

Estimating the Market Risk Premium in Regulatory Decisions:

Conditional versus Unconditional Estimates

Peter Gibbard

Working Paper no. 9, September 2013

ACCC/AER WORKING PAPER SERIES

© Commonwealth of Australia 2013

ISBN 978-1-921973-87-1

This work is copyright. Apart from any use permitted by the Copyright Act 1968, no part may be reproduced without permission of the Australian Competition and Consumer Commission. Requests and enquiries concerning reproduction and rights should be addressed to the Director of Publishing, Australian Competition and Consumer Commission, GPO Box 3131, Canberra ACT 2601.

Citation details: ACCC/AER Working Paper, September 2013

Series Note

The Australian Competition and Consumer Commission (ACCC) is an independent statutory authority formed in 1995 to administer the *Competition and Consumer Act 2010* and a range of additional legislation, promoting competition, fair trading and regulates national infrastructure for the benefit of all Australians.

The Australian Energy Regulator (AER) is a constituent part of the ACCC with an independent board.

Working papers are intended to disseminate the results of current research by ACCC staff and consultants. The aim of the series is to facilitate discussion and comment. The working papers come from across the ACCC covering competition law, economic regulation, and consumer protection. They are available at no cost in electronic format only from the ACCC website, www.accc.gov.au. To subscribe to this series, email working.papers@acc.gov.au.

The papers in this series reflect the views of the individual authors. The views expressed in the paper do not necessarily reflect the views of the ACCC or the AER.

Enquiries may be addressed to:
The Editor
ACCC/AER working paper series
Australian Competition and Consumer Commission
GPO Box 520
Melbourne Vic 3001
Email: workingpapers@acc.gov.au

About the author

Peter Gibbard currently holds the position of Principal Economic Adviser with the ACCC/AER.

He joined the ACCC/AER in 2010, and works for the Regulatory Development Branch, providing economic consultancy advice across the ACCC and AER.

Peter has a Bachelor of Economics (First Class Honours) and a Bachelor of Laws from the University of Adelaide, an M.Phil in Economics from Oxford University and a PhD in Philosophy, focusing on mathematical logic, from the University of Michigan. Prior to joining the ACCC, he worked in the Financial Stability Division of the Bank of England in London.

Contents

Series Note	2
1. Introduction	5
2. Predictability versus unpredictability: three phases of research	8
2.1 The first phase: the unpredictability of returns and random walks	8
2.2 The second phase: from unpredictability to predictability.....	9
2.3 The third phase: a renewed ‘healthy skepticism’ about predictability	11
3. The effect on the MRP of the dividend yield, risk-free rate and volatility	15
3.1 Dividend yields and the MRP	15
3.1.1 Theory	15
3.1.2 Evidence	16
3.2 The risk-free rate and the MRP.....	18
3.2.1 Theory	18
3.2.2 Evidence	19
3.3 Volatility and the MRP	20
3.3.1 Theory	20
3.3.2 Evidence	21
4. Reasons for regulators to have practical concerns about conditional estimates.....	24
4.1 The diversity and complexity of recent models of return predictability	24
4.2 Instability in models of return predictability	26
4.3 Data mining.....	27
5. Conclusion.....	30
REFERENCES.....	31

1. Introduction

Regulators have discussed at least four different kinds of methodologies for estimating the market risk premium (MRP) for the purpose of determining regulatory prices. First, MRP estimates can be informed by survey evidence, drawing on surveys of corporate executives, academics, auditors and accountants. Second, the MRP can be calculated using dividend growth models. Third, estimates of the MRP can be obtained from historical averages of annual excess returns (equity returns less the risk-free rate). Fourth, estimates of the MRP may be specified to be *conditional* on currently available information – that is, they may be specified to be a function of information such as market volatility, dividend yields and the risk-free rate.

This paper compares the third and fourth methods, identifying the key issues in the debate between conditional estimates of the MRP and historical, unconditional estimates. This debate is closely tied to the debate about the predictability of excess returns, on account of the relationship between expected returns and the required rate of return.¹ Accordingly, in the annual report on investment returns by Dimson, Marsh and Staunton, *Credit Suisse Global Investment Returns Sourcebook 2012*, when the authors evaluate the question of whether estimates of the MRP should be conditional, they discuss the debate about the predictability of returns. In particular, they observe that the debate about predictability is ‘far from settled’:

Yet despite extensive research, this debate [about predictability] is far from settled. In a special issue of the *Review of Financial Studies*, leading scholars expressed opposing views, with Cochrane (2008) and Campbell and Thompson (2008) arguing for predictability, whereas Goyal and Welch (2008) find that ‘these models would not have helped an investor with access only to available information to profitably time the market’. Cochrane’s (2011) recent Presidential Address demonstrates the persistence of this controversy (Dimson et al., 2012, p. 36).

In their contribution to the debate, Welch and Goyal (2008) argue that, in forecasting excess returns, investors cannot do better than use a historical average. The implication that can be drawn from their study is that estimates of the MRP should not be conditioned upon current information but, instead, a historical average should be used. Dimson et al. (2012, p. 37) themselves conclude that, for ‘practical purposes’, it is ‘hard’ for predictors of equity premia to outperform a long-term historical average:

In summary, there are good reasons to expect the equity premium to vary over time. Market volatility clearly fluctuates, and investors’ risk aversion also varies over time. However these effects are likely to be brief. Sharply lower (or higher) stock prices may have an impact on immediate returns, but the effect on long-term performance will be diluted. Moreover volatility does not usually stay at

¹If markets are in equilibrium and efficient, expected returns are equal to the required rate of return. Thus if, in addition, returns are predictable on the basis of current information, then (given that expectations are rational) not only expected returns but also the required rate of return is dependent on current information.

abnormally high levels for long, and investor sentiment is also mean reverting. For practical purposes, we conclude that for forecasting the long run equity premium, it is hard to improve on extrapolation from the longest history that is available at the time the forecast is being made.

When they refer to ‘the long run equity premium’, they have in mind forecast horizons of about five years.²

While Welch and Goyal (2008) is representative of an important recent strand of the research literature, the debate on predictability, as Dimson et al. (2012) observe, is ‘far from settled’, and John Cochrane’s influential defense of predictability, in particular, ‘demonstrates the persistence of this controversy’. There is an extensive and complex literature on the predictability of equity returns; and this working paper attempts to summarise the literature by identifying key phases in research on predictability since the 1960s. Three distinct phases of the literature are identified, and are discussed in Section 2 of the paper. The transition from the first to the second phase was highlighted by Cochrane in 2001 in the first edition of his book *Asset Pricing*: Cochrane proposed that whereas the first generation of research on asset pricing had emphasised the unpredictability of returns, a ‘new generation’ of research instead supported the view that returns are predictable. The transition from the second to a third phase of research is noted by Ang and Bekaert (2007, p. 653), who suggest that ‘the literature is converging to a new consensus, substantially different from the old view’. This third phase of research called for a renewed scepticism about the predictability of returns, especially in the medium- and long-run. In the debate between the second and third phases, a critical event is the 2008 issue of *Review of Financial Studies* that is described in the quotation above from Dimson et al. (2012). The contribution to the debate by Welch and Goyal (2008) was especially influential in challenging the claims for predictability. Whereas the second phase of research provides a basis for conditional estimates of the MRP, the third phase of research questions whether the MRP can be estimated conditionally – that is, it questions whether there are better estimates of the MRP than the unconditional historical average.

This third phase of research addresses a concern of the regulated businesses that an MRP estimate based on a historical average is ‘backward-looking’. The regulated businesses have questioned whether the use of a historical average of excess returns is consistent with the Capital Asset Pricing Model (CAPM): the concern is that a historical average is backward-looking, whereas the CAPM is forward-looking. But if the study of Welch and Goyal (2008) is accepted, then the historical average can be construed as a forward-looking measure: it is forward-looking because it can be construed as a good predictor of future excess returns, and because it is not clear that there are better predictors.

While Section 2 of the paper identifies the key phases in the general debate about predictability, Section 3 examines specific debates about which explanatory variables can be used to predict excess returns. It focuses on the debates about three predictor variables – dividend yields, interest rates and volatility. Whereas Sections 2 and 3 are concerned

²Dimson et al. (2012, pp. 36-37) present evidence in support of their conclusion which relates to forecast horizons of five or fewer years.

primarily with the academic literature on predictability, Section 4 has a more practical focus: it identifies problems that may arise in practice if a regulator attempts to estimate a conditional MRP. Even if it were conceded that excess returns are predictable from some given set of variables, a regulator faces at least three practical problems with using that set of variables to estimate a conditional MRP.

