

Modeling the Persistence in Expected Returns ^{*}

Dooruj Rambaccussing[†]
University of Exeter

May 9, 2011

Abstract

The major contribution of this paper is to explicitly model the persistence in the time series of expected returns. The series of expected returns is derived from a state space representation of the present value identity relating expected returns and expected dividend (earnings) growth to the observed price dividend (earnings) ratio. The state space model is adjusted in order to include the possibility of expected returns following an autoregressive fractional integrated (ARFIMA) process which captures the persistence of the process. The new ARFIMA model performs moderately compared to the simple autoregressive process, which may be due to the presence of different regimes and structural breaks. The expected returns series is used in three applications namely in predictive regressions, analysing the relationship between consumption and discount rates and also in a market timing strategy.

Keywords: Expected Returns, Persistence, Present Value, State Space Modeling

JEL classification: G12, G14

^{*}Paper prepared for the European Financial Management PhD conference 2011. The author wishes to thank John Cochrane, James Davidson, Joao Madeira and Adam Golinski for interesting comments and suggestions. The author is also very grateful to James Davidson for the provision of Time Series Modeling 4.31 which has been extensively used in this research.

[†]Address: Room 1.69, Department of Economics, Business School, University of Exeter, Streatham Court, Rennes Drive, Exeter EX4 4PU; e-mail dr244@ex.ac.uk.

1 Introduction

In a recent paper, Kojien and Van Binsbergen (2010), hence KVB, find high persistence and near unit root properties in the expected returns series for the S&P 500. In their setting, the expected returns is derived from a state space model where expected returns and the expected dividend growth (both unobservable series in real time) are filtered from the present value relation and observed dividend growth. They assume that expected returns follow an autoregressive process of order one (AR(1)). In this paper, I attempt to model explicitly the persistence in the expected returns by assuming that the series follow an autoregressive fractionally integrated process (ARFIMA (p,d,q)). The latter is a process that accounts for the possibility of a series having long memory or long range dependence. The reduced form of the model consists of two measurement equations and one state equation. The model, optimized from S&P 500 data, yields a series for expected returns and expected dividend (earnings) growth rate. Robustness checks are performed by optimizing over two samples in time. Tests of long memory and time variation are performed on the filtered series. As a by-product of the procedure, the expected returns and expected dividend (earnings) growth rate are used as predictors for realized returns and observed dividend (earnings) growth rate. The relationship between consumption and expected returns is also investigated. Finally, the expected returns and expected (earnings) growth rate is used in the present value formulae to test whether a successful trading strategy may have been implemented by identifying whether the equity index is over or under priced.

Applications of state space models have recently been popularized in the asset pricing literature to derive series for expected returns or expected dividend growth, both of which are typically unobservable to the econometrician. The application has taken two avenues of thinking with the first one focussing more on the econometric properties of the data (see Conrad and Kaul (1988), Brandt and Kang (2004), and Pastor and Stambaugh (2009)). The second avenue involves deriving the unobserved series through a more structural approach where the present value relation between the dividend price ratio, expected dividend growth and expected returns hold (Rytchkov (2007), Cochrane (2008), Kojien and van Binsbergen (2010)). I provide a brief summary of these studies. Conrad and Kaul (1988) apply the Kalman Filter to extract expected returns from the history of realized returns. The objective was to attempt to characterize the random nature of expected returns and test whether the latter was constant. Brandt and Kang (2004) investigated the relationship between expected returns and volatility. They model conditional mean and volatility of returns as unobservable variables which follow a latent VAR model and filter them from observed returns. In the same line of thinking, Cochrane (2008)

shows that the present value VAR model can be represented in state space form. Pastor and Stambaugh (2009) use the Kalman Filter to analyze the correlation between predictors and expected return in the forecast of returns. Kojien and Van Binsbergen (2010) use the state space model to model the dynamics of cash invested and market invested dividends.

ARFIMA (p,d,q) models are specifications of time processes where a fractional order of integration is involved, usually defined by the ‘d’ parameter. As such it makes the distinction between the short range components (p and q) and the long range (d). When $0 < d < 0.5$, the model is said to exhibit long memory, and is stationary. By differencing the process d times, the process becomes a stationary I(0) process, with no memory. An important advantage of this transformation is that it provides stronger consistency results for tests on the I(0) series. An interesting overview of such processes is available in Beran(1994).

One of the main properties highlighted by the studies on expected returns is that they exhibit high persistence (Campbell (1991), Rytchchkov(2006), Kojien and Van Binsbergen (2010)). In KVB, expected returns, when modeled jointly with expected dividend growth, exhibits long memory, and most importantly non stationarity¹. This paper derives the empirical series of expected returns and dividend growth rate assuming explicitly that the expected returns series might possess long memory. We reproduce the results of KVB for comparison with the ARFIMA model. Our model specifically studies the case of an ARFIMA (1,d,0). The expected returns, which is fractionally integrated, can be represented as infinite process in the time series model (Chan and Palma (1998)). They show that series that exhibit long memory may have a finite representation where the exact likelihood can be computed recursively using the Kalman Filter. The log likelihood function of the Kalman filter is optimized to the current data set to yield the optimal parameters of the present value.

The present value approach identifies only one measurement (observed) variable which is the price dividend ratio. To counter the identification problem, the observed dividend growth may be used as another measurement equation. A potential problem encountered when using dividend growth and the price dividend ratio is that it may not be fully representative of payoffs in the presence of share repurchases. In this case, I amend the Campbell and Shiller(1989) equation by using the price earnings ratio and earnings growth in the present value relation. Two different samples are used to check for the robustness of our result. In a nutshell, I estimate the model for both the AR(1) and ARFIMA(1,d,0) using both dividend and earnings data for the time periods 1926-2008 and 1946-2008.

¹The study found that the autoregressive parameter is 0.932 with a standard error of 0.128. It should be noted that the Kalman Filter still works in the presence of unit roots. (See Brockwell and Davis 1991)

The remainder of the paper is organized as follows. Section 2 and 3 derive the present value model under an AR(1) and ARFIMA(p,d,q) respectively. Section 4 describes the data set and the results. Section 5 checks the robustness of the results and performs tests of persistence and time variation. Section 6 looks at some of the applications of the derived series. Section 7 concludes.

2 Present Value assuming an AR(1)

2.1 Present Value

In this section, the present value relationship between the price dividend ratio, expected returns and expected dividend growth is derived. I first define the key variables of the present value.

Let the rate of return be:

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

Let the Price Dividend ratio at time t be:

$$PD_t = \frac{P_t}{D_t} \quad (2)$$

Let the Dividend Growth from time t to t+1 be:

$$\Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right) \quad (3)$$

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 mean adjusted conditional expected returns and dividend growth rate are modelled as an autoregressive process as in equations 4 and 5 respectively :

$$\mu_{t+1} - \delta_0 = \delta_1(\mu_t - \delta_0) + \varepsilon_{t+1}^\mu \quad (4)$$

$$g_{t+1} - \gamma_0 = \gamma_1(g_t - \gamma_0) + \varepsilon_{t+1}^g \quad (5)$$

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 δ_0 and γ_0 characterize the unconditional mean of the expected returns and dividend growth respectively. δ_1 and γ_1 represent the autoregressive parameters. ε_{t+1}^μ and ε_{t+1}^g are the shocks to the expected returns and the dividend growth rate processes. The shocks are normally distributed:

$$\varepsilon_{t+1}^\mu \sim N(0, \sigma_\mu^2) \text{ and } \varepsilon_{t+1}^g \sim N(0, \sigma_g^2).$$

However, no restriction is placed on the covariance between ε_{t+1}^μ and ε_{t+1}^g .

The realized dividend growth rate is defined as the expected dividend growth rate and the unobserved shock ε_{t+1}^d , where by :

$$\Delta d_{t+1} = g_t + \varepsilon_{t+1}^d \quad (6)$$

ε_{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 identity (derived in appendix A.1) may be written as :

$$r_{t+1} = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t \quad (7)$$

where $pd_t = E[\log(PD_t)]$, $\kappa = \log(1 + \exp(\bar{pd})) - \rho \bar{pd}$ and $\rho = \frac{\exp(\bar{pd})}{1 + \exp(\bar{pd})}$.

To study the dynamics of the price dividend ratio, the process may be written with pd_t being the subject of the formula:

$$pd_t = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}$$

By replacing lagged iterated values of pd_{t+1} in the equation, the process may be written as :

$$pd_t = \sum_{i=0}^{\infty} \rho^i \kappa + \rho^\infty pd_\infty + \sum_{i=1}^{\infty} \rho^{i-1} (\Delta d_{t+i} - r_{t+i})$$

$$pd_t = \frac{\kappa}{1 - \rho} + \rho^\infty pd_\infty + \sum_{i=1}^{\infty} \rho^{i-1} (\Delta d_{t+i} - r_{t+i})$$

By iterating equation 7 and using assumptions 4 and 5 applying the expectations operator, the functional form of the process can be written as :

$$pd_t = A - B_1(\mu_{t+1} - \delta_0) + B_2(g_t - \gamma_0) \quad (8)$$

The values of A , B_1 and B_2 are specified in the next section.

