The Enhanced Risk Premium Factor Model & Expected Returns

Javier Estrada*
IESE Business School, Department of Finance, Av. Pearson 21, 08034 Barcelona, Spain
Tel: +34 93 253 4200, Fax: +34 93 253 4343, Email: jestrada@iese.edu

Abstract
The predictability of the equity risk premium is a central and controversial issue in finance. The Risk Premium Factor model is a recent and novel approach to forecasting the equity risk premium and the equity market's level and P/E. This article aims to overcome the main limitation of, and therefore improve upon, this novel approach. The Enhanced Risk Premium Factor model proposed here to forecast the market's return is clearly supported by the evidence. Furthermore, a model that articulates the same variables considered in the framework proposed, but that imposes no specific functional form to relate them, produces very highly correlated and unbiased forecasts of the market's return.

October, 2012

1. Introduction

The model of investor behavior proposed by Markowitz (1952) and later expanded by Sharpe (1964), Lintner (1965), and Mossin (1966) collapses into the ubiquitous CAPM. This model states that the return investors should expect/require from equity consists of two parts, a risk-free rate and a risk premium, the latter being the product of beta and the equity risk premium (ERP).[1] This last variable and its role when forecasting the equity market's return are the focus of this article.

The ERP is the additional compensation investors should expect/require for investing in relatively-riskier equity as opposed to in relatively-safer (and in the limit risk-free) government debt. The two main controversies about this variable are 1) whether it varies over time, and 2) how to estimate it. Both issues are addressed in this article.

Hassett (2010) proposes a novel approach to estimate a time-varying ERP as a function of two variables, the risk-free rate and a risk premium factor (RPF); the latter is explained in more detail below. He calls this approach the RPF model, uses it to forecast the S&P-500's level and P/E ratio, and finds that it fits the data well.

The main shortcoming of Hassett's (2010) approach is that the market's level is forecasted over the same period and with the same data used to estimate the RPF, thus

* I would like to thank Steve Hassett for his comments. Sergi Cutillas provided valuable research assistance. The views expressed below and any errors that may remain are entirely my own.
[1] Note that in the equilibrium the CAPM characterizes, required and expected returns must be the same; if they were not, then by definition the situation would not be an equilibrium.

producing *in sample* forecasts. In Hassett's (2010) model, out of sample forecasts of the market's level must necessarily assume that that the estimated RPF will remain unchanged in the future. The obvious weakness of this approach is twofold: It does not identify the variables that explain the variability of the RPF over time, and therefore it is not able to predict what the RPF will be out of sample. Furthermore, this approach aims to predict the market's level, hence its expected capital gain/loss, thus ignoring dividends which are a sizeable proportion of the market's expected return.

The main contribution of this article is to propose a two-step approach to forecast the market's total return. In the first step, the RPF is forecasted on the basis of a model proposed; in the second step, the forecasted RPF is used to forecast the market's return. The model proposed to explain the variability of (hence forecast) the RPF is based on the cyclically-adjusted P/E ratio (CAPE). Thus, in the approach suggested here the market's expected return is a function of the expected RPF, which in turn is a function of the current CAPE, and the current risk-free rate.

This proposed variation of the RPF model is referred to the here as the Enhanced Risk Premium Factor model, or ERPF model for short, and is assessed by evaluating its ability to forecast the annualized returns of the S&P-500 over 10-year holding periods. As discussed below, the correlation between the observed and forecasted returns is almost 0.9; hence, the simple framework proposed here fits the data very well.

The ERPF model posits a specific functional relationship between expected returns, the RPF (hence CAPE), and the risk-free rate. A model that articulates the same variables but without imposing a specific functional relationship among them slightly outperforms the ERPF model, producing unbiased forecasts even more highly correlated with observed returns.

The rest of the article is organized as follows. Section 2 expands the discussion of the issue at stake. Section 3 reports and discusses the evidence supporting the model proposed and its extensions. Finally, section 4 provides an assessment.

2. The Issue

The equity risk premium is an essential variable in a wide variety of financial activities such as company valuation, project evaluation, capital structure optimization, performance measurement, and forecasting, to name but a few. However, despite its central role in finance, there is substantial disagreement about how to estimate it, with different proposals producing vastly different values.

One of the controversies surrounding the ERP is whether it is constant or varies over time. Early modern financial theory, resting on the concept of market efficiency, suggests the former. In fact, in efficient markets in which prices incorporate all available information prices follow a martingale process, thus implying constant expected returns; see, for example, Samuelson (1965).

Over the very short term investor preferences and economic conditions may not change substantially, and therefore a constant ERP may not be an implausible assumption. However, over longer time periods, both preferences and economic conditions are likely to change, thus causing the ERP to change. One of the implications of consumption-based asset pricing models is that a time-varying ERP can be reconciled with efficient markets and rational behavior.

Rubinstein (1976) and Lucas (1978), among others, derive a negative relationship between the ERP and economic conditions. When the economy is healthy, consumption is high, the marginal utility of consumption is low, and the return required to give up consumption is low, thus leading to a low ERP. When the economy is depressed, a similar reasoning leads to a high ERP. In other words, the ERP and the business cycle are negatively correlated.

