THE COST OF CAPITAL, THE CAPM, AND THE MARKET PRICE OF RISK*

Glenn Boyle
NZ Institute for the Study of Competition and Regulation
Victoria University of Wellington

16 November 2004

Preliminary
Not for quotation or reproduction
Comments Welcome

* Executive Director, ISCR, PO Box 600, Wellington, New Zealand. I am grateful to Russell Investment Group Ltd for providing me with data used in this study.

THE COST OF CAPITAL, THE CAPM, AND THE MARKET PRICE OF RISK

1. Introduction

The standard version of the Capital Asset Pricing Model (CAPM) has the familiar form:

$$E[R_i] = R_f + \beta_i \{E[R_m] - R_f\} \qquad (1)$$

where R_i is the random return on asset i, R_f is the riskless rate of interest, R_m is the random rate of return on the market portfolio of risky assets, β_i is the asset i 'beta', equal to the covariance of R_i and R_m divided by the variance of R_m, and E[.] is the expectations operator. Equation (1) states that the asset i risk premium is proportional to the market risk premium where the factor of proportionality is equal to β_i. In this formulation, β_i is the quantity of asset i risk and the market risk premium $\{E[R_m] - R_f\}$ is the price of that risk.

Equation (1) is a *relative* pricing model. Specifically, it relates one market price (the risk premium on asset i) to another market price (the risk premium on the market portfolio). But the second price, if correctly measured, incorporates the first price, so there is an element of circularity in this process. Practical applications of the CAPM, such as estimating the cost of capital, typically ignore this problem - on the grounds that any individual asset is an infitesimally-small portion of the market portfolio - and treat the market risk premium as a *free parameter* to be estimated from data. But as Cochrane (2001) points out, this procedure ignores the CAPM predictions for the market portfolio itself; taking the market premium as given neglects its underlying dependence on more fundamental CAPM parameters. In this paper, I show that recognition of this link potentially eliminates the need to estimate the market risk premium directly.

The details of this argument are explored in the remainder of the paper, but an intuitive overview may be helpful at this point. Underlying the CAPM is the separation result from modern portfolio theory that all investors optimally allocate their funds between two portfolios, one risky and one riskless. Investor-specific risk attitudes determine the split between the two portfolios, but have no effect on the composition of the risky asset portfolio which depends only on the distribution of future asset returns. If all investors perceive the same returns distribution, they must then wish to hold the same portfolio of risky assets. Equilibrium in the market for these assets requires that this common portfolio be the so-called market portfolio, an observation that leads directly to equation (1). However, this process is incomplete as market clearing is imposed only on the market for risky assets; equilibrium in the market for riskless assets is ignored. Incorporating this additional condition in the model places an exact restriction on the allowable value of the market risk premium in terms of the variance of market returns. As a result, applications of the CAPM need only estimate the value of the latter parameter and not the market risk premium itself.

This result is not in itself new. Assuming quadratic utility, or exponential utility with normal returns, (e.g., Friend and Blume,1975; Huang and Litzenberger, 1988), or various technical conditions (Merton, 1980), other authors have also shown that the market risk premium is proportional to the variance of market returns. However, I demonstrate that this result is a simple consequence of riskless asset equilibrium, for all preferences defined over the mean and variance of end-of-period wealth. Perhaps more importantly, I examine the implications of the result for practical applications such as estimating the cost of capital, an issue that none of the authors above, with the partial exception of Merton, considers .

Expressing the market risk premium as a function of the variance of market returns is potentially valuable for calculating the cost of capital. Estimating the variance of returns is much easier than estimating the mean, particularly when these parameters vary through time, so the alternative approach holds out the promise of identifying risk-based shifts in the cost of capital. Unfortunately, using New Zealand data, I find that the variance of returns is too volatile to be very useful for this purpose; at times the implied cost of capital is implausibly high while at other times it is implausibly low. In my view, this is not good news for CAPM-based approaches to estimating the cost of capital: seemingly-reasonable estimates based on (1) are obtained only by ignoring an important part of CAPM content.

