NORGES BANK
INVESTMENT MANAGEMENT

01|2016

THE EQUITY RISK PREMIUM

DISCUSSION NOTE

In this note, we review the extensive theoretical and empirical evidence on one of the most important variables in financial economics – the equity risk premium (ERP).

Date 07/10/2016
ISSN 1893-966X

dn@nbim.no
www.nbim.no

SUMMARY

THE EQUITY RISK PREMIUM

In this note, we review the extensive theoretical and empirical evidence on one of the most important variables in financial economics – the equity risk premium (ERP). We describe the distribution of the realised ERP across different markets and time periods, and estimate the forward-looking (expected) ERP using a variety of models, including fundamental-based, regression and discounted cash flow models. We discuss the vast theoretical literature on the ERP and various explanations that researchers have put forward to explain the observed magnitude of the ERP and its behaviour over time.

- The ERP is defined as the excess return of equities over risk-free securities. The expected ERP is the compensation required by investors to hold risky securities, and in theoretical models it reflects the equilibrium price of non-diversifiable equity market risk. The realised ERP over a particular historical period is an estimate of the expected ERP that also includes a forecasting error.

- The average realised annual ERP, calculated as the difference between the realised return on equities and short-term bills, has been large, averaging about 7 percent in the US and 5.5 percent in other major developed markets. However, the standard deviation of the realised annual ERP has also been large, ranging from 17 to 30 percent across different markets.

- The relatively large realised ERP is difficult to explain in the context of standard neoclassical macroeconomic models, an empirical observation known in the literature as the "equity premium puzzle". The main issue is that consumption growth is too smooth to be consistent with the observed premium.

- Various risk-, behavioural- and market friction-based explanations have been suggested, but no single model has been able to unequivocally resolve the puzzle. Empirical evidence suggests that economic risk has been an important driver of the ERP, but investor behaviour and institutional factors have also likely contributed to the ERP's behaviour through time.

- Variation in fundamental metrics such as the price-dividend ratio, coupled with evidence of equity return predictability, suggests that the expected ERP may be time-varying. Forward-looking metrics may be useful indicators of future changes in required rates of return, cash flow growth, or both.

- The expected World ERP can be estimated from historical averages, fundamentals, predictive and cross-sectional regressions, and dividend discount models. Models produce heterogeneous results, but broadly agree on the current level of the expected ERP relative to history.

- The average World ERP based on data from 1970 to 2015 is 6.4 percent. Adjusting the average for repricing over the period lowers the average to 3.9 percent.

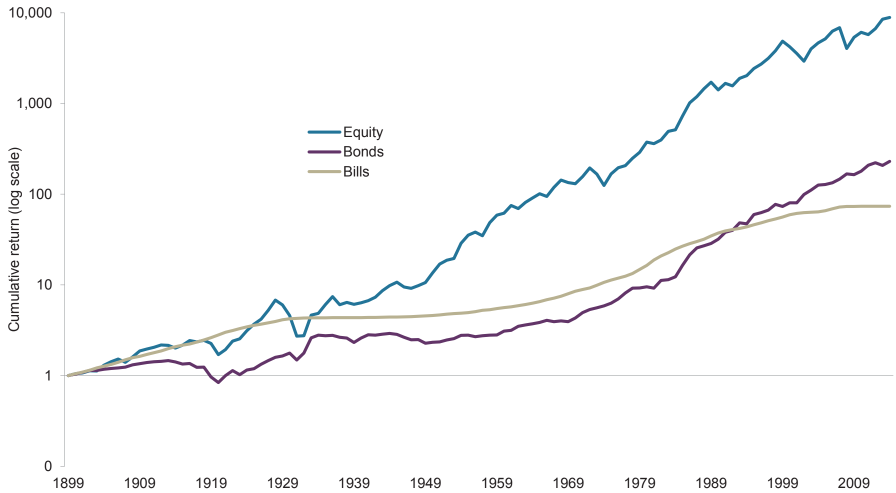
- The average World ERP estimate from various dividend discount models is 5.9 percent. These estimates may be affected by recent data bias. Cash flow growth has been exceptionally large since the end of the Global Financial Crisis in 2009, which in turn may bias upward expectations of future cash flow growth when extrapolated from historical data. In a below-average cash flow growth scenario, the estimated World ERP is 3.7 percent. Estimates of the expected ERP are also affected by the choice of proxy for the future risk-free rate. The current near-zero short-term interest rates may be a poor proxy for future short-term rates if the market expects rate increases in the future. The expected World ERP from the discount models may be closer to 4 percent if expectations of interest rate normalisation are taken into account.

- Estimates from cross-sectional and time-series models also suggest an expected World ERP of 3 to 4 percent.

I. Introduction

Investors in global equity markets have been well rewarded historically for bearing equity market risk. Over the period from 1900 to 2014, a dollar invested in global equity markets generated a return that is 38 times larger than the return of global government bonds and more than 120 times larger than the return of short-term government bills (Chart 1). While equity markets have experienced periods of extreme volatility, the spread between the return on equities and the return on government bonds has averaged 4.5 percent per year. The spread between the return on equities and the return on short-term government bills has been even larger, averaging 5.7 percent per year. This large and positive excess return of equities over bonds – termed the *equity risk premium* (ERP) – is an empirical measure of how much investors have been compensated historically for bearing equity market risk. It also underscores the tremendous wealth-building potential of equities over time and the central importance of the ERP in portfolio investment decisions.

Chart 1: Cumulative returns of global equities, bonds and bills in US dollars (1900–2014)



Source: Dimson, Marsh and Staunton (2015); Norges Bank Investment Management

The ERP is arguably one of the most important quantities in all of asset pricing from both a theoretical and a practical standpoint. As a theoretical concept, the ERP reflects the equilibrium price of equity market risk. It is the premium that investors demand to hold aggregate equity risk, which in turn affects the prices of all risky investments (Damodaran 2012). The ERP is determined by investors' collective risk aversion and the volatility of the equity market, and drives expectations of future equity market returns. It is therefore a key input in equilibrium asset pricing models like the capital asset pricing model (CAPM) and various multi-factor models. For example, the CAPM assumes that all securities are priced relative to their sensitivities (betas) to the excess return of the market portfolio. Securities with high exposure to market risk command higher expected returns and lower prices than securities with low market risk exposures. In multi-factor models like the Fama-French three-factor model, securities are again priced relative to their sensitivities to the ERP along with their sensitivities to additional risk factors for firm size and valuation. As an equilibrium concept, the ERP reflects the value-weighted average rate of return on equity required by investors, which also determines the value-weighted average cost of equity capital for firms.

From a practical standpoint, the ERP is a key input for a variety of short- and long-term investment decisions. Investors use estimates of the ERP to make long-term asset allocation decisions across equities, bonds and other asset classes, or to estimate whether a portfolio can generate enough return to meet various future obligations. Shorter-term investor views on the ERP can be used for making tactical allocations. Other stakeholders like governments and regulators use the ERP for budgeting purposes and to determine contributions to pension and healthcare funds. Firms use the ERP to estimate their cost of capital and to select among alternative projects. As equity investments have traditionally been the most important source of excess returns in financial markets, it comes as no surprise that the ERP is of central importance to both suppliers and users of financial capital.

While much progress has been made over the past 30–40 years in understanding the various risk premia in financial markets, we do not yet have a complete understanding of the ERP. The ongoing debate focuses on (1) the nature of the risk embedded in the ERP and the factors that determine the size and behaviour of the ERP over time; (2) the measurement of the expected ERP; and (3) the degree of predictability of the ERP. The ERP has been described in the literature as a "puzzle" (Mehra and Prescott 1985), a "mystery" (Kocherlakota 1996), an "anomaly" (Siegel and Thaler 1997) and a "controversy" (Jones and Wilson 2005) in response to the observed difficulty of standard macroeconomic models to explain the large historical average ERP. Researchers have proposed various statistical, risk-based and behavioural explanations for this empirical anomaly. While the debate regarding the structural determinants of the ERP is still ongoing, practitioners have come to rely on historical data, predictive regressions, discounted cash flow models and surveys to estimate the forward-looking ERP. These approaches, however, often produce diverging estimates of the ERP with large standard errors. Empirical estimates of the ERP are heavily dependent on the choice of a particular model and the data inputs used in the model. Finally, a controversy exists regarding the degree of predictability of the ERP and the ability of variables such as the dividend and earnings yields to convey information about future equity returns (Welch and Goyal 2008; Campbell and Thompson 2008).

In this note, we review the theoretical and empirical evidence on the ERP with a particular focus on the economic determinants of the ERP and the various ways that researchers and practitioners estimate the forward-looking ERP. We begin by providing the precise definition of the ERP, distinguishing between the historical (realised) ERP and the forward-looking (expected) ERP. To keep the distinction between the two clear, we refer to the historical ERP as the "realised ERP" and the forward-looking ERP as the "expected ERP" throughout the note. We survey the literature on the economic foundations of the ERP and the many explanations that researchers have put forward to explain the observed magnitude of the realised ERP. In the two empirical sections, we describe the distribution of the realised ERP across different markets and time periods and estimate the expected global ERP using a variety of models, including fundamental-based, regression and discounted cash flow models. Our goal is to illustrate the main approaches to estimating the expected ERP rather than to provide an exhaustive catalogue of all existing models.

II. Defining the ERP

The *expected* ERP can be defined as the difference between the expected future return of equities and the expected risk-free rate over a pre-specified horizon k (Duarte and Rosa 2015):

$$E_t[ERP_{t+k}] = E_t[R_{t+k}] - E_t[RF_{t+k}]$$

where $E_t[R_{t+k}]$ is the expected nominal equity market return over the time horizon k using the information available at time t and $E_t[RF_{t+k}] = RF_{t+k}$ is the nominal risk-free rate over the same time horizon, assumed to be known in advance.[1]

The *realised* ERP, on the other hand, is an ex post measure of the expected ERP (at time t) and is given by the difference between the realised equity return R_{t+k} and the risk-free rate RF_{t+k} over a given horizon k:

$$ERP_{t+k} = R_{t+k} - RF_{t+k}$$

If investors have less than perfect foresight, the realised ERP over the future horizon k will differ from the expected ERP. The difference is the forecasting error:

$$error_{t+k} = ERP_{t+k} - E_t[ERP_{t+k}]$$

As the expected ERP is unobservable at time t, we can only forecast it with an error whose size depends on how well we can forecast equity returns. Because of this forecasting error, the historical realised ERP may under- or overestimate the expected ERP. As Duarte and Rosa (2015) point out, the expected ERP is inherently a model of investor expectations, which change in response to new information not known in advance. The goal of any model of the expected ERP is to produce an estimate of the future realised ERP over the horizon k with the smallest possible forecasting error.

While the theoretical definition of the expected ERP is fairly straightforward, it is challenging to estimate the ERP in practice, for several reasons. First, commonly used proxies for the return on the market portfolio and the risk-free rate are imperfect measures of the theoretical concepts. Proxies for the market portfolio typically consist of broad-based equity indices such as the S&P 500 in the US or the MSCI World index for global stocks. In theory,

1 *Alternatively, the expected ERP can be defined as the difference between the expected* real *equity return and the expected* real *interest rate. Throughout this note, we calculate the realised and the expected ERP from nominal equity returns and nominal bill/bond returns.*

however, the market portfolio should include not only publicly traded equities but also less frequently traded assets such as private firms, property and art, as well as intangible assets such as human capital (Roll 1977). As Lustig, Nieuwerburgh and Verdelhan (2013) point out, total household wealth, which includes human capital, is much less volatile than stock market wealth. As fluctuations in the human wealth portion of total wealth are primarily driven by fluctuations in long-term bond yields rather than stock market premia, using the equity market as a proxy for total wealth may overestimate the risk premium required by investors (Lustig, Nieuwerburgh and Verdelhan 2013). Even if we restrict ourselves to the universe of publicly traded equities, there are no hard and fast rules for constructing the equity market proxy. We can choose, for example, to include only certain types of stocks in the market portfolio (e.g. large-cap stocks only or stocks traded on a particular exchange) and weight the stocks according to their market capitalisations (most common) or use some other weighting scheme such as equal weights or fundamental-based weights (Ilmanen 2011). Throughout this note, we use a broad-based market capitalisation-weighted global index as a proxy for the market portfolio.

Similarly to the market portfolio, the risk-free rate can be defined in several ways. It is most commonly proxied by the yield on short-term government bills (Treasuries), but it can also be proxied by the yield on long-term bonds (e.g. the yield on ten-year bonds). In Section III we report estimates of the historical realised ERP relative to both short-term bills and long-term bonds to illustrate the differences between the two. In Section V, where we provide estimates of the expected ERP from various dividend discount and regression models, we use short-term rates as the risk-free rate proxy. We prefer to use yields on short-term bills for the expected ERP as short-term rates are not affected by variation in the term premium and relatively less affected by variation in the credit, default and inflation premia than long-term yields.

Second, equity returns can be stated as geometric or arithmetic mean returns. An arithmetic mean return is a simple average of a series of returns, whereas the geometric mean return is the compound rate of return that measures the average performance of a portfolio over a given time period. For example, an arithmetic average return of 5 percent over ten years indicates that in any given year we expect to earn a premium of 5 percent for a total premium of 50 percent over the entire ten-year period. The corresponding geometric average return, which takes compounding into account, is $(1+0.5)^{1/10} - 1 = 4.14$ percent per year. Because of reinvestment effects, geometric average returns are always equal to or lower than arithmetic average returns. While geometric average returns provide a better indication of investment returns over multiple periods, we focus on arithmetic averages in this note, as these returns have more desirable statistical properties. In particular, if returns are uncorrelated, the arithmetic average realised ERP is an unbiased estimate of the future one-period ERP (Brennan and Schwartz 1985).

Third, the expected ERP can be estimated conditional upon variables such as the dividend yield, expected earnings, capital gains or other assumptions about the future. The underlying assumption is that variation in these variables over time is correlated with variation in the expected ERP. Such estimates of the ERP are referred to as short-term conditional forecasts (Constantinides 2002). Unconditional forecasts of the ERP, on the other hand, are calculated from historical data on the realised ERP. These estimates assume that the expected ERP is constant over time. In the empirical sections of the paper, we present estimates of the expected ERP based on both conditional and unconditional forecasts.

Finally, estimates of the expected ERP necessarily depend on the choice of model used for the ERP forecast and the inputs of the model. The simplest model of the expected ERP is the realised ERP over some historical time period (i.e. the unconditional mean forecast). Alternatively, the expected ERP can be estimated from discounted cash flow models (also known as market-implied ERPs) and various time-series and cross-sectional regression models based on stock fundamentals or macroeconomic data (conditional forecasts). In Section V, we discuss the advantages and disadvantages of each estimation approach.

To summarise, the expected ERP is a forward-looking measure of aggregate investor expectations about equity market risk. The realised ERP is an ex post measure of the risk premium that reflects both the expected ERP and a forecasting error. Unless investors have perfect foresight, the level and volatility of the realised ERP will differ from those of the expected ERP. The realised ERP, which is typically measured as the difference between the realised equity return of a broad-capitalisation equity market index and the return on short-term government bills, provides a natural starting point for discussing the main stylised facts about the expected ERP.

