Predictive Regressions of Dividend Growth and Returns

Dooruj Rambaccussing†
University of Exeter

May 7, 2010

Abstract

This paper verifies predictability of returns and dividend growth using univariate and multivariate models where returns and dividend growth are predicted by the lags of their expected values. The expected returns and expected dividend growth are derived in a time varying framework where both are modeled by an autoregressive process of order 1. The two series are decomposed from a net present value approach involving dividends. Our results show that returns and growth are very weakly predictable on their expected components. The key conclusion this paper draws in terms of empirics is that while model selection is important in the predictive regressions, much attention needs to be devoted to the correct specification of the expected returns and expected dividend growth rate values.

Key Words: Dividend Growth, Returns, State Space Modeling, Predictive Regression

JEL References: G12, G14, G17

*Preliminary and incomplete. Please do not quote without consulting the authors first.
†Corresponding author. Please send any information concerning the conference to me. Address: Room 1.69, Department of Economics, Business School, University of Exeter, Streatham Court, Renness Drive, Exeter EX4 4PU; e-mail dr244@ex.ac.uk.
1 Introduction

The predictability of returns literature has highlighted potential ratios that could possibly explain and forecast returns, namely the price dividend ratio, price earnings ratio, consumption and wealth ratio and various other factors. However, an informal opinion of the best predictor of actual returns could possibly be the lagged expected returns in itself, since expected returns might result in a self fulfilling prophecy. The actual returns on the index turns out to be the returns which agents actually expect. In the following paper, we filter out expected returns and expected dividend growth rate from a net present value structural model where both latent series are assumed to follow an autoregressive process of order one. We look at both the insample and out of sample predictability of both series using a univariate and then a multivariate setting where the actual values are determined by their expected counterparts.

I employ a structural approach for deriving expected returns and dividend growth rate. The expected returns and expected dividend growth series are filtered from realized observations based on the Kalman procedure, where the values for the expected variable are updated as a new observation of the realized value is entered in the information set. The most common way to derive the values for these variables is to use a latent variables approach. I derive the law of motion for the price dividend ratio assuming that the expected returns and the dividend growth rate follow an autoregressive process. The state space model is derived from the net present value relationship between Price Dividend ratio, expected returns and expected dividend growth. The Kalman Filter is applied to the model parameters which are optimized using the Maximum Likelihood procedure. Both components are estimated using annual data from 1900 to 2008.

A theoretical assumption of the model is that expected returns and expected dividend growth rate is best specified by a stationary autoregressive process of order one. This assumption is fundamental and stems from the data generating process literature. Any misspecification of the expected variable series will lead to weak predictability if the markets are efficient. Assuming stationarity enables an infinite order representation for both series, implying the shocks may have a slow or quick decay depending on the estimate of the autoregressive parameter. State space modeling provides an efficient estimation of the time varying unobservable components. Moreover, it is quite robust to structural breaks and does not require estimation of a large number of parameters unlike other models of deriving expected returns. The expected returns and expected dividend growth rate being unobservable, the Kalman filter is applied to extract unobserved components from the observed history of realized returns and dividend growth. The build of the state space model requires the use of state equations and measurement equations. In this case, the price dividend ratio acts as a measurement variable, since it is an observable variable.
The second part of the paper looks at the predictive ability of the models. We regress actual returns and expected returns on the past filtered observations of expected returns and expected dividend growth. We consider first univariate models where the actual returns and realized dividend growth are regressed on the past values of the lagged expected returns and dividend growth rates. We assess the in-sample predictability based on the goodness of fit measure and the out of sample accuracy based on a recursive measure of the mean squared error. We also look at forecasting from a vector autoregression. It is likely that agents actual returns and dividend growth are the cause and effect of expected returns. In this setting, a VAR model is constructed using both actual and expected series for the variable of interest.

2 Literature review

It has been widely surveyed that many factor can impact on the predictability of the expected returns and dividend growth. There are various financial ratios that have been deemed to predict actual returns in sample, namely earnings to price ratio, book to market ratio, or even macroeconomic variables such as the consumption to wealth ratio, labour to consumption ratio, housing collateral ratio, or even the cross section of risk. For an interesting literature see Lamont (1998), Baker and Wurgler (2000), Lettau and Ludvigson (2001), Menzly, Santos, and Veronesi (2004), Lustig and Van Nieuwerburgh (2005), Piazzesi, Schneider and Tuzel (2006), Polk, Thompson and Vuolteenaho (2006). Moreover, there are also interesting literature concentrating on the econometric properties of predictive regression models. This involves issues such as time varying predictability, out of sample forecasts, structural breaks and endogeneity of predictors. Interesting papers in this field involves, Bossaerts and Hillion (1999), Stambaugh (1999), Campbell and Yogo (2002), Ferson, Sarkisson and Simin (2003), Valkanov (2003), Goyal and Welch (2003 and 2006), Inoue and Killian (2004) Lewellen (2004), Ahmihud and Hurvich (2004) and Ang and Bekaert (2006).

