THE RISK FREE RATE AND THE MARKET RISK PREMIUM

Martin Lally

Associate Professor

School of Economics and Finance

Victoria University of Wellington

23 August 2012

CONTENTS

Executive Summary 3

1. Introduction 6

2. Causes and Consequences of Historically Low Government Bond Yields 6

3. Low Government Bond Yields and Regulatory Under Compensation 8

4. Possible Solutions for Under Compensation 15

4.1 Firm-Level DGM 15

4.2 Market-Level DGM 18

4.3 Long-Run Averaging 23

4.4 Choice of Methodology 26

5. Rounded Estimates of the Market Risk Premium 28

6. Investor Portfolio Reassessments and the Market Risk Premium 31

7. Conclusions 35

References 38

EXECUTIVE SUMMARY

This paper seeks to address a number of questions posed by the QCA, relating to the risk free rate and the market risk premium in Australia, and my conclusions on these matters are as follows.

The first issue was to assess CEG's claims relating to the suitability of CGS as a proxy for the risk free asset within the CAPM, or its use in regulatory situations, in the face of current CGS yields that are historically low. CEG identifies three explanations for the historically low CGS yields, comprising the current low stock of CGS, the recent 'flight to quality', and Basel III requirements, and implies that they each undercut the suitability of CGS as a proxy for the risk free asset. However the current low stock of CGS does not undercut the suitability of CGS as a proxy for the risk free asset because the CAPM imposes no requirements on the supply of the risk free asset. In addition, the recent 'flight to quality' does not undercut the suitability of CGS as a proxy for the risk free asset because the CAPM does not require that the risk free rate be invariant to such events. Finally, CEG's claim that Basel III requirements undercut the suitability of CGS as a proxy for the risk free asset might have some validity but the effect is indeterminable and would be at least partly offset by the presence of default risk in CGS. CEG imply that this 'problem' could be addressed by averaging CGS yields over a longer period. However, in the absence of any attempt by CEG to quantify the problem they allege or any ability to do so, recourse to a clearly radical alternative is not supported. Any argument for using a longer term average of CGS yields must rest on other grounds offered by CEG and these are addressed next.

The second issue was to assess CGS's claim that, regardless of the cause, Australian CGS yields are very low at the present time, that variations in these rates are strongly negatively related to variations in the market risk premium, the general Australian regulatory practice is to combine the current CGS yield with an estimate of the long-term market risk premium, this practice will then produce an estimate of the cost of equity that is too low at the present time, and therefore regulated firms will be under compensated. CEG do not present any persuasive evidence that there is a *strong* negative relationship between the ten year risk free rates and market risk premiums, and the primary evidence they do present in their Figure 8 is pre-disposed to that result. Furthermore, many Australian regulators including the QCA do not estimate the long-term market risk premium but estimate a market risk premium that reflects both prevailing and longer-term factors, and therefore CEG's argument is much less relevant to them than for regulators who estimate a purely long-term market risk premium. Furthermore, the significant issue relating to compensation is over the life of regulated assets rather than only the next few years, and therefore a process for estimating the cost of equity that is biased under some economic conditions but most accurate over the life of regulated assets might still be preferred. Finally, CEG suggests that the generally employed methodology should be abandoned only when conditions are unusual; this approach is highly subjective, may lead to variations only when they favour regulated firms, and therefore is not desirable.

The third issue, assuming that under compensation is present, was to assess whether it should be addressed through any of the three approaches suggested by CEG or by any other approach. Notwithstanding my view that under compensation at the present time is unproven and less important than under compensation over the life of the regulated assets, I have assessed the ability of CEG's three approaches to deal with any such under compensation that might exist at the present time. The first of CEG's proposals, being the Dividend Growth Model (DGM) for estimating an individual firm's cost of equity, is very similar to the DGM for estimating the market risk premium, but has the additional problems of greater exposure to fluctuations in the earnings payout rate, incentives for the firms in question to manipulate their earnings payout rate, and implicitly (and wrongly) assumes that the entire firm's activities are regulated. Consequently, I do not favour this approach. The third of CEG's proposals, in which the average risk free rate over some historical period is substituted for the current rate, suffers from a number of serious problems; these involve overestimating the cost of equity for businesses with equity betas less than 1, wrongly assuming that the widely employed MRP estimate of 6% is an estimate of the long-term MRP, ambiguity over the correct averaging period for the risk free rate, the unsubstantiated belief that variations in the MRP and the risk free rate are offsetting, the sacrifice of an observable, relevant and significant parameter, and potential spillover problems in estimating the cost of debt. I think these problems are sufficiently pronounced that this methodology should not be employed. This leaves CEG's second approach, involving using the DGM to estimate the MRP. Errors in the AMP variant rule this out of consideration, and their effect is to inflate the MRP estimate by about 1%. This point aside, the DGM is worthy of consideration but, in view of its limitations, as a complement to rather than a substitute for other approaches. I therefore favour a risk free rate matching the current rate coupled with an estimate of the market risk premium that draws upon results from various methodologies, and this corresponds to the QCA's current approach. If results from historical averaging contribute to the MRP estimate, the effect of this approach *might* be to generate an estimate of the MRP that is biased down at the present time but the mean squared error of the estimator is the more important consideration and it favours the QCA's approach. Furthermore, any downward bias at the present time is likely to be offset over the entire life of the regulated assets.

The fourth issue is whether the QCA should continue to round its MRP estimate to 1% or to some (lower) unit of rounding. There is a gain in accuracy from a lower unit of rounding but the gain is very small. Furthermore a lower unit of rounding increases the frequency of unwarranted changes in the rounded estimate, and prompts more lobbying for changes in parameter values within a methodology or changing the set of methodologies used to estimate the market risk premium. Since these disadvantages from a lower unit of rounding than 1% seem substantial, and the benefits so small, I recommend continued use of rounding to 1%.

The fifth issue is that of estimating the across-investor average period between successive portfolio reassessments. This period is not amenable to a precise estimate, but a reasonable estimate is at least one year and could easily be ten years. The choice then rests upon more pragmatic considerations, with historical data availability for the risk free rate pointing to a ten year period whilst regulatory considerations (in the form of the typical regulatory period being five years) suggest a figure of five years. On balance, I think data availability is the more significant issue, and this favours treating the across-investor average period between successive portfolio reassessments as ten years. However, with rounding to the nearest 1%, the choice of five versus ten years is very unlikely to affect the rounded result.

The final issue is whether the QCA should, at the start of each regulatory cycle, review its estimate of the across-investor average period between successive portfolio reassessments. In view of the difficulties of estimating this parameter, I do not favour so frequent a review of it.

1. Introduction

This paper seeks to address a number of questions posed by the QCA. Some of these issues relate to the risk free rate and arise from concerns raised by CEG (2012). The remaining issues relate to the market risk premium.

2. Causes and Consequences of Historically Low Government Bond Yields

CEG (2012, section 5) argues that Australian CGS yields are unusually low at the present time, at least partly because the stock of such debt is low, because the desire for low-risk liquid assets is high ('flight to quality'), and because Basel III will require banks to increase their holdings of low risk liquid assets such as CGS. Although not stated by CEG, there is a suggestion that CGS yields are therefore a biased down estimate of the risk free rate and the first of these arguments (that the low current stock of CGS induces downward bias of this type) has been previously presented by NERA (2007a, 2007b). CEG also imply that these alleged problems could be addressed by averaging CGS yields over a longer period.

To assess these claims, it is necessary to consider the context within which the risk free rate is being sought. This context is that of the Officer (1994) version of the Capital Asset Pricing Model. The CAPM embodies the concept of a risk free asset, but it does not designate any particular asset of this type. In choosing an asset to provide the risk free rate, the only explicit requirement within the CAPM is that the rate of return on that asset be free of risk. There is an implicit requirement relating to liquidity, i.e., a very illiquid asset would be unsuitable because illiquidity is (inter alia) a manifestation of high transaction costs and the CAPM assumes that there are no transactions costs. In addition, there is an implicit requirement that no investor faces restrictions upon the purchase of this asset because the model assumes that no such restrictions exist. In addition, there is an implicit requirement that investors are not attracted to or repelled from the asset for reasons other than the probability distribution on its return, because the model assumes that investors choose portfolios solely according to their return distributions.

By contrast, the CAPM does not impose any requirements whatsoever (whether explicit or implicit) relating to the *stock* of the risk free asset, i.e., it does not require the stock of the risk free asset to meet some minimum level, as CEG seems to believe. This follows from the fact that the supply of the risk free asset is exogenous to the CAPM (see Mossin, 1966, pp. 772-773; Hirshleifer, 1970, pp. 299-300). Thus, whilst a reduction in the supply of government bonds may lower their yield, it does not disqualify such bonds as a suitable proxy for the risk free asset within the context of the CAPM. Furthermore, the CAPM does not require (whether explicitly or implicitly) that the risk free rate be invariant to changes in the risk of other assets or to investors' aversion to such risks. Thus, even if the risk of equities has increased or investors have become more averse to such risks, leading to heightened demand for CGS and therefore a lower yield on them, this does not preclude CGS from being a suitable proxy for the risk free asset. Such changes in risks or risk aversion (the 'flight to quality') are simply part of the financial landscape.

Turning back to the properties that *should* be satisfied by the risk free asset, as noted above these are as follows:

(a) the return on the asset is certain
(b) the asset is liquid
(c) there are no restrictions upon the purchase of the asset by any investor
(d) investors are not attracted to or repelled from the asset for reasons other than the probability distribution on its return.