The next section explicitly derives the link between the market risk premium and the variance of market returns. In Section 3, I discuss how this might help obtain more accurate cost of capital estimates and apply this to data. The final section contains some concluding remarks.

2. CAPM and equilibrium in the market for riskless assets

Let R_i be the random return on risky asset i = 1,...,n and R_f be the certain return on a riskless asset. If λ_{ik} is expenditure on asset i by investor k = 1,...,m, and λ_{fk} is expenditure on the riskless asset, then end-of-period wealth W_k satisfies:

$$W_k = \sum_{i=1}^{n} \lambda_{ik}(1+R_i) + \lambda_{fk}(1+R_f) \qquad (2)$$

The asset expenditures must sum to initial wealth W_{0k}. Setting the latter to unity for convenience, (2) can be written as:

$$W_k = 1 + R_f + \sum_{i=1}^{n} \lambda_{ik}(R_i - R_f) \qquad (3)$$

so that:

$$E[W_k] \equiv \bar{W}_k = 1 + R_f + \sum_{i=1}^{n} \lambda_{ik}(E[R_i] - R_f) \qquad (4)$$

$$\text{var}(W_k) \equiv \sigma^2_{W_k} = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ik}\lambda_{jk}\text{cov}(R_i, R_j) \qquad (5)$$

Each investor has a utility function $v_k(\bar{W}_k, \sigma^2_{W_k})$ that depends only on the mean and variance of end-of-period wealth, and chooses the portfolio that maximises the value of this function:

$$\underset{\{\lambda_{1k} \ldots \lambda_{nk}\}}{\text{Max}} \quad v_k(\bar{W}_k, \sigma^2_{W_k})$$

The first-order conditions for this problem are:

$$\gamma_k(E[R_i] - R_f) = \sum_{j=1}^{n} \lambda_{jk}\text{cov}(R_i, R_j) \qquad i=1,...,n \qquad (6)$$

where $\gamma_k \equiv -\dfrac{\partial v_k/\partial \bar{W}_k}{2(\partial v_k/\partial \sigma^2_{W_k})}$ is investor k's marginal rate of substituting risk for return. That is, $1/\gamma_k$ is the additional mean return that investor k would need to be no worse off following a marginal increase in variance. If this is high (low), then investor k is relatively intolerant (tolerant) of risk. We can therefore interpret $1/\gamma_k$ as a measure of investor k's risk aversion.

Writing (6) in matrix form and rearranging yields the two-fund separation result of Tobin (1958), i.e., the composition of investor k's portfolio of *risky assets* is unaffected by risk attitudes (γ_k) and depends only on perceived means, variances and covariances of returns. If these parameters are the same for all investors, then they all hold the same portfolio of risky assets. Market clearing requires that this portfolio contain each asset in an amount equal to its weight in the portfolio of total invested wealth in risky assets, i.e., the *market portfolio* of risky assets. Then λ_{jk} is equal to asset j's weight in the market portfolio multiplied by investor k's total investment in risky assets ($1-\lambda_{fk}$) and equation (6) can be rewritten as:

$$\gamma_k(E[R_i] - R_f) = (1-\lambda_{fk})\text{cov}(R_i,R_m) \qquad i=1,...,n \qquad (7)$$

As (7) holds for all risky assets, it must also apply to the market portfolio:

$$\gamma_k(E[R_m] - R_f) = (1-\lambda_{fk})\sigma^2_m \qquad (8)$$

where σ^2_m is the variance of R_m. Combining (7) and (8) then yields equation (1):

$$E[R_i] = R_f + \beta_i(E[R_m] - R_f)$$

Equation (1) is, of course, the standard formulation of the CAPM, but it overlooks an important part of the underlying pricing process. Going from (6) to (7) requires that supply equal demand for each risky asset, but no corresponding requirement is imposed on the riskless asset.