2.2 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 series of μ_t and g_t . The two transition equations, are the demeaned form of equations 4 and 5:

$$\widehat{g}_{t+1} = \gamma_1 \widehat{g}_t + \varepsilon_{t+1}^g \quad (9)$$

$$\widehat{\mu}_{t+1} = \delta_1 \widehat{\mu}_t + \varepsilon_{t+1}^\mu \quad (10)$$

The two measurement equations are given by :

$$\Delta d_{t+1} = \gamma_0 + \widehat{g}_t + \varepsilon_{t+1}^d \quad (11)$$

$$pd_t = A - B_1 \widehat{\mu}_t + B_2 \widehat{g}_t \quad (12)$$

Equation 11 is the equation linking the observed dividend growth rate to the state variable of expected dividend growth. Equation 12 is the present value equation linking price dividend to expected returns and expected dividend growth rate. It is a generalization of equation 7, where $A = \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho}$, $B_1 = \frac{1}{1-\rho\delta_1}$ and $B_2 = \frac{1}{1-\rho\gamma_1}$.

Equation 12 can be rearranged into 15 such that there are only two measurement equations and only one state space equation.

$$\widehat{g}_{t+1} = \gamma_1 \widehat{g}_t + \varepsilon_{t+1}^g \quad (13)$$

$$\Delta d_{t+1} = \gamma_0 + \widehat{g}_t + \varepsilon_{t+1}^d \quad (14)$$

$$pd_{t+1} = (1 - \delta_1)A + B_2(\gamma_1 - \delta_1)\widehat{g}_t + \delta_1 pd_t - B_1 \varepsilon_{t+1}^\mu + B_2 \varepsilon_{t+1}^g \quad (15)$$

Equation 13 defines the transition (state) equation. The measurement equations are given by 14 and 15.

This can be put into a state space form as shown in appendix A.2. 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. The likelihood is optimized using the MaxBFGS procedure.

$$\Theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_\mu, \sigma_D, \rho_{g\mu}, \rho_{gd}, \rho_{\mu d}) \quad (16)$$

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 expected dividend growth is derived from the state vector update. In the case of the demeaned expected returns, the expected returns is defined as :

$$\hat{\mu}_{t-1} = B_1^{-1}(pd_t - A - B_2\hat{g}_{t-1}) \quad (17)$$

3 Present Value assuming an ARFIMA(1,d,0)

3.1 Present Value

A general ARFIMA(p,d,q) process, x_t can be written as

$$\theta(L)(1-L)^d(x_t - \bar{x}) = \Theta(L)\eta_t \quad (18)$$

where \bar{x} is the mean of x_t , $d < |\frac{1}{2}|$, η_t is a Gaussian white noise process. Equation 18 can be written as an infinite polynomial process:

$$x_t = \bar{x} + \frac{\Theta(L)}{\theta(L)}(1-L)^{-d}\eta_t = \sum_{j=0}^{\infty} \theta_j \eta_{t-j} \quad (19)$$

The state space representation of such a process is documented in **Hannan and Deistler (1988)** and **Chan and Palma (1998)**. I derive the present value using this framework based on assumption that the state variables (equations 9 and 10) are an ARFIMA(1,d,0). Hence, the conditional expected returns and dividend growth rate are modelled as mean reverting process as in equations 20.

$$\begin{aligned} \mu_t &= \delta_0 + G' C_{\mu t}, \\ g_t &= \gamma_0 + G' C_{g t}, \end{aligned} \quad (20)$$

Equation 20 relates to the mean deviation of the expected returns and expected dividend growth rate where δ_0 and γ_0 are scalars, $G' = [1 \ 0 \ 0 \ \dots]$, and $C_{\mu t}$ and $C_{g t}$ are infinite dimensional state vectors. The transition equations are:

$$\begin{aligned} C_{\mu, t+1} &= FC_{\mu t} + h_{\mu} \varepsilon_{t+1}^{\mu}, \\ C_{g, t+1} &= FC_{g t} + h_{g} \varepsilon_{t+1}^g, \end{aligned}$$

where is infinite dimensional matrix with elements

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \end{bmatrix},$$

and h_μ and h_g are infinite dimensional vectors with elements that are parameters from the infinite moving average representation of the processes μ_t and g_t :

$$h_\mu = \begin{bmatrix} 1 \\ \varphi_{\mu 1} \\ \varphi_{\mu 2} \\ \vdots \end{bmatrix}, \quad h_g = \begin{bmatrix} 1 \\ \varphi_{g 1} \\ \varphi_{g 2} \\ \vdots \end{bmatrix},$$

such that

$$\begin{aligned} \mu_t &= \delta_0 + \varepsilon_t^\mu + \varphi_{\mu 1} \varepsilon_{t-1}^\mu + \varphi_{\mu 2} \varepsilon_{t-2}^\mu + \cdots, \\ g_t &= \gamma_0 + \varepsilon_t^g + \varphi_{g 1} \varepsilon_{t-1}^g + \varphi_{g 2} \varepsilon_{t-2}^g + \cdots \end{aligned}$$

ε_t^μ and ε_t^g can be correlated.

The log linearized return is:

$$r_{t+1} = \kappa + \rho p d_{t+1} = \Delta d_{t+1} - p d_t$$

with $\bar{p}d = E(p d_t)$, $\kappa = \log(1 + \exp(\bar{p}d)) - \rho p d$, and $\rho = \frac{\exp(\bar{p}d)}{1 + \exp(\bar{p}d)}$. It can be shown that

$$p d_t = A + B'(C_{gt} - C_{\mu t}),$$

where

$$A = \frac{\kappa + \gamma_0 - \delta_0}{1 - \rho}$$

and

$$B = \begin{bmatrix} 1 \\ \rho \\ \rho^2 \\ \vdots \end{bmatrix}.$$

Thus, in estimation we can use the measurement equation:

$$\begin{aligned} \Delta d_{t+1} &= \gamma_0 + G' C_{gt} + \varepsilon_{t+1}^d \\ p d_t &= A + B' C_{gt} - B' C_{\mu t} \end{aligned}$$

and the transition equations

$$\begin{aligned} C_{g,t+1} &= FC_{gt} + h_g \varepsilon_{t+1}^g \\ C_{\mu,t+1} &= FC_{\mu t} + h_{\mu} \varepsilon_{t+1}^{\mu}. \end{aligned}$$

The optimizing vector is :

$$\Theta = (\gamma_0, \delta_0, \gamma_1, \delta_1, \sigma_g, \sigma_{\mu}, \sigma_D, \rho_{g\mu}, \rho_{gD}, \rho_{\mu D}, d_{\mu}, d_g) \quad (21)$$

The expected returns is derived from the series of the filtered series with the estimated paramaters :

$$\mu_t = [g_t - \gamma_0 + \delta_0] - G'(B')^{-1}[pd_t - A] \quad (22)$$

3.2 Kalman filter for the ARFIMA model.

In this section, the log likelihood function is derived after applying the Kalman filter to the state and measurement equations of the previous section.

Lets define $X_{g,t+1} = C_{gt}$ and $X_{\mu,t+1} = C_{\mu t}$, so the transition equations are:

$$\begin{aligned} X_{g,t+1} &= FX_{gt} + h_g \varepsilon_t^g, \\ X_{\mu,t+1} &= FX_{\mu t} + h_\mu \varepsilon_t^\mu. \end{aligned}$$

and the measurement equations are:

$$\begin{aligned} \Delta d_t &= \gamma_0 + G' X_{gt} + \varepsilon_t^d, \\ pd_t &= A + B' F X_{gt} - B' F X_{\mu t} + B' h_g \varepsilon_t^g - B' h_\mu \varepsilon_t^\mu. \end{aligned}$$

In general notation the transition and measurement equations are

$$\begin{aligned} X_{t+1} &= \bar{F} X_t + V_t, \\ Y_t &= D + \bar{G} X_t + W_t, \end{aligned}$$

where

$$\begin{aligned} \bar{F} &= \begin{bmatrix} F & 0 \\ 0 & F \end{bmatrix}, \\ \bar{G} &= \begin{bmatrix} G' & 0 \\ B' F & -B' F \end{bmatrix}, \\ D &= \begin{bmatrix} \gamma_0 \\ A \end{bmatrix}, \\ V_t &= \begin{bmatrix} h_g \varepsilon_t^g \\ h_\mu \varepsilon_t^\mu \end{bmatrix}, \\ W_t &= \begin{bmatrix} \varepsilon_t^d \\ B' h_g \varepsilon_t^g - B' h_\mu \varepsilon_t^\mu \end{bmatrix}, \end{aligned}$$

where 0 is either infinite dimensional vector or matrix of zeros.