CEG's (plausible) claim that Basel III regulations will induce a heightened demand for assets like CGS could be viewed as a violation of requirement (d), and the effect of such would be to reduce the CGS yield. Accordingly, the CGS yield would be a downward biased estimate of the risk free rate, but the magnitude of the effect is indeterminable. Furthermore it is also true that CGS violate requirement (a), because there is some default risk on CGS and the effect of this is that the CGS yield would overstate the risk free rate. In view of Australia's sovereign credit rating (AAA), this effect is likely to be very small. Thus, notwithstanding CGS's plausible claims about Basel III, it does not follow that there is any net bias in CGS as a proxy for the risk free rate and any net bias that does exist is indeterminable.

In conclusion, CEG's argument that the current low stock of CGS undercuts its suitability as a proxy for the risk free asset is invalid because the CAPM imposes no requirements of this type upon the risk free asset. In addition, CEG's argument that the recent 'flight to quality' undercuts the suitability of CGS as a proxy for the risk free asset is also invalid because the CAPM does not require that the risk free rate be invariant to such events. Finally, CEG's argument that Basel III requirements undercuts the suitability of CGS as a proxy for the risk free asset might have some validity but the effect is indeterminable and would be at least partly offset by the presence of default risk in CGS. CEG (2012, section 6.4) also imply that this 'problem' could be addressed by averaging CGS yields over a longer period. However, in the absence of any attempt by CEG to quantify the problem they allege or any ability to do so, recourse to a clearly radical alternative is not supported. Any argument for using a longer term average of CGS yields must rest on other grounds offered by CEG, and these are addressed in the next section.

3. Low Government Bond Yields and Regulatory Under Compensation

CEG (2012, sections 2-5) also argue that, regardless of the cause, Australian ten-year CGS yields are very low at the present time, that variations in these rates are strongly negatively related to variations in the ten-year market risk premium, the general Australian regulatory practice is to combine the current CGS yield with an estimate of the long-term market risk premium (of 6%), this practice will then produce an estimate of the cost of equity that is too low at the present time, and therefore regulated firms will be under-compensated.

This line of argument contains three premises and two sequential conclusions. CEG's first premise is that Australian ten-year CGS yields are very low at the present time, and this is clearly correct. CEG's second premise is that ten-year CGS yields are strongly negatively related to variations in the ten-year market risk premium. Although there is nothing in finance theory that supports (or rejects) a negative relationship between CGS yields and the market risk premium, such a relationship is plausible because the market risk premium is compensation for bearing equity risk (Merton, 1980), equity risk (volatility) seems to be greatest in depressed economic conditions (French et al, 1987, Figure 1a and 1b), and the risk free rate also tends to be lowest in depressed economic conditions (due to depressed aggregate demand and to monetary policy). However, whilst CGS yields are very low because of generally depressed world economic conditions, Australia is not experiencing depressed economic conditions. Even in relation to the probable general tendency to negative correlation, the significant issue for regulatory purposes is the *strength* of this relationship and especially its strength in respect of the ten year risk free rate and the ten year market risk premium (because CEG's analysis focuses upon these two parameters). Market volatility (and therefore the market risk premium) might be high today but volatility (and hence the market risk premium) tends to rapidly subside to normal levels (French et al, 1987, Figure 1a) and the market risk premium for the next ten years might not then be greatly increased by a temporary upsurge in it.

CEG present various types of evidence in support of their claim for a strong negative relationship between risk free rates and the market risk premium. CEG (2012, paras 42-43) cite Lettau and Ludvigson (2001) in claiming that "when the de-trended risk free rate fell the (market) risk premiums tended to rise by the same amount". However CEG do not identify any particular section of the Lettau and Ludvigson paper that supports this specific assertion.[1] Furthermore, the risk free rate used by Lettau and Ludvigson is the US 30-day Treasury Bill rate (ibid, page 825) rather than the ten-year rate. In addition the "risk premiums" referred to only changed in the opposite direction to that of the T/Bill rate over the following two years, after which they moved in the reverse direction (ibid, Table VI) contrary to CEG's claim. Furthermore, these "risk premiums" are in fact actual equity returns net of the T/Bill rate, and therefore the relationship uncovered may simply reflect market inefficiency rather than changes in risk premiums, i.e., the increases in equity returns net of the T/Bill rate subsequent to low risk free rates may reflect market undervaluation of equities at the time of the low risk free rates (when economic conditions are adverse). So, the Lettau and Ludvigson paper does *not* support the claim that a fall in the ten year risk free rate will be followed by a rise in the ten year risk premium.

CEG (2012, para 44) also cite Smithers and Co (2003, page 49) in support of the claim that the risk free rate moves inversely with the market risk premium. In turn Smithers and Co reach this view based upon the observation that the real return on stocks over the last 100-200 years has been much more stable than the real risk free rate, and they refer to this as "Siegel's Constant" (ibid, pp. 31-38). This view presumably derives from Siegel (1992, 1999), and it implies that one should estimate the market risk premium from the long-run real market return net of a current estimate of the long-run real risk free rate.[2] This methodology is one of the four approaches used by the QCA (2011, pp. 238-240). Thus the essential difference between Smithers and Co and the QCA is that the former favours one methodology (Siegel, 1992) whilst the latter considers results from a range of methodologies. All methodologies

[1] A request to CEG for clarification on this matter was unsuccessful.

[2] Long-run averaging is a suitable method for estimating a parameter only if the true value of the parameter does not significantly change over time. So, if the expected real return is more stable over time than the expected return net of the risk free rate, one should estimate the former rather than the latter from past data.

have their advantages and disadvantages and I concur with the QCA's recourse to results from a range of approaches. Furthermore, despite favourably citing Smithers and Co, CEG fail to note that that the approach favoured by Smithers and Co does not correspond to any of the approaches recommended by CEG (and these will be examined in the next section).

CEG (2012, paras 46-63) also describe the general increase in debt risk premiums on non-CGS bonds at the time of the recent fall in the risk free rate, which is uncontroversial, and they claim that "standard finance theory" would support an increase in the equity risk premium of at least that in debt risk premiums. Subsequently, CEG (2012, para 96) explain this with an example involving Victorian state government debt, for which the debt risk premium increased from 0.51% in 2011 to 0.80% in 2012. Assuming a market risk premium of 6.0% in 2011, CEG claim that the debt risk premium in 2011 of 0.51% implies a debt beta of 0.09, and coupling the same debt beta with the 2012 debt risk premium of 0.80% implies a 2012 market risk premium of 9.0%. However, these results are only true if the cost of debt is an expected rate of return and the margin over the risk free rate is compensation for only systematic risk, and both conditions are false. In particular, the cost of debt is a promised rate of return and this exceeds the expected rate of return by the expected default losses (*DF*). Furthermore, the expected rate of return on state government debt is likely to incorporate an allowance for inferior liquidity relative to CGS (*LIQ*). Thus the debt risk premium (*DRP*) can be expressed as

$$DRP = MRP\beta_d + DF + LIQ$$

where *MRP* is the market risk premium and β_d is the debt beta. Accordingly, the rise in the *DRP* on Victorian state government debt from 0.51% in 2011 to 0.80% in 2012 may have been due entirely to increases in *DF* and *LIQ*, in which case one cannot deduce that *MRP* rose. Remarkably, CEG (2012, para 55) refer to the rise in the *DRP* for state government debt and attribute this to a "heightened safety/liquidity/scarcity premium for CGS", and this seems to involve acknowledging that *DF* and *LIQ* might have risen. Thus, the evidence presented by CEG for a rise in the market risk premium is not compelling, there are credible alternative explanations, and even CEG elsewhere acknowledge these alternative explanations.

CEG (2012, paras 67-71) also generate a time-series of estimates of the market cost of equity over the past 20 years, as shown in their Figure 8, and argue that the stability in this time series in the face of considerable variation in the ten-year risk free rate implies that the market risk premium changes in an approximately offsetting fashion to the ten year risk free rate. However, in estimating this cost of equity by matching the present value of future dividends to their current market value, CEG assumes that at any point in time the market cost of equity is the same for all future years. Thus, if the current ten year risk free rate were unusually low relative to its long-term average (as is clearly the case), CEG implicitly believes that the market risk premium over the next ten years would be unusually high (relative to its long-term average) by an exactly offsetting amount. With this 'perfect-offset' assumption, CEG then generate results showing the stability of the cost of equity over time. However the 'perfect-offset' assumption necessarily leads to greater stability over time in the estimated cost of equity than would otherwise arise. Consequently this critical piece of evidence is prejudiced in favour of the result that is found.