The reconciliation of time-varying risk premiums and rationality opened the door to the exploration of economic and financial variables that could help explain the variability of the ERP. Darolles et al (2010) argue that in order to be consistent with asset pricing models of rational behavior, the variables considered must be negatively correlated to the business cycle and mean reverting in the long term. They subsequently list and discuss some of the variables that have been proposed to explain the variability of the ERP.

Explaining its variability is, in fact, a necessary step in order to forecast the ERP. To this latter purpose, at least three broad approaches can be distinguished: A historical estimation, a forward-looking estimation, and a survey-based estimation.[2] Ilmanen (2011) and Damodaran (2012) discuss at length all three approaches and provide a rather comprehensive literature review. The approach proposed by Hassett (2010) and the variation proposed in this article both fall within the second category.

2.1. The Risk Premium Factor Model

Hassett (2010) argues that a constant ERP implausibly implies that, as the risk-free rate increases, investors are willing to accept an increasingly smaller premium *as a proportion of the risk-free rate*. To illustrate, recall that the required return on the market (R) can be expressed as the sum of the risk-free rate and the ERP; that is,

$$R = y + \text{ERP}. \tag{1}$$

where y denotes the risk-free rate, typically approximated with the yield to maturity on a government bond. Assume a constant ERP of (say) 6%. Then, it follows from (1) that the

[2] Another possible approach, far less used, is to derive the ERP from a theoretical model of investor behavior; see, for example, the seminal article by Mehra and Prescott (1985).

required return on the market would be 8% when the risk-free rate is 2%, and 12% when the risk-free rate is 6%. Note that the constant ERP is three times the risk-free rate (and three quarters of the market's return) in the first case, but only one time the risk-free rate (and half of the market's return) in the second case.

Hassett (2010) suggests that the ERP should not be a constant number but a constant *proportion* of the risk-free rate and calls this proportion the risk premium factor (RPF). Under this assumption, the ERP and the required return on the market are respectively given by

$$\text{ERP} = \text{RPF}\cdot y\,, \tag{2}$$

$$R = y + \text{ERP} = y + \text{RPF}\cdot y = (1+\text{RPF})\cdot y\,. \tag{3}$$

Expressions (2)-(3) are the core of the *Risk Premium Factor model*, or *RPF model* for short.

Note that with an RPF of (say) 1.2, the required return on the market would be 4.4% when the risk-free rate is 2% (=2%+1.2·2%) and 13.2% when the risk-free rate is 6% (=6%+1.2·6%). Hence, the ERP is not constant (it is 2.4% in the first case and 7.2% in the second case) but a time-varying magnitude that fluctuates with the risk-free rate.

Hassett (2010) first estimates the RPF and then uses the estimated value of this coefficient to forecast the level of the S&P-500 *over the same period and with the same data used to estimate the RPF*.[3] When the model stops making reasonably-accurate *in sample* forecasts, the RPF is re-estimated and the new value of this coefficient is then used to make further *in sample* forecasts of the market's level.

Following this methodology over the Dec/1959-Sep/2009 period, Hassett (2010) finds only two distinct shifts in the RPF. He finds an RPF of 1.24 between Dec/1959 and Dec/1980; 0.90 between Dec/1980 and Jun/2002; and 1.48 between Jul/2002 and Sep/2009. He then uses these RPF values to forecast the annual level of the S&P-500 between 1960 and 2008, runs a regression between the forecasted and the observed values of the S&P-500, and obtains an R^2=0.895.[4]

Hassett (2010) argues that the RPF model is consistent with the framework of prospect theory developed by Kahneman and Tversky (1979). He further argues that the RPF values estimated, in the 0.90–1.48 range, are consistent with loss aversion coefficients of 1.90 and 2.48, both close to the widely-cited loss aversion coefficient of 2.25 estimated by Tversky and Kahneman (1992).

---

[3] More precisely, the RPF is estimated as the slope of a linear regression between (E/P+G) as the dependent variable and the risk-free rate as the independent variable, where E/P is the earnings yield on the S&P-500 and G is the long-term growth of earnings. For the assumptions made on G and other details of the estimation, see Hassett (2010).

[4] He also runs the model using quarterly, monthly, and daily data over the shorter Dec/1985-Sep/2009 period and reports R^2s between observed and predicted levels of the S&P of 0.806, 0.863, and 0.865.

Importantly, note that the RPF model is devised to forecast the market's level, and therefore only its capital gain/loss, not its total return. As a result, it ignores dividends, which historically have provided about half of the total return delivered by the market. The model proposed here, in turn, aims to forecast the market's *total* return.

2.2. The Enhanced Risk Premium Factor Model

As Hassett (2010) himself admits, a weakness of his approach is that the variables that determine the variability of the RPF are not identified, and, therefore, the RPF cannot be forecasted out of sample; it can only be assumed (hoped) that it will remain constant in the future. The approach proposed here overcomes both deficiencies by first suggesting a model to explain and forecast the RPF, and then using the forecasted RPF to forecast both the ERP and the market's return. Furthermore, the approach proposed here does not need some of the restrictive assumptions imposed by Hassett (2010) on his model.[5]

In order to make clear what is observed and what needs to be forecasted, let's re-write expressions (1)-(3) as

$$R_1 = y_0 + \mathrm{ERP}_1 \,, \tag{4}$$

$$\mathrm{ERP}_1 = \mathrm{RPF}_1 \cdot y_0 \,, \tag{5}$$

$$R_1 = (1+\mathrm{RPF}_1) \cdot y_0 \,, \tag{6}$$

where 0 denotes the present (the moment when the forecast is made) and 1 denotes the future (the period for which the forecast will be made). Throughout this article, R_1 denotes the annualized (total) return of the S&P-500 over 10-year holding periods.

