ESTIMATING THE MARKET RISK PREMIUM USING HISTORICAL DATA FROM MULTIPLE MARKETS

Martin Lally

School of Economics and Finance

Victoria University of Wellington*

Abstract

This paper has developed an estimator for a country's market risk premium that involves optimally combining an estimate based upon only local historical data over 100 years and the cross-country average. This paper has also compared the combined estimator to that of its two components, and the conclusions are as follows. Firstly, about 30% of the cross-country variation in estimated market risk premiums is due to cross-country variation in true market risk premiums, and therefore the combined estimator places about 30% weight upon the estimator based upon only local data. Secondly, the combined estimator has a variance that is over 30% less than that of the cross-country average and 50% less than the use of only local data; consequently, the usual practice of invoking only local data is significantly inferior to the use of this combined estimator. Furthermore, using data from the first 50 years to forecast the outcome in the last 50 years also reveals the inferiority of using only local data.

JEL: G11, G12
Keywords: Market Risk Premium, Historical Estimates

* PO Box 600, Wellington, New Zealand
Phone: +64-4-463-5998
Fax: +64-4-463-5014
Email: martin.lally@vuw.ac.nz

The helpful comments of participants at a seminar at Victoria University of Wellington are gratefully acknowledged.

1. Introduction

The market risk premium is a critical parameter in the widely used standard version of the Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1965; Mossin, 1966) and in portfolio analysis of the Markowitz type (Markowitz, 1952, 1959). In estimating this parameter for a particular market, a widely used method involves averaging over ex-post outcomes for a long period.[1] Following Ibbotson and Sinquefield's (1976) estimate for the US, such estimates have been generated for a wide range of markets and a range of alternative periods (Siegel, 1992; Dimson and Marsh, 1982; Officer, 1989; Chay et al, 1993). More recently, Dimson et al (2002) have presented estimates for 16 markets over the common period 1900-2000. Mindful that much of the variation in these estimates across markets is likely to be estimation error and these estimation errors across markets are less than perfectly correlated, this set of estimates points towards estimating the market risk premium for each market by simply averaging over the set of estimates. Furthermore, this approach has been proposed by Dimson et al (2002, p 193; 2003, pp. 13-14). However, the cross-country average suffers from the potential problem that true market risk premiums are likely to differ significantly across markets and therefore the true value for a particular market may differ from the cross-country average of the true values[2]. On the other hand, the use of only local data suffers from high standard error in the estimates. All of this suggests that one should form a weighted average of these two types of estimates. Such an approach has been adopted for estimating betas (Vasicek, 1973) and variances (Karolyi, 1993).

In light of all this, the present paper has three goals. The first is to develop an estimator of the market risk premium for an individual market that optimally combines the estimate based upon local historical data and the cross-country average; we call this the combined estimator. The second goal is to empirically assess the advantage (in terms of reduced variance in the estimator) from using the combined

[1] Many other methods have been used, including correction for inflation forecasting errors (Siegel, 1992), the constant reward-to-risk approach of Merton (1980) and the forward-looking approach of Cornell (1999). However, the focus of this paper is on averaging over ex-post outcomes.

[2] The standard version of the CAPM implies that the market risk premium is the product of the market variance and the coefficient of relative risk aversion for the representative investor in that market (Guo and Whitelaw, 2006); both phenomena may differ across markets.

estimator over that of its two components. The third goal is to assess the predictive ability of this estimator and its two components using data from the 1900-1950 period to form the estimates and data from 1951-2000 to test them.

2. Theory

Define ϕ as the average of the true market risk premiums for a set of countries, $\bar{\hat{\phi}}$ as the average of the estimated values using only local data and ϕ_j as the true market risk premium for country *j*. Assuming that we have no information about the divergence of ϕ_j from ϕ, then ϕ_j can be expressed as $\phi_j = \phi + d_j$, where d_j is a random drawing from the cross-sectional distribution of countries' true market risk premiums (with variance denoted σ_d^2). Letting e_j denote the estimation error for country *j*, it follows that the estimated market risk premium for country *j* using only local data is as follows.