- (1) In response to scepticism about predictability in the third phase of research, the recent literature has investigated a range of models of returns that is increasingly (i) diverse, and (ii) complex. If a regulator were considering conditional models of the MRP, it would be difficult for the regulator to select and implement such a model not only because of the diversity of touted models but also because of their increasing complexity.
- (2) The third-phase of research has particularly emphasised concerns about the stability of models of excess returns. A number of studies have found that the values of the parameters in the models of returns tend to change over time. If, in fact, the relationship between excess returns and a variable changes over time, it is unclear how the regulator can set the MRP as a function of that variable.
- (3) Apparently significant relationships between variables and excess returns may reflect data-mining.

The conclusion, therefore, is that the debate among researchers on predictability is, as Dimson et al. (2012, p. 36) put it, ‘far from settled’: whereas the second phase of research might be used to support the case for a conditional estimate of the MRP, the third phase of research might be used in support of an unconditional estimate. Nevertheless, there are at least three reasons why in practice regulators may have grounds for using an unconditional rather than a conditional estimate of the MRP.

2. Predictability versus unpredictability: three phases of research

2.1 The first phase: the unpredictability of returns and random walks

In his evaluation of debates about the predictability of returns over time, Cochrane (2005) distinguishes between two different phases of research on asset pricing.³ The ‘first revolution in finance’ (which, he says, peaked ‘in the early 1970s’) emphasised the ‘near unpredictability of stock returns’, whereas ‘a new generation of empirical research’ has tended to find that stock returns are predictable at least ‘over the business cycle and longer horizons’ (Cochrane, 2005, p. 389-90). While Cochrane aligns his own position with the second phase of research, he provides a helpful summary of some of the key propositions that characterise the first phase:

Stock returns are close to unpredictable. Prices are close to random walks; expected returns do not vary greatly through time... Any apparent predictability is either a statistical artifact which will quickly vanish out of sample, or cannot be exploited after transaction costs (Cochrane, 2005, p. 389).

In general, during this early phase of research, the unpredictability of stock returns was seen as a consequence of efficient – or at least near-efficient – markets. Fama (1970, p. 383) provides a helpful definition of an efficient market: ‘A market in which prices always “fully reflect” available information is called “efficient”’ (Fama (1970), p. 383). Why does market efficiency make it difficult to predict returns? In his article ‘Proof that Properly Anticipated Prices Fluctuate Randomly’, Paul Samuelson presents a theoretical account of conditions under which returns will be unpredictable. The intuition for this result is encapsulated in the following remark: ‘If one could be sure that a price will rise, it would have already risen’ (Samuelson 1965, p. 41). Cornell (1999, p.2) provides an example that illustrates this intuition:

Suppose, for example, that someone were to write a convincing book entitled *The Crash of 2000*, explaining why the new millennium will be accompanied by a dramatic drop in share prices. If the arguments were truly convincing, then investors who read the book would clearly want to sell their stock before the dawn of the millennium. Assuming that enough investors read the book and acted in accordance with its predictions, stock markets would not fall at the start of the millennium but at the time the book was widely distributed – but this would mean the predictions of the book are false.

The argument, in brief, is as follows: if markets are efficient – so that the price falls when the information in the book becomes public – then the supposition that prices are predictable implies a contradiction.

³As noted above, the first edition of Cochrane’s book appeared in 2001.

In the first phase of research, such theoretical arguments for unpredictability found support in empirical studies of stock markets. Fama (1991, p. 1578) draws attention to empirical studies of short-run correlations – correlations between daily, weekly and monthly returns. While such correlations tended to be positive, researchers concluded that there were not good statistical grounds for rejecting the assumption of constant expected returns. Fama (1991, p. 1578) summarises the empirical findings of the first phase of research:

The evidence for predictability in the early work often lacks statistical power, however, and the portion of the variance of returns explained by the variation in expected returns is so small...that the hypothesis of market efficiency and constant expected returns is typically accepted as a good working model.

2.2 The second phase: from unpredictability to predictability

In the late 1980s, however, a second body of research began to accumulate, reacting against the first phase. In his survey article on market efficiency, Fama (1991, p. 1609) provides a summary of this ‘new evidence’ on predictability:

The recent evidence on the predictability of returns from other variables seems to give a more reliable picture of the variation through time of expected returns...In contrast to the autocorrelation tests on long-horizon returns, the forecast power of D/P, E/P, and the term-structure variables is reliable for periods after the Great Depression. D/P, E/P, and the default spread track autocorrelated variation in expected returns that becomes a larger fraction of the variance in returns for longer return horizons. These variables typically account for less than 5% of the variance of monthly returns but around 25-30% of the variances of 2- to 5-year returns. In short, the recent work suggests that expected returns take large, slowly decaying swings away from their unconditional means.

Like Fama, Cochrane (2005, p. 390) emphasises the contrast between short- and long-horizon predictability. He outlines the findings of this ‘new generation’ of empirical evidence as follows:

Variables including the dividend/price ratio and term premium can in fact predict substantial amounts of stock return variation. This phenomenon occurs over the business cycle and longer horizons. Daily, weekly, and monthly stock returns are still close to unpredictable.

Whereas the second phase of research acknowledges the findings of the earlier phase – that returns are ‘close to unpredictable’ over shorter horizons – the new claim is that returns are predictable over ‘longer horizons’.

Cochrane (2005, p. 391) emphasises that predictability is consistent with efficient markets: the ‘new view of the facts need not overturn the view that markets are reasonably competitive and therefore reasonably efficient’. But how is this claim consistent with the argument

outlined in section 2.1 of this paper that efficiency entails unpredictability? How might this second phase of research respond to the intuitive argument – illustrated by Cornell’s example – that ‘If one could be sure that a price will rise, it would have already risen’?

To answer this question, it is necessary to distinguish between normal and abnormal returns. The argument in the previous section only establishes that efficient markets prevent participants from exploiting information to make *abnormal* returns. But the argument does not rule out the possibility that *normal* returns may change over time in a predictable fashion. This point is made by Peirson et al. (2006, p. 515) in the passage below. This passage is discussing the ‘random walk hypothesis’ – the hypothesis that prices are a random walk – which is one version of the hypothesis that returns are unpredictable. Peirson et al. observe that the first phase of research on predictability tied the random-walk hypothesis to market efficiency, but that, in fact, this hypothesis is not implied by market efficiency. Rather, the efficient markets hypothesis (EMH) only implies that information cannot be exploited ‘for earning *abnormal* returns’.

Evidence in support of this so-called random-walk hypothesis was later interpreted as support for market efficiency, although it is now clear that a random walk does not imply, nor is it implied by, market efficiency. The random-walk model assumes that successive price changes are independent and are identically distributed over time. Neither assumption is necessary for prices within a market to fully reflect all information contained in part price series. The EMH requires only that an analysis of prices cannot be used as the basis for earning *abnormal* returns (italics added).

Even if markets are efficient, therefore, theoretical considerations alone do not provide reason to think that returns are unpredictable. It is theoretically possible, even if markets are efficient, that *normal* returns change over time in predictable ways. A number of phase-two theoretical models explain how normal returns might evolve predictably over time. Section 3 of the paper discusses three kinds of theoretical models. The first purports to explain why excess returns may be predictable using information about the dividend yield. The second presents a similar explanation for why excess returns may be predictable from the risk-free rate. The third provides a theoretical account of the relationship between excess returns and the volatility of returns.

The following quotation from Lettau and Ludvigson (2001, p. 815) similarly conveys the sense that something of a consensus had been reached about the predictability of excess returns. They emphasize the use not only of price-dividend ratios as predictors, but also price-earnings ratios and dividend-earnings ratios.

Indeed, the forecastability of stock returns is well documented. Financial indicators such as the ratios of prices to dividends, price to earnings, or dividends to earnings have predictive power for excess returns over a Treasury-bill rate.

This suggestion by Lettau and Ludvigson that predictability is uncontroversial was made over a decade ago. Since that article appeared, a new, third phase of research has emerged, reigniting the controversy about whether returns are predictable.

2.3 The third phase: a renewed ‘healthy skepticism’ about predictability

The emergence of a third phase of research, in opposition to the second phase, is apparent in the July 2008 issue of *The Review of Financial Studies*. The third phase of research is exemplified by the lead article in the issue, Welch and Goyal (2008), ‘A Comprehensive Look at the Empirical Performance of Equity Premium Prediction’. In this article, Welch and Goyal use US data up to the end of 2005 to evaluate the predictability of excess returns. They consider a wide range of variables that have been claimed in the academic literature to forecast excess returns, including the dividend yield, stock variance and the risk-free rate (the Treasury-bill rate). Their conclusion is that these models of excess returns have performed poorly.

Welch and Goyal distinguish between ‘in sample’ (IS) and ‘out of sample’ (OOS) performance of forecasting models. To understand this distinction, it may be helpful to consider the following passage in Brooks (2008, p. 245), which insists on the importance of OOS forecast performance:

In-sample forecasts are those generated for the same set of data that was used to estimate the model’s parameters. One would expect the ‘forecasts’ of a model to be relatively good in-sample, for this reason. Therefore a sensible approach to model evaluation through an examination of forecast accuracy is not to use all of the observations in estimating the model parameters, but rather to hold some of the observations back. The latter sample, sometimes known as the *holdout sample*, would be used to construct out-of-sample forecasts.