Importantly, note from (5) that the *expected* ERP depends on the *expected* RPF and the *observed* risk-free rate, the first two over the 10-year period for which the market's return is forecasted and the last one at the time the forecast is made. Similarly, note from (6) that the market's *expected* return depends on the *expected* RPF and the *observed* risk-free rate.

It follows from (6) that $\mathrm{RPF}_1 = R_1/y_0 - 1$, which provides a way to generate a time series of observed RPFs; that is,

$$\mathrm{RPF}_{1t} = R_{1t}/y_{0t} - 1 \,, \tag{7}$$

where t indexes time. Once a time series of observed RPFs is obtained in this fashion, the next step consists of finding a model that can explain the observed behavior of this variable. Such a

[5] These assumptions include a discounted cash flow model with constant growth, earnings being a good proxy for cash flow, and a 2.6% constant annual growth of real GDP.

model can then be used to produce an out of sample forecast of the RPF over the period for which an out of sample forecast of the market's return is needed.

It is argued here that the RPF is not constant but an inverse function of risk in the equity market.[6] In other words, an increase (decrease) in risk leads investors to lower (raise) their RPF. Risk, in turn, is evaluated from a practitioner-oriented perspective, particularly that of value-oriented investors. Thus, rather than assessing risk with volatility or beta, it is suggested here that that risk is essentially a function of price relative to fundamentals. More precisely, the higher stock prices are relative to fundamentals, the higher is the risk borne by investors, and the lower is the RPF.

Marks (2011), for example, argues that "most investors think quality, as opposed to price, is the determinant of whether something's risky. But high quality assets can be risky, and low quality assets can be safe. It's just a matter of the price paid for them … Elevated popular opinion, then, isn't just the source of low return potential, but also of high risk." He also argues that when "greed goes to excess, security prices tend to be too high. That makes prospective return low and risk high. The assets in question represent mistakes waiting to produce a loss …" Similarly, Richards (2012) argues that the "more expensive stocks … are, the more risky they are."

Many variables have been proposed in the literature to assess the relationship between price and value. The focus here is on the cyclically-adjusted P/E ratio (CAPE); that is, the ratio between the level of the market (in our case, the S&P-500) at a given point in time and the average earnings per share over the previous T (in our case, 10) years. As is well known, averaging earnings over several years produces a smoothed multiple that is not heavily influenced by the typical short-term fluctuations of earnings over the business cycle. This approach, originally devised by Graham and Dodd (1934), was more recently popularized by Campbell and Shiller (1988).

The model proposed here, then, is implemented by first producing a time series of observed RPFs using expression (7) and then running the regression

$$\text{RPF}_{1t} = \alpha + \beta \cdot \text{CAPE}_{0t} + u_t\,, \qquad (8)$$

where u_t is an error term. The estimated α and β coefficients from this regression are then used to produce an out of sample forecast of the RPF with the expression

$$\text{E}(\text{RPF}_{1t}) = a + b \cdot \text{CAPE}_{0t}\,, \qquad (9)$$

[6] In fact, although Hassett (2010) argues that the RPF is constant, as already discussed he finds that it does change (albeit infrequently) over time. He does not, however, propose a model to explain those changes, which is one of the contributions of this article.

where 'E' denotes an expected value, and a and b are the sample estimates of α and β. The forecast produced from (9) is then used to produce out of sample forecasts of both the ERP and the (total) return of the equity market with the expressions

$$\mathrm{E}(\mathrm{ERP}_{1t}) = (a+b\cdot\mathrm{CAPE}_{0t})\cdot y_{0t}\,, \tag{10}$$

$$\mathrm{E}(R_{1t}) = y_{0t} + \mathrm{E}(\mathrm{ERP}_{1t}) = [1+\mathrm{E}(\mathrm{RPF}_{1t})]\cdot y_{0t} = (1+a+b\cdot\mathrm{CAPE}_{0t})\cdot y_{0t}\,. \tag{11}$$

Expressions (9)-(11) are the core of the *Enhanced Risk Premium Factor model*, or *ERPF model* for short, proposed here.

3. Evidence

This section evaluates the ability of the ERPF model to forecast the equity market's return. In order to make the results comparable to those discussed by Hassett (2010) the data and sample period are those he uses, only slightly updated to reflect the additional data available since the publication of his article. Thus, the sample period considered here goes from the end of 1959 through the end of 2011.

Monthly stock returns over the Jan/1960-Dec/2011 period are based on the S&P-500 and account for both capital gains/losses and dividends. The CAPE is based on the value of the S&P at the end of each month divided by the average earnings per share over the previous 120 months.[7] The risk-free rate is approximated with the yield to maturity on 10-year US Treasury Notes at the end of each month. The goal of the model proposed and the whole analysis is to forecast the ERP and ultimately the annualized (total) return of the S&P over 10-year holding periods.