Also, lets define

$$\begin{aligned} Q &= \begin{bmatrix} h_g h_g' \sigma_g^2 & h_g h_\mu' \rho_{\mu g} \sigma_\mu \sigma_g \\ h_\mu h_g' \rho_{\mu g} \sigma_\mu \sigma_g & h_\mu h_\mu' \sigma_\mu^2 \end{bmatrix}, \\ R &= \begin{bmatrix} \sigma_d^2 & B' h_g \rho_{dg} \sigma_d \sigma_g - B' h_\mu \rho_{d\mu} \sigma_d \sigma_\mu \\ B' h_g \rho_{dg} \sigma_d \sigma_g - B' h_\mu \rho_{d\mu} \sigma_d \sigma_\mu & (B' h_g)^2 \sigma_g^2 + (B' h_\mu)^2 \sigma_\mu^2 - 2B' h_g B' h_\mu \rho_{\mu g} \sigma_g \sigma_\mu \end{bmatrix}, \\ S &= \begin{bmatrix} h_g \rho_{dg} \sigma_d \sigma_g & h_g B' h_g \sigma_g^2 - h_g B' h_\mu \rho_{\mu g} \sigma_\mu \sigma_g \\ h_\mu \rho_{d\mu} \sigma_d \sigma_\mu & h_\mu B' h_g \rho_{\mu g} \sigma_\mu \sigma_g - h_\mu B' h_\mu \sigma_\mu^2 \end{bmatrix}. \end{aligned}$$

Under initial condition:

$$X_1 = 0,$$

the Kalman equations are:

$$\begin{aligned}\Delta_t &= \bar{G}\Omega_t\bar{G}' + R \\ \Theta_t &= \bar{F}\Omega_t\bar{G}' + S \\ \Omega_{t+1} &= \bar{F}\Omega_t\bar{F}' + Q - \Theta_t\Delta_t^{-1}\Theta_t' \\ \hat{X}_{t+1} &= \bar{F}\hat{X}_t + \Theta_t\Delta_t^{-1}\left(Y_t - D - \bar{G}\hat{X}_t\right)\end{aligned}$$

The log likelihood function is then given by:

$$\ell = 2\pi^{-n} \left(\prod_{j=1}^n \det \Delta_j \right)^{-1/2} \exp \left(-\frac{1}{2} \sum_{j=1}^n (Y_j - \hat{Y})' \Delta_j^{-1} (Y_j - \hat{Y}) \right)$$

with $\hat{Y}_t = D + \bar{G}\hat{X}_t$.

4 Results

4.1 Data

The optimization model makes use of earnings, dividend, price and consumption data from 1926-2008. The data is retrieved from Robert Shiller’s website. The returns series is derived by taking the log difference of the price series. Earnings and Dividend growth are computed as the logarithm of the ratio of the observation at time $t+1$ and at t .

4.2 Optimization of State Space Models

In this section, I present the results of the two models defined in the previous sections for the sample period 1926-2008 and 1946-2008 for both the dividends and earnings.

Price Dividend Ratio/Dividend Growth								
Parameters	1926-2008				1946-2008			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	PARAM	SE	PARAM	SE	PARAM	SE	PARAM	SE
γ_0	0.021	0.014	0.003	0.081	0.019	0.012	0.018	0.001
δ_0	0.055	0.019	0.029	0.074	0.046	0.020	0.05	0.001
γ_1	0.11	0.119	0.165	0.210	0.395	0.203	0.326	0.108
δ_1	0.921	0.050	0.158	0.031	0.929	0.049	0.118	0.033
d_1	–	–	0.475	0.050	–	–	0.457	0.606
σ_g	0.052	0.016	0.109	0.001	0.05	0.014	0.048	0.193
σ_μ	0.02	0.010	0.044	0.001	0.015	0.009	0.044	0.030
σ_d	0.092	0.055	0.007	0.001	0.014	0.040	0.007	0.043
$\rho_{g\mu}$	0.576	0.088	0.212	0.001	0.62	0.126	0.345	0.494
$\rho_{\mu d}$	–0.046	0.001	–0.166	0.001	–0.055	0.685	–0.002	0.250
Log-Likelihood	–103.79		–110.45		–130.07		–126.83	

Table 1:

Estimation of AR(1) and ARFIMA(1,d,0) model for dividend data. The parameters optimized are from the previously defined parameter sets 16 (AR(1)) and 21 (ARFIMA(1,d,0)) over the two periods. The ‘d’ parameter for the expected dividend growth process is set to zero.

Price Earnings Ratio/ Earnings Growth								
Parameters	1926-2008				1946-2008			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	PARAM	SE	PARAM	SE	PARAM	SE	PARAM	SE
γ_0	0.023	0.007	0.010	0.038	0.006	0.035	0.02	0.042
δ_0	0.089	0.044	0.079	0.054	0.048	0.068	0.05	0.052
γ_1	0.082	0.001	0.089	0.049	0.101	0.128	0.457	0.001
δ_1	0.878	0.001	0.101	0.042	0.927	0.048	0.382	0.001
d_1	–	–	0.363	0.045	–	–	0.499	0.001
σ_g	0.197	0.001	0.228	0.014	0.218	0.023	0.031	0.001
σ_μ	0.057	0.001	0.101	0.012	0.043	0.017	0.024	0.001
σ_e	0.116	0.001	0.184	0.020	0.016	0.110	0.045	0.001
$\rho_{g\mu}$	0.87	0.001	0.829	0.065	0.929	0.043	0.237	0.001
$\rho_{\mu e}$	0.126	0.001	-0.198	0.069	-0.364	0.273	0.284	0.001
Log-Likelihood	-35.27		-113.03		-39.99		-123.61	

Table 2:

Estimation of AR(1) and ARFIMA(1,d,0) model for earnings data. In this table, the measurement variables, price dividend ratio and dividend growth, are replaced by the price earnings ratio and the earnings growth respectively.

According to the log likelihood value, the ARFIMA(1,d,0) tends to perform better than the AR(1) except for the 1946-2008 dividend sample. Generally, the ARFIMA model is superior when using earnings. The result is fairly simple to understand within the present value framework. The price earnings ratio and the price dividend ratio share a common level of persistence (i.e. an autoregression on the price dividend ratio and price earnings ratio produces nearly the same autoregressive parameter). However the observed dividend growth is much more persistent than earnings growth. According to identity 12, the expected returns in earnings equations should have a higher degree of persistence which is well suited to the ARFIMA(1,d,0).

The optimized results have the same near unit root properties as in KVB for the autoregressive processes. The unit root is found in both samples with the price dividend ratio and also when the price earnings ratio is used. Interestingly, the memory component (d) is high in almost all four models. The short range component (autoregressive part) of the ARFIMA tends to be lower. Dividend growth tends to have the an equal short run parameter over both periods. The parameters of the dividend growth equation changes only marginally as to the adoption of the ARFIMA model. The sample 1946-2008 is associated with an equal improvement in the autoregressive parameter for both expected dividend growth and expected earnings growth.

For both samples, the variation of the expected and realized dividend and earnings growth rate tend to be higher than that of expected returns. The expected earnings growth appears to vary much more than the dividend growth for both samples and both models. The correlation between expected earnings growth and expected returns tends to be stronger than that between dividend growth and expected returns. Some of these findings can be confirmed from tables 6,7,8 and 9 in the appendix documenting the descriptive features of the filtered series. The mean values of the expected returns and expected dividend growth rate are close to the observed. The 1926-2008 earnings growth rate, however are exceptionally high. Interestingly, tests of stationarity $I(0)$ and non stationarity $I(1)$ show that the AR(1) models tend to be closer to being non-stationary, unlike ARFIMA models. The Robinson-Lobato p-values show that upon the adoption of the ARFIMA model, the p-values signalling rejection of fractional alternatives, tend to be lower than the AR(1) model, hence putting the case forward for a fractional process.

5 Robustness Checks

5.1 Robustness over time

Robustness checks were performed over the two samples on the derived expected returns series. The ARFIMA(1,d,0) model is not robust, with regards to this time discrepancy. However, this may be due to different regimes within the two samples which lead to substantial variation in the ARFIMA process. Based on the latent nature of the expected returns, a method to check for robustness is to see whether expected returns across the different time periods exhibit high correlation. The pairwise correlation are reported in tables 10 and 11. In the case of the AR(1), most of the pair wise correlations are above 0.8. Both the price dividend ratio and the price earnings ratio exhibit a 0.99 correlation over the two different sample sizes.

The ARFIMA model tends to exhibit low correlation over time. However, earnings and dividend measures tend to demonstrate the same level of correlation in a specific sample. It is clear that the 1929-36 depression may have had led to higher expected returns. A Monte Carlo experiment was performed by using the same sample size (87 observations) and a different regime where expected returns are set to be higher. The Monte Carlo results, reported in table 12, show that the ARFIMA parameters tend to exhibit a higher variance than the simple AR process. This may account for the improper point estimation where the optimal point of the likelihood function falls into a local optimum.

5.2 Univariate Models

In this section, the expected returns and expected dividend (earnings) growth rates are modeled according to their initial exogenously determined econometric specifications. The statistical properties presented before were when expected returns was modeled jointly with the dynamics of dividend (earnings) growth and the price dividend (earnings) ratio. We present the results for the series for the individual samples in tables 13,15,14 and 16 in the appendix.

The results show that the ARFIMA tends to perform worse than the AR(1) in the case of the expected returns (tables 13 and 15). The memory parameters tend to be unstable in both specifications of the autoregressive process. The ARFIMA tends to generally display a lower R-squared. The best ARFIMA model is from the use of the earnings data for the sample 1946-2008. One advantage of the ARFIMA

model is that it removes dependence in the residual and reduces the ARCH effects. Interestingly, the univariate models for the dividend and earnings growth show more promise for the ARFIMA model. The linear fit of the model is improved greatly and the models are free from any serial correlation and conditional heteroscedasticity. The good fit of the model is perfectly clear within the present value approach. A lower fit in either the dividend growth rate or expected returns would improve the fit of the other variable, so that the persistence in the price dividend ratio is restored.