Welch and Goyal (2008) find that, of the variables that have been argued to predict excess returns, many produced poor IS forecasts. Moreover, they found that most variables that performed well IS performed poorly OOS. The conclusion of Welch and Goyal (2008, p. 1456) is:

Most models are no longer significant even in sample (IS), and the few models that still are usually fail simple regression diagnostics...Most models have poor out-of-sample (OOS) performance, but not in a way that merely suggests lower power than IS tests. They predict poorly late in the sample, not early in the sample...Therefore, although it is possible to search for, to occasionally stumble upon, and then to defend some seemingly statistically significant models, we interpret our results to suggest that a healthy skepticism is appropriate when it comes to predicting the equity premium, at least as of early 2006. The models do not seem robust’ (underlining added).

This ‘healthy skepticism’ about predictability characterises the third phase of research on stock returns. When assessing the forecasting power of various variables, Welch and Goyal (2008) compare the forecasting performance of these variables to the forecasts generated from a simple historical average. They find that the simple historical average performs no worse than the various variables they consider:

OOS, most models not only fail to beat the unconditional benchmark (the prevailing mean) in a statistically or economically significant manner, but underperform it outright’ (Welch and Goyal, 2008, p. 1504).

This quotation is especially relevant to the question of whether the MRP should be measured using an unconditional historical average or whether, instead, the regulator should use a conditional MRP. When regulators use a historical average to estimate the MRP as part of a CAPM calculation of the cost of equity, regulated businesses have argued that it is inconsistent to use a backward-looking historical average in the forward-looking CAPM framework. The above quotation from Welch and Goyal (2008), however, indicates how the historical average can be interpreted as a forward-looking measure of the risk premium. The authors find that the unconditional historical average performs no worse than other conditional estimates for the purpose of forecasting future excess returns. The historical average of excess returns can be construed as a forward-looking estimate of the MRP, therefore, because it performs relatively well as a predictor of future excess returns.

Welch and Goyal (2008) is not the only article in the June 2008 edition of *Review of Financial Studies* to question phase-two findings of predictability. A similar conclusion is articulated in another article in that edition, Boudoukh et al. (2008), ‘The Myth of Long-Horizon Predictability’. They question the conclusion of phase-two research that long-horizon predictability is considerably better than predictability over short-horizons

In their criticism of phase-two research, Welch and Goyal (2008) especially emphasize the importance of OOS tests and their concerns that the proposed models of equity returns are unstable. The problem of model instability is discussed in more detail in Section 4.2. Another related criticism of phase-two research stresses problems with data-mining. Thus, for example, on the basis of a study of stock returns in a variety of countries, Bossaerts and Hillion (1999, p. 417) concluded that when formal model selection criteria are used – so as to avoid data mining – models of excess returns perform poorly out-of-sample:

All this indicates poor out-of-sample predictability. ...Except for Japan, we find that none of the t-ratios of the slope coefficient in the SUR regression of out-of-sample outcomes onto predictions are significant. Consequently, *we fail to find that models chosen on the basis of formal selection criteria have any external validity.* (italics in original)

The Bossaerts and Hillion (1999) paper will be discussed again in Section 4.3 below, which elaborates on the problems arising from data mining.

A further criticism of phase-two research draws attention to technical statistical challenges that arise when testing and estimating models of equity returns. The explanatory variables typically used in these models include dividend yields, book-to-market values and interest-rates, which are highly persistent variables. Zhu (2013) explains how, in such circumstances, ordinary least squares (OLS) estimation can give rise to ‘spurious regressions’, generating ‘finite sample bias’.⁴ In such a case, standard tests of significance are invalid (Zhu, 2013, p. 194). To solve this problem, Zhu constructs an alternative method of estimation – his ‘jackknife’ technique – which, he argues, produces better estimates than OLS estimates.⁵ In order to illustrate the contrast between his new technique and OLS estimates, Zhu regresses excess returns against dividend yields and short-term interest rates. He finds that whereas standard OLS estimates ‘convey a well-celebrated message – strong return predictability’, this finding of predictability ‘vanishes completely’ when his preferred jackknife technique is used:

This strong statistical evidence of predictability, however, vanishes completely after removing finite-sample biases...It indicates that the finite-sample bias explains the bulk of apparent predictability. These empirical results cast doubt on the conclusions drawn in earlier studies regarding the predictive power of the dividend yield and the short rate. (Zhu 2013, p. 211)

Torous et al. (2004) is another paper which emphasises the technical statistical challenges arising from the fact that the explanatory variables in many models of equity returns are highly persistent. They also develop an estimation technique to address this challenge and, like Zhu (2013), they obtain very different results to those in phase-two studies:⁶

We find very little evidence of predictability at horizons greater than 1 year in the entire 1926-1994 sample period or in either the pre-1952 or post-1952 subsamples. In the post-1952 subsample, however, a variety of the explanatory variables, including the dividend yield as well as the term spread and the short-term rate of interest, are seen to forecast returns at the relatively short horizon of less than 12 months. In other words, contrary to previous evidence, we find reliable evidence of predictability at short horizons rather than long horizons. (Torous et al., 2004, pp. 939-940)

Other phase-three research has similarly questioned the predictability of returns over longer horizons. For example, Ang and Bekaert (2007, p. 696) find that:

Our results suggest that predictability is mainly a short-horizon, not a long-horizon, phenomenon.

⁴For another paper that emphasises the ‘spurious regression’ problem, see Ferson et al. (2003).

⁵For another technique designed to solve this problem, see Stambaugh (1999).

⁶For most of the empirical studies quoted in this paper, the dependent variable is future *excess* returns – the difference between the return and the risk-free rate. However, in Torous et al. (2004), the dependent variable is not future excess returns, but rather future returns.

In the context of an ACCC/AER regulatory decision, short-horizon predictability – predictability over horizons of a year or less – is less relevant than predictability over longer horizons. Why? Suppose that the regulated businesses have argued that, when calculating the regulatory cost of capital, the MRP should be set as a function of a particular variable. Suppose, moreover, that the regulatory period is five years. To make out a case for conditioning the MRP on that variable, the businesses must show that the variable can be used to forecast excess returns over a horizon of five years. If the empirical evidence only shows that the variable can forecast excess returns over a shorter horizon – over a year, a quarter, or a month – then the business has not established a case for conditioning the MRP on that variable. The conclusions of Torous et al. (2004) and Ang and Bekaert (2007, p. 696), therefore, which argue that predictability is, at best, ‘a short-horizon, not a long-horizon, phenomenon’, provide evidence that the conditional MRP is inappropriate for estimating the regulatory cost of capital, at least where the regulatory period is five years.

Ang and Bekaert (2007, p. 653) suggest that ‘the literature is converging to a new consensus, substantially different from the old view’. They are referring to the development of the third phase of research on return predictability. Their paper, as well as the findings in Bossaerts and Hillion (1999), Welch and Goyal (2008), Zhu (2013) and Torous et al. (2004), all represent contributions to this new phase of research, which questions the phase-two confidence in the predictability of returns.

3. The effect on the MRP of the dividend yield, risk-free rate and volatility

Whereas Section 2 of this paper described the broad claims made by phase-three studies about challenges in predicting excess returns, the current section will focus on the performance of specific variables as predictors of excess returns. Section 3.1 will point to empirical research that explores the performance of models that use the dividend yield to forecast excess returns. Section 3.2 examines the use of the risk-free rate in such forecasting models. Section 3.3 is concerned with the relationship between volatility and excess returns.

3.1 Dividend yields and the MRP

3.1.1 Theory

As discussed in Section 2, the second phase of research on predictability particularly emphasised the relationship between the MRP and the dividend yield. Such researchers have argued that a lower dividend yield implies a forecast of lower excess returns. Campbell and Cochrane (1999) have attempted to provide a theoretical account of this relationship. Their explanation is summarised in Cochrane (2005) as follows:

A natural explanation for the predictability of returns from price/dividend ratios is that people get less risk averse as consumption and wealth increase in a boom, and more risk averse as consumption and wealth decrease in a recession. (Cochrane 2005, p. 467).

Cochrane's argument can be put as follows. In a boom, the dividend yield (the inverse of one of the 'price/dividend ratios' referred to by Cochrane) tends to be low, as the price tends to be high. But in a boom 'people get less risk averse as consumption and wealth increase in a boom'. Therefore they require a lower premium for a given level of risk. Thus there is a positive correlation between dividend yields and the risk premium.

Cochrane (2005) notes that risk aversion does not vary with the *level* of consumption but rather it depends on consumption relative to a recent trend – relative to 'an accustomed standard of living'.