3.1. The ERPF Model and Expected Returns

The analysis starts at the beginning of 1960 by looking back 10 years and running the regression in (8) with monthly data over the Dec/1949-Dec/1959 period. The purpose of this regression is to forecast the RPF over the Dec/1959-Dec/1969 period. Monthly values of the dependent variable are calculated as shown in (7). Importantly, recall that each RPF is calculated on the basis of an observed risk-free rate and a return of the market over the subsequent 10 years. Hence, expression (8) relates *future* RPFs with observed CAPEs, and can therefore be used to forecast *expected* RPFs on the basis of observed CAPEs.

[7] In order to calculate CAPE at the end of 1959, earnings per share are collected since the beginning of 1950. Furthermore, because the first regression with CAPE as independent variable is run at the beginning of 1950 (more on this below), earnings per share are actually collected since the beginning of 1940.

After running the regression in (8) over the Dec/1949-Dec/1959 period, similar regressions are run at the end of each month with a rolling window of 120 observations until the data is exhausted. The last regression is run over the Dec/1991-Dec/2001 period. Note that the RPF at that time is based on the risk-free rate observed at the end of 2001 and the annualized return of the market over the Dec/2001-Dec/2011, which exhausts the data available.

As shown in the first row of Exhibit 1, across the 505 regressions run in this fashion, the average intercept is 3.16, the average slope −0.12, and the average R-squared 0.51.[8] Thus, the first step of the analysis suggests that CAPE does in fact explain a substantial proportion of the variability of the RPF, and, therefore, can be a useful tool to forecast future values of this variable. This is a substantial improvement over the informal stories used by Hassett (2010) to explain the two breaks he observes in the RPF. In fact, the results just discussed make clear why the RPF varies over time: Because valuation (hence risk) changes over time.

Exhibit 1: Forecasting the RPF, the ERP, and the Market's Return – The ERPF Model

This exhibit shows the average intercept, slope, and R^2 across 505 regressions between the RPF and three explanatory variables, CAPE, P/E, and D/P. The RPF is calculated as shown in (7). The equity market is represented by the S&P-500 index. CAPE is based on 10-year trailing earnings; P/E and D/P are based on 1-year trailing earnings and dividends. The first regression is run over the Dec/1949-Dec/1959 period and the last one over the Dec/1991-Dec/2001 period. The exhibit also shows, over the Dec/1959-Dec/2001 period, the correlation between observed and forecasted RPFs (Rho-1), the latter estimated from (9); the correlation between observed and forecasted ERPs (Rho-2), the latter estimated from (10); the correlation between observed and forecasted 10-year annualized returns (Rho-3), the latter estimated from (11); and the mean forecasting error (MFE) calculated as the average difference between observed and forecasted annualized returns.

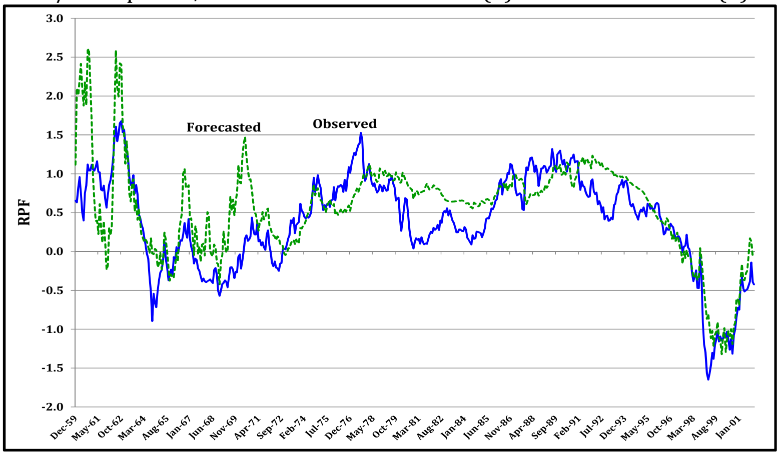
Variable	Intercept	Slope	R^2	Rho-1	Rho-2	Rho-3	MFE (%)
CAPE	3.16	−0.12	0.51	0.79	0.76	0.88	−1.42
P/E	2.10	−0.09	0.34	0.55	0.59	0.81	−1.44
D/P	−1.17	57.16	0.42	0.72	0.67	0.84	−1.41

Each of the regressions estimated was subsequently used to forecast an *out of sample* RPF using expression (9), which produced a time series of 505 RPFs over the Dec/1959-Dec/2001 period. The minimum and maximum RPFs estimated were −1.32 and 2.61, with an average of 0.55, thus indicating a range somewhat wider than that found by Hassett (2010). Figure 1 plots the time series of observed and forecasted RPFs and suggests a close relationship between these two variables. Such close relationship is confirmed by a correlation of 0.79 shown in Exhibit 1 (Rho-1) and suggests that (9) can be a useful tool to forecast the RPF.

[8] At the 5% level of significance, the intercept and the slope are significant in 83.0% and 94.3% of the 505 regressions run.