Our paper focusses on predictability through a pass through model with the Price Dividend ratio as the key measurement variable that. The Price Dividend ratio inherits the interesting property of being in line with the net present value framework assuming rational expectations. Papers discussing predictability of returns and dividend growth using this approach include Fama and French (1988), Campbell and Shiller (1991), Timmermann (1993, 1996), Lewellen (2004), Cochrane (2008), Koijen and Van Binsbergen (2010). The interesting feature of the Price Dividend ratio as a predictor of returns lies in the net present value approach. It can be shown from the Campbell and Shiller (1988) log linearised form that as long as the expected returns and dividend growth process are stationary, deviations of the price dividend ratio from its mean might either predict returns or dividend growth.
Our model uses the fact that the net present value of the dividend price ratio consists of expected dividend growth and expected returns. Our model is inspired by the state space modeling framework of Koijen and Van Binsbergen (2010). In order to reach to this model, we shall model the latent variables (expected dividend growth rate and expected returns) as an autoregressive process. We shall then derive the implied dynamics of the Price Dividend ratio. We explain the methodology for deriving the expected returns and dividend growth rate in the next part of this paper and also in the appendix.

3 Methodology

3.1 Net Present Value Model

In this section, we derive the net present value relationship between the Price Dividend Ratio and expected returns and expected dividend growth. Interestingly the series is developed from a theoretical assumption that both expected returns and dividend growth rate follows an autoregressive process of order 1. We start by defining some standard equations in the literature and then we derive the Campbell and Shiller (1988) log linearized model:

The rate of return is defined as

$$ r_t = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) $$

The Price Dividend ratio is defined as

$$ PD_t = \frac{P_t}{D_t} $$

The Dividend Growth rate is defined as

$$ \Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right) $$

One of the important assumptions that we put forward for the process of expected returns and dividend growth concerns the order of the process. The intuitive idea concerning the functional form of the process is that there should be in theory near to the data generating process. However, this endeavour of finding a best model is hectic and involves a lot of data mining. We shall assume our own functional form of the model. The mean adjusted conditional expected returns and dividend growth rate are modelled as an autoregressive process as in equations 4 and 5 respectively:

$$ \mu_{t+1} - \delta_0 = \delta_1 (\mu_t - \delta_0) + \varepsilon_t^{\mu} $$

$$ \delta_0 $$
\[ g_{t+1} - \gamma_0 = \gamma_1 (g_t - \gamma_0) + \varepsilon_{t+1}^g \]

where \( \mu_t = E_t(r_{t+1}) \) and \( g_t = E_t(g_{t+1}) \)

Equation 4 and 5 relates to the mean deviation of the expected returns and expected dividend growth rate where \( \delta_0 \) and \( \gamma_0 \) represents the unconditional mean of the expected returns and dividend growth respectively. \( \delta_1 \) and \( \gamma_1 \) represent the autoregressive parameters. \( \varepsilon_{t+1}^r \) and \( \varepsilon_{t+1}^g \) represents the shocks to the expected returns and the dividend growth rate processes. \( \varepsilon_{t+1}^r \sim N(0, \sigma_r^2) \) and \( \varepsilon_{t+1}^g \sim N(0, \sigma_g^2) \). However, we do not implement any restrictions between the covariance of \( \varepsilon_{t+1}^r \) and \( \varepsilon_{t+1}^g \) because a shock to the expected return process might actually affect the dividend growth process as well.

The realized dividend growth rate are defined as the expected dividend growth rate and expected returns and the unobserved shock \( \varepsilon_{t+1}^d \), where by:

\[ \Delta d_{t+1} = g_t + \varepsilon_{t+1}^d \]

\( \varepsilon_{t+1}^d \) and \( g_t \) are assumed to be orthogonal to each other. \( E(\varepsilon_{t+1}^d, g_t) = 0 \).

The Campbell and Shiller (1988) log linearized return equation (derived in appendix 1) may be written as:

\[ r_{t+1} = \kappa + p d_{t+1} + \Delta d_{t+1} - p d_t \]

where \( p d_t = E[\log(PD_t)] \), \( \kappa \) is an arbitrary constant defined as \( log(1 + \exp(pd)) - \rho pd \) and \( \rho = \frac{\exp(pd)}{1 + \exp(pd)} \).

The equation can be further be reduced to:

\[ r_{t+1} = \kappa + p d_{t+1} + \Delta d_{t+1} - p d_t \]

To study the dynamics of the price dividend ratio, the process may be written with \( p d_t \) being the subject of the formula:

\[ p d_t = \kappa + p d_{t+1} + \Delta d_{t+1} - r_{t+1} \]

The full derivation of the of the price dividend ratio model is explained in appendix 1.

By replacing lagged iterated values of \( p d_{t+1} \) in the equation, the process may be written as:
\[
\begin{align*}
    \rho^t i = \sum_{i=0}^{\infty} \rho^i \kappa + \rho^\infty \rho d i + \sum_{i=1}^{\infty} \rho^{i-1} (\Delta d_{t+i} - r_{t+i}) \\
    \rho d t = \frac{\kappa}{1 - \rho} + \rho^\infty \rho d i + \sum_{i=1}^{\infty} \rho^{i-1} (\Delta d_{t+i} - r_{t+i})
\end{align*}
\]

3.2 State Space Model

The state space model makes use of a transition equation and a measurement equation. The Kalman Filter best illustrates the dynamics of the estimates of \( \mu_t \) and \( \gamma_t \). The model parameters are estimated before making the forecasts. The maximum likelihood estimator is used to obtain the parameters of the Kalman filter. The Maximum likelihood is optimized using the MaxBFGS procedure.

