more complicated method would not be of any value, so the simpler procedure would be better.

Relevering the unlevered equity beta and calculating the cost of equity based on the target debt level

Plug the estimated unlevered beta (e.g., the .782 calculated on page 43) into equation (17') to yield the target leverage relevered beta. Plug the relevered beta into the CAPM equation (6) using the r_f and $E(r_m)$-r_f estimates from Section 1a. and 1b. to arrive at the target leverage cost of equity. For our sample division with an unlevered beta of .782 and a target debt ratio of 45% (also assuming the December 1993 values for r_f = 6.5% and $E(r_m)$-r_f = 7.2 (see page 30) and assuming that T*=.2) we have

$$\beta_{eL} \text{ (target debt ratio)} = .782[1+(1-.2)(.45/.55)] = 1.29$$

and

$$r_{eL} \text{ (target debt ratio)} = 6.5\% + 1.29(7.2\%) = 15.8\%$$

If we assume that, at a 45% debt ratio, the division as a stand alone company would have debt with required rate of return, r_d, of 10%, and would have a marginal tax rate of 35%, we have WACC:

$$\text{WACC} = .45(10\%)(1-.35) + .55(15.8\%) = 11.6\%$$

We have been assuming values for the cost of debt in all our examples so far. The next section of the note takes up the issue of estimating the cost of debt for companies, divisions, etc.

3. The after-tax cost of debt to use in the WACC

Equation (5) assumes there are only two types of corporate securities in the capital structure: straight vanilla long-term coupon debt and similarly flavored equity. This is a crude simplification of corporate capital structures where the debt could be short-term, long-term, floating, fixed, denominated in dollars or in a foreign currency; traded on U.S. market, a Eurobond market, or foreign bond market; straight coupon bullet principal debt, sinking fund debt, deep discount or even zero coupon debt, senior or subordinated, secured with specific collateral or general obligation debentures. The equity of the corporation could have different categories of voting rights, different categories of marketability. There is a world of hybrid securities which are some combination of debt and equity: e.g., preferred stock, convertible debt, warrants, convertible preferred stock, different types of money market preferred stocks, MIPS, etc. The different features of the instruments cause them to have different required rates of return (the r's we are estimating).

It is conceivable to have a different term in equation (5) for each distinct type of security in the capital structure and estimate the required rate of return on each type of security, summing up the market value-weighted average of these rates to arrive at WACC. But for most purposes this does not really suit our needs. The WACC is an attempt to develop a hurdle rate for different projects/divisions/corporations based on their financial structure. While we could estimate a very complicated capital structure for a specific firm with many different categories of claims, we could not hope to do this at the project or divisional level.

Let's characterize a simpler method to estimate the WACC. If you have a firm with many different debt categories on its balance sheet, pretend that the firm repurchased all these securities and issued an equivalent market value of straight coupon debt. What would the riskiness of the debt be and what would the required rate of return on the debt be? Use this as the r_d in the WACC formula.

What about short-term vs. long-term debt? Suppose a cororation uses revolving debt from a corporation (which is really floating rate rather than fixed, not short-term vs. long-term)? Should we plug in the current floating rate in for r_d? No. Refer to the discussion of the appropriate notion of the short term risk free rate to use in the CAPM: we don't just want to know the debt rate for the next period because we are discounting distant cash flows. So we want to know the expected future r_d's. These rates will be imbedded in the equivalent risk longer term debt rate.

In theory, we should be able to apply a model like the capital asset pricing model to debt securities based on their beta risk. Unfortunately, debt betas have never been shown to do a very good job at predicting bond expected rates of return (even less so than with stocks). If we know the target debt ratio of the division/corporation how can we estimate the riskiness and the required rate of return on the debt of the division/corporation? What I outline next we will not be able to do in detail for our cases, but it does represent common practice in estimating r_d's.

One theoretical aside: we should point out the distinction between the *yield to maturity* on a bond and the *expected rate of return* on the bond. If a bond is risky, its expected rate of return is always *less than* its yield to maturity. The yield to maturity is the discount rate that equates the present value of the bond's *promised* payments to its market price; the expected rate of return is the discount rate that equates the present value of the bond's *expected* payments to its market price. A risky bond is risky because there is some probability it will not pay its promised payments to the holder; since there is no probability the holder will receive more than bond's promised payments, the expected value of the future payments must be smaller than (or equal to, if the bond is riskless) the promised value of the future payments. Hence, the expected rate of return is always less than the yield to maturity on the bond.