$$\hat{\phi}_j = \phi_j + e_j = \phi + d_j + e_j$$

We now consider one of these (randomly selected) countries (country 1). Defining *k* as the weighting applied to $\hat{\phi}_1$, the combined estimator for the market risk premium of country 1 is then as follows

$$k\hat{\phi}_1 + (1-k)\bar{\hat{\phi}} \qquad (1)$$

and the error associated with this estimator for ϕ_1 is as follows.

$$k\hat{\phi}_1 + (1-k)\bar{\hat{\phi}} - \phi_1$$

For a randomly selected country 1, this error is mean zero. So, in choosing the weighting *k*, an appropriate goal would be to minimise the variance of this error. Let *N* denote the number of countries used in forming the cross-country average. Recognising that the random variables $d_1 \ldots d_N$ are mutually independent, that d_j and e_j are independent for each *j*, and that the random variables $e_1 \ldots e_N$ are correlated with

correlation coefficient denoted ρ, the variance of this combined estimator is as follows.

$$Var\left[k\hat{\phi}_1 + (1-k)\bar{\hat{\phi}} - \phi_1\right] = Var\left[k(\phi + d_1 + e_1) + (1-k)\frac{1}{N}\sum_{j=1}^{N}(\phi + d_j + e_j) - (\phi + d_1)\right]$$

$$= Var\left[d_1\left(k + \frac{1-k}{N} - 1\right) + (1-k)\frac{1}{N}\sum_{j=2}^{N} d_j + e_1\left(k + \frac{1-k}{N}\right) + (1-k)\frac{1}{N}\sum_{j=2}^{N} e_j\right]$$

$$= \sigma_d^2\left[\left(k + \frac{1-k}{N} - 1\right)^2 + \frac{(1-k)^2}{N^2}(N-1)\right] + \sigma_e^2\left[\left(k + \frac{1-k}{N}\right)^2 + \frac{(1-k)^2}{N^2}(N-1)\right]$$

$$+ 2\rho\sigma_e^2\left[\left(k + \frac{1-k}{N}\right)\left(\frac{1-k}{N}\right)(N-1)\right] + \rho\sigma_e^2\frac{(1-k)^2}{N^2}\left[(N-1)^2 - (N-1)\right]$$

$$= \sigma_d^2(1-k)^2\left(1 - \frac{1}{N}\right) + \sigma_e^2\left[\frac{1}{N} + k^2\left(1 - \frac{1}{N}\right)\right] + \rho\sigma_e^2\left(1 - k^2\right)\left(1 - \frac{1}{N}\right) \qquad (2)$$

To find the value for k that minimises this function, we set the first derivative to zero and this implies the following[3].

$$k = \frac{\sigma_d^2}{\sigma_d^2 + \sigma_e^2(1-\rho)} \qquad (3)$$

Implementation of this model requires estimates for σ_d^2, σ_e^2 and ρ. The first of these cannot be directly estimated, but can be deduced from the cross-sectional distribution of the locally-based estimates of the market risk premiums for the N countries. Defining V as the expectation of the cross-sectional sample variance in these locally-based estimates of the market risk premiums, it follows that

$$V = E\left[\frac{\sum_{j=1}^{N}(\hat{\phi}_j - \bar{\hat{\phi}})^2}{N-1}\right]$$

$$= \frac{1}{N-1}\sum_{j=1}^{N} E(\hat{\phi}_j - \bar{\hat{\phi}})^2$$

[3] The second derivative is positive, and therefore the second order condition for a minimum is satisfied.

$$= \frac{N}{N-1} E(\hat{\phi}_i - \bar{\hat{\phi}})^2$$

$$= \frac{N}{N-1} E\left[\hat{\phi}_i - \frac{1}{N} \sum_{j=1}^{N} \hat{\phi}_j \right]^2$$

$$= \frac{N}{N-1} E\left[\phi + d_i + e_i - \frac{1}{N} \sum_{j=1}^{N} (\phi + d_j + e_j) \right]^2$$

$$= \frac{N}{N-1} E\left[d_i \left(\frac{N-1}{N} \right) + e_i \left(\frac{N-1}{N} \right) - \frac{1}{N} \sum_{j \neq i} d_j - \frac{1}{N} \sum_{j \neq i} e_j \right]^2$$