There are two transition equations, one governing the dividend growth rate and the other one governing the mean return:

\[
\begin{align*}
    \gamma_{t+1} &= \gamma_1 \gamma_t + \epsilon_{t+1}^\gamma \\
    \mu_{t+1} &= \delta_1 \mu_t + \epsilon_{t+1}^\mu
\end{align*}
\]

the two measurement equations are given by:

\[
\begin{align*}
    \Delta d_{t+1} &= \gamma_0 + \gamma_t + \epsilon_{t+1}^d \\
    pd_t &= A - B \mu_t + B \gamma_t
\end{align*}
\]

Equation 10 can be rearranged into 12 such that there are only two measurement equations and only one state space model.

\[
\begin{align*}
    \gamma_{t+1} &= \gamma_1 \gamma_t + \epsilon_{t+1}^\gamma \\
    \Delta d_{t+1} &= \gamma_0 + \gamma_t + \epsilon_{t+1}^d
\end{align*}
\]

\[
\begin{align*}
    pd_{t+1} &= (1 - \delta_1) A - B_2 (\gamma_1 - \delta_1) \gamma_t + \delta_1 pd_t - B_1 \epsilon_{t+1}^\mu + B_2 \epsilon_{t+1}^\gamma
\end{align*}
\]

Equation 13 defines the transition (state) equation. The measurement equation relates the observable variable to the unobserved variables. In our case this is given by equation 14 and 15. \( A \) is equal to \( \frac{\alpha}{1 - \rho} + \frac{\gamma_0 - \delta_0}{1 - \rho} \), \( B_1 = \frac{1}{1 - \rho \gamma}, B_2 = \frac{1}{1 - \rho \gamma} \).

Equation 14 and 15 relate to the measurement equation. This can be put into a state space form as shown in appendix. Since all the equations are linear, we
can implement the Kalman Filter and obtain the likelihood which is maximized over the following vector of parameters.

\[ \Theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_D, \rho_{g\mu}, \rho_{gD}, \rho_{\mu D}) \]

The individual elements of the state and measurement vectors are given in appendix 2.

The filtered series for the expected dividend growth is just taken to be the first element for the state vector \( X_t \) (Refer to appendix 2 for a more detailed explanation). The state vector is derived according to the following update:

\[ X_{t|t-1} = FX_{t-1|t-1} \] (16)

In the case of the demeaned expected returns, the expected returns is defined as:

\[ \hat{\mu}_{t-1|t-1} = B_1^{-1}(pd_t - A - B_2\hat{g}_{t-1|t-1}) \] (17)

### 3.3 Predictive Accuracy

In this section we assess the predictive accuracy of the different forecasting models. We assess the forecasting accuracy and possibly the underlying behavior of the expected dividend growth rate and returns. In the first case, we look at the expected returns as a function of past lags. We look at both a univariate and multivariate setting. In the univariate setting, we just posit lag orders of the expected returns. The predictability concept, as in the forecasting literature is built around the concept of that a variable is predictable from other factors based on the information set available at that time. In this setting we regress actual returns and dividend growth on their past lags Equations 18 and 19 illustrate the model we are dealing with.

\[ r_t = \beta_0 + \sum_{i=1}^{p} \beta_i \mu_{t-i} \] (18)

\[ \Delta d_t = \beta_0 + \sum_{i=1}^{p} \beta_i g_{t-i} \] (19)

It is worth noting that for the econometric modeling both \( g_{t-i} \) and \( \mu_{t-i} \) are defined as \( \mu_{t-i|t-i} \) and \( g_{t-i|t-i} \) from the state space framework. In the above framework, actual returns is determined by the lags of expected returns. The
optimal lag $p$ may be chosen by any of the model selection criteria. In our case we make use of the Schwartz Information criteria. However, as well as that, we try different combinations of regressors in the model. We initially set $p = 5$ and choose among 64 competing models ($2^5$). we analyze the insample predictability of the best model.

The expected and realized variable of interest may determine each other over different time lags. A bivariate vector autoregression is put forward to account for such a possibility:

$$Y_t = c + \sum_{i=1}^{p} A_{t-i} Y_{t-i}$$  \hspace{1cm} (20)

we use this generic form for both returns and dividend growth rate. In the case of returns, $Y_t = [r_t, \mu_t]$. $A$ is the matrix of coefficients and for a particular lag, it is a $2 \times 2$ matrix. $p$ is usually the VAR order and in our case, we estimated the model for $p = 1, 2$ and $3$. 

4 Results

We first report the results of the optimization of the state space model defined by equations 13,14 and 15. The model was estimated for the sample 1899 to 2008. We shall discard the first observation when retrieving the expected dividend growth and the expected returns.