III. The realised ERP

1. The US equity risk premium

Historical data provide us with a wealth of evidence that equities have earned a substantial premium over government bills and bonds. In Table 1, we report summary statistics for the realised ERP in the US based on some of the most well-known datasets in the literature from 1871 to 2015. The realised arithmetic average ERP relative to Treasury bills is 7 to 8 percent per year with a standard deviation of 18–20 percent, translating to a standard error of 2 percent for the point estimates. The geometric average return ranges from 5 to 6 percent, suggesting that a dollar invested in equities in 1927 earned a 200–300 times higher compound return than a dollar invested in Treasury bills over the period from 1927 to 2015. The large size of the realised ERP underscores that equities have been a tremendous source of excess returns over the past century.

Table 1: Realised ERP relative to bills in the US

Data	Period	Arithmetic mean ERP	Geometric mean ERP	Std dev	Std error	Sharpe
CRSP	1927–2015	7.8	6.2	18.7	2.0	0.42
Shiller	1871–2012	7.0	5.3	17.7	1.5	0.40
Damodoran	1928–2015	7.9	5.9	20.1	2.1	0.39
Ibbotson	1926–2012	7.1	5.1	20.1	2.2	0.35

Source: CRSP database; Robert Shiller data available at: www.econ.yale.edu/~shiller/data/ie_data.xls; Aswath Damodaran data available at: http://www.stern.nyu.edu/~adamodar/pc/datasets/histretSP.xls; Ibbotson Associates (2013); NBIM calculations. The equity return in the CRSP data consists of a market capitalisation-weighted index of all stocks traded on the NYSE, Nasdaq and Amex exchanges. The market portfolio in the Shiller and Damodaran data is the S&P 500 index; the one in the Ibbotson Associates data consists of large-cap stocks. The risk-free rate is proxied by the return on three-month Treasury bills in the CRSP, Damodaran and Ibbotson data and by the return on one-year Treasury bills in the Shiller data. The equity premium in the CRSP, Damodaran and Ibbotson data is calculated from nominal equity returns and Treasury bills; the equity premium in the Shiller data is calculated as the difference between real equity returns and real Treasury returns. The Sharpe ratio is calculated as the realised arithmetic mean ERP divided by the realised ERP's standard deviation.

The average realised ERP relative to ten-year Treasury bonds is also positive, but smaller than the average realised ERP relative to short-term Treasury bills. As reported in Table 2, equities outperformed Treasury bonds by 6.2 percent on average on a year-on-year basis and 3.8 percent on an annual compound basis over the period from 1928 to 2015. The smaller realised ERP over bonds reflects the larger realised average return of bonds over the period, which had an average compound return of 5 percent compared to 3.4 percent for bills. The volatility of the excess equity return over bonds is also 1.4 percentage points higher than the excess return over bills, consistent with the fact that the return of long-term bonds has historically been more volatile than the return of Treasury bills. Because of the higher volatility of long-term bonds, the realised ERP relative to bonds has a lower Sharpe ratio than the realised ERP over bills.

Table 2: Realised ERP relative to bills and bonds in the US based on Damodaran data

	Period	Arithmetic mean return	Geometric mean return	Std dev	Std error	Sharpe
Equities	1928–2015	11.4	9.5	19.8	2.1	
3-month bills	1928–2015	3.5	3.4	3.1	0.3	
10-year bonds	1928–2015	5.2	5.0	7.8	0.8	
ERP (equity vs. bills)	1928–2015	7.9	5.9	20.1	2.1	0.39
ERP (equity vs. bonds)	1928–2015	6.2	3.8	21.5	2.3	0.29

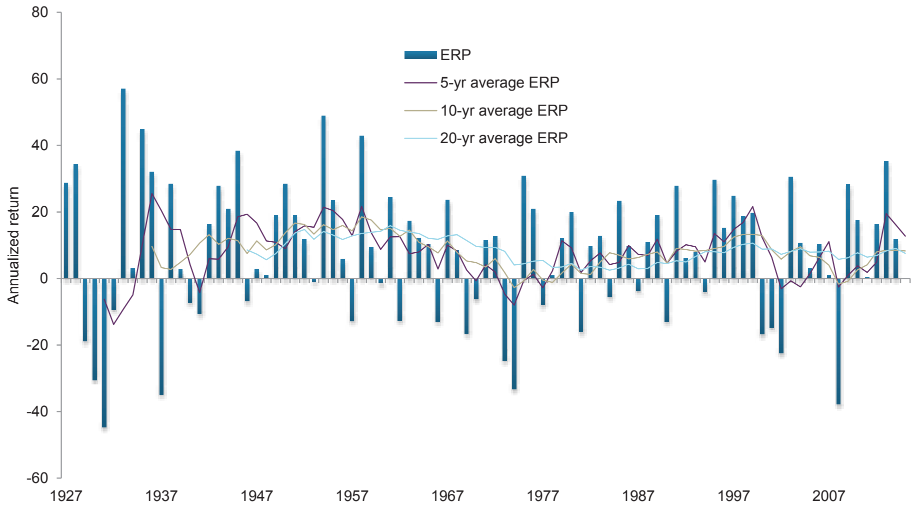
Source: Aswath Damodaran data available at: http://www.stern.nyu.edu/~adamodar/pc/datasets/histretSP.xls; NBIM calculations. The return of equities is proxied by the return of the S&P 500 index. The Sharpe ratio is calculated as the realised arithmetic mean ERP divided by the realised ERP's standard deviation.

While large on average, the realised ERP has also exhibited substantial variability over time. In Chart 2 below, we show the annual realised ERP relative to bills in every year from 1927 to 2015 as well as five-, ten- and 20-year rolling averages calculated from monthly CRSP data. The 20-year average realised ERP ranges from 2.5 to 15.9 percent. The average realised

ERP peaked after the end of WWII, during the bull market of the 1990s and more recently after the end of the Global Financial Crisis. The 20-year average was particularly low during the Great Depression of the 1930s and following the oil price shocks of the 1970s, when the US economy experienced a period of low growth and high inflation, resulting in especially low equity returns and large short-term Treasury rates. The large variability of the realised ERP suggests that the expected ERP may be time-varying as well, if the price and quantity of non-diversifiable macroeconomic risk vary over time or, alternatively, if investors are subject to behavioural biases such as overconfidence and "irrational exuberance" (Shiller 2005). As we discuss in Section IV, a large body of literature has examined the time-varying properties and predictability of the ERP, suggesting that forward-looking stock-specific and macroeconomic indicators may be able to forecast the future performance of the stock market. The observed variation in metrics such as dividend yields, which has been shown to correlate with future equity returns, suggests that the expected ERP may not be constant over time.

The significant variability of the realised ERP over time makes it very difficult to extrapolate whether the realised average premiums of 7–8 percent reported in Table 1 will persist in the future. Even though these estimates are made based on nearly 90 years of data, the standard errors of the estimates are substantial. A standard error of 2 percent translates into a 95 percent confidence for the point estimate of +/- 4 percent. Thus, the true arithmetic mean ERP may be anywhere between 3 and 12 percent. Shorter time samples result in even wider confidence intervals. As well emphasised in the literature, investors should not expect the expected ERP to remain constant over time or necessarily be as large as the realised historical average ERP (Cochrane 2011; Ilmanen 2011; Fama and French 2002).

Chart 2: Average realised ERP relative to bills in the US, 1927–2015



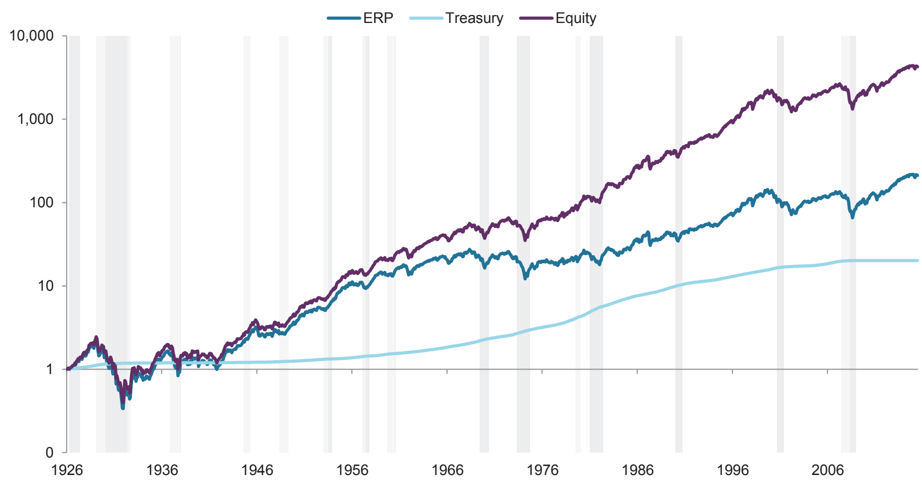
Source: CRSP database; Norges Bank Investment Management

It is also important to emphasise that the magnitude of the realised ERP is driven by both the return of equities and the return of government bills, which tend to perform differently under different economic scenarios. In Chart 3 below, we illustrate the link between the performance of US equities, Treasury bills, the realised ERP and economic growth. The shaded regions in the chart correspond to economic recessions.

First, equity returns and the realised ERP appear to be correlated with the business cycle, whereas the return of Treasury bills exhibits no strong correlation. The average equity return is 14.6 percent during economic expansions and -3.0 percent during economic recessions, whereas the average Treasury bill returns are a fairly constant 3 to 3.5 percent on average throughout this period. The realised ERP is therefore significantly larger during economic booms than during economic contractions. Economic recessions are also associated with significantly higher ERP volatility. Expectations about the business cycle are therefore an important determinant of the ERP.

Second, Treasury bill returns have significantly affected the size of the realised ERP at times, most notably during the 1970s when the US economy experienced double-digit inflation rates and high nominal interest rates, and more recently in the period since 2007 when expansionary monetary policy has led to record-low Treasury rates. By virtue of its effect on Treasury rates, monetary policy may affect the size of the ERP via the risk-free rate component of the equation.

Chart 3: Cumulative performance of US equities, Treasury bills and the ERP



Source: CRSP database; Norges Bank Investment Management. Shaded regions represent NBER recessions.

2. The global realised ERP

Many observers attribute the large realised premium in the US to the success of the US economy and the US equity market (Ilmanen 2011). Can the large historical realised ERP in the US be an artefact of survivorship bias? The evidence from global financial markets suggests that many other markets have posted large realised returns, albeit smaller than the US equity premium. In Table 3, we report summary statistics for the realised equity premia relative to short-term bills for 20 equity markets and three regions (Europe, World and World ex US) from 1900 to 2014 (Panel A) and from 1946 to 2014 (Panel B) based on the extensive studies of global equities and bonds of Dimson, Marsh and Staunton (2015). The realised arithmetic average

World ERP of 5.7 percent is about 2 percentage points lower than the realised ERP for the US, whereas the geometric average is 1.3 percentage points lower than that for the US. In the post-WWII period, we observe significantly less spread in excess stock market returns between the US and the rest of the World. The World compound average premium including the US of 5.8 percent is comparable to the 5.5 percent premium for the World ex US. During this period, European equity markets performed on a par with the US equity market, posting some of the largest realised returns in the sample. Germany, the Netherlands, France and Japan performed significantly better than the World market average in the post-war period. This is not surprising given that these countries experienced tremendous turmoil during the first half of the 20th century and their equity markets recovered from a very low base at the start of the post-WWII period.

Table 3: Realised ERP (relative to bills) in different countries and regions, 1900–2014

Panel A: 1900–2014

	Arith-metic average	Geo-metric average	Median	Std dev	Std error	Sharpe	Min	Max
Australia	8.1	6.6	10.4	17.5	1.6	0.47	-44.4	49.2
Belgium	5.4	3.0	2.2	23.9	2.2	0.23	-49.7	125.0
Canada	5.6	4.2	6.6	16.9	1.6	0.33	-34.7	49.1
Denmark	5.0	3.1	2.3	20.5	1.9	0.24	-50.3	95.3
Finland	9.5	5.9	5.4	29.9	2.8	0.32	-53.3	159.2
France	8.7	6.1	4.2	24.2	2.3	0.36	-43.1	85.7
Germany	9.7	5.9	5.5	31.2	2.9	0.31	-44.7	131.4
Ireland	5.8	3.5	2.6	21.3	2.0	0.27	-66.3	72.0
Italy	9.5	5.7	9.0	31.6	2.9	0.30	-48.6	150.3
Japan	9.3	6.1	2.8	27.7	2.6	0.34	-48.3	108.6
Netherlands	6.5	4.4	5.2	22.5	2.1	0.29	-51.4	126.7
New Zealand	5.9	4.4	6.1	18.1	1.7	0.33	-58.3	97.3
Norway	5.9	3.1	3.0	26.1	2.4	0.22	-55.0	157.1
Portugal	9.2	4.6	2.0	33.9	3.2	0.27	-75.1	141.2
South Africa	8.4	6.3	5.7	21.7	2.0	0.38	-33.9	106.2
Spain	5.5	3.4	2.9	21.6	2.0	0.25	-39.3	98.1
Sweden	5.9	3.9	7.4	20.5	1.9	0.29	-40.8	64.6
Switzerland	5.3	3.7	4.9	18.7	1.7	0.28	-37.0	54.8
UK	6.1	4.3	5.3	19.7	1.8	0.31	-54.6	121.8
US	7.5	5.6	10.1	19.6	1.8	0.38	-44.1	56.6
World	5.7	4.3	6.5	17.0	1.6	0.33	-41.9	68.6
WorldExUS	5.2	3.5	4.3	18.6	1.7	0.28	-45.1	80.9
Europe	5.2	3.4	4.1	19.3	1.8	0.27	-48.3	76.0

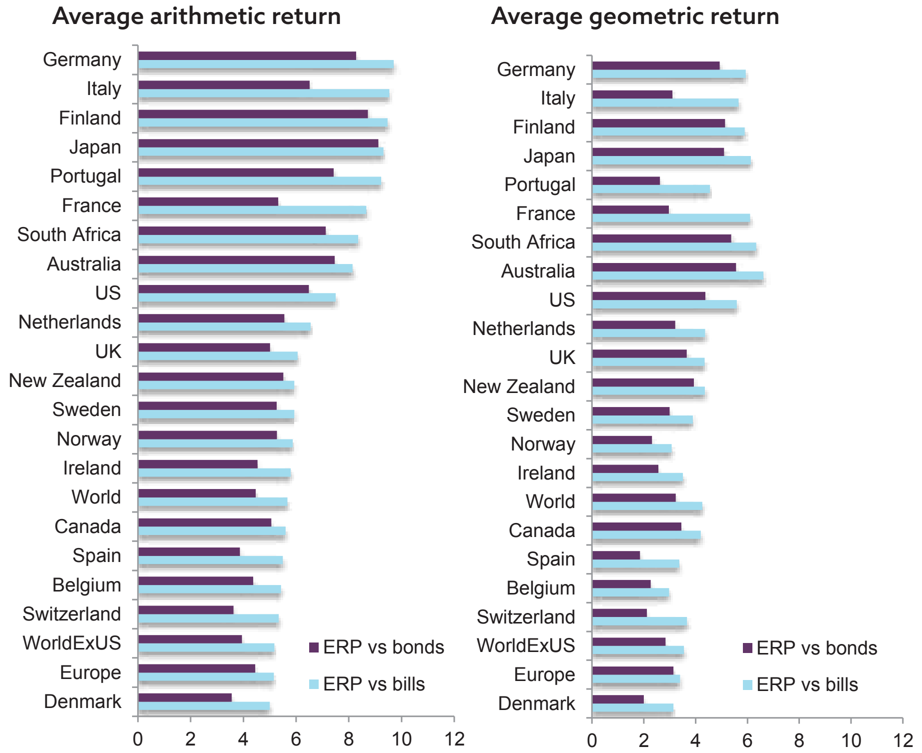
Panel B: Post-WWII (1946–2014)