To illustrate the point that the 'perfect-offset' assumption dramatically dampens variation over time in the estimated market cost of equity, suppose the market dividends in the most recent year are denoted *D* and, at any point in time, are expected to grow at 5% per year in perpetuity. Suppose further that the long-run average for the ten-year risk free rate is 5% and therefore any deviations from this give rise to the expectation of a reversion back to 5%. Suppose further that the market risk premium does not vary from 6%, so that any variations in the risk free rate from its long-run average do not induce countervailing changes in the market risk premium. Suppose further that the current ten-year risk free rate is unusually high at 7%, and therefore is expected to revert to 5% in ten years time. The current market cost of equity is then 13% and is expected to revert to 11% in ten years time. Accordingly the market value of equities *S* will be as follows:

$$S = \frac{D(1.05)}{1.13} + ... + \frac{D(1.05)^{10}}{(1.13)^{10}} + \frac{E(S_{10})}{(1.13)^{10}}$$

$$= \frac{D(1.05)}{.13-.05}\left[1-\left(\frac{1.05}{1.13}\right)^{10}\right] + \frac{\left[\frac{D(1.05)^{11}}{.11-.05}\right]}{(1.13)^{10}}$$

Per $1 of *D* this is $15.22. By contrast the process used by CEG to estimate the market cost of equity for the next ten years (*k*) assumes that all subsequent values for *k* are equal:

$$S = \frac{\$1(1.05)}{1+k} + \frac{\$1(1.05)^2}{(1+k)^2} + ... = \frac{\$1(1.05)}{k-.05}$$

Substituting S = \$15.22 into the last equation, the resulting estimate for k of 11.9%, and this is below the true value of 13% because CEG assume k is the same for all future years. The process is now repeated with a current ten-year risk free rate of 3%, which is expected to revert to 5% in ten years. So, with a market risk premium of 6%, the current market cost of equity is 9%, which is expected to revert to 11% in ten years. Following the same process as above, CEG's approach would then estimate the current market cost of equity at 10.2%. Thus the true current market cost of equity has varied from 9% to 13% whilst the estimate of it using CEG's methodology has varied from only 10.2% to 11.9% despite the fact that the market risk premium has not changed as the risk free rate has changed. So, if one observes little variation over time in the cost of equity estimated through CEG's approach, one cannot conclude that the market risk premium moves inversely with the risk free rate; most of the explanation for the stability in the estimated cost of equity arises from the assumption that, at any point in time, the cost of equity is the same for all future years.

CEG's third premise is that the general practice of Australian regulators is to estimate the long-term market risk premium (at 6%). However, this belief is not correct. For example, the QCA estimates the MRP at 6% by considering results from four different methodologies and only two of these involve long-run historical data (QCA, 2011, pp. 238-240). In addition, the AER (2012, Attachments, pp. 120-136) estimates the MRP at 6% using results primarily from two different methodologies and only one of these involves long-run historical data. Of course, this does not imply that the figure of 6% is an unbiased estimate of the market risk premium at the present time; the significant weight on results based upon historical averaging may induce a downward bias in depressed economic conditions. However, the effect of this point is that any downward bias at the present time would be less significant than otherwise.

CEG's final conclusion is that any underestimate of the cost of equity at the present time leads to under compensation for regulated firms. However the critical feature of compensation is that it should be provided over the life of the regulated assets rather than over each regulatory cycle within the life of the assets, and there are trade-offs involved here, i.e., an estimation process that yields the best results over the life of regulated assets might yield inferior results over any regulatory cycle, and should therefore still be preferred. To illustrate this point, suppose that there are two methods of estimating the cost of equity and that both use the observed risk free rate and correctly estimate the equity beta (at 1) but differ in their methodology for estimating the market risk premium. So, the difference between the two methods lies only in their estimate of the market risk premium. The first method estimates the long-run market risk premium, and does so accurately, but the true market risk premium fluctuates around this long-run value with equal probability of equalling its long-run value (6%) when economic conditions are normal, exceeding it (7%) when economic conditions are depressed, and being less than it (5%) when economic conditions are buoyant. So, if economic conditions are depressed, the market risk premium will be underestimated by 1%, and if economic conditions are buoyant, it will be overestimated by 1%. The method is therefore biased under some economic conditions but unbiased over the life of the regulated assets and has a standard deviation in its estimation errors of 0.7%. The second methodology for estimating the market risk premium is unbiased at each point in time but is equally likely to produce an estimate that equals the true value, exceeds it by 3%, and is less by 3%. This method is unbiased at each point in time, and therefore also unbiased over the life of the regulated assets, but the standard deviation of the estimation errors is 2.0%, i.e., three times that of the first method. So, both methods are unbiased over the life of regulated assets, the first (but not the second) method is biased under some economic conditions, and the first method has a considerably lower standard deviation over the life of the regulated assets than the second method. Since accuracy over the life of the regulated assets is more important than accuracy over shorter periods the first method is superior. The example here is not intended to represent the merits of historical averaging versus a forward-looking method. However it does demonstrate that one should not reject a method for estimating the market risk premium merely because it is biased under some economic conditions. The more important consideration is the accuracy of the method over the life of the regulated assets.

In addition to these points, CEG (2012, para 150) suggests that the generally employed methodology should be abandoned only when conditions are unusual. This is a highly subjective approach, and may lead to departures from the usual approach generally only when it favours regulated firms because they will be motivated to lobby for change only in these circumstances. In this event, an upward bias would be imparted to the estimated cost of equity over the life of the regulated assets. To illustrate this point, suppose that the market risk premium has equal probability of equalling its long-run value (6%) when economic conditions are normal, exceeding it (7%) when economic conditions are depressed, and being less than it (5%) when economic conditions are buoyant. Assume also that the methodology for estimating the long-run market risk premium accurately estimates this long-run value. So, if economic conditions are depressed, the market risk premium will be underestimated by 1%, and if economic conditions are buoyant, it will be overestimated by 1%. The method is therefore biased under some economic conditions but unbiased over the life of the regulated assets. By contrast, if this methodology is abandoned when economic conditions are depressed, in favour of a method that produces an unbiased estimate of the market risk premium in these circumstances (at 7%), but otherwise the long-run estimate of 6% is used, then the average result from this approach will be 6.3% and therefore upward bias over the life of the regulated assets will occur.

In summary, CEG argue that CGS yields are very low at the present time, that variations in these rates are strongly negatively related to variations in the market risk premium, that the general Australian regulatory practice is to combine the observed CGS yield with an estimate of the long-term market risk premium, that this practice will then produce an estimate of the cost of equity that is too low at the present time, and therefore regulated firms will be under-compensated. However, CEG do not present any persuasive evidence that there is a *strong* negative relationship between the ten year risk free rates and market risk premiums, and the primary evidence they do present in their Figure 8 is pre-disposed to that result. Furthermore, the QCA and the AER do not estimate the long-term market risk premium but estimate a market risk premium that reflects both prevailing and longer-term factors, and therefore CEG's argument is much less relevant to them than for regulators who estimate a purely long-term market risk premium. Furthermore, the significant issue relating to compensation is over the life of regulated assets rather than the next few years, and therefore an estimator for the cost of equity that is biased under some economic conditions but most accurate over the life of regulated assets might still be preferred. Finally, CEG suggests that the generally employed methodology should be abandoned only when conditions are unusual; this approach is highly subjective, may lead to variations only when they favour regulated firms, and therefore is not desirable.

4. Possible Solutions for Under Compensation

4.1 Firm-Level DGM

CEG (2012, section 7) presents three possible approaches to estimating the cost of equity that allegedly overcome the under-compensation problem discussed in the previous section, and these are now examined. The first of these approaches is the DGM applied to each of six Australian regulated businesses, which estimates the cost of equity consistent with the current share price, the current dividend level, and estimates of future expected dividends per share. For each company, CEG estimates the future expected dividends per share using Bloomberg forecasts for the first two years followed by a long-run growth rate of either 2.5% (expected future inflation) or 6.6% (expected future GDP growth rate). Across the six companies the average cost of equity varies from 10.87% to 14.59% according to whether the expected dividend growth rate is 2.5% or 6.6% respectively (CEG, 2012, section 4.4).[3]

This methodology has the advantage of reflecting current market conditions but is subject to a number of difficulties. CEG (2012, para 155) refers to the possible lack of credible short-term dividend forecasts. However there are more serious concerns. Firstly, the estimated cost of equity for a company is assumed to be the same for all future years. Thus, if the current ten year risk free rate were unusually *low* relative to its long-term average (as is the case) and therefore could be expected to be higher in ten years' time, this methodology implicitly assumes that the equity risk premium for the company over the next ten would be unusually *high* relative to its long-term average by an exactly offsetting amount. This 'perfect-offset hypothesis' is implausible and, since the current risk free rate is unusually low, will overestimate of the cost of equity for the next ten years.[4]

To illustrate this point, suppose that the current ten year risk free rate is 3.8%, the company's equity risk premium over the next ten years is 6.4% and therefore the current cost of equity over the next ten years is 10.2%. Since the risk free rate is unusually low, the rate expected in ten years should be higher and we assume it equals its long-term average of (for example) 6%. In addition, since the risk free rate is expected to rise, the company's equity risk premium might be expected to fall, and we therefore assume it is expected to fall to its long-

[3] In respect of the share prices, current dividend level and Bloomberg forecasts, these are not presented by CEG and a request to CEG for the computational spreadsheets to check this data and the calculations was unsuccessful.

[4] Evidence of its implausibility will be discussed in the next section.

term average of (for example) 6%. In addition, the expected growth rate in dividends is 5% per year in perpetuity. It follows that the current share value of the company (*P*) is as follows:

$$P = \frac{D(1.05)}{1.102} + ... + \frac{D(1.05)^{10}}{(1.102)^{10}} + \frac{E(P_{10})}{(1.102)^{10}}$$

$$= \frac{D(1.05)}{.102-.05}\left[1-\left(\frac{1.05}{1.102}\right)^{10}\right] + \frac{\left[\frac{D(1.05)^{11}}{.12-.05}\right]}{(1.102)^{10}}$$

where *D* is the dividends per share in the most recent year. Per \$1 of *D*, the current share value is then \$16.99. By contrast, the DGM approach assumes that the market cost of equity *k* is the same for all future years. Consequently, with *P* = \$16.99 as above, the DGM model would estimate the company's cost of equity as the solution to the following equation:

$$\$16.99 = \frac{\$1(1.05)}{1+k} + \frac{\$1(1.05)^2}{(1+k)^2} + ... = \frac{\$1(1.05)}{k-.05} \qquad (1)$$

Solving this equation yields an estimate for *k* of 11.2%, which is assumed to hold for all future years. This is 1.0% above the actual cost of equity for the first ten years of 10.2%, and the error arises from assuming the same cost of equity for all future years when the rate actually differs over future years.