Recognising that the random variables d_j and e_j are independent, and that $d_1 \ldots d_N$ are mutually independent, it follows that

$$V = \frac{N}{N-1} \left[\sigma_d^2 \left(\frac{N-1}{N} \right)^2 + \frac{\sigma_d^2}{N^2} (N-1) \right] + \frac{N}{N-1} \left[\sigma_e^2 \left(\frac{N-1}{N} \right)^2 + \frac{\sigma_e^2}{N^2} (N-1) \right]$$

$$+ \frac{N}{N-1} \rho \sigma_e^2 \left[-2 \left(\frac{N-1}{N} \right) \frac{1}{N} (N-1) + \frac{(N-1)^2 - (N-1)}{N^2} \right]$$

$$= \sigma_d^2 + \sigma_e^2 - \rho \sigma_e^2 \qquad (4)$$

Solving this equation for σ_d^2 yields the following result.

$$\sigma_d^2 = V - \sigma_e^2 (1 - \rho) \qquad (5)$$

Estimates for V, σ_e^2 and ρ arise from the set of estimates for the market risk premiums of the N countries. If the true value for σ_e^2 is considered to be the same across all markets, then one should average across the estimates for the individual markets. Similarly, if the true value for ρ is considered to be the same across all pairs of markets, then one should average across all such pairs of estimates.

3. Comparison of Estimators

The previous section has determined the variance of the combined estimator, the optimal value for k and procedures for estimating other relevant parameters. This section now seeks to estimate this variance and to compare it with the variances arising from use of the cross-country average (k = 0) and from the use of only local data (k = 1). To do this, we consider the 16 markets examined in Dimson et al (2002). Table 1 reports the estimates for these markets (column 1) along with their standard errors (column 2), with the data drawn from Dimson et al (2002, Tables 18-1…33-1). In respect of V, an unbiased estimate will arise from the cross-sectional sample variance in the estimated market risk premiums in the first column, and this is .00040. In respect of σ_e^2, a smoothed estimate arises by averaging over the estimates in the second column of Table 1, and this yields .00048. Finally, ρ is estimated by averaging over the time-series correlations between all pairs of markets, using 101 years of data in each case, and the result is .40 (Dimson et al, 2002, Table 8-3). Following equation (5), the estimate of σ_d^2 is then as follows.

$$\hat{\sigma}_d^2 = .00040 - .00048(1 - .40) = .00011$$

So, cross-country variation in true market risk premiums constitutes 28% of the cross-sectional variation in estimated market risk premiums, with the balance due to estimation error. Following equation (3), the optimal value for k matches this, as follows.

$$k_0 = \frac{.00011}{.00011 + .00048(1 - .40)} = .28$$

Following equation (2), the variance of the estimate using the optimal value of k is then .00022. By comparison, the variance arising when only local data are utilised (k = 1) is .00044 and that arising when only the cross-country average is utilised (k = 0) is .00032. Thus, using the optimal value for k lowers the variance by 50% compared to using only local data and by over 30% compared to using the cross-country average. These results and the underlying parameter estimates are summarised in Table 2.

Using the optimal value for k of k_0 = .28, the resulting estimates of the market risk premium for each country are shown in the last column of Table 1. The range of estimates is from .057 (Denmark) to .076 (Germany); using only local data, the estimates range instead from .034 to .105.

The results here are of course sensitive to parameter estimates, and the most problematic estimate is that for ρ. Consequently, sensitivity testing is conducted upon this parameter. Since the estimate invoked above is .40, we consider possible true values of .20, .40 and .60. In respect of σ_e^2, we use the average estimate in Table 2 (.00048) as both an estimate and as the true value. In respect of V, the true value varies with ρ in accordance with equation (4). Thus, we assume a true value for σ_d^2 corresponding to the estimate in Table 2 (.00011). In conjunction with the assumed true value for σ_e^2 and the range of true values for ρ, a set of true values for V arises in accordance with equation (4). The estimates for V are assumed to match these true values. Using these estimates for V along with $\hat{\sigma}_e^2$ = .00048 and $\hat{\rho}$ = .40 yields estimates for σ_d^2 in accordance with equation (5). Using these estimates for σ_d^2 along with $\hat{\sigma}_e^2$ = .00048 and $\hat{\rho}$ = .40 yields estimates for k_0 (denoted $\hat{k}_0$) in accordance with equation (3). By contrast, the true values for σ_d^2, σ_e^2 and ρ yield different estimates for k_0 in accordance with equation (3). The variance of the combined estimator is then determined using both $\hat{k}_0$ and k_0, along with true values for all other parameters. These results are shown in Table 3 along with the underlying estimates of the parameter values and the assumed true values.