To determine the implications of imposing riskless asset equilibrium, return to equation (8) and note that this can be written as

$$\lambda_{fk} = 1 - \frac{(E[R_m] - R_f)/\sigma_m^2}{(1/\gamma_k)} \qquad (9)$$

which expresses investor k's demand for the riskless asset as a function of risk aversion and market portfolio characteristics. Intuitively, $(E[R_m] - R_f)/\sigma_m^2$ is the rate at which the market portfolio trades off risk and return while $(1/\gamma_k)$ is the rate at which investor k is *willing* to make this tradeoff. If, for example, the former exceeds the latter, then investor k desires more risk than is offered by the market portfolio alone and thus borrows at rate R_f to finance a larger holding of that portfolio, i.e., $\lambda_{fk} < 0$.

Equilibrium in the riskless asset market requires that total borrowing equal total lending, i.e., $\sum_{k=1}^{m} \lambda_{fk} = 0$.[1] Applying this to (9) yields:

$$E[R_m] - R_f = \left(\sum_{k=1}^{m} \gamma_k/m\right)^{-1} \sigma_m^2$$

$$= (1/\gamma)\, \sigma_m^2 \qquad (10)$$

where $\gamma \equiv \sum_{k=1}^{m} \gamma_k/m$ is the average value of γ_k, i.e., $1/\gamma$ is the average risk aversion of all investors. Thus, equation (10) states that riskless asset equilibrium constrains the market risk premium to a value equal to the product of market risk and market risk aversion.

To understand (10) intuitively, suppose that it is not satisfied, e.g., $(E[R_m] - R_f) < (1/\gamma)\, \sigma_m^2$. Then the rate at which the market portfolio offers to trade off risk and return is

1 This assumes the riskless asset is in zero net supply. Nothing substantative in what follows is lost by allowing for a positive net supply.

below the rate required by investors to make this tradeoff. Investors will thus wish to substitute from the market portfolio to the riskless asset, i.e., there is excess demand for the riskless asset. With supplies fixed, equilibrium is re-established by a rise in ($E[R_m] - R_f$) until the excess demand is eliminated. This occurs only when the available risk-return tradeoff is equal to the required tradeoff, i.e., when (10) holds. More succinctly, any violation of (10) implies excess demand or supply in the riskless asset market, so equilibrium in that market requires that (10) be satisfied. In equilibrium, the market risk premium ($E[R_m] - R_f$) must equal the market price of risk (σ_m^2/γ).

Substituting equation (10) into (1) yields:

$$E[R_i] = R_f + \beta_i\, (\sigma_m^2/\gamma) \qquad (11)$$

which relates the expected excess return on asset i to beta and the market price of risk.

The link between (1) and (11) is worth emphasizing. Equation (1) requires only that risky asset markets clear, and is thus a partial equilibrium statement that relates one endogenous price variable ($E[R_i] - R_f$) to another ($E[R_m] - R_f$). Equation (11) takes this a step further by requiring that supply also equal demand for the riskless asset, thereby allowing $E[R_m]-R_f$ to be expressed in terms of underlying exogenous paramaters and transforming (1) into a general equilibrium statement.

Friend and Blume (1975) and Merton (1980) also report an equation similar to (11), and Huang and Litzenberger (1988) explicitly derive it in the context of either quadratic utility or exponential utility with normal returns. However, none explicitly makes the link to riskless asset equilibrium, and only Merton recognises its implications for practical applications such as estimating the cost of capital. The latter issue is the subject of the remainder of this paper.

3. Estimating the cost of capital

Suppose one wishes to estimate a firm's cost of equity capital. The usual approach estimates the market risk premium $E[R_m]-R_f$ as a free parameter and uses this in equation

(1). However, as section 2 demonstrates, the market risk premium is *not* a free parameter; rather it is an endogenous function of underlying CAPM parameters, as described by (10). Thus, applications of the CAPM can, in principle, use either (1) or (11), estimating either the market risk premium or the market price of risk. Although the latter approach is theoretically superior, the only relevant consideration in practical situations is the reliability of estimates. More precisely, is it better to estimate the market risk premium ($E[R_m] - R_f$) or the market price of risk (σ_m^2/γ)?