	Arith-metic average	Geo-metric average	Median	Std dev	Std error	Sharpe	Min	Max
Australia	7.3	5.2	9.5	20.6	2.5	0.36	-44.4	49.2
Belgium	4.3	2.3	4.9	19.6	2.4	0.22	-49.7	41.9
Canada	6.1	4.7	6.2	16.6	2.0	0.37	-34.7	47.6
Denmark	6.7	4.1	2.7	24.0	2.9	0.28	-50.3	95.3
Finland	9.6	5.2	6.1	33.1	4.0	0.29	-53.3	159.2
France	9.4	6.5	6.7	25.2	3.0	0.37	-43.1	77.6
Germany	11.7	7.9	8.5	31.3	3.8	0.38	-44.7	131.4
Ireland	8.4	4.9	9.8	26.4	3.2	0.32	-66.3	72.0
Italy	8.6	4.7	9.2	31.5	3.8	0.27	-48.1	150.3
Japan	11.7	7.9	6.9	30.7	3.7	0.38	-43.7	108.6
Netherlands	8.3	6.2	7.6	21.2	2.5	0.39	-51.4	72.0
New Zealand	6.1	3.7	6.1	22.4	2.7	0.27	-58.3	97.3
Norway	7.8	3.8	6.8	31.6	3.8	0.25	-55.0	157.1
Portugal	10.0	3.8	4.6	38.9	4.7	0.26	-75.1	141.2
South Africa	7.9	6.0	4.7	20.9	2.5	0.38	-31.1	82.3
Spain	7.5	4.7	5.8	25.0	3.0	0.30	-39.3	98.1
Sweden	9.3	6.9	11.4	22.4	2.7	0.41	-40.8	64.6
Switzerland	7.8	5.8	8.6	20.7	2.5	0.38	-37.0	54.8
UK	8.0	5.7	7.6	23.2	2.8	0.34	-54.6	121.8
US	7.9	6.5	10.5	17.2	2.1	0.46	-38.2	48.7
World	7.2	5.8	9.0	16.8	2.0	0.43	-41.9	46.3
WorldExUS	7.3	5.5	8.0	19.4	2.3	0.38	-45.1	53.4
Europe	7.6	5.7	8.0	20.0	2.4	0.38	-48.3	64.4

Source: Dimson, Marsh and Staunton (2015) ; Norges Bank Investment Management

The global realised equity premia calculated relative to long-term government bonds are smaller but largely consistent with the premia relative to short-term bills as shown in Chart 4. The World arithmetic average realised ERP relative to bonds is 4.5 percent and the World geometric average realised ERP is 3.2 percent over the period from 1900 to 2014. For every country and region in the sample, the realised average compound ERP relative to bonds was lower than the corresponding realised ERP relative to bills, indicating that long-term bonds have historically been a source of higher returns than bills in most markets. The realised equity premia over bonds, however, are still large and significant.

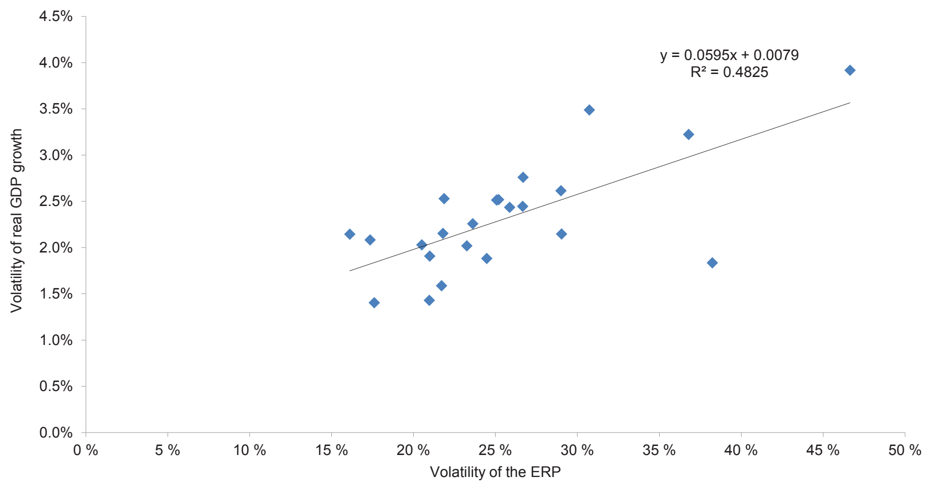
Chart 4: Global ERPs relative to short-term bills and long-term bonds, 1900–2014



Source: Dimson, Marsh, and Staunton (2015); Norges Bank Investment Management

While investments in equities have proved rewarding over the long run, relative to both short- and long-term bonds, they have been accompanied by significant variability of returns. The volatility of the realised ERP has ranged from 16.9 percent for Canada to 33.9 percent for Portugal over the full sample. The volatilities of the premiums in Germany, France, Italy and Japan are also very large, whereas the realised US ERP's volatility of 20 percent is one of the lowest in the sample. The cross-sectional dispersion in the volatility of the realised ERP appears to be positively correlated with the volatility of economic growth, as illustrated in Chart 5 below. The chart presents a scatter plot of the realised ERP volatility of the countries in the Dimson, Marsh and Staunton database from 1970 to 2014 against the volatility of the countries' real GDP growth over the same period. This strong positive association suggests that equity premia in countries with more stable and predictable economies may be relatively less risky than those in countries with high uncertainty about economic growth. This is consistent with research by Lettau, Ludvigson and Wachter (2008), which suggests a positive relationship between changing equity risk premiums in the US (as proxied by the dividend-price ratio) and shifting volatility in US GDP growth.

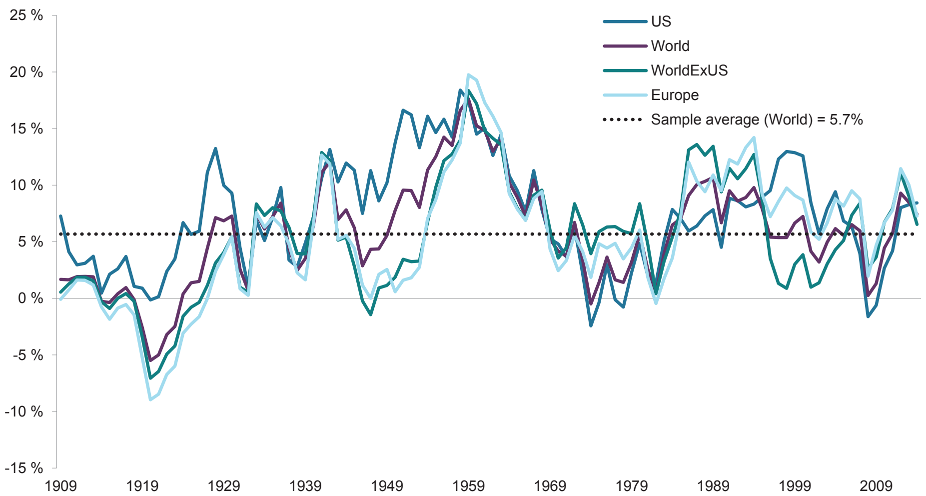
Chart 5: Volatility of the ERP in 20 countries vs volatility of GDP growth, 1970–2014



Source: Dimson, Marsh and Staunton (2015); USDA International Macroeconomic Data Set; Norges Bank Investment Management

Similarly to the realised ERPs in individual markets, the realised World ERP exhibits strong variability over time. As illustrated in Chart 6 below, the ten-year rolling average realised World ERP has varied from about -6 to 18 percent since 1910. Part of this volatility can be attributed to extreme events. As many European markets experienced war, hyperinflation and economic depressions during the first half of the 20th century and only began to recover after WWII, the volatility of the Worldwide premium in the early part of the sample was significantly higher than its volatility in the last 30 years. As Ilmanen (2011) points out, the possibility of such extreme events may have been "priced" in the US equity returns, thereby resulting in higher-than-expected returns when these risks did not materialise. Barro (2006) and Wachter (2013) present empirical evidence that the lack of rare macroeconomic risks may have contributed to the relatively large realised premium in the US. The global evidence, however, suggests the large premium in the US and other "successful" markets cannot be attributed solely to survivorship bias. As Dimson, Marsh and Staunton (2011) argue, the large premia earned by equity markets appear to be a universal phenomenon.

Chart 6: Ten-year rolling average ERP, 1910–2014



Source: Dimson, Marsh and Staunton (2015); Norges Bank Investment Management

3. Decomposing the equity return

In addition to time-varying risk premiums, survivorship bias and the possibility of extreme events that did not materialise, significant changes in equity valuations can also render historical averages of the realised ERP a poor proxy for the expected ERP. In fact, several authors attribute part of the large historical equity premiums in the US and elsewhere to "windfall gains" and luck, emphasising that equity markets have experienced upward repricing and unexpected capital gains during the second half of the 20th century. Grinold, Kroner and Siegel (2011) show that the average P/E ratio of the S&P 500 index, measured as the end-of-year price divided by trailing 12-month earnings, grew from 11.3 in 1925 to 18.5 in 2010, contributing about 0.6 percent of the equity return per year. Such upward repricing of equities, whereby investors become willing to pay a higher price per unit of company fundamentals, may not continue in the future. Similarly, Ibbotson and Chen (2003) perform a decomposition of the equity return into income growth, capital gains and changes in valuation, and show that the historical realised ERP is closer to 4 percent if it is adjusted for unexpected capital gains. Fama and French (2002) show that the realised equity premium in the US between 1951 and 2002 was about 2 percentage points higher than the premium implied by fundamentals over this period. They attribute this difference to changes in discount rates, which resulted in large unexpected capital gains for equity investors.

We illustrate the Fama-French (2002) argument in Table 4, where we compare the realised ERPs for 18 countries and four geographical regions with the estimated (unconditional) expected ERP from dividend yields and dividend growth over the period from 1970 to 2015. The data for this analysis come from MSCI, which provides pricing and equity fundamental data for major developed markets since 1970. The Fama-French idea is as follows. By definition, the average equity return can be expressed as the sum of the average dividend yield and the capital gain (price appreciation):

$$Ave(R_t) = Ave\left(\frac{D_t}{P_{t-1}}\right) + Ave\left(\frac{P_t}{P_{t-1}} - 1\right)$$

where $Ave(D_t/P_{t-1})$ is the average dividend yield and $Ave(P_t/P_{t-1}-1)$ is the average change in the equity price. The decomposition reflects the fact that equity returns stem from two sources: the dividends paid over a particular period and the price appreciation or depreciation (capital gain or loss) over the period. If the dividend-price ratio D_t/P_t is mean-reverting (i.e. D_t/P_t has a constant unconditional mean), the growth rate in price (P_t) should be well approximated by the growth in dividends (D_t). In order for the dividend-price ratio to have a constant mean, both the numerator (D_t) and the denominator (P_t) should grow, on average, by the same amount. If so, the unconditional average equity return $Ave(R_t)$ can be modelled as:

$$Ave(R_t) = Ave\left(\frac{D_t}{P_{t-1}}\right) + Ave\left(\frac{D_t}{D_{t-1}} - 1\right)$$

where we have replaced the average price appreciation with the average dividend growth. In the absence of permanent changes in the dividend-price ratio (no repricing), the dividend growth should provide an estimate of the unconditional expected capital gain component of the equity return. As aggregate dividend-price ratios may be driven by variations in discount rates (Cochrane 2011), unanticipated changes in the dividend-price ratio can result in unanticipated capital gains.

The results in Table 4 indicate that for the majority of countries in the sample the average realised ERP was larger than the implied ERP from the dividend growth model with no repricing. The average World total equity return of 11.5 percent per year (column [4]) is 2.5 percentage points larger than the sum of the average dividend yield of 3.1 percent (column [1]) and the average dividend growth of 5.9 percent (column [2]) over this period. This difference of 2.5 percentage points can be attributed to the repricing of dividends. For the Asia-Pacific region, this repricing is even more extreme: 4.5 percentage points of the total ERP of 6.7 percent cannot be explained by average dividend yields and dividend growth. With the exception of Australia, the ERPs implied from dividends (column [7]) are consistently lower than the realised ERPs. Price appreciation has consistently outpaced dividend growth, resulting in larger-than-expected capital gains. It is difficult to predict whether this trend of price appreciation is going to persist in the future. The results suggest that unexpected valuation changes may distort historical averages and the realised historical average ERP may be a poor proxy of future equity excess returns. The implied ERP from dividend growth can possibly serve as a better unconditional estimate of the mean ERP than the realised mean ERP, especially given the implied ERP's lower volatility and lower estimation error.

To summarise, the empirical evidence on the realised ERP indicates that equities have earned substantially larger returns than government bills and bonds. The unconditional arithmetic mean of the realised World ERP over Treasury bills is about 5–6 percent per year. The premium has varied significantly across countries and over time. The time-varying properties of the premium, together with the observed upward valuation changes and low interest rates in the past 30 years, make it difficult to extrapolate whether the historical premium is going to persist in the future, although it seems reasonable to conclude that, if anything, the expected ERP is lower rather than higher than the historical average.

Table 4: Realised average ERPs vs ERP estimates from dividend growth models (1970-2015)

	Average							Standard error	
	Div yield [1]	Div growth [2]	Price ret [3]	Equity ret [4] = [1]+[3]	Risk-free [5]	Realised ERP [6] = [4]-[5]	Implied ERP [7] = [1]+[2]-[5]	Realised ERP	Implied ERP
Australia	4.4	8.1	7.4	11.8	8.0	3.8	4.4	3.3	1.8
Austria	2.6	3.5	7.4	10.0	4.8	5.2	1.3	5.1	2.7
Belgium	6.3	5.8	7.6	13.9	6.0	7.9	6.1	3.8	2.6
Canada	3.1	6.5	7.8	11.0	6.1	4.9	3.6	2.9	1.7
Denmark	2.9	9.0	14.1	16.9	8.1	8.8	3.7	4.9	3.5
France	4.1	6.8	8.8	12.9	6.3	6.7	4.6	4.2	1.7
Germany	3.5	5.8	7.2	10.7	4.9	5.8	4.4	4.0	2.4
Hong Kong	3.9	10.4	19.2	23.1	5.0	18.0	9.2	7.3	1.9
Italy	3.0	7.0	8.3	11.3	8.4	2.9	1.7	4.8	3.5
Japan	1.6	3.7	8.1	9.7	3.6	6.1	1.7	4.3	1.7
Netherlands	4.5	4.8	8.3	12.8	4.8	8.0	4.5	3.5	1.4
Norway	3.3	11.0	12.6	15.9	7.2	8.6	7.1	6.0	4.1
Singapore	2.4	7.4	12.5	15.0	3.8	11.2	6.1	7.0	1.9
Spain	5.5	6.6	7.7	13.3	7.5	5.8	4.7	4.3	2.0
Sweden	3.4	12.5	14.8	18.2	6.5	11.8	9.4	4.8	2.6
Switzerland	2.4	7.4	7.5	9.9	3.4	6.5	6.4	3.3	2.0
UK	4.5	7.1	8.1	12.6	7.2	5.4	4.5	3.0	1.2
US	3.2	5.8	8.2	11.4	4.9	6.5	4.0	2.6	1.2
Asia-Pac	2.1	4.3	8.8	10.9	4.2	6.7	2.2	4.2	1.5
Europe	4.0	7.4	8.7	12.7	6.1	6.6	5.3	2.9	1.2
America	3.2	5.8	8.2	11.4	5.0	6.4	4.0	2.6	1.2
World	3.1	5.9	8.4	11.5	5.1	6.4	3.9	2.6	1.0

Source: MSCI World data; FactSet; IMF; Norges Bank Investment Management

IV. Explanations for the ERP

While there is consensus regarding the existence of a large historical realised ERP, there is less agreement on what drives the magnitude and variability of the premium through time. Researchers have proposed a number of theories that may account for the magnitude of the realised ERP, and by extension the expected ERP, including statistical, risk-based and behavioural explanations as well as explanations based on market frictions such as taxes, clientele effects, liquidity constraints and limited equity market participation. In this section, we review the main theoretical evidence on the determinants of the ERP. Research in this area was given impetus by the Mehra and Prescott (1985) observation that the large historical ERP in the US was inconsistent with standard formulations of consumption-based risk models. Much of the literature since then has tried to reconcile theory with the stylised facts of the realised ERP time series by modifying the assumptions of macroeconomic models or seeking explanations outside the standard framework.