Table 1: Optimisation of state Space model: Sample 1900-2008

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>0.0138</td>
<td>0.0108</td>
</tr>
<tr>
<td>$\delta_0$</td>
<td>0.0524</td>
<td>0.016</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.0717</td>
<td>0.1996</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.9459</td>
<td>0.0401</td>
</tr>
<tr>
<td>$\sigma_g$</td>
<td>0.0926</td>
<td>0.058</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>0.0139</td>
<td>0.0064</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.0688</td>
<td>0.0772</td>
</tr>
<tr>
<td>$\rho_{\mu}$</td>
<td>0.4802</td>
<td>0.4706</td>
</tr>
<tr>
<td>$\rho_{\mu D}$</td>
<td>-0.3815</td>
<td>1.3919</td>
</tr>
</tbody>
</table>

The results show that the unconditional mean returns turns out to be 5.2 % where as the annual dividend growth is somewhat less namely 1.4 %. the standard errors are relatively small in both cases. Interestingly, when we look at the autoregressive parameters, we find that roughly the expected dividend growth has a low autoregressive parameter implying, that expected dividend growth rate in itself is not cannot be predicted from past lags. However, the interesting result comes from the autoregressive coefficient of the expected returns which tends to depict high persistence. It is worth mentioning that this contrasts the findings on the actual market. Actual returns are highly nonpersistent where as expected returns are higly persistent.

4.1 Expected Returns and Expected Dividend Growth

In this section, we discuss the statistical features of the derived expected returns and dividend growth. The summary statistics for the series are given in table 4 in the appendix. Stationarity and long memory tests are also reported in table 5.

It is interesting to note that the mean expected returns series and expected dividend growth rate are lower that than their realized figures. However, the mean variance ratio for the expected returns tends to be higher than that of
the actual returns. In both cases, the filtered series appear smoother and less varying than the actual returns. It can be easily confirmed from the figure D and D. Interestingly the distribution properties between the realized and expected series tend to be different in terms of variation at the tails (Figures D D, D and D). In both series, the realized series exhibit a negative skewness while the expected returns displays a positive skew. Interestingly the kurtosis also provides interesting conclusions that tend to adhere to the property that agents have smooth expected returns and dividend growth. Both expected series show high skewness implying that the variance is explained by extreme deviations.

The stationarity tests show results consistent with theory. The null hypothesis of stationarity is not rejected in most cases. However we find that the stationarity results for the expected returns in terms of both the KPSS and the Robinson-Lobato test tend to be weak. The respective p-values are 0.414 and 0.1 respectively. The non stationarity test results show that there is high evidence for stationarity in the actual returns, actual and expected dividend growth. However the expected returns is purely non stationary according to the various tests. It further explains the findings of the previous I(0) tests of the KPSS and Robinson-Lobato test. Although the test show that there are non stationary models involved, we make out that it is a near unit root process as show b by the autoregressive coefficient in table 5.

Another interesting statistic that we report in tables 6 and 7 are the correlations with past values over five years between the expected and realized series. The correlation of the returns with its own past lags are quite low, although we witness a stronger negative correlation with two years lags, implying mean reversion. In the case of expected returns, we witness strong positive correlation, as expected given that the series is persistent. surprisingly, actual returns and past lags of expected returns are negatively correlated. We should expect some positive correlation at longer lags since agents behaviors will make the expected returns equal to the actual returns. However, the most interesting feature is understood from the correlation between expected returns and past lags of actual data. Accordingly, we a direct negative contemporaneous relationship between the two series. However, we find only positive relationships between the past actual returns. In other words, actual returns for the year 2, 3, 4 and 5 are important in forming actual returns. The above results points out two possible areas of attention. First of all, it points out that agents may be forming the expected returns figure based on past actual values. Indirectly, there is the learning mechanism at work, where agents update their own information about the prior belief of expected returns using actual returns. Secondly, the results suggest an evidence of mean reversion for expected returns itself.
4.2 Predictive accuracy

In this section, we present the results on predictive accuracy of a univariate and multivariate econometric model where actual returns and expected returns are determined by the corresponding expected values. Firstly, we report the R-square of the best specified (with respect to the lag order p) of equations 18 and 19.

Our findings for the insample predictive accuracy model shows that the more lags that are added to the predictive regression, the lower the adjusted R-squared tends to get over time. Figure D shows this feature. As we increase the number of lags in the model, the predictive accuracy tends to decrease. We also optimized the model according to the Schwartz Information criteria. The interesting phenomenon is that as we increase the number of lags in the phenomenon, the best model turns out to include a single regressor which is the last lag of the model. Using 5 lags the best model turns out to be lagged expected returns 5 year before with an intercept term. Both coefficients appear to be statistically significant. The results are reported in table 8. The R-squared is 0.03, which tends to explain 3 % of the variation.

In the case of predictive ability of dividend growth, a high $R^2$ is witnessed at 0.20. Interestingly the Schwartz information criterion finds one and two lags of expected returns are reasonable predictors of dividend growth (table 9). All the regressors with the exception of the intercept term are constant. We also consider adding up to 5 lags of expected dividend growth in the model. Unlike the case of expected returns, the adjusted R-squared increases more or less when the number of regressors is increased (Figure D). However, we notice that the model may contain some nonsignificant variables as we keep on adding them.

4.3 Out of Sample Forecast

In the case of the out of sample forecast, we need to go through the model selection process again in sample and then chose the best model for forecasting. In the case of the actual returns, only the coefficient with lag 5 was deemed to be sufficient. Dividend growth in sample was better predicted with the first and fourth lag of expected dividend growth. However, we need to note that the science of using the best model insample to forecast multiple step ahead is quite arbitrary when we are dealing with a small size such as ours. Hence, we also report the mean squared error for the different specifications for $p = \{1,2,3,4,5\}$. The average mean squared error was for the different horizons in figures D and D.