Secondly, this methodology assumes that the current share price of the company matches the present value of future dividends per share. Consequently, if the current share price exceeds the present value of future dividends, then the estimate for the cost of equity that arises from this methodology will be too low.[5] Similarly, if the current share price is below the present value of future dividends, then the estimate for the cost of equity that arises from this methodology will be too high. To illustrate the possible extent of the errors, suppose that the current share price of the company is 25% below the present value of future dividends. This would reduce the left-hand side of equation (1) by 25%, and solving for *k* in this new situation would then yield an estimate of 13.2%. This contrasts with the estimate of 11.2% if the current equity value of the company matched the present value of future dividends, and

[5] Smithers and Co (2003, page 49) make this same point. CEG (2012, para 44) cite them in support of a different point (on the same page) whilst ignoring their concern about the DGM on this matter.

therefore the cost of equity would be overestimated by 2.0% as a result of this point (as well as a further 1.0% as a result of the previous point).

Thirdly, the DGM methodology is error-prone in the presence of fluctuations in the firm's earnings retention rate. For example, consider a firm with a cost of equity of 10% per year in perpetuity, a current annualised dividend level of $1 per share, and an expected growth rate in dividends per share of 5% per year in perpetuity (arising from the Bloomberg forecasts for the next two years and the expected long-run GDP growth rate). The share price *P* would then be as follows:

$$P = \frac{\$1(1.05)}{1.10} + \frac{\$1(1.05)^2}{(1.10)^2} + ... = \frac{\$1(1.05)}{.10 - .05} = \$21$$

Application of the DGM methodology would then accurately estimate the firm's cost of equity at 10%, by solving the following equation:

$$\$21 = \frac{\$1(1.05)}{1+k} + \frac{\$1(1.05)^2}{(1+k)^2} + ... = \frac{\$1(1.05)}{k - .05}$$

Now suppose instead that the firm decides to reduce its retention (and hence investment) level over the next five years, and that the effect of this will be to raise its expected dividends (relative to the above path) to $1.60 per share for each of the next five years, followed by subsequent reductions in expected dividends (relative to the above path). Suppose also that the cancelled investments are NPV neutral. Since the investments are NPV neutral, the share price of $21 would not be affected by this new policy. In addition the expected long-run GDP growth rate would still be 5%. However the Bloomberg dividend forecasts for the next two years would now be $1.60 for each year. So, application of the DGM with Bloomberg's dividend forecasts for the next two years followed by a long-run expected growth rate of 5% would yield the following

$$\$21 = \frac{\$1.60}{1+k} + \frac{\$1.60}{(1+k)^2} + \frac{\$1.60}{(1+k)^3} + \frac{\$1.60(1.05)}{(1+k)^4} + ...$$

Solving this equation yields an estimate for *k* of 12%, which is 2% above the correct value of 10%. So the DGM methodology coupled with Bloomberg's dividend forecasts for the next two years followed by a long-run expected growth rate in GDP could produce very significant errors in estimating the firm's cost of equity. The source of the problem is the fact that the dividends per share of $1.60 arise from a *temporary* reduction in the firm's retention rate and therefore are not a suitable base from which to project subsequent dividends.

An alternative scenario of this general type would involve the firm paying dividends in excess of its free cash flow and borrowing to meet the cash flow shortfall. If the borrowing required here were sufficiently large to progressively raise the firm's leverage, then such a strategy would not be indefinitely sustainable, and therefore dividends will fall at some point. This scenario characterises the behaviour of a number of UK water utilities (Armitage, 2012). The DGM methodology would again not allow for the future dividend reduction and would therefore again overestimate the firm's cost of equity.

Fourthly, and consequent upon the previous point, if this DGM approach were used to estimate the cost of equity for regulated firms, the firms in question would have a very strong incentive to manipulate their retention rates for the purpose of increasing their estimated cost of equity.

Fifthly, the methodology produces an estimated cost of equity for the company and therefore will not accurately estimate the cost of equity of the regulated activities of the company if they represent only part of its activities. Furthermore, since non-regulated activities tend to have higher risk, the estimated cost of equity for the company will tend to overestimate that for its regulated activities, and the AER (2012, Attachments, pp. 159-161) notes that the six companies analysed by CEG have unregulated activities.

4.2 Market-Level DGM

The second approach considered by CEG (the AMP method) is to use the DGM model to estimate the cost of equity for the market in aggregate, using an expected growth rate in dividends of 6.6% and a dividend yield of 5.68%, and the result is 12.28% (CEG, 2012, Table 4 and section 4.3). The prevailing ten-year risk free rate of 3.77% is then deducted from this to yield the estimated market risk premium of 8.52%. This estimate is then substituted into the CAPM in the usual way, along with the prevailing risk free rate of 3.77% and the estimated equity beta of 0.8, to produce an estimate for the cost of equity for a regulated business.

Clearly CEG intend that the expected growth rate in dividends of 6.60% applies immediately and therefore the value of equities (S) is represented as follows:

$$S = \frac{D(1+g)}{1+k} + \frac{D(1+g)^2}{(1+k)^2} + ... = \frac{D(1+g)}{k-g}$$

where D is the dividends in the most recent year, g is the expected growth rate, and k is the market cost of equity capital. Solving this equation for k then yields

$$k = \frac{D}{S}(1+g) + g \qquad (2)$$

Substituting CEG's parameter values into equation (2) then yields a value for k of 12.65%, and deduction of the risk free rate of 3.77% then yields an estimate for the market risk premium of 8.89% rather than the figure of 8.52% claimed by CEG.[6]

This error aside, the methodology has the advantage of reflecting current market conditions but is subject to a number of difficulties. Firstly, CEG's assumption that the expected long-term growth rate in dividends per share (for existing companies) will match that for GDP is indefensible. If it were true then, since the expected long-term growth rate in all dividends from all companies exceeds that for dividends per share in existing companies (due to new share issues net of buybacks and also to the formation of new companies), the expected long-term growth rate in all dividends from all companies would exceed that for GDP, and therefore dividends in absolute terms would eventually exceed GDP in absolute terms. This is impossible. So, instead, it would be more reasonable to assume that the long-term growth rate in dividends for all companies will match that for GDP (to ensure that the ratio of dividends to GDP does not eventually reach zero or exceed 1). It follows that the expected growth rate in dividends per share for existing companies will be less than that for GDP, to reflect the impact of new share issues (net of buybacks) and the formation of new companies.

[6] CEG's error was to have overlooked the (1+ g) component in equation (2).

Bernstein and Arnott (2003) argue for subtracting 2% to deal with both of these points, based upon two comparisons. The first comparison is of real growth in dividends per share with real GDP growth over the last century, for a range of countries; the latter grew about 2% faster than the former (ibid, Table 1). However this comparison will exaggerate the relevant adjustment in the presence of a declining dividend payout rate, which has characterised at least the US market. Their second comparison is of the growth in market capitalisation with the growth in a capitalisation-weighted price index, using US data since 1925; the former grew about 2% faster than the latter. However, this comparison will exaggerate the relevant adjustment when market capitalisation grows simply due to listings from foreign firms and from previously unlisted US firms. Both points suggest that the correct adjustment is less than 2%. If we deduct 1% from the expected long-term growth rate in GDP, the estimate for the expected growth rate in dividends would then be 5.6%, and substitution of this into equation (2) followed by deduction of the risk free rate yields an estimated market risk premium of 7.82% rather than 8.89%. If the deduction is instead 1.5%, to yield an expected growth rate in dividends of 5.1%, then the estimated market risk premium falls further to 7.3%.

Secondly, the cost of equity for the market is assumed to be the same for all future years. Thus, if the current ten year risk free rate were unusually high relative to its long-term average, and therefore could be expected to be lower in ten years' time, this methodology implicitly assumes that the market risk premium over the next ten will be unusually low relative to its long-term average by an exactly offsetting amount. This 'perfect-offset' hypothesis is implausible, and even stronger than CEG's explicitly stated view that risk free rates and market risk premiums are negatively (but imperfectly) correlated. It is also testable using the time-series of risk free rates coupled with CEG's estimates of the market risk premium that are based on the 'perfect-offset' hypothesis (CEG, 2012, Figure 8). Figure 8 shows that, in 1994, CEG's estimate of the market cost of equity was about 10.5%, matching the contemporaneous (and unusually high) ten year risk free rate, and therefore implying a market risk premium of zero. Clearly, an estimate for the market risk premium of zero is implausible and it suggests that CEG's belief that, at any point in time, the market cost of equity over the next ten years will match the expected rate thereafter can be rejected. The much more plausible hypothesis is that, in 1994, the market cost of equity over the following ten years was larger than over subsequent years and therefore larger than CEG's estimate of 10.5%. Accordingly, in 1994, the market risk premium over the following ten years was positive rather than zero. Thus, if the perfect-offset hypothesis should be rejected in 1994 when the risk free rate was unusually high, it should also be rejected in 2012 when the risk free rate was unusually low. So, in 2012, the market cost of equity over the following ten years would be less than over subsequent years, and therefore less than CEG's estimate. Accordingly, the estimated market risk premium over the next ten years is less than CEG's estimate.