5.3 Tests of Persistence and Time variation

In the following section, we detail some of the tests that were performed on the new series. First we test whether the expected returns is a long memory process (i.e. $d > 0$). In terms of the definition of ‘persistence’, this is the strong form case. Secondly, we test the weaker form of persistence, where the autoregressive coefficient and the ‘d’ parameter are set to zero. The same test may be applied to the expected dividend growth series, which involves looking at the autoregressive coefficient only. We also test for time variation in expected returns and expected dividend (earnings) growth rate. The null hypothesis in this case is that each autoregressive parameter and the standard error are equal to zero.

The tests involve computing the likelihood ratio under alternative (L_1) and the null (L_0), and the following likelihood ratio test is applied:

$$LR = 2(L_1 - L_0)$$

The likelihood ratio is distributed as $\chi^2(k)$ where k represents the number of restrictions.

The tests are performed exclusively on the ARFIMA(1,d,0) specification and are reported in table 3.

Null Hypotheses	1926-2008 PD	1946-2008 PD	1926-2008 PE	1946-2008 PE
Memory $d = 0$	3368	2155	623.49	2690
Persistence Tests:				
$H_0 : \delta_1 = d = 0$	6478	3465	579	4153
$H_0 : \gamma_1 = 0$	22	12	14	458
Time Variation tests:				
$H_0 : \delta_1 = d = \sigma_\mu = 0$	17507	61797	317	4919
$H_0 : \gamma_1 = \sigma_g = 0$	1070	1806	130	684

Table 3:
Tests on Time Variation and Persistence.

The number of parameters in the system implies that the log likelihood test tends to vary across the different samples and the different measure of the leverage variable. However the results clearly demonstrate that the null hypothesis is being rejected. in all cases. Expected returns do appear to exhibit long memory. There appears to be persistence in both the filtered returns and filtered dividend growth rate series. However the former exhibits a higher statistic close to the rejection of the null, implying that there is a higher degree of persistence. The expected returns series shows that there is enough joint evidence of a non zero d and the autoregressive parameter δ_1 . Tests for time variation show that both the expected returns and dividend growth rate tend to vary over time. However a naive comparison of the test statistic shows that expected returns exhibit more variation over time.

6 Applications

In this section, we provide three applications for the filtered returns and dividend (earnings) series. In the first application, we test for insample predictability. In this setting the filtered series are regressed on the realized values and the accuracy is measured by the R-squared. In the second application, we look at the effect of expected returns (as a proxy for discount rates) on consumption and consumption growth. Lastly, we test whether a trading strategy may be implemented by looking whether prices revert to their present value. We use the series for expected returns and expected dividend (earnings) growth to construct the present value.

6.1 Insample Predictability

The insample predictability of the realized series by the filtered series is reported in the tables 4 and 5. The following forecasting equations were run in the case of realized values of returns and dividend growth:

$$\begin{aligned}
 r_t &= \phi_o + \phi_1 \mu_{t-1}^{AR} + \varepsilon_t \\
 r_t &= \phi_o + \phi_1 \mu_{t-1}^{ARFIMA} + \varepsilon_t \\
 r_t &= \phi_o + \phi_1 pd_{t-1} + \varepsilon_t
 \end{aligned}$$

$$\begin{aligned}\Delta d_t &= \phi_o + \phi_1 g_{t-1}^{AR} + \varepsilon_t \\ \Delta d_t &= \phi_o + \phi_1 g_{t-1}^{ARFIMA} + \varepsilon_t \\ \Delta d_t &= \phi_o + \phi_1 pd_{t-1} + \varepsilon_t\end{aligned}$$

	AR(1)	ARFIMA(1,d,0)	pd _{t-1} /pe _{t-1}
1926-2008 PD	0.049	0.015	0.050
1946-2008 PD	0.101	0.088	0.066
1926-2008 PE	0.042	0.002	0.055
1946-2008 PE	0.072	0.079	0.090

Table 4:

Goodness of fit for the returns equation. The figures in the table show the respective R-squared from each regressor.

For the dividend growth the following functional models were assumed:

	AR(1)	ARFIMA(1,d,0)	pd _{t-1} /pe _{t-1}
1926-2008 PD	0.020	0.01	0.002
1946-2008 PD	0.015	0.026	0.008
1926-2008 PE	0.014	0	0.001
1946-2008 PE	0.027	0.081	0.002

Table 5:

Goodness of Fit for the dividend(earnings) growth equation

Returns are better forecast by the price dividend and price earnings ratio. The expected returns are marginally weaker than the ratios. The autoregressive process tends to work for the 1946 sample. A higher level of predictability is noticed for the period, 1946-2008. There is no apparent predictability for the expected dividend growth rate. The ARFIMA model performs relatively well for the sample 1946-2008 for the price earnings ratio.

6.2 Consumption and Expected Returns

The second application of both series is to see the reaction of consumption growth to a shock in expected returns. There is a wide theoretical literature linking the

time series properties of consumption and discount rates (See Campbell (2003) and Cochrane (2010) for an overview). Consumption and discount rates are counter cyclical to each other. When discount rates (expected returns) are low, consumption is high. To test this relationship, a simple regression regressing expected returns on logarithm of consumption.

$$\ln C_t = \alpha + \beta\mu_t + v_t \quad (23)$$

We also produced the impulse response for a first order vector autoregression model, with consumption growth (defined as ΔC_t) and expected returns as the endogenous variables.

$$Y_t = A + BY_{t-1} + v_t$$

where $Y_t = [\Delta C_t \quad \mu_t]'$, $\alpha = [\alpha_1 \quad \alpha_2]'$, $B = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{bmatrix}$, $v_t = \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$

The above system is estimated and impulse response functions are plotted as a result to the expected returns process. The result from equation 23 resulted in a negative relationship between consumption and expected returns. The parameters range from -3.51 to the extreme case of -8.13. No definite distinction between expected returns under the ARFIMA(1,d,0) and AR(1) was found.

The impulse response functions (plots 6.2 and 6.2) show that there is a higher persistence in consumption growth after a shock to the AR(1) expected returns. The expected return series from the ARFIMA model has already accounted for the long memory components and as such, shocks are damped after each lag. This modest finding may be reconciled with business cycle theories, where the frequency of a cycle is shown to be four years.

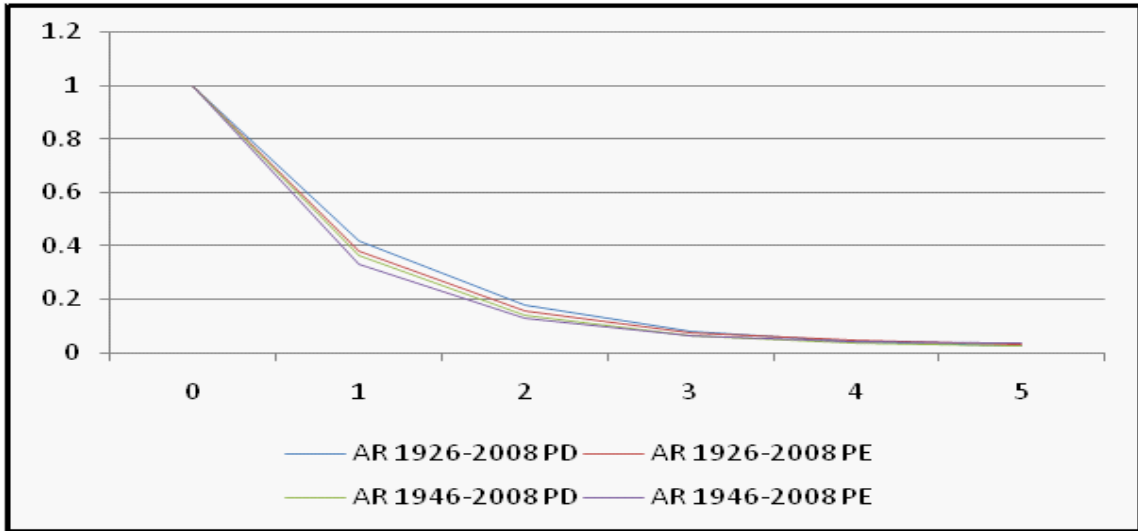


Figure 6.2: Impulse Response for Consumption Growth based on Expected Returns-AR(1)