There are two good reasons to favour the latter approach. First, as discussed at length by Merton (1980), Black (1993) and Campbell et al (1997), it is much easier to estimate the variance of returns than it is to estimate expected returns. The essence of their argument is that the precision of the variance estimate increases with the number of observations while the precision of the expected return estimate increases only with the length of the data series. In other words, a good estimate of variance can be obtained even with a short time series so long as the data are sufficiently high-frequency, but the only way to get a similarly good estimate of the mean is to have a long time series. Consequently, given the usual constraints on available data, variance estimates will be considerably more accurate than expected return estimates.

The second reason for favouring (11) follows from the first. In calculating the cost of capital, the relevant distribution of returns is the *conditional* distribution, since it is this that describes the current risk outlook: if risk is high at a particular date, then the market risk premium, and hence the cost of capital, should also be high at that date. More precisely, if one wishes to use equation (1) to estimate the cost of capital, then an estimate of the current market risk premium is required (i.e., the conditional mean of excess market returns); equation (11) requires, instead, an estimate of current risk (i.e., the conditional variance of market returns). In the case of (1) however, because of the long time period needed to estimate expected returns, the best one can feasibly do is obtain a single estimate of the unconditional market risk premium. Consequently, applications of (1) are unable to incorporate variation over time in the market risk premium and thus reflect the current risk environment. By contrast, equation (11) is potentially able to identify this variation because of the shorter time series required for estimating variance.

Of course, there is also a significant disadvantage to using (11): the parameter γ is unobservable. However, it may be possible to estimate this fairly accurately. First, it seems reasonable to assume, at least as a first approximation, that $1/\gamma$ is a constant; as Campbell and Viceria (2002) point out, there have been large increases in per capita consumption and wealth in the last 100 or so years, but no corresponding trends in risk premia or interest rates consistent with investors having changed their attitudes towards relative risks. Second, if the unconditional market premium can be accurately estimated from available data, then the unconditional version of (10) can be used to estimate the constant $1/\gamma$ applicable to a given market. That is, given estimates of $\{E[R_m] - R_f\}$ and σ_m^2 from a long time-series, one can rearrange (10) to solve for $1/\gamma$. This can then be used in a conditional version of (11) to estimate the current cost of capital.

For example, suppose one obtains, from 100 years of data, estimates of $\{E[R_m] - R_f\}$ and σ_m^2 equal to 0.06 and 0.03 respectively. Then the implied value of $1/\gamma$ is two and, from (11), the current (date t) expected return on asset i is $R_{ft} + 0.06\beta_{it}$, where the t subscripts denote date t values.

To provide a concrete illustration and assessment of this process, I use equation (10) to estimate an annual series for the New Zealand market price of risk over the last 30 years. This requires, first, estimation of γ as described above, and, second, annual estimates of σ_m^2.

To calculate $1/\gamma$, I use the recent study of Lally and Marsden (2004) on NZ returns during the 1931-2002 period. Their estimates of the unconditional values of $\{E[R_m] - R_f\}$ and σ_m^2 imply a value of $1/\gamma$ equal to 1.42.[2]

Turning to σ_m^2, I use monthly real stock returns on the NZ stockmarket to calculate a moving average variance of returns for the 34 years from January 1970 to December 2003.[3]

---

2 Lally and Marsden (2004) report estimates of $\{E[R_m] - R_f\}$ and σ_m^2 equal to 0.074 and 0.059 respectively. However, their calculation of the former is with respect to bonds rather than bills. For the countries examined in Dimson et al (2002), this understates the market premium by approximately one percentage point. Hence, I use $1/\gamma = 0.084/0.059 = 1.42$.