1. The consumption-based capital asset pricing model (CCAPM) and the Mehra and Prescott (1985) equity premium puzzle

A general equilibrium risk model like the consumption-based capital asset pricing model (CCAPM) is a natural starting point for the question about the "right" size of the premium. The CCAPM is a generalisation of the traditional CAPM that recognises the intertemporal dimension of portfolio choice. The key idea of the model is that households trade financial assets in order to smooth consumption over time and across states of the economy. Assets that pay off in future states when consumption is low are more desirable and more highly valued than assets that pay off when future consumption is high and any additional unit of consumption provides less utility. In this framework, representative agents' willingness to substitute between consumption today and future consumption (their marginal rate of intertemporal substitution) along with their risk aversion determines the price of assets with uncertain future payoffs. Under the assumptions of homogenous constant risk aversion preferences and complete frictionless markets, the expected return on an asset becomes proportional to its "consumption" beta. The quantity of stock market risk is given in equilibrium by the covariance of the excess stock return with consumption growth, while the price of risk is the coefficient of relative risk aversion of the representative investor (Campbell 2003).

In their seminal paper, Mehra and Prescott (1985) set out to test the empirical validity of the standard CCAPM. In their framework, the representative agent has time-separable constant relative risk aversion (CRRA) preferences of the form:

$$U(c) = \frac{c^{1-a} - 1}{1 - a} ,$$

where c denotes consumption, a is the coefficient of relative risk aversion, and its reciprocal a^{-1} represents the intertemporal rate of substitution, which determines the agent's willingness to substitute between consumption today and consumption in future periods. This particular utility function has the advantage of being independent of the scale of the economy and the initial wealth endowments. However, it also assumes that a change in the agent's preferences to smooth consumption across time also affects the agent's preferences to smooth consumption across states of the economy, an assumption that has subsequently been revisited (Constantinides, 1990; Epstein and Zin 1989, 1991). In equilibrium, it can be shown that the expected excess return on an asset ($R - rf$) is proportional to the coefficient of relative risk aversion (a) and the covariance of the asset return (R) with per capita consumption growth (Δc):[2]

$$E[R - rf] \approx aCov(\Delta c, R)$$

Based on per capita consumption growth and equity return data for the US from 1890 to 2011 by Robert Shiller, the covariance of the equity return with per capita consumption growth is 0.002. For commonly accepted risk

2 For a complete derivation of the CCAPM, see Mehra and Prescott (1985), Campbell (2003) and Cochrane (2008).

aversion coefficients of 1 to 3 (Hansen and Singleton 1983; Chetty 2006), the consumption-based model predicts an expected ERP of 0.2 to 0.6 percent. These values are an order of magnitude smaller than the actual realised ERP of 6 percent. Similarly, Mehra and Prescott (1985) found that the ERP should not exceed 0.5 percent per year under very liberal assumptions. The inability of the consumption-based CAPM to predict the size of the ERP has become known in the literature as the "equity premium puzzle".

The CCAPM cannot explain the size of the ERP, largely because the volatility of consumption growth is too low. Consumption is much smoother than equity returns, and one needs to assume an implausibly large coefficient of risk aversion to reconcile the realised ERP with the low covariance of consumption growth and equity returns. For the historical average premium of 6 percent, one needs to assume a risk aversion coefficient of 30 or more to calibrate the model to the historical data. Such high levels of risk aversion are not only rejected by experimental evidence, but also predict implausibly high risk-free rates in the range of 20–25 percent per year (Cochrane 2008). Campbell (2003) shows that this specification of the CCAPM is also not consistent with global equity data from MSCI for 12 different equity markets. The implied coefficients of risk aversion for the majority of markets are even higher than those for the US.

2. Risk-based explanations

Several authors have tried to find a solution of the equity premium puzzle within the CCAPM framework by modifying some of the assumptions of the Mehra and Prescott (1985) specification. Epstein and Zin (1989) and Weil (1989) suggest a different utility function, wherein the rate of intertemporal substitution is not reciprocal to the coefficient of risk aversion as in the CRRA utility function. In the Epstein-Zin framework, agent preferences for consumption substitution across time are independent of their preferences across states of the economy. By separating time and risk preferences, Epstein and Zin essentially introduce an additional degree of freedom in the model and show that a large equity risk premium does not necessarily require a low intertemporal rate of substitution as under CRRA preferences.

Bansal and Yaron (2004) also use Epstein-Zin preferences in a richer economic environment to show that risks related to varying growth prospects and fluctuating economic uncertainty can help resolve the ERP puzzle. The authors suggest a process of consumption growth that consists of a small persistent expected growth rate component ("long-run consumption risk") and a time-varying economic uncertainty component. The long-run risk poses difficulty for Epstein-Zin type agents as they desire an early resolution of uncertainty. Variation in future expected growth rates thus can lead to large changes in the marginal rate of substitution of the representative agent, resulting in large equity risk premia. The authors show that the historical ERP can be quantitatively justified by the model using more plausible risk aversion parameters. While the paper purports to resolve the ERP puzzle, it has been criticised on the basis of making some implausible predictions. In particular, the Bansal-Yaron model counterfactually predicts strong persistence in dividend growth not observed in actual data and implies extremely low yields and negative term premia on inflation-linked bonds (Beeler and Campbell 2012).

Constantinides (1990), Abel (1990) and Campbell and Cochrane (1999) develop an alternative set of utility functions where the agent's utility depends not only on the per-period personal consumption level but also on the agent's past consumption habits or on the consumption level of others. Agents experience a loss of utility if their consumption falls below what they are used to consuming (their "internal habit") or below what their peers consume (their "external habit"). In this framework, utility is no longer derived from growth in personal consumption but rather from ratios of consumption relating current consumption to one of the habit benchmarks above. If so, it is possible that small variations in personal consumption translate into more volatile relative-to-habit levels. This class of models can potentially account for the observed ERP, but has also been criticised on the grounds of ad-hoc assumptions and as leading to several unusual implications. As Ljungqvist and Uhlig (1999) show, it is possible for agents with habit formation preferences to experience significant permanent welfare gains through a one-time lowering of consumption gains, as subsequent gains relative to the new lower habitual level would quickly make up for the initial loss of utility. For example, a one-time reduction in consumption of 10 percent can lead to a welfare gain of 16 percent, a pattern that is not consistent with observed human behaviour. More recently, Campbell and Cochrane (2015) have responded to this criticism by showing that a drop in consumption in Ljungqvist and Uhlig's model may not lead to welfare gains under alternative assumptions.

A promising recent line of research has focused on agent heterogeneity and alternative measures of household consumption in explaining the ERP puzzle. This line of research attempts to relax the unrealistic assumption of "homogeneous agents" in the standard CCAPM and recognise that different agents may have different utility functions based on their wealth endowments. Ait-Sahalia, Parker and Yogo (2004) propose a utility function that distinguishes between consumption of basic goods and consumption of luxury goods and show that the risk aversion coefficient implied by luxury good consumption is significantly larger than the one implied by aggregate consumption in the standard CCAPM. The implication is that the higher risk aversion to luxury good consumption of wealthy individuals, who also own the largest portion of the stock market, drives the average equity return. Savov (2011) on the other hand suggests that reported aggregate consumption may underestimate the volatility of actual consumption and that an alternative measure of consumption based on household garbage appears to produce the volatility necessary to match the observed equity premium. Similarly, Kroencke (2016) attributes the deficiency of the CCAPM to the failure of reported consumption to measure consumption risk correctly. He argues that statisticians optimally filter out observable consumption to produce a reported aggregate consumption series that is free from measurement error but in the process artificially smooth out consumption growth. He shows that an "unfiltered" consumption series is better able to explain the equity premium puzzle in the context of the classical CCAPM with more reasonable risk aversion. Moreover, his proposed "unfiltered" consumption series can serve as a unifying explanation for why Savov's (2011) garbage series and Jagannathan and Wang's (2007) fourth-quarter to fourth-quarter consumption series provide better estimates of the consumption risk of stocks.

Guvenen (2009) and Malloy, Moskowitz and Vissing-Jorgensen (2009) have proposed models that attribute the large equity premium to limited stock market participation. In the Guvenen (2009) model, the large premium arises from the unequal distribution of risks among stockholders and non-stockholders in the economy. Non-stockholders, who have low intertemporal rates of substitution, use only the bond market to smooth out fluctuations in their labour income, which in turn concentrates risk in the stock market. As a result, stockholders require a "large" premium to bear this risk. Moreover, non-stockholders (less wealthy households) have a stronger desire to smooth consumption than stockholders (wealthy individuals) in recessions, which in turn furthers the countercyclicality of the premium. Wealth inequality may thus account for both the magnitude and the behaviour of the premium over time. Using the long-run consumption risk framework of Bansal and Yaron (2004), Malloy, Moskowitz and Vissing-Jorgensen (2009) present empirical evidence that stockholder consumption growth is larger than aggregate consumption growth. The risk borne by stockholders, rather than by the "average" agent in the economy, appears to drive the large equity risk premium. By emphasising the different risks borne by stockholders and non-stockholders, the Guvenen (2009) and Malloy, Moskowitz and Vissing-Jorgensen (2009) models provide a plausible explanation for the large realised premia in recent decades. Rising wealth inequality and increased concentration of stock market wealth contribute to large premia in these frameworks.

Constantinides, Donaldson and Mehra (2002) propose a modification of the CCAPM by introducing borrowing constraints in the context of a life-cycle investor model. The key insight is that young investors are different from older investors. Young investors face high wage uncertainty and relatively low correlation of stock returns with wage income, which makes equities attractive for hedging fluctuations in wages. Older investors, on the other hand, face the opposite situation: their wage uncertainty is low and their consumption is relatively more correlated with stock market returns. The implication is that in a frictionless market young investors should hold more equities than bonds relative to older investors. However, because of borrowing constraints, young investors are effectively shut out of capital markets, and equity prices are thus determined by older investors, leading to low bond premia and higher equity risk premia. Borrowing constraints and heterogeneity of investors can thus also account for the size of the ERP. The implication of this model is that demographic changes may have an impact on long-term equity returns. Ageing populations can lead to rising equity risk premia in the long term, putting downward pressure on equity prices in the short term.

Rietz (1988) and the "peso" problem literature offer a different approach to resolving the ERP puzzle, namely that the large historical realised ERP may be due to priced-in catastrophic risk that actually did not materialise during the sample period. Examples of such risks include economic depressions, war, destruction, natural disasters and epidemic outbreaks. The Rietz (1988) model uses the Mehra-Prescott (1985) CCAPM set up, but adds a "disaster" state scenario to the model which has a low probability of occurring but an extremely high marginal utility of consumption. The existence of such a disaster state can then account for the large premiums required by investors to hold risky assets. While it is theoretically possible, Donaldson and Mehra (2008) question the plausibility of Rietz's argument by pointing out that it predicts an inverse relationship between the probability of rare events and interest rates, which does not seem to match the historical record. For example, the perceived probability of a depression is likely to have been high after the end of WWII, while interest rates remained relatively low.

More recent work on catastrophic risk, however, has brought renewed interest in this type of models, especially in light of the large swings in equity markets during the Global Financial Crisis. Barro (2006) and Barro and Ursua (2012) present a more general disaster risk scenario model than Rietz (1988) that is better calibrated with the historical record. The model relies on agents with Epstein-Zin preferences and disaster scenarios drawn from history (i.e. actual drops in economic output). The model predicts results consistent with the average equity premium with a reasonable coefficient of relative risk aversion. Nakamura, Steinsson, Barro and Ursua (2013) show that rare disaster models can stand up to incorporating a more realistic process for consumption dynamics that unfolds over multiple years rather than in an instantaneous drop as in the early models. In a comprehensive survey article, Tsai and Wachter (2015) argue that rare disaster risk can not only explain the ERP puzzle but also anomalies in bond markets, responding to a frequent criticism of the rare event literature that it cannot explain the unusually low bond premia. Overall, this line of research suggests that while rare events may not be fully able to account for the observed premia in asset markets, they may be an important contributing factor.

3. Behavioural explanations

Proponents of behavioural finance argue that investors have a tendency to depart from the assumed rational behaviour underlying the assumptions of the standard CCAPM, thereby leading to pricing anomalies. One potential behavioural explanation for the ERP puzzle is the myopic loss aversion hypothesis of Benartzi and Thaler (1995). The authors argue that investors suffer from a "loss aversion" bias – i.e. they dislike losses more than they like gains. If people are myopic loss averters and adjust their portfolios at least annually, they will require high premia to hold equities, as the probability of the stock market underperforming risk-free assets over short horizons is rather high. This myopic loss aversion has received some experimental support in Gneezy, Kapteyn and Potters (2003). In line with people's tendency to avoid short-term losses, they find that more frequent feedback information about the performance of an investment portfolio and more frequent rebalancing result in less risk taking and higher risk aversion. Conversely, market prices are higher if investors do not trade as frequently and have less frequent information about portfolio performance. Market interactions do not seem to eliminate these attitudes towards avoiding short-term losses.

A related theory of investor behaviour dubbed "prospect theory" argues that, in addition to being myopic loss averters, investors exhibit narrow framing, another well-known bias in decision-making under uncertainty in experimental settings (Barberis, Huang and Santos 2001). Investor utility is defined over equity gains and losses rather than over total wealth, with the extent of loss sensitivity depending on the investor's prior portfolio performance. If an investor has experienced equity losses (relative to a benchmark) in the past, the investor is relatively more sensitive to incurring additional losses. Conversely, investors who have "made money" in the past become relatively less risk-averse. By incorporating myopic loss aversion and narrow framing into a CCAPM, the authors show that a large ERP can be consistent with a smooth consumption series. The myopic loss aversion hypothesis and prospect theory are reminiscent of the habit formation models of Constantinides (1990) and Campbell and Cochrane (1999) mentioned before. They offer plausible explanations for the ERP, but strict aggregation of these types of preferences may not hold true (Donaldson and Mehra, 2008).