We cannot tie risk aversion to the *level* of consumption and wealth, since that increases over time while equity premia have not declined. Thus, to pursue this idea, we must specify a model in which risk aversion depends upon the level of consumption or wealth relative to some 'trend' or the recent past...Perhaps we get used to an accustomed standard of living so a fall in consumption hurts after a few years of good living, even though the same level of consumption might have seemed very pleasant if it arrived after years of bad times. This thought can at least explain the perception that recessions are awful events, even though a recession year may be just the second or third best year in human history rather than the absolute best (Cochrane 2005, p. 467).

In summary: a low dividend yield is associated with relatively high stock prices, which tend to occur in a boom; but in booms risk aversion tends to be low, which ensures that the MRP is low.⁷

3.1.2 Evidence

In the second phase of research on predictability, a number of empirical studies were presented that purported to establish a positive relationship between dividend/price ratios and future returns. A seminal paper on the empirical relationship between dividend yields and the MRP is Fama and French (1988), which claimed to discover an important distinction between short-term and long-term returns. While dividend yields do not explain returns over short horizons, they do, according to Fama and French, explain returns over longer horizons.

Our tests confirm existing evidence that the predictable (expected) component of returns is a small fraction of the short-horizon return variances. Regressions of returns on yields typically explain less than 5% of monthly or quarterly return variances. More interesting, our results add statistical power to the evidence that the predictable component of returns is a larger fraction of the variation of long-horizon returns. Regressions of returns on D/P often explain more than 25% of the variances of two- to four-year returns. (Fama and French, 1988, p. 4)

While Fama and French (1988) provide evidence that equity *returns* are predictable from dividend yields, Fama and French (1989) provide similar evidence that *excess returns* are predictable from dividend yields. Thus Fama and French (1988, p. 14) conclude:

The results for excess returns are similar to those for real returns... Thus the variation in expected real stock returns tracked by dividend yields is also present in the expected premiums of stock returns over one-month bill returns.

Cochrane (2005), which was cited in section 2.2 as exemplifying the second phase of research on predictability, reaches a similar conclusion to Fama and French (1988, 1989). Updating their study using a sample from 1947 to 1996, Cochrane (2005, p. 392) reports the following results from regressing excess returns on the dividend price ratio:

The one-year horizon 0.15 R^2 is not particularly remarkable. However, at longer and longer horizons larger and larger fractions on return variation are forecastable. At a five-year horizon 60% of the variation in stock returns is forecastable ahead of time from the price/dividend ratio.

⁷Cochrane (2005, p. 393) provides an intuitive explanation for why dividend yields might be a good predictor of long-term returns but a poor predictor of short-term returns. The explanation, Cochrane proposes, is that dividend yields are a 'slow-moving variable': 'If daily returns are very slightly predictable by a slow-moving variable, that predictability adds up over long horizons. For example, you can predict that the temperature in Chicago will rise about 1/3 degree per day in the springtime. This forecast explains very little of the day-to-day variation in temperature, but tracks almost all of the rise in temperature from January to July. Thus the R^2 rises with horizon'.

Lettau and Ludvigson (2001), which was also referred to in section 2.2 as an example of the second phase of research, comes to the same conclusion about the contrast between long and short horizon forecasts of the excess returns using dividend yields:

For example, at a horizon of one year, the dividend yield displays little predictive power for returns, the \bar{R}^2 is negligible and the coefficient estimate is not significantly different from zero. The dividend yield only becomes a significant forecaster at a return horizon of six years...Thus, consistent with existing evidence, the dividend yield is a strong forecaster of long-horizon returns, but has little capacity to forecast short- and medium-horizon returns. (Lettau and Ludvigson, 2001, pp. 839-40)

In the third phase of research, this conclusion that ‘the dividend yield is a strong forecaster of long-horizon returns’ has been questioned. As part of their general attack on predictability of returns – or at least predictability over longer horizons – the papers cited in section 2.3 (Welch and Goyal (2008), Boudoukh et al. (2008), Ang and Bekaert (2007), Torous et al. (2004) and Zhu (2013)) express, in particular, concerns about the use of the dividend yield as a predictor.

Over the past decade, a number of other papers have expressed similar concerns about the use of dividend yields to forecast excess returns. For instance, Welch and Goyal (2008) build upon their earlier paper Goyal and Welch (2003), which examined in particular the use of dividend yields to forecast excess returns. Goyal and Welch (2003) develop a diagnostic to evaluate the performance of a forecasting model. Using this diagnostic, they conclude that dividend ratios⁸ (including the dividend yield) are poor predictors of future excess returns out-of-sample.⁹ Moreover, the poor predictive ability of the dividend ratios is not simply the consequence of adding recent data. Rather, dividend ratios have always been a poor predictor of returns:

Our diagnostic shows that dividend ratios’ presumed equity premium forecasting ability was a mirage, apparent even before the 1990s. Despite good in-sample predictive ability for annual equity premia prior to 1990, §4 shows that dividend ratios had a poor out-of-sample forecasting ability even then....Thus, our paper concludes that the evidence that the equity premium has *ever* varied predictably with past dividend ratios has always been tenuous: A market-timing trader could not have taken advantage of dividend ratios to outperform the prevailing moving average – and should have known this. By assuming that the equity premium was

⁸Goyal and Welch use the following terminology to distinguish between the two key dividend ratios: ‘the total dividends paid by all stocks (D(t)), divided by the total stock market capitalization, either at the beginning of the year (the dividend yields, P(t-1)) or at the end of the year (the dividend-price ratio, P(t))’ (Goyal and Welch, 2003, p. 639).

⁹According to Goyal and Welch (2003), while the dividend yield does retain some forecasting power, it is only over forecasting horizons that are longer than regulatory periods: ‘Only on horizons greater than about 5 to 10 years does the Cochrane’s (1997) accounting identity (that dividend yields have to predict long-run dividend growth or market returns) begin to dominate the self-predictive properties of the dividend yield’ (Goyal and Welch, 2003, p. 640).

‘like it always has been’, a trader would have performed at least as well as in most of our samples’ (Goyal and Welch, 2003, p. 640).

This quotation provides support for estimating the MRP using a historical average, rather than an estimate conditional on the dividend yield. Goyal and Welch (2003) find that, for the purpose of forecasting excess returns, dividend yields perform no better than historical averages:

Neither the dividend-yield nor the dividend-price ratio had both the in-sample and out-of-sample performance that should have lead one to believe that it could outperform the simple prevailing equity premium average in an economically or statistically significant manner (Goyal and Welch, 2003, p. 653)..

Whereas Lettau and Ludvigson (2001) found that the dividend yield was a good predictor of excess returns over a six year horizon, their later paper Lettau and Ludvigson (2005, pp. 598-599) expresses scepticism about the power of dividend yields to predict returns. Note, however, they emphasise that the data since the middle of the 1990s is especially responsible for undermining the power of the dividend yield as a predictor:

The log dividend-price ratio has little power for forecast [*sic*] aggregate stock market returns from one to six years in this sample. These results differ from those reported elsewhere, primarily because we have included the last few years of stock market data in the sample....The extraordinary increase in stock prices in the late 1990s substantially weakens the statistical evidence for predictability by $d_t - p_t$ that had been a feature of previous samples.

Taken together, Goyal and Welch (2003), Welch and Goyal (2008), Ang and Bekaert (2007), Torous et al. (2004), Zhu (2013) and Lettau and Ludvigson (2005) provide evidence for scepticism about the performance of dividend/price ratios as a predictor of returns or excess returns, at least for the forecast horizons relevant to a five-year regulatory period.

3.2 The risk-free rate and the MRP

3.2.1 Theory

While the discussion in Section 3.1 above was specifically concerned with the relationship between the MRP and the dividend yield, it suggests a more general account of the movement in the MRP across time. In particular, Section 3.1 provides an explanation of why the MRP might be expected to be counter-cyclical – to rise in a slowdown and fall in a boom. Fama and French (1989) argue for such a counter-cyclical movement of the MRP, and Cochrane (2005, p. 392) summarises their conclusion as follows: ‘Expected returns vary with the business cycle; it takes a higher risk premium to get people to hold stocks at the bottom of a recession’. According to Cochrane, not only the MRP but also the dividend yield moves systematically with the business cycle, giving rise to a correlation between the two. But similar reasoning might be applied to other variables which, like the dividend yield, move

with the business cycle. For example, McKenzie and Partington (2013, pp. 25-26) address the question of whether a countercyclical movement of the MRP might give rise to a negative relationship between the risk-free rate and the MRP: if the MRP were counter-cyclical and if the risk-free rate were pro-cyclical, the implication would be that the MRP and the risk-free rate have a negative correlation. While this is perhaps the simplest theoretical argument for a negative correlation between the risk-free rate and the MRP, there are a variety of other arguments for such a correlation (Breen et al., 1989, p. 1178-9).