The different lags of the expected returns tend to have the same trend in the mean squared error. The model with $p =1$ and 2 tends to perform better over
time. As expected for the year 2008 (a structural break) the model tends to perform much worse, exemplified through a large increase in the mean squared error. In the case of the dividend growth, 2008 does not seem to be a bad period for forecasting. Indeed for that particular period, the mean squared error actually fell. Very interesting trends can be witnessed. The best model in the initial years (expected dividend growth with lags 1 and 4) tend to perform very badly afterwards. On the other hand the worst model tend to perform the best. A high level of uncertainty in terms of the best model can be witnessed.

4.4 VAR models

In this section, we report the results from the vector autoregression. First we shall comment on the model selection and predictive ability of the models estimated. Among the three different lags, we find that predictability is best achieved by looking at a model with 2 lags. The Adjusted R-squared is 0.04, which is quite average in the predictability literature. The number of lags postulated by the AIC and SBC shows that the optimal number of lags should be three. However it can be contested on the statistical significance of the regressors. Most of the regressors in the VAR(3) are statistically insignificant. Very surprisingly, we do not find any statistically evident relationship driving expected returns to actual returns. However, in the VAR(3), we do find that the opposite does happen, in the sense that expected returns are driven by actual returns. The results are reported in table 2.
Table 2: Results from VAR model with realized and expected returns: Sample 1900-2008

<table>
<thead>
<tr>
<th></th>
<th>P =1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R_t$</td>
<td>$\mu_t$</td>
<td>$R_t$</td>
<td>$\mu_t$</td>
<td>$R_t$</td>
<td>$\mu_t$</td>
</tr>
<tr>
<td>C</td>
<td>0.057**</td>
<td>0.027**</td>
<td>0.110**</td>
<td>0.005</td>
<td>0.099**</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.006)</td>
<td>(0.030)</td>
<td>(0.002)</td>
<td>(0.032)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$R_{t-1}$</td>
<td>0.070</td>
<td>-0.039</td>
<td>0.059</td>
<td>-0.016</td>
<td>0.078</td>
<td>-0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(0.026)</td>
<td>(0.099)</td>
<td>(0.009)</td>
<td>(0.103)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>$\mu_{t-1}$</td>
<td>-0.153</td>
<td>0.441**</td>
<td>-0.365</td>
<td>0.700**</td>
<td>-0.345</td>
<td>0.881**</td>
</tr>
<tr>
<td></td>
<td>(0.336)</td>
<td>(0.086)</td>
<td>(0.928)</td>
<td>(0.088)</td>
<td>(1.051)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>$R_{t-2}$</td>
<td>-0.219*</td>
<td>-0.005</td>
<td>-0.217*</td>
<td>-0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.009)</td>
<td>(0.102)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{t-2}$</td>
<td>-0.625</td>
<td>0.226**</td>
<td>-1.050</td>
<td>0.073</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.937)</td>
<td>(0.089)</td>
<td>(1.194)</td>
<td>(0.065)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_{t-3}$</td>
<td></td>
<td></td>
<td>0.071</td>
<td>0.0681**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.103)</td>
<td>(0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{t-3}$</td>
<td></td>
<td></td>
<td>0.535</td>
<td>0.0355</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.981)</td>
<td>(0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj R- Squared</td>
<td>-0.012</td>
<td>0.199</td>
<td>0.043</td>
<td>0.812</td>
<td>0.031</td>
<td>0.936</td>
</tr>
<tr>
<td>F-statistic</td>
<td>2.191</td>
<td>115.9</td>
<td>1.566</td>
<td>258.5</td>
<td>1.681</td>
<td>30.828</td>
</tr>
<tr>
<td>Akaike</td>
<td>-3.703</td>
<td>-3.811</td>
<td>-5.097</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwartz</td>
<td>-3.554</td>
<td>-3.561</td>
<td>-4.745</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** statistical significance at the 1 % level; * denotes significance at the 5 % level
In the case of the actual dividend growth, we find nearly the same evidence from returns (table 3. We find that the important factor driving both actual dividend growth and expected dividend growth are the lagged dividend growth. Lagged realized dividend growth impact on present expected dividend growth. The only instance when expected dividend growth impacts on actual dividend growth is when we use the VAR(3) process. In that case, implicitly an increase in the expected dividend growth rate will move the actual dividend growth rate in the opposite way. Interestingly in both cases the VAR models tend to outperform the univariate models.

Table 3: Results from VAR model with realized and expected dividend growth: Sample 1900-2008