A different line of behavioural research suggests that irrational noise traders in the market create additional risks for "true" investors, thereby causing asset prices to be lower than suggested by fundamentals (De Long, Shleifer, Summers and Waldman 1990; Pontiff 1996). Because of this "noise trader risk", risk premia in the market are higher than they would be under market efficiency. More recent research has focused on the effects of heterogeneous beliefs on asset prices. Scheinkman and Xiong (2003) show how differences in beliefs among agents can lead to stock market overvaluation in the presence of short sale constraints. Bhamra and Uppal (2014) solve a model where agents have heterogeneous beliefs of the "catching up with the Joneses" type (agents compare consumption to external benchmarks). Collin-Dufresne, Johannes and Lochstoer (2016) present a model where individual belief formation is age-dependent. Young agents suffer from a "this time is different" bias: they put relatively less weight on events they have not personally experienced and relatively more weight on recent events. This bias leads to small deviations from rational behaviour but can amplify the effects of macroeconomic shocks on asset prices.

4. Other explanations

Another class of explanations for the ERP focuses on characteristics of equity and Treasury markets not captured by standard pricing models such as liquidity constraints, transaction costs and various secular trends.

Transaction costs or liquidity constraints may prevent investors from fully smoothing out consumption as assumed in many models. As investors are subject to various costs in equity markets, part of the ERP may constitute compensation for transaction costs or illiquidity premia. If investors expect to incur large costs for liquidating positions, they will demand a high ex ante premium. While the majority of the liquidity literature focuses on the cross-section of stocks, some studies suggest that variations in liquidity premia may contribute to the aggregate ERP. Bekaert, Harvey and Lundblad (2007) suggest that differences in liquidity across emerging markets can partially explain differences in equity premia. Gibson and Mougeot (2004) document that the liquidity risk premium in the aggregate stock market index has a time-varying component related to the probability of future recession. Variations in the size of the liquidity premium can explain some of the time variation in the excess return of the S&P 500. Swan (2000) also presents some evidence that equity transaction costs may be priced in the equity premium. Secular trends such as the growing popularity of low-cost ETFs among retail investors, competition in the mutual fund industry and increased stock market participation may reduce the impact of liquidity costs and lower the equity premium.

Some evidence suggests that Treasury bills and bonds may actually carry a negative liquidity premium (i.e. a convenience yield) as investors may place a value on holding Treasury securities above the securities' cash flows. Krishnamurthy and Vissing-Jorgensen (2012) trace the demand for convenience yields in US Treasuries and show that the negative correlation between the corporate bond spread and the debt-to-GDP ratio can be attributed to variation in the convenience yield of Treasuries rather than variation in the corporate default risk. When supply of debt is low, investors bid up the prices of Treasuries and lower their yields. They estimate that Treasury securities may command a convenience yield of up to 1 percent.

Secular trends such as globalisation of financial markets and increased participation in global equity markets may also have contributed to the large equity premia. Fama and French (2002) surmise that much of the large upward repricing that the US equity market experienced during the second half of the 20th century can be attributed to increased equity market participation. As Damodaran (2012) points out, more flows into the equity market from other asset classes or geographical regions could account for some of the variation in the ERP. Such explanations, of course, describe only the consequence of investor behaviour on the premium and cannot explain the driving forces of increased flows. Yet, they underscore the fact that the realised premium may be a poor proxy for the expected premium if such trends subside or reverse.

Another potential driver identified in the literature regarding the future of the ERP is globalisation and market integration. Capital mobility and free trade have given rise to greater integration across equity markets, allowing for more diversification and greater risk-sharing. Stulz (1999) argues that globalisation provides for better international risk sharing, which in turn reduces the overall risk in equity markets. On the corporate side, globalisation makes it less costly for firms to borrow in different markets, thereby reducing their cost of capital. Studies by Foerster and Karolyi (1999) and Bekaert, Harvey, Lundblad and Siegel (2011) present some empirical evidence for a structural reduction in equity risk across various markets, suggesting that investors will earn a lower ERP in the future if these trends persist. In addition, He and Krishnamurthy (2013) show that the behaviour of financial intermediaries such as banks and financial institutions can also affect the size of the ERP. In the presence of equity capital constraints, financial intermediaries increase their equity investments in response to government policy that increases the availability of equity capital (capital injections) or reduces borrowing costs. In times of crisis, equity injections are particularly effective in normalising risk premiums in the market by reducing financial intermediaries' equity capital constraints.

5. Summary

While much progress has been made in understanding the drivers of the ERP, no single model has been able to fully capture the ERP's complex behaviour. Relaxing the assumptions of constant risk aversion in the standard CCAPM helps bring the model predictions closer to reality, but as Donaldson and Mehra (2008) emphasise, many of these alternative utility functions lack axiomatic underpinnings. While economic risk appears to be an important driver of the premium, investor behaviour and institutional factors have also likely contributed to the ERP's behaviour through time.

V. The expected ERP

As structural demand-driven models cannot fully account for the observed magnitude of the ERP, practitioners have come to rely on a variety of reduced-form predictive models to estimate the expected ERP. In particular, valuation measures such as dividend yields and earnings yields have long been used as indicators of the relative "richness" of equity markets. In this section, we discuss some of the main approaches to estimating the expected ERP and the advantages and disadvantages of each approach.

We illustrate the various estimation approaches based on global MSCI equity data from 1970 to 2015. The sample consists of 18 developed markets included in the original MSCI World index when it was first constructed in 1970 and for which we have uninterrupted data since inception. For some of the models, we also rely on earnings forecasts from FactSet Estimates and GDP growth data from the USDA International Macroeconomic Data Set. Data on risk-free rates are collected from the IMF, Bloomberg, the OECD and FactSet. In addition to estimates by country, we also report regional estimates for Asia, Europe and America, as well as a global estimate (World). The goal of this empirical exercise is not to test the performance of the models in a robust statistical sense, but to illustrate different approaches used to estimate the expected ERP.

1. Long-term averages

As discussed previously, the historical realised average ERP is the simplest and easiest-to-compute estimate of the expected ERP. In Table 5, we provide averages and standard errors of the realised ERP over the sample period, as well as averages by decade from 1970 to 2015. The global realised ERP is 6.4 percent over this period, consistent with averages based on longer historical samples. The standard error of the realised World ERP is substantial, averaging 2.6 percent over the period. The unconditional mean ERP estimates vary by country, ranging from 2.9 percent for Italy to 18 percent for Hong Kong. As discussed in Section III, these unconditional estimates may be a poor proxy for the expected ERP, especially over shorter horizons. They are volatile and imprecise, depend heavily on the historical window over which they are calculated, assume that the future will remain exactly as in the past, and may be plagued by survivorship bias. Yet, because of their simplicity, unconditional mean forecasts remain an important "benchmark" for more complex models. In fact, as Welch and Goyal (2008) show, a number of conditional models based on predictive regressions fail to outperform naïve unconditional mean forecasts out-of-sample.

Table 5: Unconditional mean ERPs by country and region (1970–2015)

	Ave	Std error	Average by decade				
			1970s	1980s	1990s	2000s	2010s
Australia	3.8	3.3	-4.5	9.5	4.6	2.3	7.2
Austria	5.2	5.1	0.9	7.8	6.6	4.7	5.6
Belgium	7.9	3.8	-0.2	16.6	10.2	-4.9	21.0
Canada	4.9	2.9	3.1	4.9	4.2	4.6	8.5
Denmark	8.8	4.9	1.7	10.7	7.1	6.2	21.4
France	6.7	4.2	-1.7	16.4	8.1	-0.4	11.5
Germany	5.8	4.0	-1.7	8.5	10.9	-0.2	13.0
Hong Kong	18.0	7.3	32.6	22.5	14.1	7.1	14.1
Italy	2.9	4.8	-12.6	19.8	5.3	-3.1	4.0
Japan	6.1	4.3	10.9	16.7	-7.8	1.3	11.6
Netherlands	8.0	3.5	-0.2	14.6	14.1	-2.3	15.0
Norway	8.6	6.0	2.8	20.9	2.4	7.0	9.9
Singapore	11.2	7.0	28.3	7.7	2.8	8.1	10.4
Spain	5.8	4.3	-8.1	16.4	11.0	2.0	6.5
Sweden	11.8	4.8	0.9	22.9	11.9	6.2	17.5
Switzerland	6.5	3.3	-0.8	6.2	16.9	0.1	10.7
UK	5.4	3.0	4.1	10.7	6.9	-3.1	9.7
US	6.5	2.6	0.3	7.4	14.8	-3.9	15.9
Asia-Pac	6.7	4.2	10.5	16.9	-5.4	2.4	10.7
Europe	6.6	2.9	0.9	12.8	9.8	-0.8	11.4
America	6.4	2.6	0.5	7.3	14.2	-3.4	15.3
World	6.4	2.6	2.0	11.7	8.1	-1.7	13.5

Source: Factset, IMF, OECD, Bloomberg; Norges Bank Investment Management

2. Implied ERPs from present-value models

A second frequently used approach to estimate the expected ERP is motivated by the dividend discount model (DDM):

$$P_t = \sum_{k=1}^{\infty} \frac{E_t(D_{t+k})}{(1 + rf + ERP)^k}$$

where the current market price of equity P_t is equal to the expected stream of dividends D_{t+k} discounted to the present at the cost of equity capital rate $rf+ERP$. The cost of equity consists of the sum of the risk-free rate rf, which reflects the time value of money, and the required risk premium on equities ERP, which reflects the riskiness of equities. The expected ERP can then be calculated as the value of the term ERP that makes the projected stream of future dividends equal to the current equity market price . This ERP is referred to as the "market-implied" ERP.

Dividend discount models are relatively straightforward to implement. As they model the ERP from market prices, they are consistent with no arbitrage (assuming assets are fairly priced), provide timely and fast-moving estimates of the ERP, and do not rely exclusively on historical data. However, estimates tend to be sensitive to the inputs of the model and especially sensitive to the assumed growth rates of future cash flows. Small changes in the expected growth rates of cash flows can lead to vastly different risk premia. The growth rate can be assumed constant (the Gordon constant dividend growth model) or it can vary over time (multi-stage growth models). Alternative measures of cash flows such as earnings, free cash flows and residual income can be also used in the denominator as dividends may be a poor proxy of expected cash flows from holding equities in certain markets (Ilmanen 2011). We illustrate the DDM estimation using six alternative specifications.

(1) Dividend yield (D/P)

The simplest specification of the DDM is the Gordon constant dividend growth model, which assumes future dividends grow at a constant rate g in perpetuity. The DDM formula simplifies to:

$$P_t = \frac{D_t}{rf + ERP - g}$$

The expected ERP is then given by:

$$ERP = \frac{D}{P} + g - rf\,.$$

If we assume that dividends grow at the risk-free rate in perpetuity (g equals rf) then the dividend-price ratio provides a direct estimate of the expected ERP. The dividend yield is a classic predictor variable for the ERP and has been used extensively as a proxy of the ERP in a variety of applications (Ilmanen 2011; Damodaran 2012).

(2) Earnings yield – rf (the so-called "Fed model"3)

The so-called "Fed model" is a variation of the constant dividend growth model that uses the difference between the earnings yield and the risk-free rate as an estimate of the expected ERP. In this case, the Gordon constant dividend model can be expressed as:

$$ERP = \frac{E}{P} - rf$$

where E is per-share company earnings. We use 12-month trailing earnings for the value of E.

(3) Expected earnings yield – rf ("expected Fed")

The same as (2) above but using consensus expected earnings from FactSet Estimates instead of trailing earnings. Including expected earnings may

3 The spread between equity earnings yields and bond yields as a measure of the richness of equity markets is known among practitioners as the "Fed model", but is not officially endorsed by the Federal Reserve as a valuation model. As pointed out by Asness (2003), comparing earnings yield (a real variable) and bond yields (a nominal variable) may be inappropriate in the presence of time-varying inflation.

better capture the forward-looking nature of the DDM, but the ERP estimates will necessarily depend on the quality of analysts' earnings forecasts.

(4) Shiller earnings yield – rf

The same as (2) and (3) above, but using lagged ten-year average earnings to adjust for business cycle variations (Shiller 2005). Because cyclically-adjusted earnings ameliorate the effects of extreme business cycle movements, they may provide better estimates of company performance over longer horizons.

(5) Sum of the parts model (SOP)

The sum of the parts model (SOP) of Ferreira and Santa–Clara (2011) is an example of a "building block" model of the ERP, where the three sources of equity returns – income, dividend (earnings) growth and the valuation change (change in P/E or P/D) – are modelled separately. From the definition of an equity return, the SOP model in its dividend growth version can be written as:

$$R = \frac{D}{P} + g + \Delta P/D$$

where R is the equity return, D/P is the expected dividend yield at the end of the period, g is the growth rate of dividends and $\Delta P/D$ is the change in the multiple. In the absence of a valuation change ($\Delta P/D$), the SOP is equivalent to the Gordon steady-state model. Similarly to Ferreira and Santa-Clara, we estimate a steady-state version of the SOP model where we assume that the expected dividend-price ratio is equal to the current one, the growth rate of dividends is modelled as a moving average of the prior 20-year average dividend growth, and there is no change in the multiple ("SOP Const"). We also estimate two additional versions of the SOP model: (1) dividends grow at their 20-year rolling average and the price multiple declines or increases linearly to its ten-year moving average ("SOP MeanRev") and (2) dividends grow at below the historical rate (1 percent) with no change in the price multiple ("SOP Low growth"). In the low-growth scenario, we assume that the expected dividend growth is lower than the average dividend growth over the prior 20 years. Such an assumption can be justified on the basis of a possible regime change in dividend growth rates. Dividend growth has been exceptionally large since the end of the Global Financial Crisis and these high rates may bias the 20-year average dividend growth rate upwards. By excluding the effects of the recent high dividend growth rates on the 20-year dividend growth average, we provide a more conservative estimate of the ERP consistent with a regime of lower-than-historical dividend growth rates. Moreover, Gordon (2012, 2014) suggests that future long-term economic growth may be lower than in the past due to structural issues that major developed economies currently face, including ageing populations, plateauing educational attainment, inequality, debt overhang, globalisation and environmental issues. These factors have the potential to reduce output growth to 0.9 percent per year, and per-capita real income growth to as low as 0.4 percent for 99 percent of the US population.

(6) Multi-stage DDMs

Constant growth models are simple and intuitive, but come at the cost of reduced flexibility. In reality, companies go through life cycles during which cash flow growth rates may vary substantially. For example, early-stage companies tend to grow faster than mature companies. Multi-stage growth models attempt to model explicitly the expected path of growth rates. They typically assume that high initial growth rates decline to a lower steady-state growth rate in the long run.