Figure 1: Observed and Forecasted RPFs

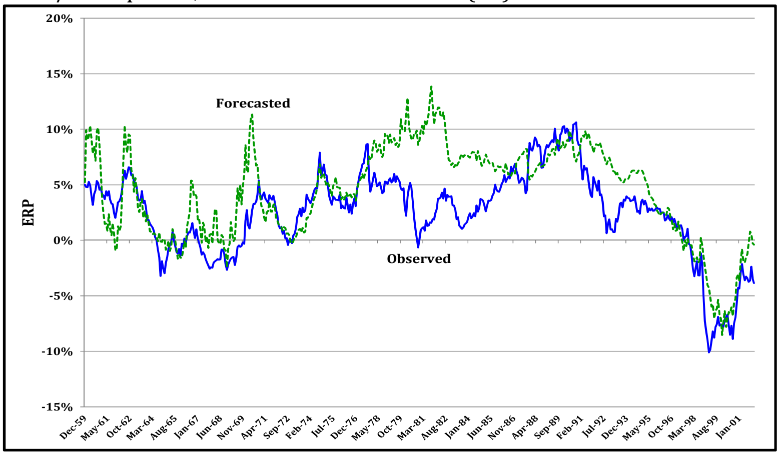
This figure shows observed and forecasted RPFs over the Dec/1959-Dec/2001 period, the former estimated from (7) and the latter from (9).



The time series of forecasted RPFs, together with a time series of observed risk-free rates, can be used to forecast *out of sample* ERPs using expression (10). The observed and forecasted ERPs over the Dec/1959-Dec/2001 period are depicted in Figure 2.[9] The correlation of 0.76 between these two variables shown in Exhibit 1 (Rho-2) confirms the close relationship between them and suggests that (10) can be a useful tool to forecast the ERP.[10]

Figure 2: Observed and Forecasted ERPs

This figure shows observed and forecasted ERPs over the Dec/1959-Dec/2001 period, the latter estimated from (10).



Finally, the time series of forecasted RPFs and observed risk-free rates can be used to forecast *out of sample* returns using expression (11).[11] Figure 3 plots the observed and

[9] Consistent with the whole analysis, observed ERPs are calculated on a forward-looking basis. To illustrate, the observed ERP at the end of Dec/2001 is based on the differential total return between stocks (the S&P) and bonds (10-year Treasury Notes) over the Dec/2001-Dec/2011 period.

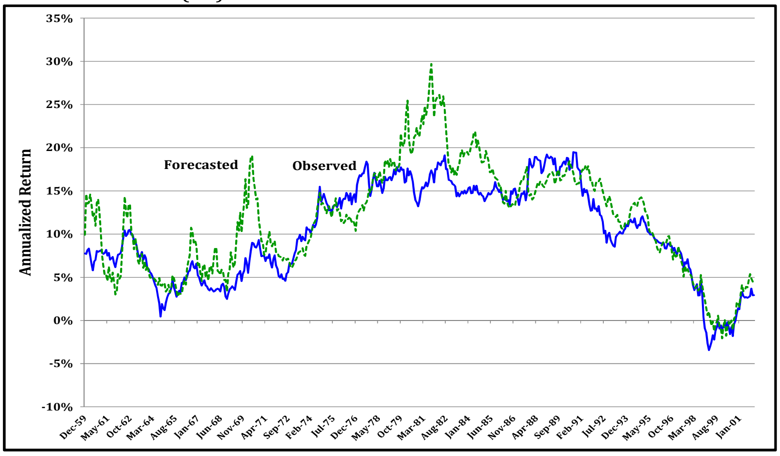
[10] The big gap between observed and forecasted ERPs in the middle of the figure coincides with the period of double-digit yields of 10-year Treasury notes, between late 1979 and late 1985.

[11] As before, and consistent with the whole analysis, observed returns are calculated on a forward-looking basis. To illustrate, the observed return at the end of Dec/2001 is the return over the Dec/2001-Dec/2011 period.

forecasted 10-year annualized returns of the market between Dec/1959 and Dec/2001. The very close relationship between these two variables suggested by the figure is confirmed by a correlation of 0.88 (Rho-3) and a mean forecasting error of −1.42% (MFE) reported in Exhibit 1. This MFE indicates that, on average, the model slightly overestimates realized returns.

Figure 3: Observed and Forecasted Returns

This figure shows observed and forecasted 10-year annualized returns of the S&P-500 over the Dec/1959-Dec/2001 period, the latter estimated from (11).



These results clearly suggest that the ERPF model proposed here can be a useful tool to forecast the market's return. Interestingly, if the model's forecasting ability is evaluated over the shorter period beginning in 1986, on which Hassett (2010) seems to focus, the correlation between observed and forecasted returns increases to a remarkable 0.95, with an MFE of just −0.48%. The stronger relationship between these two variables over this shorter sample period is also obvious by a simple inspection of Figure 3.

3.2. Slight Variations of the Model

The model proposed explains the variability of the RPF as a function of risk in the equity market, the latter assessed by price relative to fundamentals and quantified by CAPE. Although as just discussed the data strongly supports this model, it is clear that other multiples could be used for the same purpose; two of the most widely used are P/E and D/P.

The second and third rows of Exhibit 1 show that neither of these multiples performs as well as CAPE when forecasting the RPF (hence the ERP and the market's return), although D/P comes rather close. Across the 505 regressions between RPF and D/P the average R^2 is 0.42; the out of sample RPFs forecasted from these regressions have a correlation of 0.72 with the observed RPFs. Furthermore, the ERPs forecasted from an expression similar to (10), obviously replacing CAPE by D/P, have correlation of 0.67 with the observed ERPs. Finally, the market returns forecasted from an expression similar to (11), again replacing CAPE by D/P, have a correlation of 0.84 with the observed returns. In short, D/P performs neither significantly worse than, nor as well as CAPE, when forecasting the RPF, the ERP, and the market's return.