The central row of Table 3 shows the situation when ρ is correctly estimated, whilst the first (last) row shows the results when it is overestimated (underestimated). Overestimation of ρ by .20 induces overestimation of k_0 by .19, whilst underestimation of ρ by .20 induces underestimation of k_0 by .34. Whilst these errors in estimating k_0 are substantial, they exert only trivial effects upon the variance of the combined estimator, i.e., the true variance using $\hat{k}_0$ is never significantly larger than that using k_0. Thus, even substantial errors in estimating ρ do not exert a material impact upon the true variance of the combined estimator. This is because errors in

estimating ρ affect the true variance of the combined estimator only via k, and the effects upon k are substantial but the relationship between k and the variance of the combined estimator is quadratic, and therefore the variance of the combined estimator is not significantly affected over a wide range of values for k around the optimal value k_0. Thus, although there is considerably uncertainty about the true value for ρ in this study, this parameter is not critical to the results.

4. Prediction

Dimson et al (2002) provide estimates of market risk premiums using data for both 1900-1950 and 1951-2000. This permits an out-of-sample comparison of various estimators, by using data from 1900-1950 to estimate the market risk premiums for the 16 markets and then compare the resulting estimates to the outcomes for the period 1951-2000. Table 4 reports the Dimson et al estimates for each of these markets for these two subperiods. We consider the following three predictors of the Dimson et al estimates for 1951-2000: the Dimson et al estimate using country-specific data from 1900-1950, the Dimson et al estimates using the 1900-1950 data in conjunction with equation (1) and the optimal value for k, and the average of the Dimson et al estimates for 1900-1950 (.056).

To estimate the optimal value for k following equations (3) and (5), and using data from 1900-1950, we require estimates of the standard errors σ_e^2 for each market. However, Dimson et al (2002) do not provide such estimates for that period. So, for any market, we assume that the estimated standard deviation of annual returns ($\hat{\sigma}^2$) is the same in both the first 50 years (1900-1950) and in the full 100 years. Accordingly, for each market, the estimated standard error for the first 50 years is related to that for the full 100 years as follows.

$$\hat{\sigma}^2_{e(50)} = \frac{\hat{\sigma}^2}{49} = \frac{\hat{\sigma}^2}{99}\left[\frac{99}{49}\right] = \hat{\sigma}^2_{e(100)}\left[\frac{99}{49}\right]$$

Having estimated $\sigma^2_{e(100)}$ at .00048 for each market, it follows that the estimate for $\sigma^2_{e(50)}$ for each market is .00097. In respect of V, the cross-sectional data for 1900-

1950 shown in Table 4 generate an unbiased estimate of .00110. In respect of ρ, we adopt the earlier estimate of .40. Substitution of these results into equation (5) and then (3) yields an estimate for the optimal weight *k* of .47.

In summary then, our three estimators are the Dimson et al estimate using country-specific data from 1900-1950, the Dimson et al estimates using the 1900-1950 data in conjunction with equation (1) and the optimal value for *k* of .47, and the average of the Dimson et al estimates for 1900-1950 (.056), i.e.,

$$\hat{\phi}_1, \quad .47\hat{\phi}_1 + .53(.056), \quad .056$$

Each of these estimators gives rise to a set of estimation errors, and the RMSE of these estimation errors for each of these three estimators are .038, .027 and .025. Thus, using only the data from a market to estimate its market risk premium is markedly inferior to utilising the cross-country average or an appropriate weighted average of the two. Such results are broadly consistent with the conclusion reached earlier in this paper.