Following closely the framework of Pastor, Sinha and Swaminathan (2008) and Li, Ng and Swaminathan (2013), we estimate the following multi-stage DDM:

$$P_t = \sum_{k=1}^{T} \frac{FE_{t+k}(1 - b_{t+k})}{(1 + r_e)^k} + \frac{FE_{t+T+1}}{r_e(1 + r_e)^T}$$

where FE_{t+k} is the earnings forecast for year $t+k$; b_{t+k} is the plowback rate (1 minus the dividend payout rate); and r_e is the implied cost of equity. The number of different growth stages is set at T=15 and cash flows after stage 15 are assumed to grow at a constant rate equal to the firm's cost of equity (i.e. no excess returns over the long run). For years one and two, we use consensus earnings forecasts from FactSet Estimates; for the remaining years up to year 15, earnings grow at rates that revert to a steady-state long-run rate g:

$$g_{t+k} = g_{t+k-1} \exp\left[\log\left(\frac{g}{g_2}\right)/T\right]$$

$$FE_{t+k} = FE_{t+k-1}(1 + g_{t+k})$$

We make three alternative assumptions for the terminal growth rate g, the most important input in the model. As in Li, Ng and Swaminathan (2013), we first set the rate equal to an expanding average of GDP growth rates for a given country (model "Multi-stage DDM GDP"). Second, we set it equal to the more conservative long-term average growth of World EPS earnings (model "Multi-stage DDM EPS"). As Arnott and Bernstein (2003) point out, the growth rate of corporate earnings since 1900 has lagged behind GDP growth, and using average GDP growth rates as terminal growth rates for corporates may be an overstatement. Third, we estimate a model where the growth rate declines to the steady-state long-term GDP growth rate over a period of five years ($T=5$) rather than 15 years.

We model the plowback rate as 1 minus the dividend payout in the most recent year, and after that as:

$$b_{t+k} = b_{t+k-1} - (b_1 - b)/T$$

where plowback rates are assumed to decrease linearly to a steady-state plowback rate b, which is set equal to the ratio of the long-term growth rate of earnings g and the cost of capital . The assumption is that, in the long run, firms do not earn excess profits above their return on equity and that the long-term growth rate g is the product of firms' return on investment (ROI) and the steady-state plowback rate.

Estimates of the implied ERPs from dividend discount models are reported in Table 6. The models are generally consistent in their predictions across countries but the magnitude of the estimates varies across models. As expected, the Gordon growth model, which assumes dividends grow at the risk-free rate, provides the lowest estimate for the World ERP, whereas the SOP Const, the SOP MeanRev and the multi-stage models, which assume significantly higher growth rates, provide two to three times larger estimates. As expected, the SOP Low-growth model, which assumes that future dividend growth is lower than its 20-year historical average, provides a more conservative estimate of the expected ERP (World ERP of 3.4 percent). The World ERP varies from 2.7 to 8.7 percent across the different models, averaging 5.9 percent. The multi-stage DDM that models earnings growth over a five-year period provides lower estimates than the versions of the DDM that model growth over 15 years, underscoring the sensitivity of these models to the growth rate assumptions. Overall, the current World ERP forecast from DDMs is on a par with the unconditional mean forecasts (Table 5), and hence somewhat above unconditional means adjusted for repricing (Table 4).

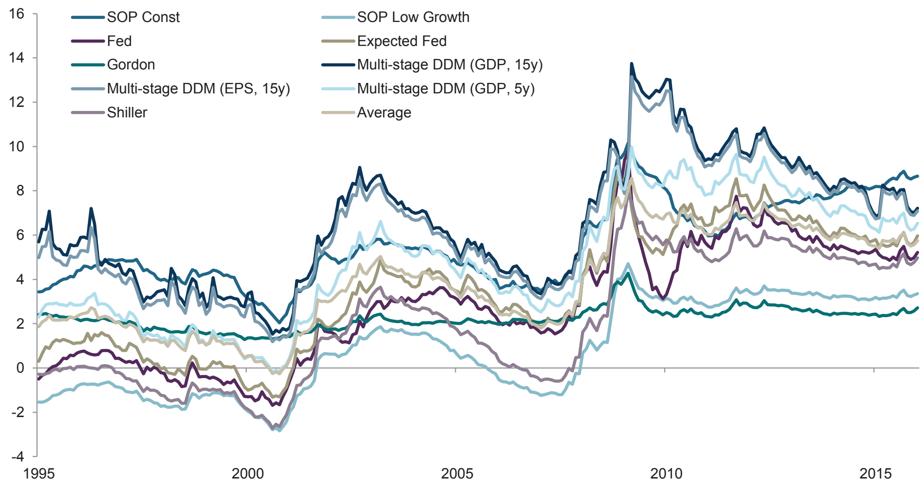
Table 6: Implied ERPs from dividend discount models (January 2016)

	Gordon	Fed	Expected Fed	Shiller	SOP Const	SOP Mean-Rev	SOP Low growth	Multi-stage DDM (GDP, 15 year)	Multi-stage DDM (EPS, 15 year)	Multi-stage DDM (GDP, 5 year)	Ave all models
Australia	5.4	4.7	4.4	4.9	9.9	7.9	4.2	5.1	5.1	4.7	5.6
Austria	2.0	7.4	8.3	11.6	7.9	9.4	2.1	6.5	6.3	7.3	6.9
Belgium	3.1	5.8	5.7	5.5	8.7	8.4	4.5	6.0	5.9	5.8	5.9
Canada	3.3	4.5	6.1	5.6	11.2	8.6	3.8	8.6	8.4	7.2	6.8
Denmark	1.7	3.1	4.0	1.9	14.4	13.1	1.7	5.4	5.4	4.6	5.5
France	3.3	5.3	7.2	7.0	9.5	9.6	4.6	8.3	8.3	7.7	7.1
Germany	3.0	5.5	6.8	5.5	10.1	9.6	3.1	7.5	7.5	7.1	6.6
Hong Kong	3.3	9.4	7.2	6.9	6.9	5.2	4.1	8.6	8.1	7.8	6.8
Italy	3.5	2.8	5.7	9.3	9.1	10.4	3.5	11.9	11.9	8.4	7.6
Japan	2.1	5.7	5.9	3.2	6.1	4.0	2.0	7.6	7.5	6.7	5.1
Netherlands	2.3	5.3	6.7	6.0	6.0	8.6	3.7	7.9	7.8	7.3	6.2
Norway	4.5	7.8	6.9	9.7	19.4	16.8	4.8	6.5	6.1	6.6	8.9
Singapore	4.6	7.3	7.3	7.9	11.2	7.7	4.0	8.5	7.3	7.8	7.4
Spain	5.2	7.2	8.0	10.2	15.0	14.0	6.6	10.6	10.4	9.2	9.6
Sweden	4.1	6.6	7.2	6.2	20.2	17.9	5.5	7.9	7.9	7.5	9.1
Switzerland	3.3	6.2	6.4	5.2	13.8	11.2	4.7	7.3	7.4	6.8	7.2
UK	4.4	6.6	6.7	8.8	8.0	6.2	5.5	7.1	7.0	6.9	6.7
US	2.3	4.8	5.7	4.3	8.1	7.1	2.9	6.9	6.8	6.3	5.5
America	2.9	5.9	5.8	4.1	7.1	5.0	2.7	7.0	6.9	6.3	5.4
Asia	3.6	5.9	6.7	7.1	10.6	9.5	4.7	7.2	7.0	6.5	6.9
Europe	2.3	4.8	5.7	4.3	8.2	7.1	3.0	7.7	7.7	7.2	5.8
World	2.7	5.2	6.0	5.0	8.7	7.4	3.4	7.2	7.1	6.5	5.9

Source: Factset, IMF, OECD, Bloomberg, USDA Macroeconomic data; Norges Bank Investment Management

The expected World ERP from DDMs has varied substantially over time as illustrated in Chart 7. The expected premium bottomed during the dot-com bubble in the late 1990s and peaked in 2009 at the height of the Global Financial Crisis. Since the end of the crisis, the premium has gradually declined by more than 50 percent from peak 2009 levels. This suggests that the equity market has become significantly richer over the past six years, and future realised equity returns may be lower than those realised over the period from 2010 to 2015.

Chart 7: Implied ERPs from dividend discount models: World



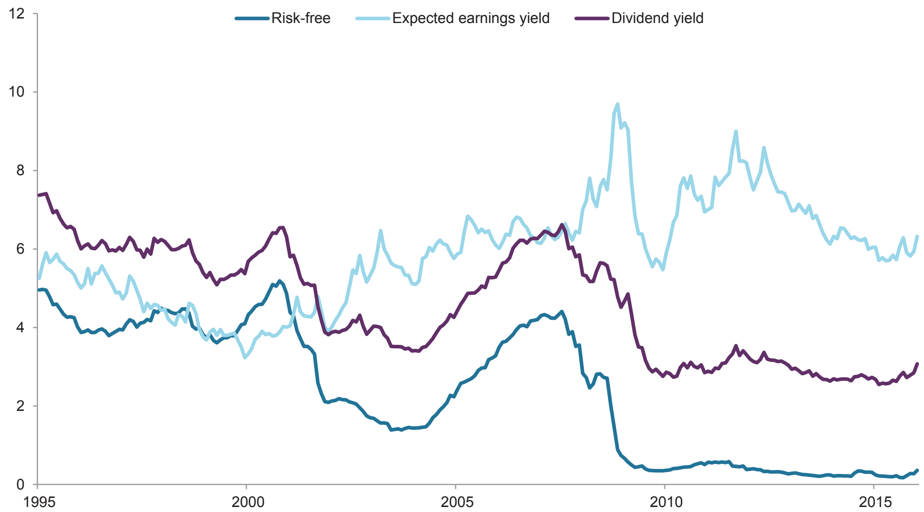
Source: Factset, IMF, OECD, Bloomberg, USDA Macroeconomic data; Norges Bank Investment Management

The current estimate for the World ERP of 5.9 percent in Table 6 should be viewed with caution for at least two reasons. First, it is calculated as the equally-weighted average of the ERP forecasts of the models in Table 6, which may suffer from the same recent-data bias. In the majority of the models, the expected cash flow growth rate is assumed to be equal to its ten-year or 20-year historical average or analyst earnings estimates, which also rely heavily on recent data. The fact that cash flow rates have been exceptionally large since the end of the Global Financial Crisis may bias these historical averages upwards. A more conservative dividend growth assumption as in the "SOP Low growth" model produces a significantly lower World ERP of 3.4 percent.

Second, estimates of the expected ERP are affected by the current exceptionally low risk-free rate environment. In all of the dividend discount models described above, it is assumed that the current interest rate (0.28 percent as at January 2016) is a good proxy for the expected future interest rate. In the absence of a structural model for the evolution of future short-term rates, it is common to assume that current short-term rates are going to persist in the future. However, if the market expects that the current near-zero short-term rates will revert to their historical averages, it may be appropriate to use a different proxy for the future risk-free rate that better reflects these expectations of mean reversion. The average short-term rate over the period from 1995 to 2015 is 2.35 percent, implying that the expected World ERP may be closer to 4 percent if we account for a potential future normalisation of interest rates.

We illustrate the decomposition of the expected ERP into the expected equity return component (as proxied by the dividend and earnings yields) and the expected interest-rate component of the ERP as proxied by the point-in-time short-term rates in Chart 8. The current earnings yield and current dividend yields are similar or slightly lower than their average values prior to the financial crisis, whereas the risk-free rate has declined from 4 percent to near zero. Because of the large decline in our proxy for the expected interest rate part of the expected ERP equation, the expected ERPs in the latter part of the sample appear to be relatively large even through equity market valuations are in line with their long-term averages.

Chart 8: World risk-free rate, expected earnings yield and dividend yield



Source: Factset, IMF, OECD, Bloomberg, USDA Macroeconomic data; Norges Bank Investment Management

3. Regression-based estimates

Another approach to estimating the expected ERP is based on time-series and cross-sectional regression models. Predictive regressions take the general form:

$$R_{t+k} - rf_{t+k} = \alpha + \beta \times Fundamental_t + \varepsilon_{t+k}$$

where $R_{t+k} - rf_{t+k}$ is the ERP over a period $t+k$, and $Fundamental_t$ is a predictor variable, usually based on equity fundamentals or macroeconomic variables known at time t.

A vast body of literature has explored the time series predictability of returns based on variables such as the dividend and earnings yield (Fama and French 1988; Campbell and Shiller 1989); book-to-market (Pontiff and Schall 1998); default spreads (Campbell 1987); stock volatility (Goyal and Santa-Clara 2003); corporate issuing activity (Baker and Wurgler 2000); the consumption-wealth ratio (Lettau and Ludvigson 2001); the cross-sectional variation in book-to-market ratios (Kelly and Pruitt 2013); technical trading rules (Neely, Rapach and Zhou 2004); and short interest (Rapach, Ringgenberg and Zhou 2016). While many of these studies find evidence of statistically significant return predictability in-sample, these methods have been questioned on the basis of their poor out-of-sample performance, lack of economic rationale for many of the predictors, and a host of issues related to statistical inference. Welch and Goyal (2008) systematically test many of these models and find that univariate predictive regression cannot "beat", out-of-sample, a naïve estimate of the ERP based on the sample mean. The main problem with the predictability approach is that the unrestricted regression coefficients tend to fit the noise in small samples, which in turn leads to overfitting of the data and instability of the estimated coefficients. More recent approaches have proposed imposing an economic structure on the regression coefficients and thereby reducing the overfitting bias (Campbell and Thompson 2008). Van Binsbergen and Koijen (2010) propose an alternative two-step approach whereby they first estimate the dynamics of the dividend-price ratio from past dividend growth and then use a maximum likelihood estimator to derive the expected return series. This modelling approach appears to explain a higher proportion of the out-of-sample variability of returns than traditional predictive regressions.

Researchers have also responded to Goyal and Welch's (2008) critique of univariate regressions by advancing models that include combinations of forecasts, diffusion indexes and controls for the time-varying properties of model parameters (regime-switching). Rapach, Straus and Zhou (2010) show that combining several individual forecasts produces a robust out-of-sample forecast of equity returns. They attribute the better performance of combination forecasts to reduced model uncertainty and coefficient instability when estimates are averaged across different models. Ludvingson and Ng (2007) employ a dynamic factor analysis on a dataset of close to 400 different economic and financial predicting variables to summarise the common variation in the data into a small set of factors (diffusion indexes). These indexes appear to capture the variation in the one-quarter ahead excess return and exhibit stable and statistically significant out-of-sample forecasting power. Dangl and Halling (2012) use a comprehensive Bayesian framework to show that models with time-varying coefficients dominate models with constant coefficients and have stronger and more stable out-of-sample predictive power. Henkel, Martin and Nardari (2011) show empirical evidence that excess equity returns are more predictable during recessions than during economic expansions. A forecasting model that allows for the different predictability of the premium across the business cycle outperforms the historical average out-of-sample. Overall, such improved strategies that address the issues of model uncertainty and parameter instability point towards stronger evidence of out-of-sample equity return predictability.

While a variety of financial and economic variables have been suggested to predict excess returns, the dividend-price ratio is perhaps the most extensively researched predictor in the literature. Studies by Campbell and Shiller (1988), Fama and French (1988) and Campbell and Shiller (1989) provide empirical support that the variation in the aggregate dividend yield may be attributed to time-varying expected returns (discount rates). The idea introduced in the work of Campbell and Shiller (1988) is that equity returns can be decomposed into two components: news about discount rates, which reflects time-varying risk aversion or investor sentiment, and news about cash flows, which is linked to company fundamentals. In an environment with no bubbles, variation in the dividend-price ratio should be correlated with either discount rate news or cash flow news, or both. Empirically, the dividend-price ratio has been found to predict discount rates better than dividend growth in US data (Cochrane 2008a), suggesting that stock prices are exposed to transitory changes in discount rates in addition to long-term fundamental risk. Recent research by Golez and Koudijs (2016) confirms that the dividend-price ratio predicts equity returns in the US and Dutch equity markets over the last four centuries, but they find that the inability of the ratio to predict cash flow growth (dividend growth) is a fairly recent phenomenon. Prior to 1945, the ratio strongly predicts dividend growth, and changes in cash flow news have a larger effect on stock prices than changes in discount rates. The extra volatility of the stock market induced by discount rate news is most prominent in the second half of the 20th century. Moreover, studies on global equity markets point towards a stronger dividend growth predictability outside the US. Rangvid, Schmeling and Schrimpf (2014) show empirically that dividend yield predicts dividend growth in many global equity markets, particularly in those where dividends are less frequently smoothed by companies and thus better reflect changes in company fundamentals. Within the cross-section of different equity markets, the more volatile dividends are, the more predictable dividend growth is.