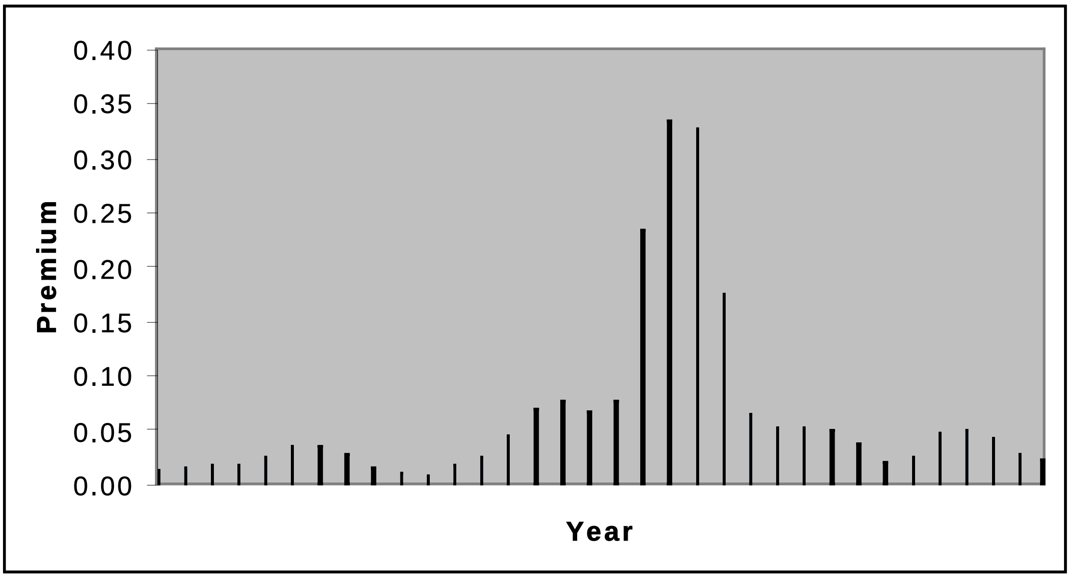
3 The returns data are obtained from the NZ Gross Index compiled by Russell Investment Group Ltd. I am grateful to Craig Ansley and Fiona Lintott for providing me with access to these data. Details of its construction can be found in Chay et al (1993). Nominal returns are deflated using movements in the CPI.

Specifically, the conditional variance in each month during this period is calculated as the sample variance of returns over the previous 36 months.[4] These monthly variances are then converted to annual figures and averaged across the 12 months of each calendar year, giving a single variance estimate for each year.

Combining these estimates of σ_m^2 with γ = 1.42 gives a time series of annual estimates of the market price of risk, and the results from this procedure are outlined in Figure 1 and Table 1. The former depicts the year-by-year variation in the market price of risk estimate; the latter summarises these data for both the full period and three sub-periods.

Figure 1
Time-Variation in the Market Price of Risk

This figure illustrates the time variation in the New Zealand market price of risk σ_m^2/γ, where σ_m^2 is the variance of market returns and $1/\gamma$ is the risk aversion parameter for the representative investor. For each month, σ_m^2 is calculated as the sample variance of returns over the previous 36-months and then converted to an annual figure. γ is set equal to 1.42, based on Lally and Marsden (2004).



4 This simple procedure is similar to that used by Officer (1973) and Merton (1980). The preferred modern method for estimating conditional variance is via a GARCH process, e.g., Bollerslev (1986). As this more complex method yields virtually identical conclusions in this case, I report only the simple approach described above.

Table 1
Time-Variation in the Market Price of Risk: Summary Statistics

This table calculates the average, maximum and minimum values of the market price of risk series appearing in Figure 1.

	Average	Maximum	Minimum
Full Sample			
1970-2003	0.064	0.336	0.009
Sub-Samples			
1970-79	0.021	0.037	0.011
1980-89	0.096	0.336	0.009
1990-2003	0.071	0.327	0.020

The primary impression from these results is one of considerable volatility in the variance of market returns. For the period as a whole, the average market price of risk is a reasonable-sounding 6.4%, but this hides significant intra-period variation. Throughout the 1970s, the maximum value was 3.7% and the average only 2.1%. In the 1980s, the price of risk ranged from less than 1% to more than 33%; the period since 1990 is similarly volatile. Although the most extreme values occur in the 1980s, the remaining periods are also characterized by unexpectedly high and low values. For the most recent year (2003), the estimated market price of risk is 2.3%, implying a very low cost of capital for most projects.

If one wishes to use the CAPM to capture time variation in the market price of risk, these results are disappointing. For instance, is it really plausible that the price of risk went from less than 1% in the early part of the 1980s to more than 30% by the end? Answering in the affirmative implies an acceptance of very large swings in the cost of capital. Similarly, does it seem reasonable that the average-risk firm (i.e., $\beta_i = 1$) in 2003 had a cost of equity only 2.3 percentage points above the riskless rate of interest?