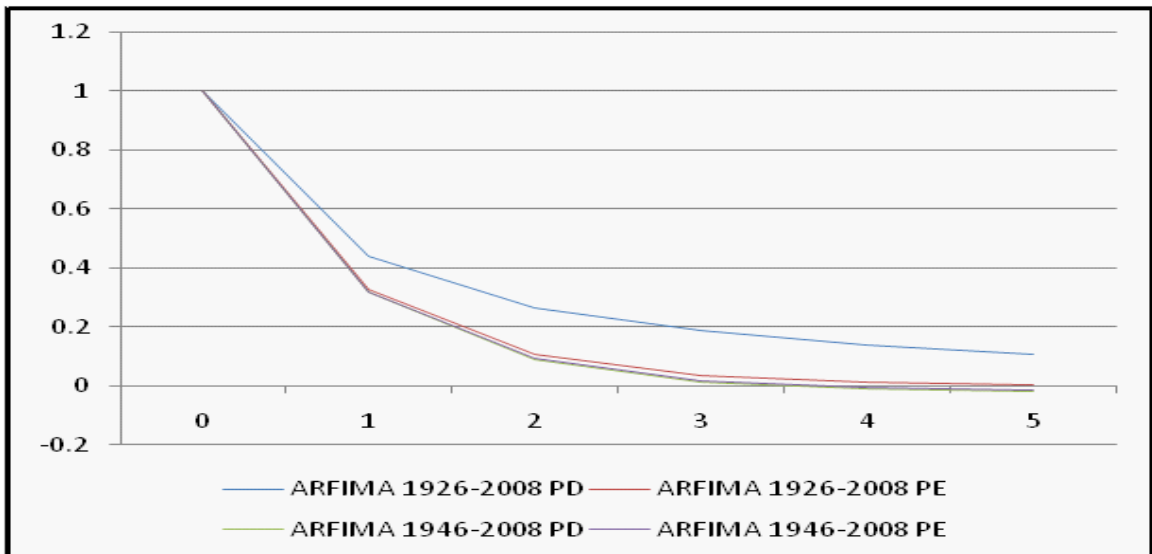


Figure 6.3: Impulse Response function for Consumption growth based on a shock in expected returns.- ARFIMA(1,d,0)

6.3 A Real-Time Trading Strategy

In this section, an application of the expected returns and expected dividend growth rate series is provided. I test whether a profitable trading strategy may be implemented by identifying whether the stock market is underpriced or overpriced. The trading strategy is an all long strategy with the choice of either going long in bonds or the equity index. If the market is overpriced, reversion towards fundamental value will imply that price will fall in the following period(s), leading to a potential capital loss if equity index is held. In this case, the trading rule is to go long on treasury bills. Likewise, if the equity index is underpriced, reversion to the fundamental value implies that there would be an increase in the price and therefore implying a positive return.

The present value is computed in real-time using the values of the derived series. As a measure of comparison, the present value is computed using the previous period realized values. The trading strategy is compared against the Buy and Hold. Two versions of the present value is assumed. The first model of the present value is the Gordon Dividend Growth model, which assumes that the dividend growth rate is constant. The second present value is the discounted next period dividend and price. The expected future price is proxied by the present price, given the random walk nature of the price.

$$PV1^{AR} = \frac{D_t(1 + g_t^{AR})}{\mu_t^{AR} - g_t^{AR}} \quad (\text{Model 1})$$

$$PV1^{ARFIMA} = \frac{D_t(1 + g_t^{ARFIMA})}{\mu_t^{ARFIMA} - g_t^{ARFIMA}} \quad (\text{Model 2})$$

$$PV2^{AR} = \frac{D_t(1 + g_t^{AR}) + P_t}{\mu_t^{AR} + 1} \quad (\text{Model 3})$$

$$PV2^{ARFIMA} = \frac{D_t(1 + g_t^{ARFIMA}) + P_t}{\mu_t^{ARFIMA} + 1} \quad (\text{Model 4})$$

$$PV^R = \frac{D_t(1 + \Delta d_t) + P_t}{r_t + 1} \quad (\text{Model 5})$$

The cumulated returns from the trading strategy (using the present value) are plotted for the respective time period and payoff variable. The results are reported in the appendix. Based on the four time periods, the Buy and Hold strategy tends

to beat the trading strategy. The only exception comes from the price earnings optimization problem for the period 1946-2008 where the ARFIMA model tends to beat the Buy and Hold over the whole period. However, it is worth mentioning that the graphical plots do not do justice to the proper performance of the trading strategy since a high market return in one period may bias the Buy and Hold strategy towards having a higher accumulated return than the trading strategy. In such a circumstance, we report a measure based on the binary outcome of whether the rule advised going towards the highest return between treasury bills and the equity index. We report the percentage of times each model was successful.

	1926-2008 PD	1926-2008 PE	1946-2008 PD	1946-2008 PE
Model 1	0.37	0.53	0.41	0.55
Model 2	0.35	0.46	0.40	0.66
Model 3	0.36	0.30	0.38	0.46
Model 4	0.47	0.57	0.24	0.60
Model 5	0.42	0.47	0.46	0.51
Buy and Hold	0.69	0.69	0.71	0.71

Interestingly in the binary measure case, none of the trading strategies manage to beat the Buy and Hold. However, compared to the graphical plots the strategies do not perform as badly. There is no definite winner in terms of the present value formulation adopted. Both versions of the present value tend to perform well for the different time periods involved. This finding of non-robustness may be due to the presence of breaks or regime switches, which invalidate the constant dividend growth theory. The ARFIMA models appears to work better than the autoregressive models. The interesting phenomenon is that the ARFIMA, by accounting for hyperbolic decay, is a smoother series, implying that unless there are huge changes in the dividend growth and realized dividends, the present value will be a smooth function over time. In such a case, it may be likely, that the ARFIMA performs better than the AR models because of its ability to capture smoother business cycle transition over time.

7 Conclusion

In this paper, it was found that the time series of annual expected returns exhibit long memory. Taking into account the nature of the process, the fractional process removes the components associated with long memory, and hence to exhibit lower serial correlation. However the results from the ARFIMA are not robust across time

and it is suspected that regime switches across the two different sample adopted might be the cause. This finding has been backed by a small Monte Carlo experiment with two regimes. The univariate specifications of the expected returns series also show that the AR(1) fares better than the ARFIMA (1,d,0).

The filtered series was used in three simple applications namely in evaluating return predictability, assessing the relationship between consumption growth and discount rate and lastly in a market timing strategy. In terms of predictability of returns, both the AR(1) and ARFIMA(1,d,0) have equal power in predicting returns, although it is marginally lower than valuation ratios. However the filtered series have stronger forecasting power for dividend and earnings growth rate. Consumption and expected returns was found to be negatively related and a simple impulse response function showed that the effect of a shock in the discount rate may last until four years on consumption growth. The results on the trading rule is that it is impossible to jointly build a mean reverting strategy by identify the over or under pricing of the equity market exante. Such a strategy performs poorly against the buy and hold.

The paper draws out many lines of research for the future. The most obvious area to consider is the possibility of different regimes within the specification of the expected returns equations. The filtered series may have power against the AR(1) and ARFIMA(1,d,0) process. In this study, the applications involved simple predictability of returns. It might be interesting to see if forecast long horizon returns from the filtered series may overtake the valuation ratios, which will cumulates noise emanating from the dividend (earnings) growth. Moreover, we can extend the time series framework to the cross sectional cost of capital approach for different industries and firms. The present value may then be compared to the risk based measures to compute the discount factor.

References

- Beran, J., 1994. *Statistics for Long Memory Processes*. Chapman and Hall, New York, NY, USA.
- Binsbergern, J. H. V., Koijen, R., 2010. Predictive regressions: A present-value approach. *Journal of Finance* 65 (4), 1439–1471.
- Brandt, M. W., Kang., Q., 2004. On the relationship between the conditional mean and volatility of stock returns: A latent var approach. *Journal of Financial Economics* 72 (2), 217 – 257.

- Brockwell, P., Davis, R, A., 1991. Time Series: theory and mehtods. Vol. Second Edition. New York.
- Campbell, J. Y., 1991. A variance decomposition for stock returns. *The Economic Journal* 101 (405), pp. 157–179.
- Campbell, J. Y., 2003. Chapter 13 consumption-based asset pricing. In: *Financial Markets and Asset Pricing*. Vol. 1, Part 2 of *Handbook of the Economics of Finance*. pp. 803 – 887.
- Chan, N. H., Palma, W., 1998. State space modeling of long-memory processes. *The Annals of Statistics* 26 (2), pp. 719–740.
- Cochrane, J., 2008. State space vs var. Chicago Booth working paper.
- Cochrane, J., 2010. Discount rates. Chicago Booth Working Paper.
- Conrad, J., Kaul, G., 1988. Time-variation in expected returns. *The Journal of Business* 61 (4), pp. 409–425.
- Hannan, E. J., Deistler, M., 1987. *The statistical theory of linear systems*. John Wiley & Sons, Inc., New York, NY, USA.
- Pastor, L., Stambaugh, R. F., Jan. 2007. Predictive systems: Living with imperfect predictors. NBER Working Papers 12814, National Bureau of Economic Research, Inc.
- Rytchkov, 2006. Filtering out expected returns and expected dividend growth. MIT Working paper.

A The Present Value Model.

Equations 1 , 2 and 3 are shown again:

$$r_t = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \quad (24)$$

$$PD_t = \frac{P_t}{D_t} \quad (25)$$

$$\Delta d_{t+1} = \log\left(\frac{D_{t+1}}{D_t}\right) \quad (26)$$

The return process can be written as

$$\begin{aligned} r_t &= \log\left(\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \cdot \frac{D_t}{D_t} \cdot \frac{D_{t+1}}{D_{t+1}}\right) \\ &\log\left(\left(\frac{P_{t+1}D_t + D_{t+1}D_t}{P_tD_{t+1}}\right) \cdot \frac{D_{t+1}}{D_t}\right) \end{aligned} \quad (27)$$

$$\log\left(\left(\frac{D_t}{P_t} \cdot \frac{P_{t+1}}{D_{t+1}} + \frac{D_t}{P_t}\right) \cdot \frac{D_{t+1}}{D_t}\right) \quad (28)$$

$$\log\left(\left(\frac{P_{t+1}}{D_{t+1}} + 1\right) \frac{D_{t+1}}{D_t} \cdot \frac{D_t}{P_t}\right) \quad (29)$$

$$\log(1 + e^{pd_{t+1}}) + \Delta d_{t+1} - pd_t \quad (30)$$

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

$$r_t \simeq \log((1 + e^{pd_{t+1}})) + \frac{\exp(\overline{pd})}{1 + \exp(\overline{pd})} + \Delta d_{t+1} - pd_t$$

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

where $\kappa = \log((1 + e^{pd_{t+1}}))$ and $\rho = \frac{\exp(\overline{pd})}{1 + \exp(\overline{pd})}$

Hence,

$$pd_t = \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1}$$

B State Space Model assuming AR(1)

In this section I reproduce the Filtering employed in KVB.