3.2.2 Evidence

Several recent empirical studies of US data have found a negative relationship between the risk-free rate and future excess returns. Lettau and Ludvigson (2001) find a negative relationship between the 'relative bill rate' and future excess returns.¹⁰ Ang and Bekaert (2007) show that three-month Treasury bills have a negative relationship with future excess returns.

When evaluating the relevance of such findings for estimating the regulatory cost of capital, however, it is important to distinguish between forecasts over short-term and longer-term horizons. The two papers cited above find that the interest rate performs well as a predictor only over short forecast horizons. Ang and Bekhaert (2007, p. 652) conclude that 'the most robust predictive variable for future excess returns is the short rate, but it is significant only at short horizons', and – in particular – it is not significant for horizons of five years. Moreover, Lettau and Ludvigson (2001, p. 840) find that when the relative bill rate is regressed against future excess returns, it is significant only for horizons of less than or equal to one year.

When setting the regulatory cost of capital for a five-year regulatory period, the regulator should be concerned with returns over the entire five years, not just for the first of the five years. The studies by Ang and Bekhaert (2007) and Lettau and Ludvigson (2001) would suggest that, given that the risk-free rate is only related to excess returns for forecast horizons that are significantly less than five years, it should not be used by regulators for setting the MRP when the regulatory period is five years. There have been instances where the consultants for the regulated businesses have, in fact, relied on Lettau and Ludvigson (2001) to justify a negative relationship between the risk-free rate and the MRP.¹¹ But they fail to mention that this correlation is only significant over horizons significantly shorter than the regulatory period, so this finding is of limited relevance for the purpose of estimating the regulatory cost of capital.

¹⁰The relative bill rate is a detrended bill rate. Campbell (1991, p. 166) explains why a detrended rate should be used: 'The short-term interest rate itself may be nonstationary over this sample period, so it needs to be stochastically detrended. The subtraction of a one-year moving average is a crude way to do this'.

¹¹ See, for example, Competition Economics Group, *Internal Consistency of Risk Free Rate and MRP in the CAPM*, March 2012, p. 8: 'there is a general consensus that the market risk premium tends to move in the opposite direction to the risk free rate – especially for material changes in the level of the risk free rate. For example, Lettau and Ludvigson find that the risk premiums tend to move in the opposite direction to the detrended government bond rate. Amongst other findings, they found a strongly statistically significant negative relationship between the de-trended US bill rates and the change in the log excess return (the variable they introduce akin to the MRP)'.

In his recent paper, Zhu (2013, pp. 209-11) finds that the risk-free rate is even a poor predictor of excess returns over short horizons. As discussed in Section 2.3, when Zhu applies his jackknife technique to correct for short-term bias, the evidence of the predictive power of the risk-free rate ‘vanishes completely’.

3.3 Volatility and the MRP

3.3.1 Theory

The theoretical relationship between the MRP and volatility can be derived from the seminal paper of Merton (1973), in which he presents his ‘intertemporal CAPM’ (ICAPM) model. In Merton’s model, the MRP changes over time. Armitage (2005, p. 82) provides a good introduction to the ICAPM, and he offers the following helpful characterisation of the model: the ICAPM ‘is a multifactor model in which an asset’s risk premium is determined by its sensitivity to state variables’.

While Merton (1973) is frequently cited in discussions of the relationship between the MRP and volatility, Merton does not specifically explore this relationship in his 1973 paper. Rather, this relationship is a focus of Merton (1980), which draws upon, and examines special cases of, his earlier 1973 paper. Under certain assumptions, Merton asserts, the MRP at a given time is a positive linear function of the conditional volatility at that time.

$$MRP_t = \theta \sigma_t^2$$

The coefficient θ depends, in turn, upon investors’ coefficients of relative risk aversion.¹²

On the one hand, a positive relationship between the MRP and volatility might seem to be intuitively plausible. After all, the MRP rewards investors for bearing risk, so if the risk increases it might be expected that the premium increases. On the other hand, the existence of such a relationship is not uniformly supported by the data (see section 3.3.2 below). In response, several authors have suggested that the positive relationship between the MRP and volatility is not as intuitively and theoretically plausible as it initially seems to be. For instance, Cornell (1999, p. 51) asserts that:

It turns out that the economic intuition that periods of high price variability should also be characterized by high stock-market returns is false.

Cornell suggests, furthermore, that the ambiguity of the relationship between returns and volatility ‘is not so surprising’ in the light of the following observation of Glosten et al. (1993, pp. 1779-80).

At first blush, it may appear that the rational risk-averse investors would require a relatively larger risk premium during times when the payoff of the security is

¹²See Merton (1980, pp. 329-330) for this result. For examples of studies which rely on Merton (1973, 1980) to derive a relationship between expected returns and conditional volatility, see Poterba and Summers (1986), p. 1143; Harvey (1989), p. 291; Dean and Faff (2001), pp. 172-73; Glosten, Jagannathan and Runkle (1993), pp.1781-82; Cuthbertson and Nitzsche (2004), p. 658.

more risky. A larger risk premium may not be required, however, because time periods which are relatively more risky could coincide with times when investors are better able to bear particular types of risk. Further, a larger risk premium may not be required because investors may want to save relatively more during periods when the future is more risky. If all the productive assets available for transferring income to the future carry risk and no risk-free investment opportunities are available, then the price of the risky asset may be bid up considerably, thereby reducing the risk premium. Hence a positive as well as a negative sign for the covariance between the conditional mean and the conditional variance of the excess return on stocks would be consistent with theory.

Despite the intuitive plausibility of the positive relationship between volatility and the MRP, and despite the influence of Merton's analysis, the relationship should be regarded as theoretically ambiguous. As Glosten et al. (1993) argue, theory alone is not sufficient to establish a positive relationship between the MRP and volatility. The question about the relationship between the MRP and volatility must be resolved by empirical evidence and not by theory.

3.3.2 Evidence

As well as pointing out the ambiguity in the theoretical relation, Cornell (1999, p. 51) also makes it clear that the empirical evidence on this relationship is decidedly mixed. Some econometric studies find the relationship to be insignificant; and those which identify a significant relationship disagree over its sign:

The literature on the relation between stock returns and the variability of returns includes contributions by Black (1976); Merton (1980); French, Schwert, and Stambaugh (1987); Poterba and Summers (1988); Breen, Glosten and Jagannathan (1989); Turner, Startz, and Nelson (1989); Nelson (1991); Campbell and Hetschel (1992); and Glosten, Jagannathan and Runckle (1993). All of these articles document the fact that the variability of returns changes over time. Unfortunately, their authors disagree as to how the changing variability is related to the risk premium. Some present findings that indicate a positive relation, others present findings that indicate a negative relation, and still others find no significant relation at all. If there is a bottom line, it is that the relation between stock returns and the variability of returns is remarkably weak...Whatever the explanation for the weak relation between the ex-post risk premium and the variability of returns, it means that return variability is not a good variable for modeling possible changes in the risk premium (underlining added).

A number of other researchers echo Cornell's conclusion that the empirical literature on the relationship between volatility and the equity premium is inconclusive. Thus Scruggs (1998, p. 575) observes that

It is surprising, however, that the empirical nature of this important relationship [between the market risk premium and conditional market variance] has not been resolved. Theory generally predicts a positive relation between the market risk premium and conditional market variance if investors are risk averse. Yet, empirical studies to date fail to agree on the sign of this important relation.

Scruggs provides a helpful summary of the findings in the empirical literature. The following table is abstracted from Scruggs (1998, p. 577), and it shows the divergent empirical findings on the relationship between the risk premium and variance.

Table 1:

Survey of Empirical Research on the Relation between the Risk Premium and Volatility¹³

Paper	Empirical relation between risk premium and market variance
French, Schwert and Stambaugh (1987)	Insignificant positive
	Insignificant positive
	Insignificant positive
Campbell (1987)	Significant negative
Harvey (1989)	Significant positive
Turner, Startz and Nelson (1989)	Significant negative
	Significant positive
	Significant positive
Baillie and DeGennaro (1990)	Significant positive
	Insignificant positive
	Insignificant positive
	Insignificant positive
Glosten, Jagannathan and Runkle (1993)	Insignificant positive
	Significant negative

In another summary of the empirical literature, Whitelaw (1994, p. 515-516) also emphasises the variety of divergent findings about the relationship between the MRP and volatility:

On a market-wide level, strong intuition suggests that risk and returns should be positively related. Consequently, researchers have searched for both a positive relation between expected returns and the conditional volatility of returns...Yet, prior empirical investigations into the contemporaneous correlation between the first two moments of stock market returns yield decidedly mixed results.

Dean and Faff (2001, p. 169-170) provide a similar characterisation of the empirical literature on the relationship between expected returns and volatility:

¹³Some papers provide multiple estimates because they may use multiple proxies, data sets, models of variance, estimation methodologies or specifications of the risk-return relation.