<table>
<thead>
<tr>
<th></th>
<th>P =1</th>
<th>P =2</th>
<th>P =3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\Delta d_t)</td>
<td>(g_t)</td>
<td>(\Delta d_t)</td>
</tr>
<tr>
<td>C</td>
<td>0.013 (0.011)</td>
<td>0.001 (0.003)</td>
<td>0.010 (0.010)</td>
</tr>
<tr>
<td>(\Delta d_{t-1})</td>
<td>0.153 (0.095)</td>
<td>0.119*** (0.028)</td>
<td>0.235** (0.089)</td>
</tr>
<tr>
<td>(g_{t-1})</td>
<td>-0.264 (0.302)</td>
<td>0.075 (0.089)</td>
<td>0.043 (0.303)</td>
</tr>
<tr>
<td>(\Delta d_{t-2})</td>
<td>-0.258** (0.096)</td>
<td>0.007 (0.031)</td>
<td>-0.291** (0.099)</td>
</tr>
<tr>
<td>(g_{t-2})</td>
<td>0.234 (0.280)</td>
<td>0.015 (0.091)</td>
<td>0.293 (0.301)</td>
</tr>
<tr>
<td>(\Delta d_{t-3})</td>
<td>-0.010 (0.098)</td>
<td>0.0003 (0.001)</td>
<td></td>
</tr>
<tr>
<td>(g_{t-3})</td>
<td>-0.571* (0.279)</td>
<td>0.003 (0.003)</td>
<td></td>
</tr>
<tr>
<td>Adj R-Squared</td>
<td>0.012</td>
<td>0.134</td>
<td>0.080</td>
</tr>
<tr>
<td>F-statistic</td>
<td>1.710</td>
<td>9.390</td>
<td>3.352</td>
</tr>
<tr>
<td>Akaike</td>
<td>-5.395</td>
<td>-5.497</td>
<td>-12.25</td>
</tr>
<tr>
<td>Schwartz</td>
<td>-5.247</td>
<td>-5.249</td>
<td>-11.90</td>
</tr>
</tbody>
</table>

** statistical significance at the 1% level; * denotes significance at the 5% level

We also perform some out of sample dynamic forecasts from the dynamic VAR models and we report the Mean Squared Error. Figures D, D, D and D to illustrate the forecasts from the VAR model. The forecasts are unlikely to reproduce the current values of the actual returns. They do not follow the same direction. Moreover, they tend to be stable for some years where as the actual returns tend to fluctuate a lot. However, the forecasts for the expected
returns tend to reproduce the trend in the expected returns. We also plot the
average mean squared error over time. We find that the VAR(3) is quite a good
model for forecasting for some periods of time. For p = 1 and 2, they tend to
be very good for some periods and then tend to become worse in others. For
the expected returns, the VAR(1) and VAR(3) are better for 2004 and 2005
respectively. After 2005, the VAR(2) tend to be the best model for forecasting
expected returns.

In the case of the dividend growth, the VAR(2) and (3) tend to be very
good. The mean squared error tends to fall over time after 2005. In the case
of expected dividend growth, there tends to be a downgrade in the forecast
models between 2004 and 2006. Afterwards, there appears to be a smoother
improvement in expected dividend growth.

5 Conclusion

In the above paper, we attempt to find to model the relationship between ac-
tual returns, and expected returns and realized dividend growth and expected
dividend growth. The conclusions of this paper rests on some important as-
sumptions for deriving the expected returns and expected dividend growth rate.
Firstly, we model both processes in the context of the net present value model.
Secondly, we assume that the expected returns and dividend growth rate can be
modeled by the an autoregressive process of order one. The second assumption
is the reason for finding weak predictability in returns while dividend growth
exhibited stronger predictability. From our data, an AR(1) process fits actual
returns badly where as it fits better the realized dividend growth rate.

It appears natural to imagine that the data generating process should be
close, in terms of both fitting and predicting, to the actual realized values.
However, in our case, the realized values do not follow an autoregressive process.
The misspecification of expected returns and dividend growth rate series in itself
would imply that the Kalman filter updates the information set through an
improper channel.

Empirically, we found evidence of weak predictability in both the univariate
and multivariate cases. The VAR produced better forecasts both insample and
out of sample for the period 2001-2008. weak predictability was witnessed with
an adjusted R-squared of 0.04 at its maximum. This study provides further areas
of research in examining the relationship between actual returns and expected
returns, subject to chosing the dynamics of the expected returns series properly.
If the correct model, or a model that tends to approximate the data generating
process of both returns and dividend growth can be selected, it is worth trying
that model for the state space specification, and see how whether the time
varying perspective will actually better predict returns.
References


A     The Net Present Value Model.

Equations 1, 2 and 3 are represented

\[ r_t = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \quad (21) \]

\[ PD_t = \frac{P_t}{D_t} \quad (22) \]

\[ \Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right) \quad (23) \]

The return process can be written as

\[
\begin{align*}
    r_t &= \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \\
    &\quad \log\left(\frac{P_t}{D_t}\right) \quad (24) \\
    &\quad \log\left(\frac{D_{t+1}}{P_t}\right) + \exp(p_d) \\
    &\quad \exp(p_d) + \frac{\exp(p_d)}{1 + \exp(p_d)} + \Delta d_{t+1} - pd_t \quad (27)
\end{align*}
\]

Assuming the log linearisation of Campbell and Shiller (1988) the returns can be written as

\[
r_t \approx \log((1 + e^{pd_{t+1}})) + \frac{\exp(pd_t)}{1 + \exp(pd_t)} + \Delta d_{t+1} - pd_t
\]

\[ r_t = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t \]

Hence,

\[ pd_t = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1} \]
B State Space Model.