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

$$\widehat{g}_{t+1} = \gamma_1 \widehat{g}_t + \varepsilon_{t+1}^g \quad (31)$$

$$\widehat{\mu}_{t+1} = \delta_1 \widehat{\mu}_t + \varepsilon_{t+1}^\mu \quad (32)$$

the two measurement equations are given by :

$$\Delta d_{t+1} = \gamma_0 + \widehat{g}_t + \varepsilon_{t+1}^d \quad (33)$$

$$pd_t = A - B\widehat{\mu}_t + B\widehat{g}_t \quad (34)$$

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

$$\widehat{g}_{t+1} = \gamma_1 \widehat{g}_t + \varepsilon_{t+1}^g \quad (35)$$

$$\Delta d_{t+1} = \gamma_0 + \widehat{g}_t + \varepsilon_{t+1}^d \quad (36)$$

$$pd_{t+1} = (1 - \delta_1)A - B_2(\gamma_1 - \delta_1)\widehat{g}_t + \delta_1 pd_t - B_1 \varepsilon_{t+1}^\mu + B_2 \varepsilon_{t+1}^g \quad (37)$$

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 $\frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho}$, $B_1 = \frac{1}{1-\rho\delta_1}$, $B_2 = \frac{1}{1-\rho\gamma_1}$.

The state equation is defined by :

$$X_{t+1} = FX_t + R\varepsilon_t$$

$$Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t$$

where $Y_t = \begin{bmatrix} \Delta d_t \\ pd_t \end{bmatrix}$

The variables of the transition equation are X_t and ε_{t+1}^x and are made up of the following elements:

$$X_t = \begin{bmatrix} \widehat{g}_{t-1} \\ \varepsilon_t^D \\ \varepsilon_t^g \\ \varepsilon_t^\mu \end{bmatrix} \quad F = \begin{bmatrix} \gamma_1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\varepsilon_{t+1}^x = \begin{bmatrix} \varepsilon_{t+1}^D \\ \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^\mu \end{bmatrix}$$

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

$$M_0 = \begin{bmatrix} \gamma_0 \\ (1 - \delta_1) * A \end{bmatrix} \quad M_1 = \begin{bmatrix} 0 & 0 \\ 0 & \delta_1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ B_2(\gamma_1 - \delta_1) & 0 & B_2 & -B_1 \end{bmatrix}$$

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

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

$$\begin{aligned} X_{0|0} &= E[X_0] \\ P_{0|0} &= E[X_t X_t'] \\ X_{t|t-1} &= F X_{t-1|t-1} \\ P_{t|t-1} &= F P_{t-1|t-1} F' + R \Sigma R' \\ \eta_t &= Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1} \\ S_t &= M_2 P_{t|t-1} M_2' \\ K_t &= P_{t|t-1} M_2' S_t^{-1} \\ X_{t|t} &= X_{t|t-1} + K_t \eta_t \\ P_{t|t} &= (I - K_t M_2) P_{t|t-1} \end{aligned}$$

C Statistical properties of Expected Returns and Growth

C.1 Expected Returns

The statistical properties of expected returns are reported in this section. It presents a first pass misspecification test as to the adequacy of the ARFIMA model. The

following tables report both the usual moments of each series under both models as well as some stationarity and nonstationarity tests.

	Price Dividend Ratio/Dividend Growth					
	1926-2008			1946-2008		
	AR(1)	ARFIMA(1,d,0)	r_t	AR(1)	ARFIMA(1,d,0)	r_t
Mean	0.062	0.064	0.080	0.063	0.065	0.073
Std. Dev	0.049	0.054	0.195	0.045	0.036	0.169
Kurtosis	3.310	3.49	2.944	3.309	4.282	3.090
Test of I(0)						
Robinson-Lobato	-0.075	1.152	-0.599	-	-0.484	0.208
P-value	0.524	0.125	0.726	-	0.686	0.42
KPSS	1.378	0.583	0.073	0.251	0.037	0.080
P-value	0.3	0.025	1	1	1	1
Test of I(1)						
ADF test	-1.11	-4.63	-8.49	-1.214	-2.672	-6.733
P-value	0.9	0.01	0.01	0.9	0.1	0.01
Phillips-Perron	-1.61	-5.81	-8.58	-1.331	-2.800	-6.823
P-value	0.9	0.01	0.01	0.9	0.1	0.01

Table 6:

Descriptive Statistics and Stationarity tests on the Expected Returns series for Dividend series. The stationarity tests are distinct from each other based on the null hypothesis. Tests of I(0) assume that the null hypothesis is in fact a stationary series. In the case of the Robinson-Lobato(1998) test, the alternative is a fractional process. Tests of I(1).are tests with null hypothesis being a an integrated series of I(1). For the I(0) and I(1) tests, the reported p-values are the rejection regions where the test statistic lies.

	Price Earnings Ratio/Earnings growth					
	1926-2008			1946-2008		
	AR(1)	ARFIMA(1,d,0)	r_t	AR(1)	ARFIMA(1,d,0)	r_t
Mean	0.087	0.081	0.080	0.096	0.060	0.073
Std. Dev	0.079	0.036	0.195	0.064	0.034	0.169
Kurtosis	4.691	3.608	2.944	4.297	4.069	3.090
Test of I(0)						
Robinson-Lobato Test	-0.120	0.537	-0.599	-0.299	-0.139	0.208
P-value	0.548	0.295	0.726	0.618	0.592	0.42
KPSS test	0.457	0.129	0.073	0.576	0.549	0.080
P-value	0.1	1	1	0.025	0.05	1
Test of I(1)						
ADF test	-1.009	-4.46	-8.49	-0.953	-1.700	-6.733
P-value	0.9	0.01	0.01	0.9	0.9	0.01
Phillips-Perron	-2.66	-4.53	-8.58	-2.207	-2.053	-6.823
P-value	0.1	0.01	0.01	0.9	0.9	0.01

Table 7:
Descriptive Statistics and Stationarity tests on the Expected Returns series using Earnings data. The resulting statistical features of the expected returns are from the optimization model with the price earnings and earnings growth.

C.2 Statistical properties of Expected Dividend Growth

The expected dividend growth properties are presented in the following tables.

	Price Dividend Ratio/ Dividend Growth					
	1926-2008			1946-2008		
	AR(1)	ARFIMA(1,d,0)	Δd_t	AR(1)	ARFIMA(1,d,0)	Δd_t
Mean	0.021	0.018	0.017	0.020	0.032	0.021
Std. Dev	0.017	0.045	0.106	0.029	0.023	0.055
Kurtosis	3.175	5.49	10.06	5.773	2.687	5.132
Test of I(0)						
Robinson-Lobato Test	-0.498	-0.294	0.793	0.246	0.723	0.719
P-value	0.691	0.616	0.786	0.403	0.235	0.236
KPSS test	0.322	0.069	0.034	0.066	0.147	0.113
P-value	1	1	1	1	1	1
Test of I(1)						
ADF test	-5.958	-6.582	-7.471	-5.158	-3.289	-4.811
P-value	0.01	0.01	0.01	0.01	0.025	0.01
Phillips-Perron	-8.167	-7.42	-7.474	-5.243	-3.344	4.887
P-value	0.01	0.01	0.01	0.01	0.025	0.01

Table 8:
Descriptive Statistics and Stationarity tests on the Expected Dividend Growth rate.

	Price Earnings Ratio/ Earnings growth					
	1926-2008			1946-2008		
	AR(1)	ARFIMA(1,d,0)	Δe_t	AR(1)	ARFIMA(1,d,0)	Δe_t
Mean	0.087	0.081	0.007	0.012	0.024	0.013
Std.Dev	0.079	0.036	0.226	0.069	0.012	0.217
Kurtosis	4.69	3.574	6.613	6.238	4.451	8.899
Test of I(0)						
Robinson-LobatoTest	-0.119	0.515	-1.395	-0.498	0.165	-1.536
P-value	0.548	0.303	0.918	0.691	0.434	0.938
KPSS test	0.159	0.013	0.049	0.122	0.093	0.152
P-value	1	1	1	1	1	1
Test of I(1)						
ADF test	-2.213	-4.463	-6.486	-4.739	-2.761	-4.698
P-value	0.9	0.01	0.01	0.01	0.1	0.01
Phillips-Perron	-2.66	-4.532	-6.506	-7.327	-3.708	-5.034
P-value	0.05	0.01	0.01	0.01	0.01	0.01

Table 9:
Descriptive Statistics and Stationarity tests on the Expected Earnings Growth rate.