Although financial theory predicts that there should be a positive relationship between expected return and its variance, researchers cannot find consensus even on estimates of the sign of the relationship, let alone predict its magnitude.

More recent authors continue to emphasise the inconclusive character of the empirical literature on the relationship between volatility and expected returns. For instance, Cornell's conclusion that 'the relation between stock returns and the variability of returns is remarkably weak' is cited with approval in Armitage (2005, p. 88). Bollerslev et al. (2009, p. 4465) provide a similar description of the empirical literature:

The classical intertemporal CAPM model of Merton (1973) is often used to motivate the existence of a traditional risk-return tradeoff in aggregate market returns. Despite an extensive empirical literature devoted to the estimation of such a premium, the search for a significant time-invariant expected return-volatility tradeoff type relationship has largely proved elusive.

When Bollerslev et al. (2009) talk of 'a significant time-invariant expected return-volatility tradeoff' they mean a significant time-invariant positive relationship between expected returns and volatility.

On the basis of these observations of Cornell (1999), Scruggs (1998), Whitelaw (1994), Dean and Faff (2001), Armitage (2005) and Bollerslev et al. (2009), it can be concluded that there is no consensus in the empirical literature that there is a robust positive relationship between the MRP and volatility. It should be noted that in response to this conclusion, at least some researchers have attempted to devise more complex models of expected returns, in which various measures of volatility are used to model returns: see, for example, Guo and Savickas (2006) and Bollerslev et al. (2009). These models represent a considerable departure, however, from the simple relationship between volatility and the MRP that is typically presented in submissions by the regulated businesses. As will be discussed below in Section 4.1, given the diversity and complexity of the conditional models of the MRP in the current academic literature, it is unclear how regulators can (i) make an evidence-based choice of a particular conditional model of the MRP and (ii) implement such models.

4. Reasons for regulators to have practical concerns about conditional estimates

Sections 2 and 3 above summarise the debate about the predictability of excess returns. But even if it were conceded that excess returns are to some degree predictable, it would not follow that regulators should use a conditional estimate of the MRP. There are two kinds of reasons why, despite some evidence of predictability, regulators might have reason to avoid conditional models of the MRP. First, if markets are not perfectly efficient, return predictability does not imply a conditional MRP.¹⁴ Second, there are reasons why, in practice, regulators may face difficulties in applying the findings of economic research to their regulatory decision making. It is the second kind of reason that is the focus of this section. Sections 4.1 to 4.3 below specify three reasons why conditional estimates may potentially be especially problematical for regulators.

4.1 The diversity and complexity of recent models of return predictability

The current research literature on return predictability is characterised by a diversity of distinct and complex models of excess returns.

The third phase of the research literature has re-evaluated the claims of the second phase of research on predictability. For some researchers, such as Welch and Goyal (2008), this re-evaluation has taken the form of a broad scepticism about claims of return predictability. For other researchers, however, this re-evaluation has not taken the form of a denial of

¹⁴This observation is a commonplace of the academic literature, and is articulated by Peseran and Timmermann (1995, pp. 1201-2) in the passage below. They note that there is an interpretation of return predictability – their ‘second interpretation’ – according to which markets are inefficient, and the MRP is constant despite the predictability of excess returns:¹⁴

Many recent studies conclude that stock returns can be predicted by means of publicly available information...However, the economic interpretation of these results is controversial and far from evident. First, it is possible that the predictable components in stock returns reflect time-varying expected returns, in which case predictability is, in principle, consistent with an efficient stock market. A second interpretation takes expected returns as roughly constant and regards predictability of stock returns as evidence of stock market inefficiency...Inevitably, all theoretical attempts at interpretation of excess return predictability will be model-dependent, and hence inconclusive (see Fama (1991)).

Peseran and Timmermann’s quotation refers to Fama (1991, p. 1577), which makes the same point:

In brief, the new work says that returns are predictable from past returns, dividend yields, and the various term-structure variables...This means, however, that the new results run head-on into the joint-hypothesis problem: Does return predictability reflect rational variations through time in expected returns, irrational deviations of price from fundamental value, or some combination of the two?

predictability, but rather the investigation of a range of novel – and generally more complex – models of predictability.

The following sample of six recent studies of predictability illustrates (i) the influence of phase-three research on recent academic literature and (ii) the complexity and the diversity of the models formulated as a response to phase-three research.

1. Rapach et al. (2010) open by recognising the current scepticism about return predictability, citing phase-three research, including Bossaerts and Hillion (1999), Goyal and Welch (2003) and Welch and Goyal (2008). Their response to the problems identified by phase-three research is to move away from ‘individual forecasts’ – forecasts based on a single observed variable. Instead, they use fifteen different variables to estimate fifteen individual forecasts. They recommend using a forecast based on a combination of these fifteen individual forecasts.

2. Timmermann (2008) similarly motivates his paper by pointing to the scepticism about predictability articulated in Bossaerts and Hillion (1999), Goyal and Welch (2003) and Pesaran and Timmermann (1995). He suggests that nevertheless there may be ‘local predictability: ‘most of the time stock returns are not predictable, but there appear to be pockets in time where there is modest evidence of local predictability’ (Timmermann, 2008, p. 17). If an econometrician runs a range of models, he or she may be able to make local predictions if it is possible to measure which models are working well at different points of time: the econometrician must obtain ‘some indication of when different models produce valuable forecasts and when they fail to do so e.g. in the form of a real-time monitoring system tracking how reliable the forecasts have been over the recent time’ (Timmermann, 2008, pp. 16-17).

3. Cooper and Priestley (2009) also begin by acknowledging the skepticism about return predictability that arises from phase-three research, citing the papers by Bossaerts and Hillion (1999) and Goyal and Welch (2003). They respond by suggesting a novel variable for predicting excess returns – the output gap (the deviation of industrial production from trend).

4. Like the previous three studies, Pettenuzzo et al. (2012) opens by pointing to the phase-three research of Bossaerts, Hillion, Goyal and Welch. They suggest that the performance of forecasts can be improved by introducing constraints on the forecasting models – in particular, the conditional mean of the equity premium is constrained to be non-negative, and the conditional Sharpe ratio is constrained to lie within certain bounds.

5. Bollerslev et al. (2009) examine the relationship between volatility and the risk premium. They open their discussion by pointing to the failure of the empirical literature to establish such a relation, concluding that ‘the search for a significant time-invariant expected return-volatility tradeoff type relationship has largely proved elusive’ (Bollerslev et al. 2009, p. 4465). This encourages the authors to investigate a more complex relationship between expected returns and volatility. Rather than explaining returns by using a straightforward measure of volatility, they focus on the difference between two measures of volatility: their central conclusion is that ‘the difference between the “model free” implied and realized

variances is able to explain a nontrivial fraction of the variation in quarterly stock returns over the 1990-2007 sample period' (Bollerslev et al. 2009, p. 4465).

6. Like Bollerslev et al. (2009), Guo and Savickas (2006, p. 43) begin by pointing to empirical research which 'failed to uncover a positive risk-return relation in the stock market across time'. Their response is to model expected returns not as a function of a single measure of volatility, but rather as a function of two distinct measures of volatility, idiosyncratic volatility and stock market volatility. They find that whereas these two measures of volatility 'individually have negligible forecasting power in the in-sample regression, they jointly provide a significant predictor of excess stock market returns' (Guo and Savickas, 2006, p. 43).

These six papers all were published in or after 2006, and provide a representative sample of recent attempts to model excess returns. Responding to scepticism about predictability in the phase-three literature, these papers offer novel – and frequently more complex – model specifications. As a result of the phase-three literature, there is a considerable range of novel and complex models of excess returns in the academic literature. In this literature, there is no consensus – or anything approaching a consensus – on the appropriate set of methodologies for modeling future excess returns. Thus if a regulator were considering providing a time-varying, conditional model of the MRP, it is unclear how the regulator would make an evidence-based selection on the basis of this literature. Given the diversity and complexity of forecasting models, it would be difficult for the regulator not only to select but also to implement a model of the changes in the MRP over time.

4.2 Instability in models of return predictability

A number of studies have found instability in models of return predictability – that is, the models tend to change over time.

A parameter in a model is said to be unstable if it changes over time. The third phase of research on return predictability has emphasized the instability of models of excess returns. This instability provides a reason for a regulator to avoid conditional models of the MRP – to avoid setting the MRP as a function of some variable. If that function is unstable over time, it is difficult, if not impossible, for the regulator to measure accurately how the MRP should be adjusted in response to changes in the variable. This problem was alluded to in Section 2.3 (in the discussion of Welch and Goyal (2008)) and in section 4.1 (in the discussion of Timmerman (2008)), but it is a sufficiently important problem to warrant a more detailed discussion.