To illustrate CEG's overestimation of the market risk premium for the next ten years, suppose that the current ten year risk free rate is 3.8%, the market risk premium over the next ten years is 6.4% and therefore the current cost of equity over the next ten years is 10.2%. Since the risk free rate is so low, the rate expected in ten years should be higher and we assume it equals the long-term average of (for example) 6%. In addition, since the risk free rate is expected to rise, the market risk premium might be expected to fall, and we therefore assume it is expected to fall to its long-term average of (for example) 6%. In addition, we assume an expected long-run growth rate in dividends of 5%. It follows that the current value of equities is as follows:

$$S = \frac{D(1.05)}{1.102} + ... + \frac{D(1.05)^{10}}{(1.102)^{10}} + \frac{E(S_{10})}{(1.102)^{10}}$$

$$= \frac{D(1.05)}{.102-.05}\left[1-\left(\frac{1.05}{1.102}\right)^{10}\right] + \frac{\left[\frac{D(1.05)^{11}}{.12-.05}\right]}{(1.102)^{10}}$$

Per $1 of initial dividends, the current equity value is then $16.99. Substitution of this equity value into the DGM equation (2) along with the expected growth rate of 5% then yields an estimated cost of equity of 11.2%, and deduction of the current risk free rate of 3.8% then yields an estimated market risk premium of 7.4%. This is 1.0% above the actual market risk premium of 6.4% for the first ten years, and the error arises from assuming the same cost of equity for all future years when the rate actually differs over future years.

Thirdly, this methodology assumes that the current value of the market matches the present value of future dividends. Consequently, if the current value of the market exceeds the present value of future dividends, then the estimate for the market cost of equity (and hence the market risk premium) that arises from this methodology will be too low. Similarly, if the current value of the market is below the present value of future dividends, then the estimate for the market cost of equity (and hence the market risk premium) that arises from this methodology will be too high. To illustrate the possible extent of the errors, suppose that the current value of the market is 25% below the present value of future dividends. In addition, consistent with CEG, suppose that the expected growth rate in dividends is 6.60%, the current dividend yield is 5.68%, and the current ten year risk free rate is 3.77%. These parameters in conjunction with equation (2) imply that the market risk premium is estimated at 8.89%. However, if the current value of the market matched the present value of future dividends, the estimate of the market risk premium would be 7.37%, and therefore it would have been overestimated by 1.52% as a result of the market valuation error.

Fourthly, the DGM methodology is error-prone in the presence of short-term fluctuations in the market's earnings retention rate. For example, suppose the market cost of equity is 10% per year in perpetuity, the expected growth rate in dividends per share is 5% per year in perpetuity (matching the expected long-run expected GDP growth rate), and the dividends in the most recent year were \$1b. Suppose also that the risk free rate is 4% in perpetuity, and therefore the market risk premium is 6%. Using the first three of these parameters, the current value of equities would then be as follows:

$$S = \frac{\$1b(1.05)}{1.10} + \frac{\$1b(1.05)^2}{(1.10)^2} + ... = \frac{\$1b(1.05)}{.10 - .05} = \$21b$$

Substitution of this value for *S*, along with the current dividend level *D* and the expected growth rate in dividends *g*, into the DGM equation (2) would then accurately estimate the market cost of equity at 10%, and deduction of the risk free rate of 4% would then yield an accurate estimate of the market risk premium of 6%. Now suppose instead that firms in aggregate lowered their earnings retention rate in the most recent year and that the effect of this was to raise the current dividend level from \$1b to \$1.3b, at the expense of future dividends (relative to the above path). Suppose also that the effect of this change was NPV neutral, so that the current value of equities would be lower by \$0.3b. So, application of the DGM in equation (2) with *g* still estimated from the expected long-run growth rate in GDP (at 5%) would yield the following estimate of the market cost of equity:

$$k = \frac{\$1.3b}{\$20.7b}(1.05) + .05 = .116$$

Deduction of the risk free rate of 4% would then yield an estimate of the market risk premium of 7.6%. Since the true market risk premium is 6%, the DGM has overestimated it by 1.6%. The source of the problem is the fact that the higher current dividends of $1.3b arise from a *temporary* reduction in firms' retention rates and therefore are not a suitable base from which to project subsequent dividends.

Fifthly, the DGM combines the current dividend level of firms (which reflects the current earnings retention rate) with an expected long-run growth rate in dividends per share for existing companies that is based upon the expected long-run growth rate in GDP, and the latter is based upon historical averaging and therefore upon the historical average earnings retention rate (assuming plausibly that the growth rate in GDP is affected by the level of corporate investment). Thus, if the earnings retention rate has fallen over time, so that the current level is below its historical average, then estimating the expected long-run growth rate in GDP from its historical average will over estimate this parameter and therefore overestimate the market risk premium.

4.3 Long-Run Averaging

The third approach considered by CEG is to invoke the CAPM along with an estimate of the long-run market risk premium (6%) and an estimate of the long-run risk free rate of 5.99%, with the latter based upon averaging results over the entire period since the RBA adopted inflation-targeting (June 1993). Although neither of these average parameter values would necessarily match their current values, CEG argues that variations across time are largely offsetting and therefore the resulting cost of equity from their proposed approach is more reliable than the generally employed methodology amongst Australian regulators, which involves the current risk free rate and an estimate of the long-term average MRP.

This proposal is subject to a number of difficulties, as follows. Firstly, even if all CEG's claims about this approach were true, it would only produce an accurate estimate for the cost of equity for a company with a beta of 1. For businesses with equity betas less than 1, CEG's approach will overestimate the cost of equity because the overestimate in the risk free rate will exceed the underestimate of the risk premium.

Secondly, the approach assumes that the estimate for the market risk premium of 6% is an estimate of the long-term average market risk premium. However, in respect of some Australian regulators, this belief is not correct. For example, the AER (2012, Attachments, pp. 128-136) bases its estimate of 6% upon both historical average excess returns and forward-looking evidence such as surveys whilst the QCA (2011, pp. 238-240) bases its estimate of 6% upon the results from four different methodologies and only two of these involve long-run historical data with the other two being forward-looking methods. Thus, even if one viewed the reliance upon long-run historical data by both regulators as an attempt to estimate the long-term market risk premium, it cannot be said that their estimates of the market risk premium are entirely of this kind.

Thirdly, CEG's proposed methodology requires a judgement about the historical period over which to average the risk free rate and different judgements will produce different results. CEG propose using the period from June 1993, when the RBA adopted inflation targeting (of 2-3%). However the CGS yields in the first five years after this were high by comparison with subsequent rates; in particular, they reached 11% in the first five years whereas they did not subsequently exceeded 7.3% even during the prolonged world-wide boom from 2002-2007 (CEG, 2012, Figure 11). A possible explanation is that there was considerable scepticism amongst investors within the first few years of inflation targeting that inflation would be constrained to 3% and the CGS yields gradually subsided as the RBA's credibility grew and scepticism subsided. If so, then the averaging should be done from about 1998 and the result would then be significantly lower. Furthermore it could be argued that even this period from about 1998 is over-represented by unusually good economic conditions and this imparts an upward bias to the resulting estimate of the long-run risk free rate. It is therefore not clear which historical period should be used for averaging the risk free rate, and therefore it is not clear what the result from this methodology should be.

Fourthly, CEG's proposed methodology rests on the belief that variations in the risk free rate and the market risk premium are largely offsetting over time, and the principal evidence offered by CEG in support of this claim is the relatively stable behaviour over time of the estimated market cost of equity (as shown in CEG's Figure 8). However, as noted earlier, the estimated costs of equity shown in Figure 8 are obtained by assuming that, at any point in time, the market cost of equity is the same for all future years, and this assumption produces a smoother time series in the estimated cost of equity than would otherwise be the case. Furthermore, as noted previously, this assumption underlying Figure 8 can be tested by observing that the model gives rise to an estimated market risk premium of zero in 1994; this outcome is not plausible and therefore suggests that the underlying assumption is not plausible.

Fifthly, CEG's proposed methodology sacrifices a relevant, critical and observable parameter (the current risk free rate) in order to offset alleged errors in another parameter (the market risk premium). However, the preceding three points above indicate that the benefits from this approach are exaggerated or nebulous. Thus, the proposed methodology involves a clear cost without a clear benefit, and is therefore not favoured.

Sixthly, although CEG does not comment upon the cost of debt, their proposal to use the historical average risk free rate in the process of estimating the current ten-year cost of equity raises the question of whether the same historical average risk free rate would be used in the process of estimating the current ten-year cost of debt, and therefore whether the historical average debt risk premium would also be used in estimating the current ten-year cost of debt. CEG do not answer these questions but a regulated business that relies upon CEG's proposal to estimate the cost of equity does answer these questions, and favours using the historical average risk free rate along with the current debt risk premium to estimate the current ten-year cost of debt (Aurora, 2012). This is pure cherry-picking, and leads to overestimation of the current ten-year cost of debt. For example, suppose the current ten-year risk free rate is 3.8%, the current ten-year debt risk premium is 3.6%, the historical average ten-year risk free rate is 6.0%, and the historical average ten-year debt risk premium is 2.0%. The current ten-year cost of debt would then be 7.4% but Aurora would have overestimated it at 9.6% by combining the higher of the two risk free rates (the historical average of 6.0%) along with the higher of the two debt risk premiums (the current premium of 3.6%). Even if Aurora had used both the historical average risk free rate (6.0%) and the historical average debt risk premium (2.0%), they would still have overestimated the current ten-year debt risk premium at 8.0%. The appropriate parameters to use in estimating the current cost of debt are the current risk free rate and the current debt risk premium, because the former is observable and the latter can be estimated with a high degree of precision. Any argument in favour of using a historical average risk free rate in estimating the cost of *equity*, because the MRP is difficult to estimate, has no relevance to estimating the cost of debt.