As shown by (10)-(11), the model proposed forecasts the ERP and the market's return on the basis of a current risk-free rate. A potential shortcoming of this approach is that the risk-free rate may be unusually (and perhaps artificially) high or low at the moment of making a forecast of the market's return, which may be made for a subsequent long period such as 10 years.[12] In fact, it is no coincidence that the model's worst performance when forecasting the market's return is in times of double-digit bond yields.

One way to deal with this volatility is to smooth the risk-free rate by taking an average of the rates observed over the previous years. However, replacing the current risk-free rate in expressions (10)-(11) with a moving average over the previous 3, 5, or 10 years does not improve the model's forecasting ability in a substantial way. Over the Dec/1959-Dec/2001 period, the correlations between forecasted returns from (11) and observed returns are 0.89, 0.89, and 0.86 for 3-year, 5-year, and 10-year moving averages of the risk-free rate, in all cases very close to the correlation of 0.88 based on current risk-free rates. Hence, there seems to be no gain in forecasting ability from smoothing the risk-free rate.

3.3. An Unrestricted Model

The ERPF model proposed here, as well as the RPF model proposed by Hassett (2010), imposes a specific functional relationship between the market's expected return, the RPF, and the risk-free rate given by $R_1=(1+\text{RPF}_1)\cdot y_0$. For this reason, it is worth asking whether a model that relates these three variables without imposing a specific functional relationship improves upon the forecasting ability of the ERPF model.

To that purpose, the regression

$$R_{1t} = \gamma + \delta\cdot\text{CAPE}_{0t} + \theta\cdot y_{0t} + u_t \qquad (12)$$

was run 505 times with a rolling window of 120 observations, the first time over the Dec/1949-Dec/1959 period and the last one over the Dec/1991-Dec/2001 period, which exhausted the data available. The average sample estimates of γ, δ, and θ (0.24, −0.01, and −0.13) are reported in Exhibit 2, together with the average R^2 (0.69) and adjusted R^2 (0.68).[13] Thus, this model explains, on average, almost 70% of the variability in returns.

---

[12] To illustrate, in August and September, 1981, the yield on 10-year Treasury Notes was above 15%, but five years later the same yield was under 7.5%. Currently, the Fed is pursuing a policy of artificially low interest rates and the yield on 10-year Notes is under 2%.

[13] At the 5% level of significance, the intercept and the slopes of CAPE and y are significant in 100%, 90.7%, and 53.7% of the 505 regressions run.

Exhibit 2: Forecasting the Market's Return – Unrestricted Models
This exhibit shows the average intercept, slope, R^2, and adjusted R^2 ($AdjR^2$) across 505 regressions between the market's observed 10-year annualized return and two explanatory variables, CAPE and the 10-year yield on Treasury Notes (y). The equity market is represented by the S&P-500 index. CAPE is based on 10-year trailing earnings. The first regression is run over the Dec/1949-Dec/1959 period and the last one over the Dec/1991-Dec/2001 period. The exhibit also shows, over the Dec/1959-Dec/2001 period, the correlation between observed and forecasted 10-year annualized returns (Rho) and the mean forecasting error (MFE) calculated as the average difference between observed and forecasted annualized returns.

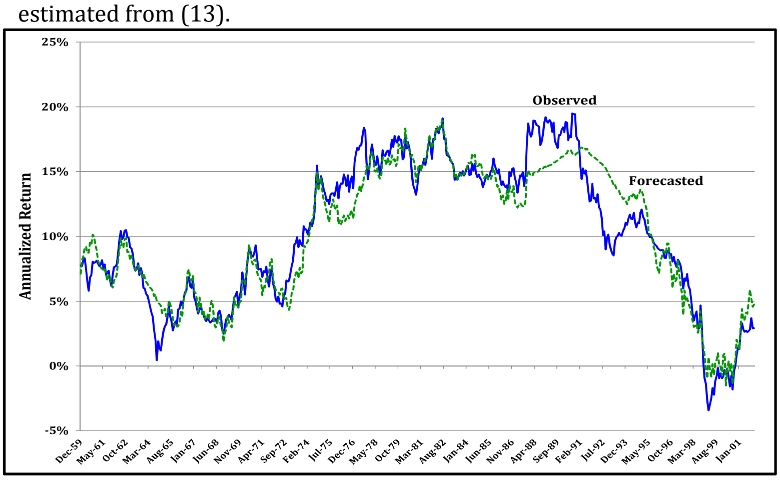
Variables	Intercept	CAPE	y	R^2	$AdjR^2$	Rho	MFE (%)
CAPE, y	0.24	−0.01	−0.13	0.69	0.68	0.95	0.07
CAPE	0.26	−0.01	N/A	0.64	0.64	0.94	0.04

Each of the regressions run was subsequently used to forecast an out of sample return with the expression

$$E(R_{1t}) = c + d \cdot \text{CAPE}_{0t} + e \cdot y_{0t}\,, \tag{13}$$

where c, d, and e are the sample estimates of γ, δ, and θ, which produced a time series of 505 returns over the Dec/1959-Dec/2001 period. Figure 4 plots the time series of observed and forecasted 10-year annualized returns and suggests a very close relationship between these two variables. Such close relationship is confirmed by a whopping correlation of 0.95 and a negligible mean forecasting error of 0.07%, both reported in Exhibit 2, and suggests that (13) can be a very accurate tool to forecast the market's return.