Further evidence on this issue arises from conducting a cross-sectional regression of the Dimson et al estimates for 1950-2000 ($\hat{\phi}_{2j}$) on those for 1900-1950 ($\hat{\phi}_{1j}$). The regression model is as follows.

$$\hat{\phi}_{2j} = a + b\hat{\phi}_{1j} + e_j$$

If cross-sectional variation in $\hat{\phi}_j$ is entirely attributable to cross-sectional variation in true market risk premiums, then *a* = 0 and *b* = 1. By contrast, if cross-sectional variation in $\hat{\phi}_j$ is entirely attributable to estimation error, then *a* will be equal to the population mean of the true market risk premiums for 1950-2000 (ϕ_2) and *b* = 0. It follows that if cross-sectional variation in $\hat{\phi}_j$ is partly attributable to cross-sectional variation in true market risk premiums and partly to estimation error, then $0 \le a \le \phi_2$ and $0 \le b \le 1$. Conducting the cross-sectional regression yields $\hat{a}$ = .068 (standard error .009) and $\hat{b}$ = .126 (standard error .137). This clearly rules out the first

hypothesis, and therefore the practice of using only data from one market in estimating the market risk premium from that market. Again, this is consistent with the conclusions reached earlier in the paper.

5. Conclusions

This paper has developed an estimator for a country's market risk premium, which involves optimally combining an estimate using only local historical data for 100 years and the cross-country average. This paper has also compared the combined estimator to that of its two components, and the conclusions are as follows. Firstly, about 30% of the cross-country variation in estimated market risk premiums is due to cross-country variation in true market risk premiums, and therefore the combined estimator places about 30% weight upon the estimator based upon only local data. Secondly, the combined estimator has a variance that is 50% less than that arising from using only local data and over 30% less than that of the world average. Consequently, the usual practice of invoking only local data is significantly inferior to the use of a combined estimator. Furthermore, using data from the first 50 years to forecast the outcome in the last 50 years also reveals the inferiority of using only local data.

Table 1: Estimated Market Risk Premiums for Sixteen Markets

	$\hat{\phi}$	$\sigma(\hat{\phi})$	$\hat{\phi}_c$
Australia	.075	.019	.068
Belgium	.049	.021	.061
Canada	.056	.018	.063
Denmark	.034	.017	.057
France	.074	.021	.068
Germany	.105	.029	.076
Ireland	.048	.017	.061
Italy	.091	.030	.073
Japan	.090	.033	.072
Netherlands	.066	.021	.066
S Africa	.075	.020	.068
Spain	.042	.020	.059
Sweden	.073	.022	.068
Switzerland	.041	.019	.059
UK	.058	.017	.063
US	.069	.020	.066
Mean	.065	.022	.065

This table shows estimates of the market risk premiums for 16 markets using local data for each market ($\hat{\phi}$), the estimated standard errors of these estimates ($\sigma(\hat{\phi})$) and estimates that arise from optimally combining $\hat{\phi}$ and the cross-country average of these estimates.

Table 2: The Variances of Competing Estimators

N	$\hat{V}$	$\hat{\sigma}_e^2$	$\hat{\rho}$	$\hat{\sigma}_d^2$	k_0	$Var(estimator)$		
						$k = 1$	$k = k_0$	$k = 0$
16	.00040	.00048	.40	.00011	.28	.00044	.00022	.00032

This table shows the variance of the estimator for the market risk premium when using only local data (k = 1), when using the cross-country average (k = 0) and when optimally combining these two estimators ($k = k_0$).

Table 3: The Impact of Errors in Estimating the Correlation Coefficient

ρ	σ_e^2	σ_d^2	V	$\hat{V}$	$\hat{\rho}$	$\hat{\sigma}_e^2$	$\hat{\sigma}_d^2$	$\hat{k}_0$	k_0	$\sigma^2 \vert \hat{k}_0$	$\sigma^2 \vert k_0$
.20	.00048	.00011	.00049	.00049	.40	.00048	.00020	.41	.22	.00022	.00020
.40	.00048	.00011	.00040	.00040	.40	.00048	.00011	.28	.28	.00022	.00022
.60	.00048	.00011	.00030	.00030	.40	.00048	.00001	.03	.37	.00040	.00038

This table shows the effects of errors in estimating ρ (the correlation coefficient between the estimated market risk premiums for a pair of markets) on the variance of the estimator for the market risk premium that seeks to optimally combine the estimate based upon only local data and the cross-country average. For a range of values for ρ, the table shows the resulting variance of the estimator ($\sigma^2 \vert k_0$) and also the variance arising from an imperfect estimate of ρ and therefore an imperfect estimate of this optimal weight ($\sigma^2 \vert \hat{k}_0$).