In Goyal and Welch (2003, p. 653), the diagnosis for the poor out-of-sample forecasting performance of dividend yields is that the underlying relationships are unstable:

The primary source of poor predictive ability is parameter instability. The estimated dividend-price ratio autoregression coefficient has increased from about 0.4 in 1945 to about 0.9 in 2002.

Bossaerts and Hillion (1999, p. 407) similarly suggest that model instability accounts for the poor OOS predictability of excess returns:

The poor external validity of the prediction models that formal model selection criteria chose indicates model nonstationarity: the parameters of the ‘best’ prediction model change over time.

The following passage provides more detail about the changes in the model of excess returns over time:

The discrepancy between the regression results of Period I (entire sample) and Period II (1/70-8/80) indicates the presence of model nonstationarity. The sign of the regression coefficients is almost always the same across the two periods, but the magnitude often differs dramatically (with Period II generating the highest values). Pesaran and Timmerman (1995) also document an increase in predictability of U. S. stock returns in the 1970s. (Bossaerts and Hillion, 1999, p.412)

Bossaerts and Hillion cite the results reported in Pesaran and Timmerman (1995, p. 1225), which also emphasise the instability in models of excess returns:

Also there does not seem to be a robust forecasting model in the sense that the determinants of the predictability of stock returns in the U.S. seems to have undergone important changes throughout the period under consideration. The timing of the episodes where many of the regressors get included in the forecasting model seems to be linked to macroeconomic events such as the oil price shock in 1974 and the Fed’s change in its operating procedures during the 1979 to 1982 period.

Even Lettau and Ludvigson (2001, p. 844), a study cited by the regulated businesses in support of return predictability, emphasises the changes in the performance of such models over time:

Although our findings of out-of-sample predictability are particularly strong relative to those of some other studies, we caution that our results do not imply forecastability in all episodes. It is clear, for example, that the last five years have been marked by highly unusual stock market behavior, as prices relative to any sensible divisor have reached unprecedented levels.

Model instability has been a focus of phase-three research over the past decade. Two other recent papers that particularly emphasise the challenges for return predictability arising from model instability are Pesaran and Timmermann (2002) and Paye and Timmermann (2006).

4.3 Data mining

Findings of predictability may reflect data mining rather than genuine relationships between variables and future excess returns. Data mining is a particular problem for research on return predictability.

What is data mining? Data mining (which is also referred to as ‘data dredging’ and ‘data snooping’) may be intentional or unintentional. The following is an example of unintentional data mining. Suppose that twenty different econometricians are attempting to ascertain the determinants of variable Y. Suppose each econometrician examines Y’s relationship to a different variable: the first tests the relationship between Y and variable X1, the second the relationship between Y and variable X2, and so forth. Suppose that, in fact, Y is not related to any of the twenty variables X1, X2...X20. Nevertheless, there is a good chance that at least one of the twenty econometric tests will be a ‘false positive’ – that is, even though there is no relationship between Y and the tested variable, the relationship will be found to be statistically significant at a 5 per cent level. Suppose, then, that one econometrician finds a significant relationship and the other nineteen do not. Suppose, moreover, that when making decisions about what to publish, academic journals look particularly favorably on articles that purport to discover new relationships of significance. As a consequence, the econometrician who finds a significant relationship is the only one that publishes his results (i.e., the nineteen econometricians who did not find a significant relationship are unable to publish their results). Then the body of academic literature on the determinants of Y will be misleading. In general, data mining refers to multiple uses of a given data set – that is, it refers to the ‘mining’ of a dataset.

Data mining may also be intentional. Suppose that an econometrician is cynically attempting to establish a relationship between variables X and Y but, in fact, there is no relationship between the two variables. Suppose that the econometrician attempts to estimate the relationship by testing a large range of different model specifications (e.g. different start dates, end dates, frequencies, proxies, specifications of outliers, functional forms, selections of other variables etc.) until he or she obtains a statistically significant relationship between X and Y. This exemplifies intentional data mining.

Data mining undermines findings of statistical significance. Suppose, for example, that when a dependent variable Y is regressed on X, the relationship between the two variables is found to be ‘statistically significant at a 5 per cent level’. The standard interpretation of this claim is that the probability that there is no relationship between the two variables is no more than 5 per cent. Suppose, however, that the econometrician obtains this finding only after regressing Y on twenty different variables, but only reports the model found to be ‘statistically significant’. Then the finding of ‘statistical significance’ is undermined: it cannot be maintained that the probability that there is no relationship between the two variables is no more than 5 per cent. Verbeek (2008, pp. 58-59) puts this point as follows:

In general, data snooping refers to the fact that a given set of data is used more than once to choose a model specification and to test hypotheses. You can imagine, for example, that, if you have a set of 20 potential regressors and you try each one of them, it is quite likely to conclude that one of them is significant, even though there is no true relationship between any of these regressors and the variable you are explaining.

Researchers in financial economics have long recognised that data mining is a particular problem for studies on return predictability. Thus in his article on market efficiency, Fama (1991, p. 1577) observes that:

We should also acknowledge that the apparent predictability of returns may be spurious, the result of data-dredging and chance sample-specific conditions.

He goes on to explain why data mining may be especially problematical in empirical research on return predictability:

Inference is also clouded by an industry-level data-dredging problem. With many clever researchers, on both sides of the efficiency fence, rummaging for forecasting variables, we are sure to find instances of “reliable” return predictability that are in fact spurious (Fama, 1991, p. 1585).

In an article entitled ‘Data-Snooping Biases in Tests of Financial Asset Pricing Models’, Lo and MacKinlay (1990, p. 432) make a similar point. Data mining is a particular problem for research on return predictability because – at least in part – there are a large ‘number of published studies performed on [a] single data set’:

We can expect the degree of such biases to increase with the number of published studies performed on any single data set – the more scrutiny a collection of data is subject to, the more likely will interesting (spurious) patterns emerge. Since stock market prices are perhaps the most studied economic quantities to date, tests of financial asset pricing models seem especially susceptible.

Their article argues that data mining renders ‘standard tests of significance’ invalid (p. 434). Sullivan, Timmermann and White (1999) investigate the potential for data mining to produce spurious trading rules – rules that appear to identify significant predictors of future excess returns, even though, in fact, they lack ‘predictive power’:

If enough trading rules are considered over time, some rules are bound by pure luck, even in a very large sample, to produce superior performance even if they do not genuinely possess predictive power over asset returns (Sullivan, Timmermann and White, 1999, p. 1649).

The potential for data mining creates a challenge for the regulator when it attempts to evaluate econometric studies that purport to provide a basis for a conditional MRP. It may be difficult, therefore, for the regulator to make an evidence-based selection of a conditional MRP model.

5. Conclusion

This paper explores methods for estimating the MRP for the purpose of regulatory pricing, focusing, in particular, on a comparison between conditional estimates and estimates based on historical averages. It points out the connection of this question to the debate about the predictability of excess returns, and surveys three phases of the literature on the predictability debate. The paper shows that the third phase might be used to support a historical estimate of the MRP. According to the third-phase of research, when forecasting excess returns, it is difficult to do better than a historical average; therefore the historical average can be construed as a good forward-looking estimate of the MRP.

The debate among researchers on predictability is, as Dimson et al. (2012, p. 36) put it, ‘far from settled’. But, even if it were conceded that excess returns are, to some degree, predictable from a given set of variables, the regulator faces at least three practical problems with using that set of variables to estimate a conditional MRP.

- (1) In response to skepticism about predictability in the third phase of research, the recent literature has developed a range of models that is increasingly (i) diverse, and (ii) complex. If a regulator were considering conditional models of the MRP, it would be difficult for the regulator to make an evidence-based selection of the appropriate model not only because of the diversity of touted models but also because of their increasing complexity. It would also be difficult to implement many of these models.
- (2) The third phase of research has particularly emphasised concerns about the stability of models of excess returns. A number of studies have found that the values of the parameters in the models of returns tend to change over time. Given the high degree of instability in models of excess returns, it is unclear how the regulator can set the MRP as a function of some specific variable. In particular, it is unclear how the regulator would ascertain how much the MRP should be adjusted in response to movements in that variable.
- (3) Apparently significant relationships between variables and excess returns may reflect data mining.