Figure 4: Observed and Forecasted Returns (II)
This figure shows observed and forecasted 10-year annualized returns of the S&P-500 over the Dec/1959-Dec/2001 period, the latter estimated from (13).



Because the risk-free rate turns out to be non-significant in many of the 505 regressions run, it makes sense to ask whether dropping it from (12)-(13) would imply a substantial loss in terms of the model's ability to forecast the market's return. Thus the model

$$R_{1t} = \lambda + \mu \cdot \text{CAPE}_{0t} + u_t \tag{14}$$

was run 505 times with a rolling window of 120 observations, the first time over the Dec/1949-Dec/1959 period and the last one over the Dec/1991-Dec/2001 period, which exhausted the data available. The average sample estimates of λ and μ (0.26 and −0.01) are reported in Exhibit 2, together with the average R^2 (0.64) and adjusted R^2 (0.64).[14] These last two figures suggest that the explanatory power of (14) is only marginally lower than that of (12). In other words, there does not seem to be a substantial loss of explanatory power from dropping the risk-free rate from the model.

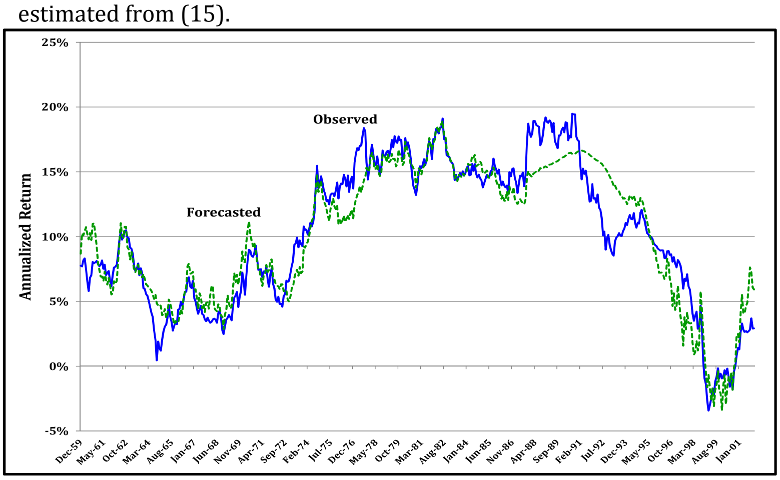
More importantly, if each of the 505 regressions from (14) is used to forecast an out of sample return over the Dec/1959-Dec/2001 period with the expression

$$E(R_{1t}) = f + g \cdot \text{CAPE}_{0t}\,, \tag{15}$$

where f and g are the sample estimates of λ and μ, the forecasts produced are just as accurate as those from (13). Figure 5 plots the time series of observed and forecasted 10-year annualized returns from (15) and suggests a very close relationship between these two variables. Such close relationship is confirmed by a correlation of 0.94 and a mean forecasting error of 0.04%, both reported in Exhibit 2. Hence, these results suggest that (15) is a more parsimonious model than (13) and with essentially the same ability to forecast the market's return.

Figure 5: Observed and Forecasted Returns (III)

This figure shows observed and forecasted 10-year annualized returns of the S&P-500 over the Dec/1959-Dec/2001 period, the latter estimated from (15).



3.4. Which Model Then?

Figures 3-5 plot the observed 10-year annualized returns of the market over the Dec/1959-Dec/2001 period, each with a different set of forecasts of those returns. Exhibit 3 summarizes some relevant information about the relationship between the observed and the forecasted returns in these three figures.

[14] At the 5% level of significance, a and b are significant in 100% and 90.3% of the 505 regressions run.

Exhibit 3: Forecasting the Market's Return – Summary

This exhibit shows the expressions used to produce the forecasted returns in Figures 3-5; the correlation between observed and forecasted 10-year annualized returns (Rho); and the mean forecasting error (MFE) calculated as the average difference between observed and forecasted annualized returns. The sample period stars in Dec/59 and ends in Dec/2001. The equity market is represented by the S&P-500 index. CAPE is based on 10-year trailing earnings and the risk-free rate (y) is approximated with the yield on 10-year Treasury Notes. The exhibit also shows the p-value of two null hypotheses, the intercept equal to 0 and the slope equal to 1, in both cases in regressions between observed and forecasted annualized returns.

Figure	Forecasting Expression	Rho	MFE (%)	Intercept=0?	Slope=1?
3	(11): $E(R_{1t}) = (1+a+b\cdot CAPE_{0t})\cdot y_{0t}$	0.88	−1.42	0.00	0.00
4	(13): $E(R_{1t}) = c+d\cdot CAPE_{0t}+e\cdot y_{0t}$	0.95	0.07	0.22	0.15
5	(15): $E(R_{1t}) = f+g\cdot CAPE_{0t}$	0.94	0.04	0.78	0.64

As the exhibit shows, and as already discussed, the correlation between observed and forecasted returns is higher when the forecasts are produced from (13) and (15) than when they are produced from (11). Hence, the unrestricted forecasts based on CAPE and the risk-free rate have a slight edge in terms of forecasting ability. That being said, the correlation of 0.88 produced by the ERPF model is obviously high enough to give the model its due credit.