In Tables 7 and 8 below, we test the predictability of the World ERP and World dividend growth for our sample of global equities. The models use log dividend yield, the past average dividend growth and the past average size of the ERP over alternative estimation windows as predictors. We use a panel data regression approach based on 9,500 rolling month and country observations. To address potential autocorrelations induced by our use of rolling observations, we report standard errors that are adjusted for fixed effects across time and country in addition to asymptotic Newey-West errors. The results in Table 8 generally point towards in-sample predictability of equity returns. Consistent with the literature, the coefficient on dividend yield has the expected positive sign and is statistically significant under alternative standard error specifications. High dividend yield implies large risk premia. In addition, we find some evidence of mean-reversion in returns, as evidenced by the negative statistically significant coefficients on the past ERP. The R-squares of the regressions, however, are low, indicating that the models can explain only a small portion of the variation in the premia. The standard errors of the estimates are relatively large, again emphasising that regression-based point estimates are relatively uncertain and should be viewed with caution.

Unlike many of the studies that find no dividend growth predictability in US equity data, we find that high dividend yields are associated with low future dividend growth. The dividend yield coefficients have the expected negative signs and are statistically significant over alternative estimation windows. The R-square for the one-year model is fairly large at 16 percent, suggesting that the variation in dividend yields, along with variation in past dividend growth and past equity returns, has substantial explanatory power for cash flow growth. Consistent with the study of Rangvid, Schmeling and Schrimpf (2014), the stylised fact of no dividend growth predictability may thus not hold true for the global aggregate. Collectively, the results in Tables 7 and 8 suggest that variation in the dividend yield may be correlated with both discount rate and cash flow risk. Moreover, the regression results suggest that the current point estimates for the World ERP range from 3.9 to 5.2 percent depending on the estimation window. These estimates are lower than the estimates from the DDMs in Table 6.

Table 7: Predictive regressions for the ERP

The table below reports panel regression betas for the one-year ERP (total equity return of the country MSCI index minus local government bill rate) of 18 countries from 1972 to 2015. The data frequency is monthly. The past n-year ERP represents the percentage change in the respective MSCI country-specific ERP over the prior n years. The past n-year dividend growth represents the change in dividends paid by the country-specific MSCI index over the prior n years. The one-year trailing dividend yield (log) represents the natural log of the sum of the dividends paid by the country-specific MSCI index over the past year divided by the current index price. T-stats are calculated from standard errors that have been adjusted for autocorrelation and heteroskedasticity using four alternative methods: (1) a Newey-West covariance estimator with 12 lags; (2) White standard errors that are robust to within-cluster correlation across time (cluster variable is month); (3) White standard errors that are robust to within-cluster correlation across countries (cluster variable is country); (4) standard errors that are robust to within-cluster correlation across both time and countries as described in Thompson (2011).

Lookback Window (# years)		Inter-cept	Divi-dend yield	Past n-year div growth	Past n-year ERP	R-sq	Deg of free-dom
n=1	Estimate	-0.004	0.077	-0.100	0.016	0.02	9,499
	t-stat						
	Newey-West	-0.30	9.65	-3.69	0.75		
	Clustered by country	-0.24	4.94	-2.28	0.63		
	Clustered by time	-0.27	8.57	-3.31	0.65		
	Clustered by country and time	-0.19	4.47	-2.00	0.48		
n=2	Estimate	0.012	0.062	0.016	-0.109	0.03	9,283
	t-stat						
	Newey-West	1.01	7.52	0.38	-3.60		
	Clustered by country	1.06	4.97	0.21	-2.13		
	Clustered by time	0.89	6.69	0.33	-3.16		
	Clustered by country and time	0.73	4.27	0.18	-1.84		
n=3	Estimate	0.017	0.056	0.057	-0.185	0.03	9,067
	t-stat						
	Newey-West	1.45	6.90	0.98	-4.87		
	Clustered by country	1.43	4.11	0.38	-2.07		
	Clustered by time	1.27	6.16	0.88	-4.31		
	Clustered by country and time	1.02	3.63	0.35	-1.92		
n=4	Estimate	0.033	0.060	-0.093	-0.208	0.03	8,851
	t-stat						
	Newey-West	2.80	7.45	-1.37	-4.87		
	Clustered by country	2.36	3.85	-0.45	-1.95		
	Clustered by time	2.47	6.68	-1.24	-4.35		
	Clustered by country and time	1.82	3.51	-0.43	-1.84		
n=5	Estimate	0.059	0.046	0.008	-0.429	0.05	8,635
	t-stat						
	Newey-West	5.19	5.71	0.12	-8.68		
	Clustered by country	3.54	2.89	0.04	-3.33		
	Clustered by time	4.61	5.12	0.11	-7.74		
	Clustered by country and time	2.98	2.65	0.04	-3.16		

Table 8: Predictive regressions for dividend growth

The table below reports panel regression betas for the one-year dividend growth of 18 countries from 1972 to 2015. The data frequency is monthly. The one-year dividend growth represents the rolling 12-month percentage change in dividends paid by the 18 MSCI country-specific equity indices. The past n-year dividend growth represents the change in dividends paid by the country-specific MSCI index over the prior n years. The one-year trailing dividend yield (log) represents the natural log of the sum of the dividends paid by the country-specific MSCI index over the past year divided by the current index price. The past n-year equity return represents the percentage change in the respective MSCI country-specific gross price over the prior n years. T-stats are calculated from standard errors that have been adjusted for autocorrelation and heteroskedasticity using four alternative methods: (1) a Newey-West covariance estimator with 12 lags; (2) White standard errors that are robust to within-cluster correlation across time (cluster variable is month); (3) White standard errors that are robust to within-cluster correlation across countries (cluster variable is country); (4) standard errors that are robust to within-cluster correlation across both time and countries as described in Thompson (2011).

Lookback Window (# years)		Inter-cept	Divi-dend yield	Past n-year div growth	Past n-year return	R-sq	Deg of free-dom
n=1	Estimate	0.088	-0.041	0.088	0.175	0.16	9,499
	t-stat						
	Newey-West	15.91	-11.39	5.91	15.43		
	Clustered by country	3.80	-2.58	1.77	9.65		
	Clustered by time	14.45	-10.39	5.82	13.59		
	Clustered by country and time	3.73	-2.54	1.76	8.23		
n=2	Estimate	0.084	-0.039	0.007	0.242	0.14	9,283
	t-stat						
	Newey-West	14.18	-10.60	0.29	15.69		
	Clustered by country	3.50	-2.44	0.09	7.13		
	Clustered by time	12.85	-9.64	0.26	13.83		
	Clustered by country and time	3.42	-2.40	0.09	6.52		
n=3	Estimate	0.096	-0.045	-0.045	0.228	0.09	9,067
	t-stat						
	Newey-West	14.42	-10.97	-1.47	11.37		
	Clustered by country	3.57	-2.52	-0.53	4.20		
	Clustered by time	12.93	-9.89	-1.31	10.04		
	Clustered by country and time	3.48	-2.47	-0.51	3.95		
n=4	Estimate	0.108	-0.048	-0.151	0.222	0.07	8,851
	t-stat						
	Newey-West	15.65	-11.31	-3.48	9.48		
	Clustered by country	3.52	-2.44	-1.60	2.92		
	Clustered by time	14.03	-10.16	-3.07	8.39		
	Clustered by country and time	3.45	-2.40	-1.46	2.80		
n=5	Estimate	0.115	-0.052	-0.108	0.170	0.05	8,635
	t-stat						
	Newey-West	16.60	-11.94	-2.12	6.55		
	Clustered by country	3.61	-2.50	-1.06	2.20		
	Clustered by time	14.86	-10.74	-1.87	5.83		
	Clustered by country and time	3.54	-2.47	-0.95	2.10		

A second regression-based approach explores the predictability of the ERP from cross-sectional pricing models such as the CAPM or various multi-factor models. We illustrate this approach by estimating a version of the CAPM-based cross-sectional predictive model of Polk, Thompson and Vuolteenaho (2006). The key insight of the model is that the price of market risk can be estimated from the cross-sectional variation in proxies of expected returns (e.g. the dividend yield) and CAPM betas across stocks. In particular, the expected ERP (λ) is derived from the cross–sectional relation:

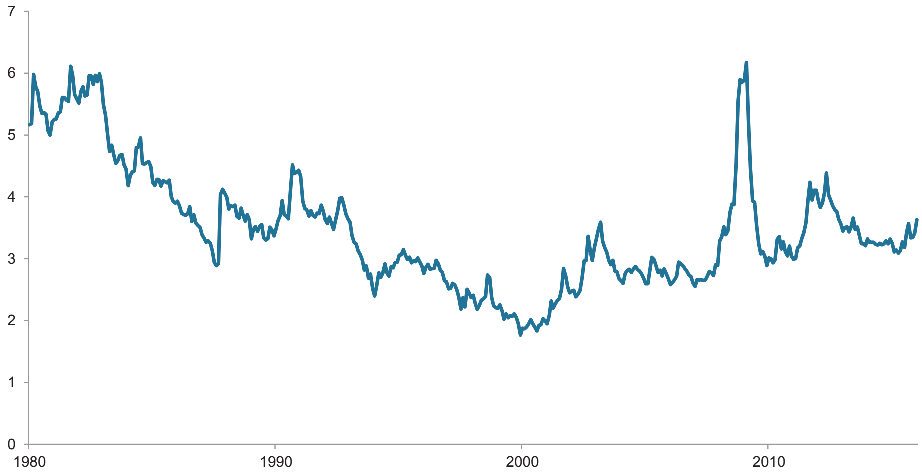
$$ERP\_proxy_i = \lambda \cdot CAPM\_beta_i$$

where ERP_proxy_i is a variable known to be associated with the expected ERP and $CAPM\_beta_i$ is an estimated sensitivity to the realised World ERP over some prior time period. The goal is to find the number λ that makes the asset exposures to market risk as close as possible to the assets' expected returns as proxied by the variable on the left-hand side of the equation. The underlying assumptions are that the equilibrium pricing model holds in the cross-section, the beta sensitivities and the market price of risk remain constant over time, and the left-hand side variable is correlated with expected returns. The advantage of this approach is that it makes use of a variety of asset prices in forecasting the market premium. The disadvantage is that it imposes ex ante assumptions about the distribution of expected risk premia across assets.

We derive a forecast of the expected World ERP based on the Polk, Thompson and Vuolteenaho (2006) approach in two steps. First, for every country in the sample, we estimate the countries' betas relative to the realised World ERP over rolling ten-year windows. Second, we regress the cross-section of country dividend yields in the following month on the estimated country betas to derive the parameter λ, which equals the expected World market price of risk in the CAPM setting. The use of dividend yields on the left-hand side is motivated by the Gordon constant dividend growth model and the regression results in Table 7 that suggest a link between dividend yields and future equity returns. Results of the estimation are reported in Chart 9.

The current expected World premium using this approach is 3.5 percent, an estimate that is 2 to 3 percentage points lower than the estimates from DDMs in Table 6. Relative to history, the current premium is smaller than the premia in the 1980s but larger than the premia in the 1990s when World stock markets were at all-time highs. The premium peaked during the Global Financial Crisis and has declined by 3 percentage points since February 2009. These results suggest that the current dispersion of dividend yields across countries implies a more average World ERP. Alternative model specifications such as making the betas conditional on business cycle variables may yield different results.

Chart 9: Implied World ERP from cross-sectional regressions



Source: Factset, IMF, OECD, Bloomberg, USDA Macroeconomic data; Norges Bank Investment Management

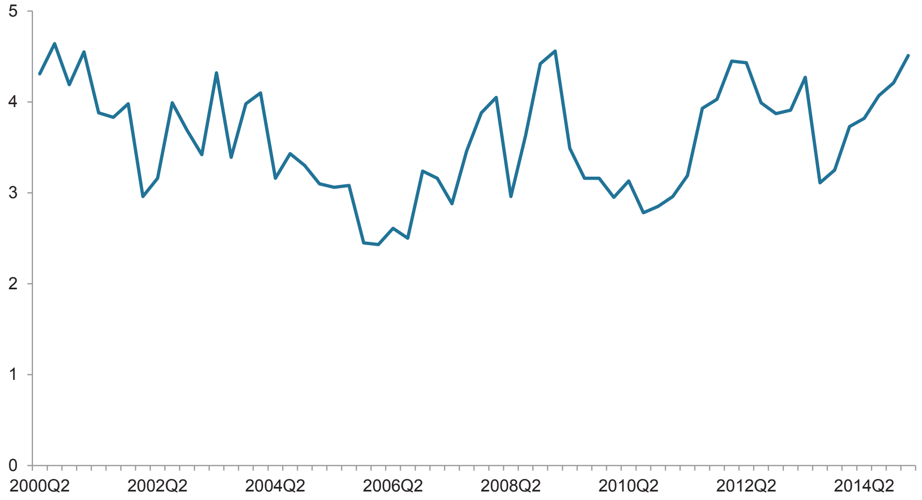
4. Survey data

In addition to quantitative models, practitioners often rely on survey data for the expected ERP. In surveys, the investigator asks academics, financial market participants, analysts or corporate executives for their subjective expectation of the ERP. Surveys are appealing because they extract investors' forward-looking views on the ERP or what ERP values corporate executives use in practice to estimate companies' cost of capital. Critics of surveys, however, argue that survey data are weak and noisy as the answers of survey respondents may be affected by how the question is framed. Survey respondents can also provide wide differences of opinion and answers may be affected by the collective views of the constituent base. For example, Welch (2000) finds that, because of academics' heavy reliance on the historical data approach, their estimates of the ERP do not change very often.

Greenwood and Shleifer (2014) analyse six popular investor surveys and make the surprising finding that investor expectations are negatively correlated with model-based ERP forecasts (the dividend-price ratio). Expectations tend to be positively correlated with equity returns, the current level of the stock market, and flows into equity mutual funds. These findings suggest that investor surveys are not simply noise but rather reflect widely shared beliefs of expected returns. Moreover, they cast doubt on the explanation that the expected ERP is determined solely by time-variation in required rates of return as assumed in theoretical models.

In Chart 10, we show the expected ten-year ERP in the US from the 2015 edition of the Graham and Harvey (2015) survey among corporate CFOs. The average expected risk premium of the S&P 500 index over the next ten years is 4.5 percent, an increase of about 0.5 percent over the prior year. In a wider survey among economics professors, analysts and managers of companies, Fernandez, Oritz and Acin (2015) report estimates of the ERP of about 5 to 5.5 percent in major developed markets. Compared to the previous edition of the survey in 2013, market participants reported a reduction in the expected risk-free rate and relatively unchanged views on the market premium.