D Robustness checks

D.1 Correlation over time

The following tables report the correlation coefficient of the different series for both the autoregressive and fractionally autoregressive series over the different samples.

	$\mu_t^{PD,I}$	$\mu_t^{PE,I}$	$\mu_t^{PD,II}$	$\mu_t^{PE,II}$
$\mu_t^{PD,I}$	1	0.85	0.99	0.84
$\mu_t^{PE,I}$	0.85	1	0.85	0.99
$\mu_t^{PD,II}$	0.99	0.85	1	0.85
$\mu_t^{PE,II}$	0.84	0.99	0.85	1

Table 10:

Correlation of the Expected Returns from AR(1) specification. The above shows the correlation coefficient from the expected returns from the dividend (pd) and earnings (pe) specifications. The first period (*I*) denotes the sample period 1926-2008, while the second period(*II*)

	$\mu_t^{PD,I}$	$\mu_t^{PE,I}$	$\mu_t^{PD,II}$	$\mu_t^{PE,II}$
$\mu_t^{PD,I}$	1	0.88	0.56	0.34
$\mu_t^{PE,I}$	0.88	1	0.26	0.04
$\mu_t^{PD,II}$	0.56	0.26	1	0.85
$\mu_t^{PE,II}$	0.34	0.04	0.85	1

Table 11:

Correlation of the Expected Returns from ARFIMA(1,d,0) specifications.

D.2 Monte Carlo Estimation in the presence of regimes

Compared to the expected returns, the monthly expected dividend and earnings growth rate (tables 8, 9 and Model 3) show no major higher unconditional mean than the realized values. However, for the annual frequency, a higher unconditional mean is witnessed. The stationarity and non stationarity tests depict the dividend growth as I(0) as expected. Observed earnings growth is nonstationary as shown by the I(0) and I(1) tests in the third column of table 9. This is due to the end of sample structural break which occurs in the last three months of 2008. When these observations are discarded, the series is I(0).

	Mean	SD	Skewness	Kurtosis
Intercept	0.06	0.41	-0.36	8.39
AR1	0.84	0.10	-3.48	25.7
Intercept	-1.46	6.78	0.20	17.9
ARFIMA d	0.47	0.49	-0.08	1.33
AR1	0.39	0.44	-0.01	1.31

Table 12:

Monte Carlo Experiment for 2 regimes with 83 observations. In the following experiment, a two regime switching model is set up, with regime 1 having an AR(1) process with 0.06 as intercept term and an autoregressive coefficient of 0.9. The second regime has an autoregressive coefficient of 0.8 with an intercept term of 0.09.

D.3 Univariate Models

	1926-2008				1946-2008			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	Coeff	p-val	Coeff	p-val	Coeff	p-val	Coeff	p-val
Intercept	0.002	0.5	0.201	0	0.003	0.48	0.027	0
AR(1)	0.94	0	-0.221	0	0.943	0	0.937	0
ARFIMA(d)			0.541	0			-0.123	0.05
Goodness of Fit	0.87		0.194		0.87		0.651	
Autocorrelation	4.766	0.44	4.135	0.53	4.29	0.51	4.742	0.45
Neglected ARCH	17.02	0	16.51	0	14.01	0.02	9.156	0.103
NH - Joint Parameters	0.554	1	1.265	0.05	0.717	0.2	1.07	0.1
NH - conditional variance	0.242	1	0.860	0.01	0.447	0.075	0.78	0.01
NH - Intercept	0.244	0.2	0.038	0.01	0.295	0.2	0.128	1
NH - Ar1	0.061	1	0.111	1	0.129	1	0.128	1
NH - d			0.155	1			0.209	1

Table 13:

Univariate regressions on Expected Returns sample using Price Dividend and Dividend Growth Specification. The specification for the AR(1) model is $(1 - \varphi_1 L)\mu_t^{AR(1)} = \varphi_0 + \eta_t$. For the ARFIMA model, the specification is $:(1 - \varphi_1 L)^d = \varphi_0 + \eta_t$. NH stands for the Nyblom-Hansen test which tests whether there are excessive variation in the estimated parameters. Neglected ARCH is a first pass misspecification test of conditional heteroscedasticity in the model.

	1926-2008				1946-2008			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	Coeff	p -val	Coeff	p -val	Coeff	p -val	Coeff	p -val
Constant	0.011	0.42	0.081	0	0.012	0.44	0.020	0.07
AR(1)	0.834	0.	0.804	0	0.851	0	1.050	0
ARFIMA(d)			-0.272	0.24			-0.111	0.01
Goodness of Fit	0.61		0.364		0.622		0.851	
Autocorrelation	8.86	0.11	3.22	0.67	14.88	0.25	5.112	0.402
ARCH	38.43	0	4.303	0.507	11.49	0.49	5.098	0.404
NH - Joint Parameters	0.823	0.2	0.925	0.2	1.127	0.05	0.968	0.2
NH - conditional variance	0.501	0.05	0.593	0.05	0.547	0.05	0.467	0.075
NH - Intercept	0.295	0.2	0.102	1	0.444	0.075	0.147	1
NH - Ar1	0.203	1	0.049	1	0.585	0.05	0.142	1
NH- d			0.051	1			0.212	1

Table 14:
Univariate regressions on Expected Returns 1946-2008

	1926-2008				1946-2008			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	Coeff	p -val	Coeff	p -val	Coeff	p -val	Coeff	p -val
Constant	0.018	0	0.015	0.01	0.012	0	0.010	0
AR(1)	0.100	0.46	0.191	0	0.382	0.05	0.695	0
Goodness of Fit	0.01		0.036		0.145		0.484	
Autocorrelation	4.157	0.527	3.852	0.571	3.619	0.61	6.156	0.291
ARCH	6.437	0.266	15.32	0	39.86	0	10.49	0.062
NH- Joint Parameters	0.871	0.1	1.154	0.05	1.063	0.05	1.040	0.05
NH- conditional variance	0.704	0.025	0.927	0.01	0.927	0.01	0.769	0.01
NH - Intercept	0.287	0.2	0.051	1	0.110	1	0.280	0.2
NH - Ar1	0.116	1	0.067	1	0.074	1	0.331	0.2

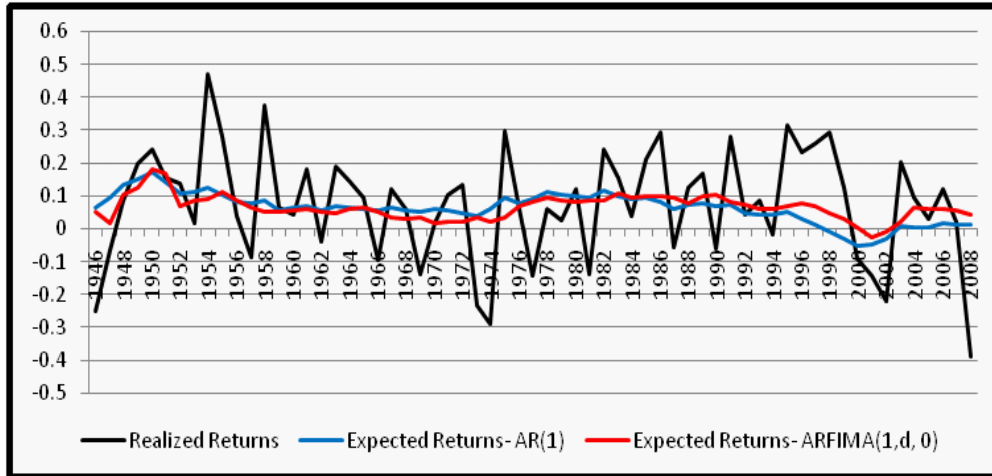
Table 15:
Univariate regressions on Dividend Growth 1926-2008

	Monthly				Annually			
	AR(1)		ARFIMA(1,d,0)		AR(1)		ARFIMA(1,d,0)	
	Coeff	p -val	Coeff	p -val	Coeff	p -val	Coeff	p -val
Constant	0.021	0	0.018	0.12	0.011	0.21	0.025	0
AR(1)	-0.024	0.86	0.601	0	0.042	0.82	0.683	0
Goodness of Fit	0.001		0.359		0.002		0.467	
Autocorrelation	3.627	0.46	5.380	0.25	4.022	0.55	8.2406	0.14
ARCH	3.460	0.48	3.215	0.52	4.78	0.44	21.5726	0.
NH- Joint Parameters	0.410	1	0.669	1	0.597	1	0.955	0.07
NH- conditional variance	0.266	0.2	0.289	0.2	0.401	0.07	0.843	
NH - Intercept	0.060	1	0.267	0.2	0.094	1	0.071	1
NH - Ar1	0.058	1	0.293	0.2	0.015	1	0.068	1