In this section, we describe the Kalman filter procedure. The model has been coded using Ox 5. From the paper, there are two measurement equation and one transition equation. Equations 13, 14 and 15 can be written in this form:

\[ X_t = FX_{t-1} + \varepsilon_t \]

\[ Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t \]

where \( Y_t = \begin{bmatrix} \nabla d_t \\ pd_t \end{bmatrix} \)

The variables of the transition equation are \( X_t \) and \( \varepsilon_{t+1} \) and are made up of the following elements:

\[ X_t = \begin{bmatrix} g_{t-1} \\ \varepsilon_{t}^D \\ \varepsilon_{t}^g \\ \varepsilon_{t}^\mu \\ \varepsilon_{t+1}^\mu \\ \varepsilon_{t+1}^d \\ \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^D \end{bmatrix} \]

\[ F = \begin{bmatrix} \gamma_1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ R = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

The parameters of the measurement equation include parameters of the net present value model to be estimated. These are defined as:

\[ M_0 = \begin{bmatrix} \gamma_0 \\ (1 - \delta_1) * A \end{bmatrix} \]

\[ M_1 = \begin{bmatrix} 0 & 0 \\ 0 & \delta_1 \end{bmatrix} \]

\[ M_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ B_2(\gamma_1 - \delta_1) & 0 & B_2 & -B_1 \end{bmatrix} \]

The variance covariance matrix from the state space model is given by:

\[ \Sigma = \text{var} \begin{bmatrix} \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^\mu \\ \varepsilon_{t+1}^d \\ \varepsilon_{t+1}^D \end{bmatrix} = \begin{bmatrix} \sigma_g^2 & \sigma_{g\mu} & \sigma_{gd} \\ \sigma_{g\mu} & \sigma_{\mu}^2 & \sigma_{D\mu} \\ \sigma_{gd} & \sigma_{D\mu} & \sigma_D^2 \end{bmatrix} \]

The Kalman Filter procedure is given by the following equations:

\[ X_{0|0} = E[X_0] \]

\[ P_{0|0} = E[X_0 X_0'] \]

\[ X_{t|t-1} = FX_{t-1|t-1} \]

\[ P_{t|t-1} = FP_{t-1|t-1} F' + R \Sigma R' \]

\[ \eta_t = Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1} \]

\[ S_t = M_2 P_{t|t-1} M_2' \]

\[ K_t = P_{t|t-1} M_2' S_t^{-1} \]

\[ X_{t|t} = X_{t|t-1} + K_t \eta_t \]

\[ P_{t|t} = (I - K_t M_2) P_{t|t-1} \]
The parameters to be optimized are:

$$\Theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_D, \rho_g, \rho_D, \rho_D)$$
C Results and Statistical Tests

<table>
<thead>
<tr>
<th></th>
<th>$R_t$</th>
<th>$\mu_t$</th>
<th>$\Delta d_t$</th>
<th>$g_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.057</td>
<td>0.047</td>
<td>0.013</td>
<td>0.003</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.185</td>
<td>0.040</td>
<td>0.114</td>
<td>0.036</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.612</td>
<td>1.084</td>
<td>-0.684</td>
<td>9.607</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.341</td>
<td>4.019</td>
<td>7.694</td>
<td>98.17</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>7.330</td>
<td>26.09</td>
<td>109.6</td>
<td>43206</td>
</tr>
</tbody>
</table>
Table 5: Tests of Stationarity on Expected Returns and Expected Dividend Growth: Sample 1900-2008

<table>
<thead>
<tr>
<th></th>
<th>$R_t$</th>
<th>$\mu_t$</th>
<th>$\Delta d_t$</th>
<th>$g_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stationarity test of I(0)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson-Lobato</td>
<td>-0.922{&lt;0.822}</td>
<td>0.216 {&lt;0.414}</td>
<td>-1.276 {&lt;0.899}</td>
<td>0.429 {&lt;0.334}</td>
</tr>
<tr>
<td>KPSS test</td>
<td>0.037{&lt;1}</td>
<td>0.427 {&lt;0.1}</td>
<td>0.035 {&lt;1}</td>
<td>0.300 {&lt;1}</td>
</tr>
<tr>
<td>Lo’s RS</td>
<td>0.869 {&lt;0.95}</td>
<td>1.012 {&lt;0.9}</td>
<td>0.782 {&lt;0.995}</td>
<td>1.039 {&lt;0.8}</td>
</tr>
<tr>
<td><strong>Stationarity test of I(1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Augmented Dickey-Fuller</td>
<td>-9.204 {&lt;0.01}</td>
<td>-2.071 {&lt;0.9}</td>
<td>-8.897 {&lt;0.01}</td>
<td>-9.633 {&lt;0.01}</td>
</tr>
<tr>
<td>Phillips-Perron</td>
<td>-9.300 {&lt;0.01}</td>
<td>-2.065 {&lt;0.9}</td>
<td>-8.923 {&lt;0.01}</td>
<td>-9.722 {&lt;0.01}</td>
</tr>
<tr>
<td>DF-GLS</td>
<td>-7.583 {&lt;0.01}</td>
<td>-2.091 {&lt;1}</td>
<td>-8.783 {&lt;0.01}</td>
<td>-9.676 {&lt;0.01}</td>
</tr>
<tr>
<td>P</td>
<td>1.186 {&lt;0.01}</td>
<td>4.538 {&lt;1}</td>
<td>1.263 {&lt;0.01}</td>
<td>1.874 {&lt;0.01}</td>
</tr>
</tbody>
</table>
Table 6: Correlation of Actual and Expected Returns: Sample 1900-2008