Table 4: Sub-period Estimates of the Market Risk Premiums for Sixteen Markets

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_c$
Australia	.083	.078	.063
Belgium	.040	.055	.051
Canada	.060	.060	.057
Denmark	.027	.040	.048
France	.083	.058	.063
Germany	.115	.083	.072
Ireland	.020	.073	.046
Italy	.102	.067	.069
Japan	.106	.100	.070
Netherlands	.032	.102	.049
S Africa	.060	.081	.057
Spain	.020	.064	.046
Sweden	.048	.100	.054
Switzerland	.010	.068	.043
UK	.033	.080	.049
US	.054	.086	.055
Mean	.056	.075	.056

This table shows estimates of the market risk premiums for 16 markets using local data for each market for 1900-1950 ($\hat{\phi}_1$), local data for each market for 1950-2000 ($\hat{\phi}_2$), and estimates that arise from optimally combining $\hat{\phi}_1$ and the cross-country average of these estimates ($\hat{\phi}_c$).

REFERENCES

Chay, J., Marsden, A. and Stubbs, R. 1993, 'Historical Rates of Return to Equities, Bonds, and the Equity Risk Premium: New Zealand Evidence', *Pacific Accounting Review*, vol.5, pp.27-46.

Cornell, B. 1999, *The Equity Risk Premium*, John Wiley & Sons, New York.

Dimson, E. and Marsh, P. 1982, 'Calculating the Cost of Capital', *Long Range Planning*, vol.15, pp.112-120.

_____ and Staunton, M. 2002, *Triumph of the Optimists*, Princeton University Press.

_____ 2003, 'Global Evidence on the Equity Risk Premium', *Journal of Applied Corporate Finance*, vol. 15 (4), Fall, pp. 27-38.

Guo, H. and Whitelaw, R. 2006, 'Uncovering the Risk-Return Relation in the Stock Market', *The Journal of Finance*, vol. 61, pp. 1433-1463.

Ibbotson, R. and Sinquefield, R. 1976, Stocks, Bonds, Bills and Inflation; Year by Year Historical Returns (1926-1974), Journal of Business, 11–47.

Karolyi, A. 1993, 'A Bayesian Approach to Modelling Stock Return Volatility for Option Valuation', *Journal of Financial and Quantitative Analysis*, vol. 28, pp. 579-594.

Lintner, J. 1965, 'The Valuation of Risky Assets and the Selection of Investments in Stock Portfolios and Capital Budgets', *Review of Economics and Statistics*, vol.47, pp.13-37.

Markowitz, H. 1952, 'Portfolio Selection', *Journal of Finance*, vol.7, pp.77-91.

_____ 1959, *Portfolio Selection*, Yale University Press, New Haven.

Merton, R. 1980, 'On Estimating the Expected Return on the Market', *Journal of Financial Economics*, vol. 8, pp. 323-61.

Mossin, J. 1966, 'Equilibrium in a Capital Asset Market', *Econometrica*, vol.24, pp.768-783.

Officer, R. 1989, 'Rates of Return to Equities, Bond Yields and Inflation Rates: An Historical Perspective', in Ball, R., Finn, F., Brown, P. and Officer, R. editors, *Share Markets and Portfolio Theory*, 2nd edition, University of Queensland Press.

Sharpe, W. 1964, 'Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk', *Journal of Finance*, vol.19, pp.425-442.

Siegel, J. 1992, 'The Equity Premium: Stock and Bond Returns since 1802', *Financial Analysts Journal*, Jan-Feb, pp. 28-38.

Vasicek, O. 1973, ‘A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas’, *Journal of Finance*, vol. 28, pp. 1233-1239.