REFERENCES

- Ang, Andrew and Geert Bekaert (2007), 'Stock Return Predictability: Is it There?', *Review of Financial Studies*, 20(3), pp. 651-707.
- Armitage, Seth (2005), *The Cost of Capital: Intermediate Theory* (Cambridge: Cambridge University Press).
- Baillie, Richard and Ramon DeGennaro (1990), 'Stock Returns and Volatility', *Journal of Financial and Quantitative Analysis* 25(2), pp. 203-214.
- Bollerslev, Tim, George Tauchen and Hao Zhou (2009), 'Expected Stock Returns and Variance Risk Premia', *Review of Financial Studies* 22(11), pp. 4463-4492.
- Bossaerts, Peter and Pierre Hillion (1999), 'Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?', *Review of Financial Studies* 12(2), pp. 405-28.
- Boudoukh, Jacob, Matthew Richardson and Robert Whitelaw (2008), 'The Myth of Long-Horizon Predictability', *Review of Financial Studies*, 21(4), July, pp. 1577-1605.
- Breen, William, Lawrence Glosten and Ravi Jagannathan (1989), 'Economic Significance of Predictable Variations in Stock Index Returns', *Journal of Finance*, 44(5), pp. 1177-1189.
- Brooks, Chris (2008), *Introductory Econometrics for Finance*, 2nd ed. (Cambridge, Cambridge University Press).
- Campbell, John (1987), 'Stock Returns and the Term Structure', *Journal of Financial Economics* 18, pp. 373-99.
- Campbell, John (1991), 'A Variance Decomposition for Stock Returns', *Economic Journal*, Vol. 101, No. 405, pp. 157-179.
- Campbell, John and John Cochrane (1999), 'By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior', *Journal of Political Economy*, 107(2), pp. 205-51.
- Campbell, John and Ludger Hentshel (1992), 'An Asymmetric Model of Changing Volatility in Stock Returns', *Journal of Financial Economics*, 31, pp. 281-318.
- Campbell, John and Samuel Thompson (2008), 'Predicting Excess Stock Returns out of Sample: Can Anything Beat the Historical Average?', *Review of Financial Studies*, 21(4), July, pp. 1509-31.
- Competition Economics Group (2012), *Internal Consistency of Risk Free Rate and MRP in the CAPM*, March.
- Cochrane, John (2005), *Asset Pricing*, 2nd ed. (Princeton: Princeton University Press).

- Cochrane, John (2008), 'The Dog that did not Bark: A Defence of Return Predictability', *Review of Financial Studies*, 21(4), pp. 1533-75.
- Cochrane, John (2011), 'Discount Rates', *Journal of Finance*, 66(4), pp. 1047-1108.
- Cooper, Ilan and Richard Priestley (2009), 'Time-Varying Risk Premiums and the Output Gap', *Review of Financial Studies*, 22(7), pp. 2801-2833.
- Cornell, Bradford (1999), *The Equity Risk Premium: The Long-Run Future of the Stock Market* (New York: Wiley).
- Cuthbertson, Keith and Dirk Nitzsche (2004), *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*, 2nd ed. (Chichester: John Wiley & Sons).
- Dean, Warren and Robert Faff (2001), 'The Intertemporal Relationship between Market Return and Variance: An Australian Perspective', *Accounting and Finance*, 41, pp. 169-196.
- Dimson, Elroy, Paul Marsh and Mike Staunton (2012), *Credit Suisse Global Investment Returns Sourcebook 2012* (Zurich: Credit Suisse Research Institute).
- Fama, Eugene (1970), 'Efficient Capital Markets: A Review of Theory and Empirical Work', *Journal of Finance* 25(2), pp. 383-417.
- Fama, Eugene (1991), 'Efficient Capital Markets: II', *Journal of Finance* 46(5), pp. 1575-1617.
- Fama, Eugene and Kenneth French (1988), 'Dividend Yields and Expected Stock Returns', *Journal of Financial Economics*, 22, pp. 3-25.
- Fama, Eugene and Kenneth French (1989), 'Business Conditions and Expected Returns on Stocks and Bonds', *Journal of Financial Economics*, 25, pp. 23-49.
- Person, Wayne, Sergei Sarkissian and Timothy Simin (2003) 'Spurious Regressions in Financial Economics?', *Journal of Finance*, 58(4), pp. 1393-1413.
- French, Kenneth, William Schwert and Robert Stambaugh (1987), 'Expected Stock Returns and Volatility', *Journal of Financial Economics* 19, pp. 3-29.
- Glosten, Lawrence, Ravi Jagannathan and David Runkle (1993), 'On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks', *Journal of Finance* 48(5), pp. 1779-1801.
- Goyal, Amit and Ivo Welch (2003), 'Predicting the Equity Premium with Dividend Ratios', *Management Science*, 49(5), pp. 639-54.
- Guo, Hui and Robert Savickas (2006), 'Idiosyncratic Volatility, Stock Market Volatility, and Expected Stock Returns', *Journal of Business and Economic Statistics*, 24(1), pp. 43-56.
- Harvey, Campbell (1989), 'Time-Varying Conditional Variances in Tests of Asset Pricing Models', *Journal of Financial Economics*, 24, pp. 289-317.

- Lettau, Martin and Sydney Ludvigson (2001), 'Consumption, Aggregate Wealth and Expected Stock Returns', *Journal of Finance*, 56(3), pp. 815-849.
- Lettau, Martin and Sydney Ludvigson (2005), 'Expected Returns and Expected Dividend Growth', *Journal of Financial Economics*, 76, pp. 583-626.
- Lo, Andrew and A. Craig MacKinlay (1990), 'Data-Snooping Biases in Tests of Financial Asset Pricing Models', *Review of Financial Studies* 3(3), pp. 431-467.
- McKenzie, Michael and Graham Partington (2013), 'Report to the AER: Review of the AER's Overall Approach', report dated 14 February, 2013.
- Merton, Robert (1973), 'An Intertemporal Capital Asset Pricing Model', *Econometrica* 41(5), pp. 867-87.
- Merton, Robert (1980), 'On Estimating the Expected Return on the Market', *Journal of Financial Economics*, 8, pp. 323-61.
- Nelson, Daniel (1991), 'Conditional Heteroskedasticity in Asset Returns: A New Approach', *Econometrica* 59(2), pp. 347-370.
- Paye, Bradley and Allan Timmermann (2006), 'Instability of Return Prediction Models', *Journal of Empirical Finance* 13, pp. 274-315.
- Peirson, Graham, Rob Brown, Steve Easton, Peter Howard, Sean Pinder (2006), *Business Finance*, 9th ed. (North Ryde: McGraw-Hill).
- Pesaran, Hashem and Allan Timmermann (1995), 'Predictability of Stock Returns: Robustness and Economic Significance', *Journal of Finance* 50(4), pp. 1201-28.
- Pesaran, Hashem and Allan Timmermann (2002), 'Market Timing and Return Prediction under Model Instability', *Journal of Empirical Finance* 9, pp. 495-510.
- Pettenuzzo, David, Allan Timmermann and Rossen Valkanov (2012), 'Forecasting Stock Returns Under Economic Constraints', (7 December, 2012). Available at SSRN: <http://ssrn.com/abstract=2186574> or <http://dx.doi.org/10.2139/ssrn.2186574>.
- Poterba, James and Lawrence Summers (1986), 'The Persistence of Volatility and Stock Market Fluctuations', *American Economic Review*, 76(6), pp. 1142-1151.
- Poterba, James and Lawrence Summers (1988), 'Mean Reversion in Stock Prices', *Journal of Financial Economics* 22, pp. 27-59.
- Rapach, David, Jack Strauss and Guofu Zhou (2010), 'Out-of-sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy', *Review of Financial Studies* 23(2), pp. 821-862.
- Samuelson, Paul (1965), 'Proof that Properly Anticipated Prices Fluctuate Randomly', *Industrial Management Review* 6(2), pp. 41-49.

- Scruggs, John (1998), 'Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two Factor Approach', *Journal of Finance*, 53(2), pp. 575-603.
- Stambaugh, Robert (1999), 'Predictive Regressions', *Journal of Financial Econometrics*, 54, pp. 375-421.
- Sullivan, Ryan, Allan Timmermann and Halbert White (1999), 'Data-Snooping, Technical Trading Rule Performance, and the Bootstrap', *Journal of Finance* 54(5), pp. 1647-1691.
- Timmermann, Allan (2008), 'Elusive Return Predictability', *International Journal of Forecasting*, 24 (2008), pp. 1-18.
- Torous, Walter, Rossen Valkanov and Shu Yan (2004), 'On Predicting Stock Returns with Nearly Integrated Explanatory Variables', *Journal of Business*, 77(4), pp. 937-966.
- Turner, Christopher, Richard Startz and Charles Nelson (1989), 'A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market', *Journal of Financial Economics* 25, pp. 3-22.
- Verbeek, Marno (2008), *A Guide to Modern Econometrics* (Chichester: Wiley).
- Welch, Ivo and Amit Goyal (2008), 'A Comprehensive Look at the Empirical Performance of Equity Premium Prediction', *Review of Financial Studies*, 21(4), pp. 1455-1508.
- Whitelaw, Robert (1994), 'Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns', *Journal of Finance* 49(2), pp. 515-541.
- Zhu, Min (2013), 'Jackknife for Bias Reduction in Predictive Regressions', *Journal of Financial Econometrics*, 11(1), pp. 193-220.