Chart 10: Expected ten-year ERP from Graham and Harvey (2015) CFO survey



Source: Graham and Harvey (2015)

VI. Conclusion

In this note, we outlined some of the main empirical and theoretical evidence on arguably the most important variable in finance – the equity risk premium. The long-term average realised premium is large but also time-varying, underlying both the risks and rewards of investing in equity markets. While significant progress has been made in understanding the drivers of the premium, no single explanation can account for all of the stylised facts in the premium's historical record. The empirical evidence indicates that the premium constitutes a compensation for risk which can be expected to persist over the long run, but the historical size of the premium has likely also been affected by secular trends such as increased participation in the equity market, periods of "irrational exuberance" as in the 1990s, and more recently by loose monetary policy.

The expected ERP is typically estimated from quantitative models that assume that investors' required rates of return equal the expected premium. Popular models include unconditional mean forecasts, dividend discount models and predictive regressions. There is significant heterogeneity in the estimates of the expected World ERP across the different models. Based on dividend discount models, which assume that the current interest rate is a good proxy for the future expected short-term rate, we estimate that the expected World ERP is around 6 percent as at January 2016. If we account for the effect of the current low interest rates or put less emphasis on recent cash flow growth data, we estimate a World ERP of 3 to 4 percent. Our cross-sectional and time-series regression-based approaches also support a lower World ERP of around 3 to 4 percent. The expected premium as estimated by these models has declined significantly since the end of the Global Financial Crisis and remains near its long-term repricing-adjusted average level.

VII. References

Abel, A. (1990): "Asset prices under habit formation and catching up with the Joneses", *American Economic Review Papers and Proceedings*, 80, 38-42.

Ait-Sahalia, Y., Parker, J., Yogo, M. (2004): "Luxury goods and the equity premium", *Journal of Finance*, 59, 2959-3004.

Arnott, R., Bernstein, P. (2002): "What risk premium is 'normal'?", *Financial Analysts Journal*, 58(2), 64-85.

Asness, C. (2003): "Fight the Fed model", *Journal of Portfolio Management*, 30(1), 11-24.

Baker, M., Wurgler, J. (2000): "The equity share in new issues and aggregate stock returns", *Journal of Finance*, 55(5), 2219-2257.

Bansal, R., Yaron, A. (2004): "Risks for the long run: A potential resolution of asset pricing puzzles", *Journal of Finance,* 59(4), 1481-1509.

Barberis, N., Huang, M., Santos, T. (2001): "Prospect theory and asset prices", *Quarterly Journal of Economics*, 116(1), 1-53.

Barro, R. (2006): "Rare disasters and asset markets in the twentieth century", *Quarterly Journal of Economics*, 121(3), 823-866.

Barro, R., Ursua, J. (2012): "Rare macroeconomic disasters", *Annual Review of Economics*, 4(1), 83-109.

Beeler, J., Campbell, J. (2012): "The long-run risks model and aggregate asset prices: An empirical assessment", *Critical Finance Review*, 1(1), 141-182.

Bekaert, G., Harvey, C., Lundblad, C. (2007): "Liquidity and expected returns: Lessons from emerging markets", *Review of Financial Studies*, 20(6), 1783-1831.

Bekaert, G., Harvey, C., Lundblad, C., Siegel, S. (2011): "What segments equity markets?", *Review of Financial Studies*, 24, 3841-3890.

Benartzi, S., Thaler, R. (1995): "Myopic loss aversion and the equity premium puzzle", *Quarterly Journal of Economics*, 110(1), 73-92.

Bernstein, W., Arnott, R. (2003): "Earnings growth: The two percent dilution", *Financial Analysts Journal*, 59(5), 47-55.

Bhamra, H., Uppal, R. (2014): "Asset prices with heterogeneity in preferences and beliefs", *Review of Financial Studies*, 27(2), 519-580.

Brennan, M., Schwartz, E. (1985): "Evaluating natural resource investments", *Journal of Business*, 58(2), 135-157.

Campbell, J. (1987): "Stock returns and the term structure", *Journal of Financial Economics*, 18, 373-399.

Campbell, J. (2003): "Consumption-based asset pricing", in Constantinides, G., Harris, M., Stulz, R. (eds.) *Handbook of the Economics of Finance*, Elsevier Science.

Campbell, J., Cochrane, J. (1999): "By force of habit: A consumption-based explanation of aggregate stock market behaviour", *Journal of Political Economy*, 107(2), 205-251.

Campbell, J., Cochrane, J. (2015): "The fragile benefits of endowment destruction", *Journal of Political Economy*, 123(5), 1214-1226.

Campbell, J., Shiller, R. (1988): "Stock prices, earnings and expected dividends", *Journal of Finance*, 43(3), 661-676.

Campbell, J., Shiller, R. (1989): "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies*, 1(3), 195-228.

Campbell, J., Thompson, S. (2008): "Predicting excess stock returns out of sample: Can anything beat the historical average?", *Review of Financial Studies*, 21(4), 1509-1531.

Chetty, R. (2006): "A new method of estimating risk aversion", *American Economic Review*, 96(5), 1821-1834.

Cochrane, J. (2008): "Financial markets and the real economy", in Mehra, R. (ed.) *Handbook of the equity risk premium*, Elsevier.

Cochrane, J. (2008a): "The dog that did not bark: A defense of return predictability", *Review of Financial Studies*, 21(4), 1533-1575.

Cochrane, J. (2011): "Discount rates", *Journal of Finance*, 66(4), 1047-1108.

Collin-Dufresne, P., Johannes, M., Lochstoer, L. (2016): "Parameter learning in general equilibrium: The asset pricing implications", *American Economic Review*, 106(3), 664-698.

Constantinides, G. (1990): "Habit formation: A resolution of the equity premium puzzle", *Journal of Political Economy,* 98, 519-543.

Constantinides, G. (2002): "Rational asset prices", *Journal of Finance*, 57(4), 1567-1591.

Constantinides, G., Donaldson, J., Mehra, R. (2002): "Junior can't borrow: A new perspective on the equity premium puzzle", *Quarterly Journal of Economics*, 117(1), 269-296.

Damodaran, A. (2012): "The equity risk premiums (ERP): Determinants, estimation and implications – the 2012 edition", working paper, New York University.

Dangl, T., Halling, M. (2012): "Predictive regressions with time-varying coefficients", *Journal of Financial Economics*, 106(1), 157-181.

De Long, J., Shleifer, A., Summers, L., Waldmann, R. (1990): "Noise trader risk in financial markets", *Journal of Political Economy*, 98, 703-738.

Dimson, E., Marsh, P., Staunton, M. (2011): "Equity premia around the world", working paper, London Business School.

Donaldson, J., Mehra, R. (2008): "Risk-based explanations of the equity premium", in Mehra, R. (ed.) *Handbook of the equity risk premium*, Elsevier.

Duarte, F., Rosa, C. (2015): "The equity premium: A review of models", Federal Reserve Bank of New York Staff Report 714.

Epstein, L., Zin, S. (1989): "Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework", *Econometrica*, 57(4), 937-969.

Epstein, L., Zin, S. (1991): "Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis", *Journal of Political* Economy, 99(2), 263-286.

Fama, E., French, K. (1988): "Dividend yields and expected stock returns", *Journal of Financial Economics*, 22, 3-25.

Fama, E., French, K. (2002): "The equity premium", *Journal of Finance*, 57(2), 637-659.

Fernandez, P., Oritz, A., Acin, I. (2015): "Discount rate (risk-free rate and market risk premium) used for 41 countries in 2015: A survey", IESE working paper.

Ferreira, M., Santa-Clara, P. (2011): "Forecasting stock market returns: The sum of the parts is more than the whole", *Journal of Financial Economics*, 100(3), 514-537.

Foerster, S., Karolyi, A. (1999): "The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stocks listing in the United States", *Journal of Finance*, 54(3), 981-1013.

Gibson, R., Mougeot, N. (2004): "The pricing of systematic liquidity risk: Empirical evidence from the US stock market", *Journal of Banking and Finance*, 28, 157-178.

Gneezy, U., Kapteyn, A., Potters, J. (2003): "Evaluation periods and asset prices in a market experiment", *Journal of Finance*, 58, 821-837.

Golez, B., Koudijs, P. (2016): "Four centuries of return predictability", working paper.

Gordon, M. (1962): *The investment financing and valuation of the corporation*, Irwin.

Gordon, R. (2012): "Is U.S. economic growth over? Faltering innovation confronts the six headwinds", NBER Working Paper 18315.

Gordon, R. (2014): "The demise of U.S. economic growth: Restatement, rebuttal and reflections", NBER Working Paper 19895.

Goyal, A., Santa-Clara, P. (2003): "Idiosyncratic risk matters!", *Journal of Finance*, 58, 975-1008.

Graham, J., Harvey, C. (2015): "The equity risk premium in 2015: Evidence from the Global CFO Outlook Survey".

Greenwood, R., Shleifer, A. (2014): "Expectations of returns and expected returns", *Review of Financial Studies*, 27(3), 714-746.

Grinold, R., Kroner, K., Siegel, L. (2011): "A supply model of the equity premium. Rethinking the equity risk premium", CFA Institute.

Guvenen, F. (2009): "A parsimonious macroeconomic model for asset pricing", NBER Working Paper 15243.

Hansen, L., Singleton, K. (1983): "Stochastic consumption, risk aversion and the temporal behaviour of asset returns", *Journal of Political Economy*, 91(2), 249-265.

He, Z., Krishnamurthy, A. (2013): "Intermediary asset pricing", *American Economic Review*, 103(2), 732-770.

Henkel, S., Martin, J., Nardari, F. (2011): "Time-varying short-horizon predictability", *Journal of Financial Economics*, 99(3), 560-580.

Ibbotson, R., Chen, P. (2003): "Long-run stock returns: Participating in the real economy", *Financial Analysts Journal*, 59(1), 88-98.

Ilmanen, A. (2011): *Expected returns*, Wiley.

Jagannathan, R., Wang, Y. (2007): "Lazy investors, discretionary consumption and the cross-section of stock returns", *Journal of Finance*, 62, 1623-1661.

Jones C., Wilson J. (2005): "The Equity Risk Premium Controversy", *Journal of Investing*, 14(2), 37-43.

Kelly, B., Pruitt, S. (2013): "Market expectations in the cross-section of present values", *Journal of Finance*, 68(5), 1721-1756.

Kocherlakota, N. (1996): "The equity premium: It's still a puzzle", *Journal of Economic Literature*, 24, 42-71.

Krishnamurthy, A., Vissing-Jorgensen, A. (2012): "The demand for Treasury debt", NBER Working Paper 12881.

Kroencke, T. (2016): "Asset pricing without garbage", accepted manuscript in the *Journal of Finance*, 4 August 2016.

Lettau, M., Ludvigson, S. (2001): "Consumption, aggregate wealth and expected stock returns", *Journal of Finance*, 56, 815-849.

Lettau, M., Ludvigson, S., Wachter, J. (2008): "The declining equity risk premium: What role does macroeconomic risk play?", *Review of Financial Studies*, 21, 1653-1687.

Li, Y., Ng, D., Swaminathan, B. (2013): "Predicting market returns using implied cost of capital", *Journal of Financial Economics*, 110(2), 419-436.

Ljungqvist, L., Uhlig, H. (1999): "On consumption bunching under Campbell-Cochrane habit formation", SSE/EFI Working Paper Series in Economics and Finance 337, Stockholm School of Economics.

Ludvigson, S., Ng, S. (2007): "The empirical risk-return relation: A factor analysis approach", *Journal of Financial Economics*, 83(1), 171-222.

Lustig, H., Van Nieuwerburgh, S., Verdelha, A. (2013): "The wealth-consumption ratio", *Review of Asset Pricing Studies*, 3, 38-94.

Malloy, C., Moskowitz, T., Vissing-Jorgensen, A. (2009): "Long-run stockholder consumption risk and asset returns", *Journal of Finance*, 63, 2427-2479.

Mehra, R., Prescott, E. (1985): "The equity premium: A puzzle", *Journal of Monetary Economics*, 15(2), 145-161.

Nakamura, E., Steinsson, J., Barro, R., Ursua, J. (2013): "Crises and recoveries in an empirical model of consumption disasters", *American Economic Journal: Macroeconomics*, 5(3), 35-74.

Neely, C., Rapach, D., Tu, J., Zhou, G. (2014): "Forecasting the equity risk premium: The role of technical indicators", *Management Science*, 60(7), 1772-1791.

Pastor, L., Sinha, M., Swaminathan, B. (2008): "Estimating the intertemporal risk-return tradeoff using the implied cost of capital", *Journal of Finance*, 63(6), 2859-2897.

Polk, C., Thompson, S., Vuolteenaho, T. (2006): "Cross-sectional forecasts of the equity premium", *Journal of Financial Economics*, 81(1), 101-141.

Pontiff, J. (1996): "Costly arbitrage: Evidence from closed end funds", *Quarterly Journal of Economics*, 111, 1135-1152.

Pontiff, J., Schall, L. (1998): "Book-to-market ratios as predictors of market returns", *Journal of Financial Economics*, 49, 141-160.

Rangvid, J., Schmeling, M., Schrimpf, A. (2014): "Dividend predictability around the world", *Journal of Financial and Quantitative Analysis*, 49, 1255-1277.

Rapach, D., Ringgenberg, M., Zhou, G. (2016): "Short interest and aggregate stock returns", *Journal of Financial Economics*, 121(1), 46-65.

Rapach, D., Strauss, J., Zhou, G. (2010): "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy", *Review of Financial Studies*, 23(2), 821-862.

Rietz, T. (1988): "The equity risk premium: A solution", *Journal of Monetary Economics*, 22, 117-131.

Roll, R. (1977): "A critique of the asset pricing theory's tests. Part I: On past and potential testability of the theory", *Journal of Financial Economics*, 4(2), 129-176.

Savov, A. (2011): "Asset pricing with garbage", *Journal of Finance*, 66, 177-201.

Scheinkman, J., Xiong, W. (2003): "Overconfidence and speculative bubbles", *Journal of Political Economy*, 111, 1183-1219.

Shiller, R. (2005): *Irrational exuberance*, 2nd ed., Princeton University Press.

Siegel, J., Thaler, R. (1997): "Anomalies: The equity premium puzzle", *Journal of Economic Perspectives*, 11(1), 191-200.

Stulz, R. (1999): "Globalization, corporate finance and the cost of capital", *Journal of Applied Corporate Finance*, 12(3), 8-25.

Swan, P. (2000): "Illiquidity rather than 'risk aversion' explains the equity premium puzzle: The role of endogenous stock market trading", working paper, University of Sydney.

Thompson, S. (2011): "Simple formulas for standard errors that cluster by both firm and time", *Journal of Financial Economics*, 99, 1-10.

Tsai, J., Wachter, J. (2015): "Disaster risk and its implications for asset pricing", *Annual Review of Financial Economics*, 7, 219-252.

Van Binsbergen, J., Koijen, R. (2010): "Predictive regressions: A present-value approach", *Journal of Finance*, 65(4), 1439-1471.

Wachter, J. (2013): "Can time-varying risk of rare disasters explain aggregate stock market volatility?", *Journal of Finance*, 68, 987-1035.

Weil, P. (1989): "The equity premium puzzle and the risk-free rate puzzle", *Journal of Monetary Economics*, 24, 401-421.

Welch, I. (2000): "Views of financial economists on the equity premium and on professional controversies", *Journal of Business*, 73(4), 501-537.

Welch, I., Goyal, A. (2008): "A comprehensive look at the empirical performance of equity premium prediction", *Review of Financial Studies*, 21(4), 1455-1508.