Bond rating agencies (such as Moody's, Standard and Poor's, etc.) attempt to assess the riskiness of corporate bonds. Standard and Poor's categorizes bonds into decending quality categories AAA, AA, A, BBB, BB, B, CCC, etc. There are subcategories in each category. Bonds in the top four major categories (AAA, AA, A and BBB) are referred to as *investment grade* bonds; bonds rated below BBB are below investment grade (also called *junk bonds*). The type of risk that bond rating agencies are trying to separate out is *default risk*, the probability that a bond will subsequently default on interest and/or principal. While this is not exactly the same as the *systematic* risk of a bond, the yields to maturity on bonds are higher on average, the lower the bond rating.[29] Also the variance of bond yields within rating category is very small relative to the variance of bond yields between rating categories. As such, if we can predict, given information about the company/division/project, what the bond rating would be of the debt in the capital structure of the company/division/project, we could estimate r_d as the average current expected rate of return on the debt issues in that rating category.

There are models (generally using multiple discriminant analysis) available which try to predict the bond rating classification a corporation's bonds would receive if they were rated on the basis of financial and market variables. Important variables for predicting bond rating include the debt to capital ratio, measures of interest coverage, measures of profitability and the size of the company. A detailed discussion of bond rating models is beyond the scope of this note.[30]

If we can predict the bond rating of the debt of the firm/division/project based on our analysis of its financials, it is easy to look up the current average yield to maturity on bonds in that ratings category. Moody's, Standard and Poor's, etc., all keep statistics on these numbers. It is tempting to simply plug the average yield to maturity into equation (5) for r_d. We have to be careful, however, because the expected rate of return is less than this amount.

For highly rated corporate bonds (e.g., AAA, AA, and to some extent A) the probability of subsequent default is quite low, so the yield to maturity on these bonds may be very good estimates of the expected rates of return. As the credit quality lessens, however, the yields to maturity will be

---

[29] See BM, Section 4 of Chapter 23 for a discussion of bond ratings and default risk.

[30] BM discuss a simpler application of discriminant analysis is Chapter 30: the prediction of whether or not a firm will subsequently go bankrupt. The interested reader is directed to George Foster's text, *Financial Statement Analysis*, Chapter 14.

too large an estimate of the cost of debt. For example, after the crash of the junk bond market in 1989-1990, it was not uncommon to see junk bonds trading below 50 cents on the dollar. Their yields to maturity were huge, say 60%, 70%. Do you think that the investors expected to earn that high of a yield? Doubtful! It is difficult to estimate the expected rate of return on risky debt. There is a rule of thumb that for lower rated debt, the yield to maturity on BBB debt is a good estimate of the actual expected rate of return on lower rated debt. I don't have much faith in this rule of thumb, but there you have it.

4. Capital structure weights

We request you look at the part of Section 2 where we discuss the capital structure weights in gory detail.

5. Conclusion

There is much art and a little science in estimating the costs of capital for companies/divisions/projects. The following flowchart (to come later) outlines the steps the analyst must take to estimate the WACC for the application at hand.

References

1. Brealey, Richard and Stewart Myers. **Principles of Corporate Finance**, 4th Edition, McGraw-Hill, 1991.

2. Chen, Nai-Fu, Richard Roll and Stephen Ross. *Economic forces and the stock market*, **Journal of Business**, 1986.

3. Copeland, Tom, Tim Koller and Jack Murrin. **Valuation: Measuring and Managing the Value of Companies**, 2nd Edition, Wiley, 1994.

4. Dimson, E., *Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics*, 1979.

5. Fama, Eugene, and Kenneth French, *Dividend yields and expected stock returns*, **Journal of Financial Economics**, 1988.

6. Foster, George. **Financial Statement Analysis**, 2nd Edition, Prentice-Hall, 1986.

7. Ibbotson Associates, **Stocks, Bonds, Bills and Inflation (SBBI yearbook)**, 1994.

8. Lo, Andrew and Craig MacKinlay. *Stock prices do not follow random walks: Evidence from a simple specification test*, **Review of Financial Studies**, 1988.

9. Poterba, J. and L. Summers. *Mean reversion in stock prices: Evidence and implications*, **Journal of Financial Economics**, 1988.

10. Scholes, Myron and Joseph Williams, *Estimating betas from non-synchronous data*, **Journal of Financial Economics**, 1977.

11. Thompson, Timothy. *Teaching note: Valuation using the adjusted present value (APV) method vs. the adjusted discount rate (ADR) method: Theory*, unpublished teaching note, 1994.

12. Thompson, Timothy. *Financial leverage and the cost of equity capital: Teaching note*, unpublished teaching note, 1994.