Measurement error in the two crucial variables - γ and σ_m^2 - seems unlikely to resolve these problems. The volatility in σ_m^2 is unaffected by the use of alternative estimation periods, higher-frequency data, or more sophisticated estimation methods.[5] Similarly, any error in the estimate of γ has no effect on the volatility of the market price of risk. Of course, it is possible that, contrary to the assumption maintained above, γ is not independent of σ_m^2, but this seems likely to exacerbate matters: γ is most likely to be high (low) when market risk is high (low).

If variable measurement error is not driving the results, then only two possibilities remain. One is that the true cost of capital is subject to much higher volatility, and takes on more extreme values, than has previously been thought. The other is that the simple product of γ and σ_m^2 fails to adequately capture the market price of risk. Or, to put it another way, the relationship between market risk and expected return is considerably more complex than envisaged by the CAPM. If one is sceptical about the first possibility, then the unpalatable conclusion is that the CAPM is incapable of relating the cost of capital to the contemporaneous risk environment.

4. Concluding Remarks

I interpret the results of this paper as bad news for practical applications of the CAPM. The usual approach takes the market risk premium as given and uses this, along with other exogenous parameters, to estimate the risk premium on some other asset or project. However, this suffers from two drawbacks. First, the market risk premium is an endogenous variable in CAPM equilibrium, so applications that treat it as exogoneous are effectively ignoring part of the CAPM. Second, because the market risk premium can only be accurately estimated using a long time series of data, the estimate used in applications is likely to reflect little of the current risk environment. The second problem can, in principle, be resolved by explitly dealing with the first so that the market risk premium is linked to the variance of market returns. Unfortunately, annual estimates of this risk parameter are extremely volatile, resulting in implausible swings in the market risk premium. At the same time, the fact that market risk seems to vary through time seems at odds with an approach that ignores the effect of risk on the market risk premium.

---

[5] Although he does not discuss implications for cost of capital estimates, Merton's (1980) tables 4.7 and 4.8 indicate that similar volatility is present in US data. Thus, my results do not appear to be an artifact of NZ data.

In short, the observed volatility in market risk casts doubt on the usual approach that implicitly assumes the market risk premium is a constant. But relaxing this assumption leads to the opposite problem: extreme variation in the market risk premium resulting in alternately high and low values of the cost of capital. The widespread popularity of the CAPM in applications seems to largely reflect a willingness to ignore both empirical reality (time-variation in market risk) and theoretical consistency (the implications of the CAPM for pricing market risk).

References

Black, F., 1993. Estimating expected return. Financial Analysts Journal 49, 36-38.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327.

Campbell, J, A. Lo, and A.C. Mackinlay, 1997. The Econometrics of Financial Markets. New Jersey: Princeton University Press.

Campbell, J. and L. Viceira, 2002. Strategic Asset Allocation. Oxford: Oxford University Press.

Chay, J., A. Marsden and R. Stubbs, 1993. Historical rates of return to equities, bonds, and the equity risk premium: New Zealand evidence. Pacific Accounting Review 5, 27-46.

Cochrane,J., 2001. Asset Pricing. Princeton: Princeton University Press.

Dimson, E., P. Marsh, and M. Staunton, 2002. Triumph of the Optimists. Princeton: Princeton University Press.

Friend, I. and M. Blume, 1975. The demand for risky assets. American Economic Review 65, 900-922.

Huang, C. and R. Litzenberger, 1988. Foundations for Financial Economics. New York: Elsevier.

Lally, M. and A. Marsden, 2004. Tax-adjusted market risk premiums in New Zealand: 1931-2002. Pacific-Basin Finance Journal 12, 291-310.

Merton, R., 1980. On estimating the expected return on the market. Journal of Financial Economics 8, 323-361.

Officer, R., 1973. The variability of the market factor of the New York stock exchange. Journal of Business 46, 434-453.

Tobin, J., 1958, Liquidity preference as behavior towards risk, Review of Economic Studies, 65-86.