<table>
<thead>
<tr>
<th>Lag</th>
<th>$R_t$ vs $R_{t-j}$</th>
<th>$\mu_t$ vs $\mu_{t-j}$</th>
<th>$R_t$ vs $\mu_{t-j}$</th>
<th>$\mu_t$ vs $R_{t-j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-0.198</td>
<td>-0.198</td>
</tr>
<tr>
<td>1</td>
<td>0.069</td>
<td>0.898</td>
<td>-0.139</td>
<td>-0.161</td>
</tr>
<tr>
<td>2</td>
<td>-0.182</td>
<td>0.839</td>
<td>-0.164</td>
<td>0.147</td>
</tr>
<tr>
<td>3</td>
<td>0.077</td>
<td>0.802</td>
<td>-0.189</td>
<td>0.129</td>
</tr>
<tr>
<td>4</td>
<td>-0.069</td>
<td>0.749</td>
<td>-0.147</td>
<td>0.087</td>
</tr>
<tr>
<td>5</td>
<td>-0.100</td>
<td>0.687</td>
<td>-0.141</td>
<td>0.132</td>
</tr>
</tbody>
</table>

Table 7: Correlation of Actual and Expected Dividend Growth: Sample 1900-2008

<table>
<thead>
<tr>
<th>Lag</th>
<th>$\Delta d_t$ vs $\Delta d_{t-j}$</th>
<th>$g_t$ vs $g_{t-j}$</th>
<th>$\Delta d_t$ vs $g_{t-j}$</th>
<th>$g_t$ vs $\Delta d_{t-j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.380</td>
<td>0.380</td>
</tr>
<tr>
<td>1</td>
<td>0.071</td>
<td>0.155</td>
<td>0.107</td>
<td>-0.013</td>
</tr>
<tr>
<td>2</td>
<td>-0.015</td>
<td>-0.229</td>
<td>0.107</td>
<td>-0.087</td>
</tr>
<tr>
<td>3</td>
<td>-0.049</td>
<td>-0.079</td>
<td>-0.223</td>
<td>0.061</td>
</tr>
<tr>
<td>4</td>
<td>0.012</td>
<td>-0.066</td>
<td>-0.072</td>
<td>-0.140</td>
</tr>
<tr>
<td>5</td>
<td>-0.127</td>
<td>-0.015</td>
<td>-0.075</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Table 8: Best Model for Modeling Actual Returns InSample: Sample 1900-2008

<table>
<thead>
<tr>
<th>Regressors</th>
<th>Coefficient</th>
<th>std. error</th>
<th>pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.09948</td>
<td>0.03175</td>
<td>0.002</td>
</tr>
<tr>
<td>Expected returns (-5)</td>
<td>-1.02914</td>
<td>0.58937</td>
<td>0.084</td>
</tr>
<tr>
<td>Schwarz Criterion</td>
<td>25.17</td>
<td>R-Bar-Squared</td>
<td>0.0304</td>
</tr>
<tr>
<td>Hannan-Quinn Criterion</td>
<td>26.74</td>
<td>Sum of Squares</td>
<td>3.4322</td>
</tr>
<tr>
<td>Akaike Criterion</td>
<td>27.81</td>
<td>Residual SD</td>
<td>0.1834</td>
</tr>
</tbody>
</table>
Table 9: Best Model for modeling actual dividend growth: Sample 1900-2008

<table>
<thead>
<tr>
<th>Regressors</th>
<th>Coefficient</th>
<th>std. error</th>
<th>pval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.00012</td>
<td>0.00067</td>
<td>0.859</td>
</tr>
<tr>
<td>Expected dividend growth (-1)</td>
<td>0.46116</td>
<td>0.1622</td>
<td>0.005</td>
</tr>
<tr>
<td>Expected Dividend Growth(-2)</td>
<td>-0.3304</td>
<td>0.13917</td>
<td>0.014</td>
</tr>
<tr>
<td>Schwarz Criterion</td>
<td>369.785</td>
<td>R-Bar-Squared</td>
<td>0.2009</td>
</tr>
<tr>
<td>Hannan-Quinn Criterion</td>
<td>372.152</td>
<td>Sum of Squares</td>
<td>0.0047</td>
</tr>
<tr>
<td>Akaike Criterion</td>
<td>373.765</td>
<td>Residual SD</td>
<td>0.0068</td>
</tr>
</tbody>
</table>
D  Graphical Plots

Plot of Expected Returns and actual returns : 1900 -2008

Plot of Expected Dividend Growth rate and actual Dividend : 1900 -2008
Distribution of Actual Returns: 1900 - 2008

Distribution of Expected Returns: 1900 - 2008
Distribution of Expected Dividend Growth rate: 1900-2008

Plot of R-squared values over the different lags of the actual returns model
Plot of R-squared values over the different lags of the realized dividend growth model.
Plot of Average Mean Squared Error for out of sample models of realized returns.

Plot of Average Mean Squared Error for out of sample models of realized dividend growth.
VAR forecasts of Actual Returns:

Plot of Root Mean Squared Error for Out of Sample forecast of Returns from the different VAR models

Plot of Root Mean Squared Error for Out of Sample forecast of Expected Returns from the different VAR models
Plot of Root Mean Squared Error for Out of Sample forecast of Dividend Growth from the different VAR models

Plot of Root Mean Squared Error for Out of Sample forecast of Expected Dividend Growth from the different VAR models