Exhibit 3 also shows that expressions (13) and (15) produce forecasts that have not only a higher correlation but also a lower mean forecasting error than those produced from (11). In fact, although the mean forecasting error from (13) and (15) is not statistically different from 0, that from (11) is clearly negative (in all cases at the 5% level of significance).

In a regression between observed and forecasted returns, unbiased forecast should produce an intercept equal to 0 and a slope equal to 1. The last two columns of Exhibit 3 show the p-values on these two null hypotheses and show that although both hypotheses are rejected when the forecasts are produced from (11), they are not rejected when the forecasts are produced from (13) and (15). In other words, these last two models produce unbiased forecasts of the market's return.

Finally, note that the same approach discussed to forecast the market's return can be used to forecast the ERP. Expression (10) of the ERPF model can be used straightforwardly for this purpose, and so can the expected returns from expressions (13) and (15) net of the observed risk-free rate. The correlation between observed and forecasted ERPs based on expressions (10), (13), and (15) over the Dec/1959-Dec/2001 period are 0.76, 0.86, and 0.86, with mean forecasting errors of −1.91%, −0.42%, and −0.45%. Thus, all three models, and particularly the last two, seem to provide accurate forecasts of the ERP.

4. Assessment

Forecasting the market's return and the ERP are two of the most debated and controversial issues in finance. There is little agreement about whether the ERP is constant or time varying, how to estimate it, how to forecast the market's return, and whether returns are predictable at all. This article tackled, and hopefully shed some light on, these issues.

Hassett (2010) recently proposed a novel approach, the Risk Premium Factor (RPF) model, to estimate the ERP as a function of the RPF and the risk-free rate. The main limitation of his approach is that the RPF, which is at the heart of his model, is estimated from the same sample and with the same data used to forecast the market's level. Hence, out of sample forecasts of the market's level can only assume or hope that the estimated RPF will remain constant in the future. Furthermore, his approach aims to estimate the market's level, and therefore its expected capital gain/loss, thus ignoring dividends which are a sizeable proportion of the market's expected return.

The Enhanced Risk Premium Factor (ERPF) model proposed here overcomes the main limitation of the RPF model by proposing a two-step approach to forecast the market's return. In the first step, the RPF is forecasted as a function of risk in the equity market, measured by price relative to fundamentals and quantified by CAPE. In the second step, the forecasted RPF and a current risk-free rate are used to forecast the market's return. The evidence discussed shows that the returns forecasted from this model are very highly correlated with observed returns.

Furthermore, a model that relates the market's expected return to the expected RPF (estimated on the basis of the current CAPE) and the current risk-free rate, but that does not impose a specific functional relationship among these variables, outperforms the ERPF model. In fact, such unrestricted approach produces unbiased forecasts of the market's return whose correlation with observed returns are in excess of 0.9.

To summarize, Hassett (2010) has proposed a novel approach to think about and estimate the ERP. The main weakness of his approach is that the variable at the heart of his model, the RPF, cannot be forecasted out of sample because the factors that explain its variability over time are not identified. The approach proposed here overcomes this limitation by proposing a model to both explain and forecast the RPF. Forecasts of the market's return produced by this model are very highly correlated with observed returns, and forecasts that articulate the same variables without imposing a specific functional relationship on them produce even better results.

References

Campbell, John, and Robert Shiller (1998). "Valuation Ratios and the Long-Run Stock Market Outlook." Journal of Portfolio Management, Winter, 11-26.

Darolles, Serge, Karl Eychenne, and Stephane Martinetti (2010). "Time Varying Risk Premiums and Business Cycles: A Survey." Lyxor white paper, September, issue 4.

Damodaran, Aswath (2012). "Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2012 Edition." Stern School of Business, Working Paper.

Graham, Benjamin, and David Dodd (1934). *Security Analysis*. McGraw-Hill.

Hassett, Stephen (2010). "The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables." Journal of Applied Corporate Finance, Spring, 118-130.

Ilmanen, Antti (2011). *Expected Returns – An Investor's Guide to Harvesting Market Rewards*. Wiley.

Kahneman, Daniel, and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." Econometrica, 47, 263-292.

Lintner, John (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." Review of Economic and Statistics, 47, 13-37.

Lucas, Robert (1978). "Asset Prices in an Exchange Economy." Econometrica, 46, 1429-1445.

Markowitz, Harry (1952). "Portfolio Selection." Journal of Finance, 7, 77-91.

Marks, Howard (2011). *The Most Important Thing – Uncommon Sense for the Thoughtful Investor*." Columbia Business School Publishing.

Mehra, Rajnish, and Edward Prescott (1985). "The Equity Risk Premium: A Puzzle." Journal of Monetary Economics, 15, 145-161.

Mossin, Jan (1966). "Equilibrium in a Capital Asset Market." Econometrica, 34, 768-783.

Richards, Carl (2012). *The Behavior Gap – Simple Ways To Stop Doing Dumb Things With Money*. Portfolio Penguin.

Rubinstein, Mark (1976). "The Valuation of Uncertain Income Streams and the Pricing of Options." Bell Journal of Economics, 7, 407-425.

Samuelson, Paul (1965). "Proof That Properly Anticipated Prices Fluctuate Randomly." Industrial Management Review, 6, 41-49.

Sharpe, William (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk." Journal of Finance, 19, 425-442.

Tversky, Amos, and Daniel Kahneman (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty." Journal of Risk and Uncertainty, 5, 297-323.