4.4 Choice of Methodology

The first of CEG's proposals, being the DGM for estimating an individual firm's cost of equity, is very similar to the DGM for estimating the market risk premium, but has the additional problems of greater exposure to fluctuations in the earnings payout rate, incentives for the firms in question to manipulate their earnings payout rate, and implicitly (and wrongly) assumes that the entire firm's activities are regulated. Consequently, I do not favour this approach.

The third of CEG's proposals, in which the average risk free rate over some historical period is substituted for the current rate, suffers from a number of serious problems; these involve overestimating the cost of equity for businesses with equity betas less than 1, wrongly assuming that the widely employed MRP estimate of 6% is an estimate of the long-term MRP, ambiguity over the correct averaging period for the risk free rate, the unsubstantiated belief that variations in the MRP and the risk free rate are offsetting, the sacrifice of an observable, relevant and significant parameter, and potential spillover problems in estimating the cost of debt. I think these problems are sufficiently pronounced that this methodology should not be employed.

This leaves CEG's second approach (the AMP method), involving using the DGM to estimate the MRP. Errors in the AMP method (as described in section 4.2) rule this out of consideration, and their effect is to inflate the MRP estimate by about 1%. This point aside the DGM is worthy of consideration but, in view of its limitations, as a complement to rather than a substitute for other approaches.

The QCA's current approach involves using the current risk free rate along with an estimate of the MRP based upon results from a range of methodologies including historical averaging and the DGM applied to the MRP. I favour this approach because the observable value of the risk free rate is used and the difficulties in estimating the MRP are minimised by combining results from a set of estimators that are less than perfectly correlated. The gains from imperfectly correlated estimators of the MRP are large. For example, if two estimators are uncorrelated and have standard deviations of 2% each, then an equally-weighted average of the two will have a standard deviation of only 1.4%; with four uncorrelated estimators, the standard deviation of the equally-weighted average falls further to 1.0%.

Even if one or more of these estimators were biased at the present time, such as the historical average excess return (for the reasons claimed by CEG), inclusion of that estimator will still be warranted although the weighting on it might decline. In particular, in the presence of bias, one should choose a weighted-average of estimators whose estimation errors are smallest, and the usual expression of this is minimising mean square error (MSE).[7] Letting $\hat{T}$ denote an estimator and T the true value of the parameter being estimated, the MSE of an individual estimator is as follows:

$$\begin{aligned} MSE &= E\left[\hat{T} - T\right]^2 \\ &= E\left[\hat{T} - E(\hat{T}) + E(\hat{T}) - T\right]^2 \\ &= E\left[\hat{T} - E(\hat{T})\right]^2 + \left[E(\hat{T}) - T\right]^2 \end{aligned} \tag{3}$$

where the first term in the last equation is the variance of the estimator and the second term is the square of the bias. Suppose at the present time that the historical average excess return is biased down by 1% as an estimator of the market risk premium over the next ten years, and that its standard deviation is 2%.[8] Suppose also that a forward-looking estimator is unbiased, that it also has a standard deviation of 2%, and that the two estimators are uncorrelated. Using equation (3), the MSE of the historical average excess return is $.022^2$ whilst that of the forward-looking estimator is $.02^2$. Forming a weighted-average of the two estimators, with the weight on the first (w) chosen to minimise the MSE of the weighted-average, then w is chosen to minimise

$$\begin{aligned} MSE &= E\left[w\hat{T}_1 + (1-w)\hat{T}_2 - T\right]^2 \\ &= E\left[w(\hat{T}_1 - T) + (1-w)(\hat{T}_2 - T)\right]^2 \\ &= w^2 E\left[\hat{T}_1 - T\right]^2 + (1-w)^2 E\left[\hat{T}_2 - T\right]^2 \\ &= w^2 MSE_1 + (1-w)^2 MSE_2 \end{aligned} \tag{4}$$

[7] The MSE is the average over the squared differences between estimated value and the true value.

[8] The standard deviation of 2% is consistent with a standard deviation for annual Australian excess returns of 20% (Dimson et al, 2011, Table 13) coupled with the use of 100 years of annual excess returns.

With MSE_1 and MSE_2 as given above, MSE is minimised with w = .44, i.e., a 44% weight on the historical average excess returns and therefore a 56% weight on the forward-looking estimator. So, even if the historical average excess return were currently significantly biased in estimating the market risk premium, it would still seem to warrant significant weight in a weighted-average estimator.

An even better goal than choosing an estimator with minimal MSE for the MRP over the next several years would be to choose an estimator with minimal MSE for the MRP over the *life* of the regulated assets, i.e., under or over estimation within a single regulatory cycle would be of no great consequence relative to aggregate errors over the entire life of the regulated asset. With such a long period, short-term biases in the historic average excess return methodology are likely to wash out, and therefore the merits of historical averaging will be even greater than previously concluded.

In summary, I favour the QCA's existing approach in which the current risk free rate is coupled with an estimate of the market risk premium, and the latter is based upon results from a range of methodologies including historical-averaging of excess returns and the DGM applied to the market risk premium. If results from historical averaging contribute to the MRP estimate the effect of this approach *might* be to generate an estimate of the MRP that is biased down at the present time but the mean squared error of the estimator is the more important consideration and it favours the QCA's approach. Furthermore, any downward bias at the present time is likely to be offset over the entire life of the regulated assets.

5. Rounded Estimates of the Market Risk Premium

Since the market risk premium can only be estimated rather than observed, then any estimate must be subject to some degree of rounding. The QCA's current practice is to round to the nearest 1%, and this raises the question of whether a different rounding rule should be employed. Since the QCA's unit of rounding seems to be at the upper limit of regulatory practice, the question then becomes one of whether rounding should be undertaken to a lower level, which would be 0.5% or even 0.25%. The pros and cons of doing so are therefore considered.

A lower unit of rounding will improve accuracy, but the effects are quite small. For example, suppose the true MRP (*T*) is uniformly distributed over the interval from 4.5% to 7.5%, and the unrounded estimation error around the true value (*e*) is uniformly distributed over the interval from -1.5% to 1.5%. If the estimate of the true value (*T* + *e*) is unrounded, the resulting root mean square error (RMSE) is the standard deviation of *e*, which is 0.87%. If the estimate is rounded to the nearest 1.0%, the resulting RMSE is 0.92%:

$$\sqrt{E\left[(T+e)_{R=01} - T\right]^2} = .0092$$

If the estimate is instead rounded to the nearest 0.5%, then the RMSE is 0.88%. So, rounding to 1% reduces accuracy only marginally. Furthermore, the more important consideration is error over the life of regulated assets rather than error over a single regulatory cycle, and the errors from rounding will tend to offset over the life of the regulated assets.

These RMSE results are highly dependent upon the variance in the estimation errors (*e*). If the variance in the distribution collapses to zero (i.e., the MRP can be observed without error), then the RMSE from no rounding would be zero, that from rounding to 0.5% would be 0.13% and that from rounding to 1.0% would be 0.32%. In this case, the errors from rounding would be much greater.

A lower unit of rounding also increases the frequency of unwarranted changes in the rounded estimate, because variations in unrounded estimates arise even if the true value for the market risk premium has not changed. Furthermore, the extent of this estimation error can be estimated, at least in respect of some estimation methodologies. Consider the historical averaging of excess returns, with a standard deviation for annual excess returns (market returns net of the risk free rate) in Australia of 20% (consistent with Dimson et al, 2011, Table 13). Suppose this has been estimated using 100 years of data, at 6%, and that this matches the true value. We also suppose that annual excess returns are independent, identically distributed and normal. The average return over the next five years is then normally distributed with mean 6% and standard deviation of $.20/\sqrt{5} = .089$. With a 1% rounding rule, the next five years of excess returns will then have to be such as to raise the average to at least 6.5% or less than 5.5%, and this requires an average excess return over the next five years ($\overline{R}_5$) of at least 16.5% or no more than -4.5%. The probability P of such an event is then

$$
\begin{aligned}
P &= 1 - \Pr ob\left(-.045 \le \overline{R}_5 \le .165\right) \\
&= 1 - \Pr ob\left(\frac{-.045 - .06}{.0894} \le \frac{\overline{R}_5 - .06}{.0894} \le \frac{.165 - .06}{.0894}\right) \\
&= 1 - \Pr ob\left(-1.17 \le Z \le 1.17\right) \\
&= 0.24
\end{aligned}
$$

So, with a 1% rounding rule, there is 24% probability of the market risk premium being unjustifiably changed. With a 0.5% rounding rule, the average excess return over the next five years need only be less than 0.75% or more than 11.25% to provoke a change in the rounded estimate, and such an event has a probability of 56%. Thus, halving the rounding interval to 0.5% raises the probability of an unjustified change to the rounded estimate from 24% to 56%. Naturally these unjustified changes would quickly be reversed, and therefore the effect of a lower unit of rounding would be a sharply increased frequency in upward and downward oscillations in the estimated MRP due purely to estimation error. These conclusions cannot be automatically extrapolated to other methodologies for which such probabilities cannot be readily estimated. However there are no strong grounds to suppose that estimation errors from other methodologies are significantly less than for historical averaging of excess returns.

A lower unit of rounding also increases the benefit to lobbying for either a change in some aspect of one of the estimation methodologies employed by the QCA or a change in the set of estimation methodologies. For example, suppose the average MRP estimate across the four methodologies used by the QCA is 6.2%, and that a possible change in one aspect of estimation method would raise this average to 6.3%. With the estimate rounded to the nearest 1%, there would be no incentive for any interested party to lobby the QCA for this change. However, if the QCA rounded to the nearest 0.5%, there would then be an incentive for regulated firms to lobby for this change because successful lobbying on this matter would raise the rounded estimate from 6.0% to 6.5%.

In summary there is an accuracy gain from a lower unit of rounding for the MRP but the gain is very small, because the extent of estimation error in the MRP is so large. Furthermore a lower unit of rounding increases the frequency of unwarranted changes in the rounded estimate, and prompts more lobbying for changes in parameter values within a methodology or changing the set of methodologies used to estimate the market risk premium. Since these disadvantages from a lower unit of rounding than 1% seem substantial, and the benefits so small, I recommend continued use of rounding to 1%. In respect of other parameters, for which the degree of estimation error is different, the appropriate degree of rounding would be different.

6. Investor Portfolio Reassessments and the Market Risk Premium

The CAPM version to which all previous discussion relates is the Officer (1994) model, which assumes, inter alia, that investors select portfolios today based upon their (agreed) probability distributions for returns over some future but unspecified period (Sharpe, 1964; Lintner, 1965; Mossin, 1966). The model is then single-period in nature but it can be extended into a multi-period world of indefinite duration so long as the probability distributions for asset returns do not change over time or investors act as if this is the case (Fama, 1970). In this more realistic multi-period scenario, the model then assumes that investors select portfolios today with the intention of reassessing them after some common but unspecified period, i.e., investors have a common but unspecified investment horizon. Since this period is not specified in the model and it varies across investors, a pragmatic response would be to define this period as the across-investor average period between successive portfolio reassessments, and the QCA seeks an estimate for this period.

A useful starting point is the current turnover rate for Australian listed shares. This is approximately 1 (ASX, 2011). If every investor had the same turnover rate, then the inverse of the aggregate turnover rate (one year) would be the (common) period for which assets were held. However, if there is any variation across investors in their turnover rates, then the average holding period will exceed the inverse of the aggregate turnover rate. For example, if investors holding half of all assets have a turnover rate of 0.2 and the rest have a turnover rate of 1.8 (consistent with an aggregate turnover rate of 1), the holding periods would be 5 years and 0.56 years respectively, with an average of 2.78 years (which is almost three times the inverse of the turnover rate). Froot et al (1991, Table 1) gives the “time horizon” (the inverse of the turnover rate) for each of ten US investor categories, ranging from .03 years to 7.1 years, and the variation across individual investors would be even greater. So, returning to the Australian turnover rate, a wide range in turnover rates around the average of 1 would have to be recognised, with rates from very close to zero (for 'passive' funds) to figures in excess of 1000 for a 'day-trader'. A natural candidate for the density function is then the lognormal distribution, i.e.,

$$T = e^{\mu + \sigma Z} \tag{5}$$

where Z is the standard normal random variable, and μ and σ are the mean and standard deviation of the lognormal distribution. It follows that

$$E(T) = e^{\mu + 0.5\sigma^2}$$

Since the average turnover rate is 1, the estimate for $E(T)$ is 1 and therefore $\mu = -0.5\sigma^2$. Since the holding period H is the inverse of the turnover rate T, it follows that the expected holding period is

$$\begin{aligned} E(H) &= \int \frac{1}{T} f(Z) dZ \\ &= \int e^{-\mu - \sigma Z} \frac{1}{\sqrt{2\pi}} e^{-0.5Z^2} dZ \\ &= \int e^{0.5\sigma^2 - \sigma Z} \frac{1}{\sqrt{2\pi}} e^{-0.5Z^2} dZ \\ &= e^{\sigma^2} \int \frac{1}{\sqrt{2\pi}} e^{-0.5(Z+\sigma)^2} dZ \\ &= e^{\sigma^2} \end{aligned} \tag{6}$$

The ASX (2011, page 2) reveals that 85% of the ownership of listed Australian equities lies with Australian financial institutions and foreigners (presumably mostly institutions), and passive funds should now represent a significant portion of this.[9] If this proportion is even 25% (of 85%) and a passive fund has a turnover of 0.1, then

$$\Pr ob(T \leq 0.1) = 0.21 \tag{7}$$

[9] Frino and Gallaher (2002, page 3) estimate that 11% of the managed funds industry in Australia in 2000 was passive, and that this share was growing rapidly. Therefore the current figure is likely to be much more than 11%.

Substituting the lognormal distribution for T shown in equation (5) into equation (7) and solving for σ reveals that $\sigma = 1.46$. Substituting this value into equation (6) then reveals that the average holding period is 8.5 years.

This estimate is based upon listed equity, which is likely to constitute a portion of investors' portfolios with a lower than average holding period (due to its high liquidity and low transactions costs). Higher holding periods would therefore exist for real estate, pension entitlements, etc.[10] Consequently, the average holding period across the aggregate portfolios of investors is likely to significantly exceed any figure estimated purely from listed equities. However the period between successive portfolio reassessments is likely to be significantly less than an investor's average holding period. For example, an investor who reassesses their portfolio annually and on average replaces 25% of it would have a holding period of four years, and therefore the average holding period would be four times the period between successive portfolio reassessments. On the other hand, many transactions are not the result of portfolio reassessments, but simply savings by some investors (which require asset purchases), dis-saving by other investors (which requires asset sales), and portfolio realignments by passive funds in response to changes in the composition of the particular market index they track. Excluding these transactions would raise the average interval between portfolio reassessments. All of this reveals that it is difficult to estimate the across-investor average period between successive portfolio reassessments. However the figure is likely to exceed one year and even a figure of ten years is entirely plausible.

Since a high degree of precision seems unattainable, it would be reasonable to consider more pragmatic factors, particularly data availability, relevant regulatory practices, and the likelihood that any change from the current QCA practice of ten years would have any effect. In respect of data, long historical series of Australian risk free rates are only available for three month and ten year bonds and this limits estimation of market risk premiums using historical averaging to these options (Brailsford et al, 2008, pp. 82-84). However, the historical series of (say) five year Australian rates could be proxied for some period prior to their availability from the Australian ten year series coupled with the differential between

[10] For example, the turnover rate for housing in Australia is about 6% (Bloxham et al, 2010, Graph 1) and the transactions costs for buyers and sellers average about 7% each (Kelly, et al, 2011, page 33). By contrast, in respect of Australian listed equities, the turnover rate is 100% (ASX, 2011a) and transaction costs are limited to brokerage fees (ASX, 2011b) which are less than 1% (For eg: https://www.asbsecurities.co.nz/section91.asp).

five and ten year rates for an earlier period from another market for which both series are available.

In respect of relevant regulatory practices (of the QCA), the typical regulatory term seems to be about four-five years and the QCA also matches the term of the risk free rate within the first term of the CAPM to the regulatory cycle. Consequently, defining the market risk premium with respect to five year bonds would ensure that the risk free rate used within the market risk premium would typically match that within the first term of the CAPM, thereby avoiding both the appearance of inconsistency within the model's parameters and the resulting contentious debates on this question. By contrast, for a regulator whose policy is to use the same risk free rate in both places within the model, this issue does not arise.

Finally, in respect of whether the choice of risk free rate term has any effect upon the estimate of the market risk premium, this depends upon the impact of the choice on the unrounded estimate of the market risk premium and the extent of rounding in the estimate. The QCA (2011, page 238) considered the effect of using a five versus ten year risk free rate, and it only raised the estimate of the market risk premium by 0.1%, which was insufficient to affect the rounded result whether rounded to 1.0% or 0.5%. However, if rounded to 0.5%, the probability of changing the rounded estimate would rise from 10% to 20%. As discussed in the previous section, I favour rounding to 1% rather than a lower level and therefore the choice of (say) five versus ten year risk free rates is very unlikely to affect the estimate.

In summary, the across-investor average period between successive portfolio assessments is not amenable to a precise estimate, but a reasonable estimate is at least one year and could easily be ten years. The choice then rests upon more pragmatic considerations, with historical data availability for the risk free rate pointing to a ten year period whilst regulatory considerations (in the form of the typical regulatory period being five years) suggest a figure of five years. On balance, I think data availability is the more significant issue, and this favours treating the across-investor average period between successive portfolio reassessments as ten years. However, with rounding to the nearest 1%, the choice of five versus ten years is very unlikely to affect the rounded result.

7. Conclusions

The QCA has posed a number of questions relating to the risk free rate and the market risk premium, and my conclusions on these matters are as follows.

The first issue was to assess CEG's claims relating to the suitability of CGS as a proxy for the risk free asset within the CAPM, or its use in regulatory situations, in the face of current CGS yields that are historically low. CEG identifies three explanations for the historically low CGS yields, comprising the current low stock of CGS, the recent 'flight to quality', and Basel III requirements, and implies that they each undercut the suitability of CGS as a proxy for the risk free asset. However the current low stock of CGS does not undercut the suitability of CGS as a proxy for the risk free asset because the CAPM imposes no requirements on the supply of the risk free asset. In addition, the recent 'flight to quality' does not undercut the suitability of CGS as a proxy for the risk free asset because the CAPM does not require that the risk free rate be invariant to such events. Finally, CEG's claim that Basel III requirements undercut the suitability of CGS as a proxy for the risk free asset might have some validity but the effect is indeterminable and would be at least partly offset by the presence of default risk in CGS. CEG imply that the 'problem' could be addressed by averaging CGS yields over a longer period. However, in the absence of any attempt by CEG to quantify the problem they allege or any ability to do so, recourse to a clearly radical alternative is not supported. Any argument for using a longer term average of CGS yields must rest on other grounds offered by CEG and these are addressed next.

The second issue was to assess CGS's claim that, regardless of the cause, Australian CGS yields are very low at the present time, that variations in these rates are strongly negatively related to variations in the market risk premium, the general Australian regulatory practice is to combine the current CGS yield with an estimate of the long-term market risk premium, this practice will then produce an estimate of the cost of equity that is too low at the present time, and therefore regulated firms will be under compensated. CEG do not present any persuasive evidence that there is a *strong* negative relationship between the ten year risk free rates and market risk premiums, and the primary evidence they do present in their Figure 8 is pre-disposed to that result. Furthermore, many Australian regulators including the QCA do not estimate the long-term market risk premium but estimate a market risk premium that reflects both prevailing and longer-term factors, and therefore CEG's argument is much less relevant to them than for regulators who estimate a purely long-term market risk premium. Furthermore, the significant issue relating to compensation is over the life of regulated assets rather than the next few years, and therefore a process for estimating the cost of equity that is biased under some economic conditions but most accurate over the life of regulated assets might still be preferred. Finally, CEG suggests that the generally employed methodology should be abandoned only when conditions are unusual; this approach is highly subjective, may lead to variations only when they favour regulated firms, and therefore is not desirable.

The third issue was to assess, assuming that under compensation is present, whether it should be addressed through any of the three approaches suggested by CEG or by any other approach. Notwithstanding my view that under compensation at the present time is unproven and less important than under compensation over the life of the regulated assets, I have assessed the ability of CEG's three approaches to deal with any such under compensation that might exist at the present time. The first of CEG's proposals, being the DGM for estimating an individual firm's cost of equity, is very similar to the DGM for estimating the market risk premium, but has the additional problems of greater exposure to fluctuations in the earnings payout rate, incentives for the firms in question to manipulate their earnings payout rate, and implicitly (and wrongly) assumes that the entire firm's activities are regulated. Consequently, I do not favour this approach. The third of CEG's proposals, in which the average risk free rate over some historical period is substituted for the current rate, suffers from a number of serious problems; these involve overestimating the cost of equity for businesses with equity betas less than 1, wrongly assuming that the widely employed MRP estimate of 6% is an estimate of the long-term MRP, ambiguity over the correct averaging period for the risk free rate, the unsubstantiated belief that variations in the MRP and the risk free rate are offsetting, the sacrifice of an observable, relevant and significant parameter, and potential spillover problems in estimating the cost of debt. I think these problems are sufficiently pronounced that this methodology should not be employed. This leaves CEG's second approach, involving using the DGM to estimate the MRP. Errors in the AMP variant rule this out of consideration, and their effect is to inflate the MRP estimate by about 1%. This point aside, the DGM is worthy of consideration but, in view of its limitations, as a complement to rather than a substitute for other approaches. I therefore favour a risk free rate matching the current rate coupled with an estimate of the market risk premium that draws upon results from various methodologies, and this corresponds to the QCA's current approach. If results from historical averaging contribute to the MRP estimate, the effect of this approach *might* be to generate an estimate of the MRP that is biased down at the present time but the mean squared error of the estimator is the more important consideration and it favours the QCA’s approach. Furthermore, any downward bias at the present time is likely to be offset over the entire life of the regulated assets.

The fourth issue is that of whether the QCA should continue to round its MRP estimate to 1% or to some (lower) unit of rounding. There is a gain in accuracy from a lower unit of rounding but the gain is very small. Furthermore a lower unit of rounding increases the frequency of unwarranted changes in the rounded estimate, and prompts more lobbying for changes in parameter values within a methodology or changing the set of methodologies used to estimate the market risk premium. Since these disadvantages from a lower unit of rounding than 1% seem substantial, and the benefits so small, I recommend continued use of rounding to 1%.

The fifth issue is that of estimating the across-investor average period between successive portfolio reassessments. This period is not amenable to a precise estimate, but a reasonable estimate is at least one year and could easily be ten years. The choice then rests upon more pragmatic considerations, with historical data availability for the risk free rate pointing to a ten year period whilst regulatory considerations (in the form of the typical regulatory period being five years) suggest a figure of five years. On balance, I think data availability is the more significant issue, and this favours treating the across-investor average period between successive portfolio reassessments as ten years. However, with rounding to the nearest 1%, the choice of five versus ten years is very unlikely to affect the rounded result.

The final issue is whether the QCA should, at the start of each regulatory cycle, review its estimate of the across-investor average period between successive portfolio reassessments. In view of the difficulties of estimating this parameter, I do not favour so frequent a review of it.

REFERENCES

AER, 2012, *Final Distribution Determination Aurora Energy Pty Ltd 2012-13 to 2016-17* (www.aer.gov.au).

Armitage, S., 2012, 'Demand for Dividends: The Case of UK Water Companies', *Journal of Business Finance and Accounting*, 39, pp. 464-499.

Aurora Energy, 2012, *Revised Regulatory Proposal Supporting Information: Return on Capital Aurora Response to the AER's Draft Distribution Determination* (www.aer.gov.au).

ASX, 2011a, *Australian Cash Equity Market December 2011* (http://www.asx.com.au/documents/resources/australian_cash_equity_market.pdf).

____ 2011b, *Getting Started in Shares* (http://www.asx.com.au/documents/resources/getting_started_in_shares.pdf).

Bernstein, P. and Arnott, R. 2003, "Earnings Growth: The Two Percent Dilution", *Financial Analysts Journal*, Sept/October, pp. 47-55.

Bloxham, P., McGregor, D., and Rankin, E., 2010, 'Housing Turnover and First-Home Buyers', *Reserve Bank of Australia Bulletin June Quarter*.

Brailsford, T., Handley, J., and Maheswaren, K., 2008, 'Re-examination of the Historical Equity Risk Premium in Australia', *Accounting and Finance*, vol. 48, pp. 73-97.

CEG. 2012, *Internal Consistency of Risk Free Rate and MRP in the CAPM*, report prepared for Envestra, SP AusNet, Multinet and APA.

Dimson, E., Marsh, P. and Staunton, M., 2011, *Credit Suisse Global Investment Returns Sourcebook 2011*, Credit Suisse.

Fama, E., 1970, 'Multiperiod Consumption-Investment Decisions', *The American Economic Review*, pp. 163-174.

French, K., Schwert, G., and Stambaugh, R. 1987, 'ExpectedStock Returns and Volatility', *Journal of Financial Economics*, vol. 19, pp. 3-29.

Frino, A., and Gallagher, D., 2002, 'Is Index Performance Achievable?: An Analysis of Australian Equity Index Funds', *Abacus*, vol. 38, pp. 200-214.

Froot, K., Perold, A., and Stein, J., 1992, 'Shareholder Trading Practices and Corporate Investment Horizons', *Journal of Applied Corporate Finance*, vol. 5(2), pp. 42-58.

Hirshleifer, J. 1970, *Investment, Interest and Capital*, Prentice-Hall, Englewood Cliffs.

Kelly, J-F., Breadon, P., and Reichl, J., 2011, *Getting the Housing we Want*, Grattan Institute, Melbourne(http://www.politiquessociales.net/IMG/pdf/117_report_getting_the_housing_we_want.pdf).

Lettau, M. and Ludvigson, S., 2001, 'Consumption, Aggregate Wealth and Expected Stock Returns', *The Journal of Finance*, vol. 56, pp. 815-849.

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.

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

Mossin, J. 1966, 'Equilibrium in a Capital Asset Market', *Econometrica*, vol. 24, pp. 768-83.

NERA Economic Consulting. 2007a, *Bias in Indexed CGS Yields as a Proxy for the CAPM Risk Free Rate*, report prepared for the ENA.

________________________ 2007b, *Absolute Bias in Nominal Commonwealth Government Securities*.

Officer, R. 1994, 'The Cost of Capital of a Company under an Imputation Tax System', *Accounting and Finance*, vol. 34, pp. 1-17.

QCA, 2011, *Final Report SEQ Interim Price Monitoring for 2010/11 Part B-Detailed Assessment March 2011* (www.qca.org.au).

Sharpe, W. 1964, 'Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk', *Journal of Finance*, vol. 19, pp. 425-42.

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

_______ 1999, 'The Shrinking Equity Premium', *Journal of Portfolio Management*, vol. 26, pp. 10-17.

Smithers and Co Ltd, 2003, *A Study into Certain Aspects of the Cost of Capital for Regulated Utilities in the U.K.*, report prepared for the OFT, CAA, OFWAT, Ofgem, Oftel, ORR, and OFREG.