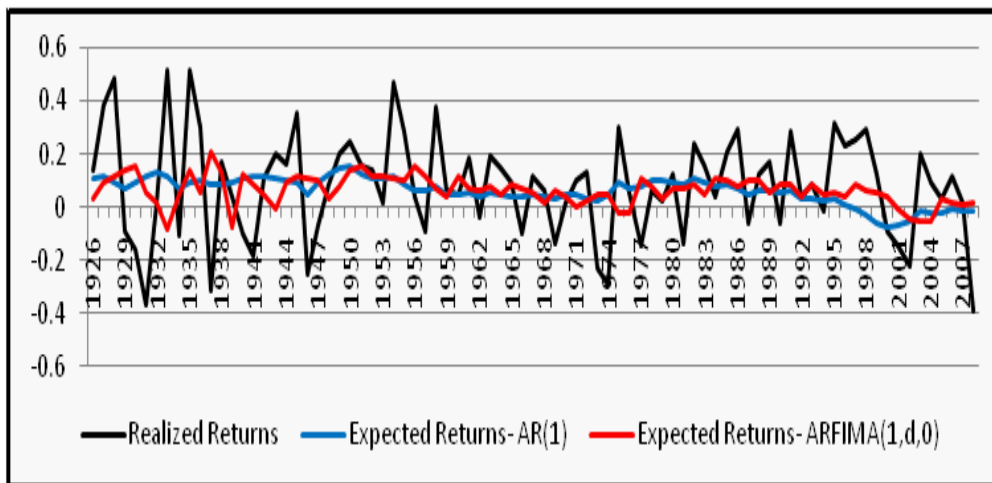
Table 16:
Univariate regressions on Expected Dividend Growth 1946-2008

E Graphical Plots

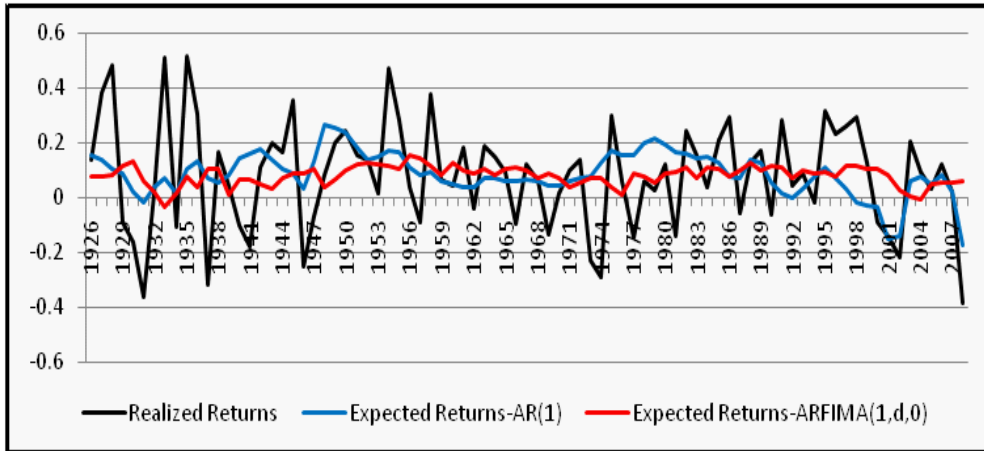
E.1 Expected Return



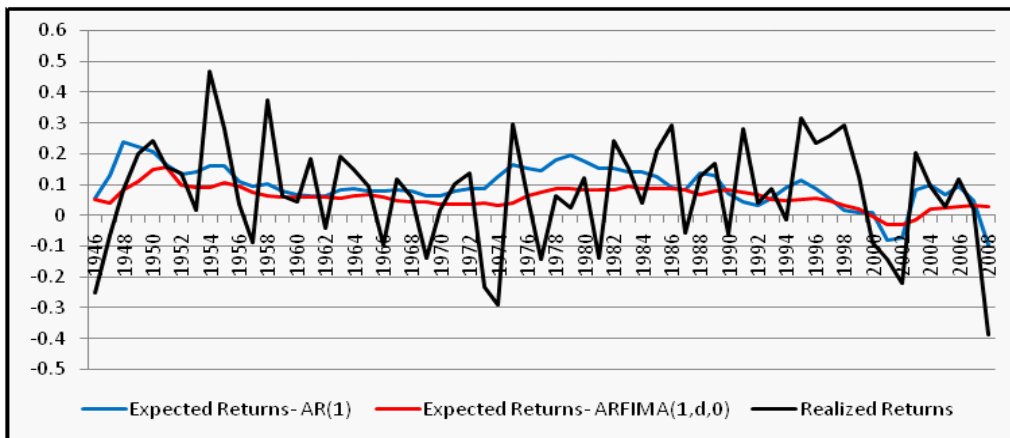
Plot of Expected Returns for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1946-2008 using dividend data.



Plot of Expected Returns for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1926-2008 using dividend data.

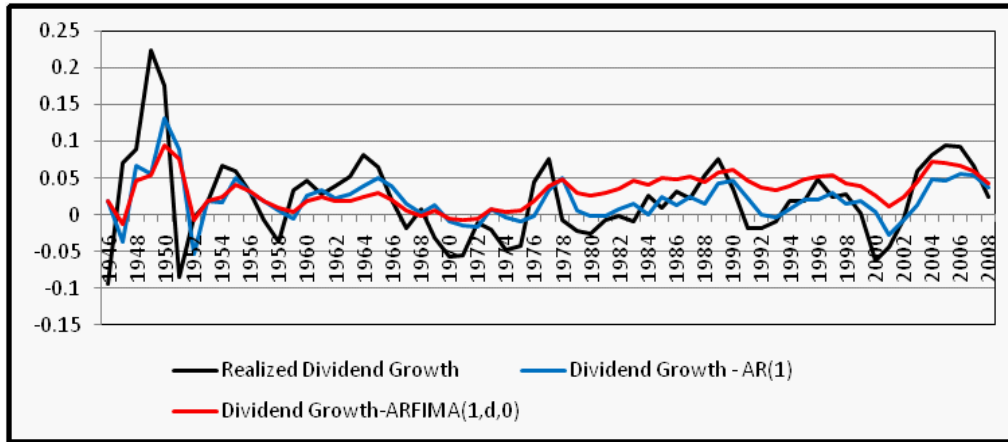


Plot of Expected Returns for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1926-2008 using earnings growth.

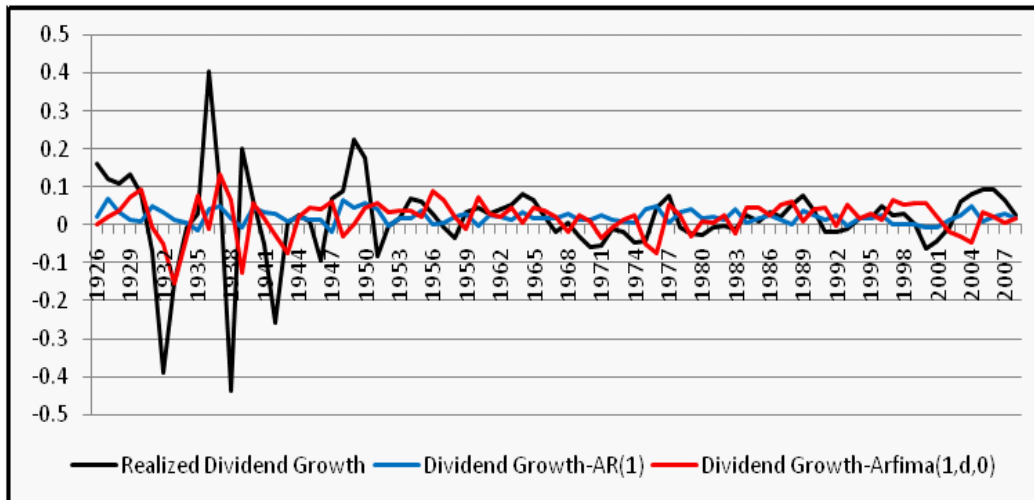


Plot of Expected Returns for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1946-2008 using earnings data.

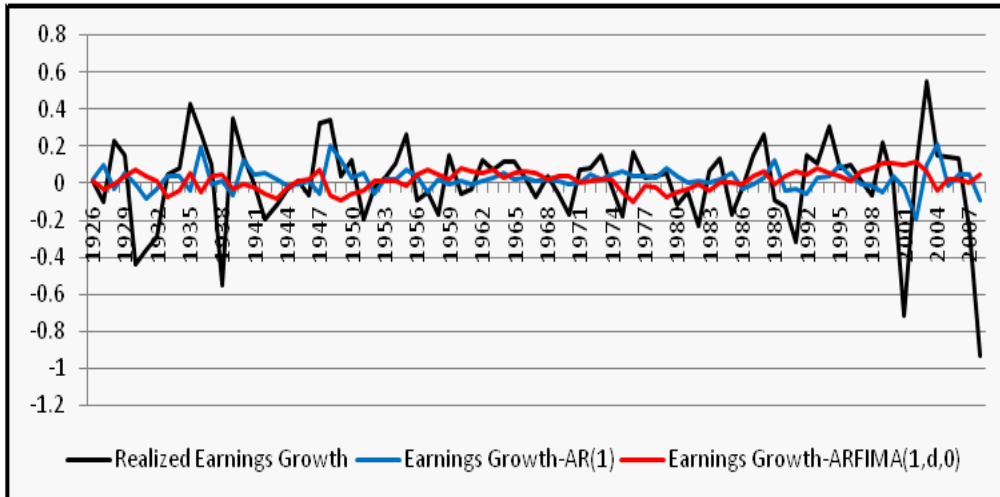
E.2 Dividend and Earnings growth



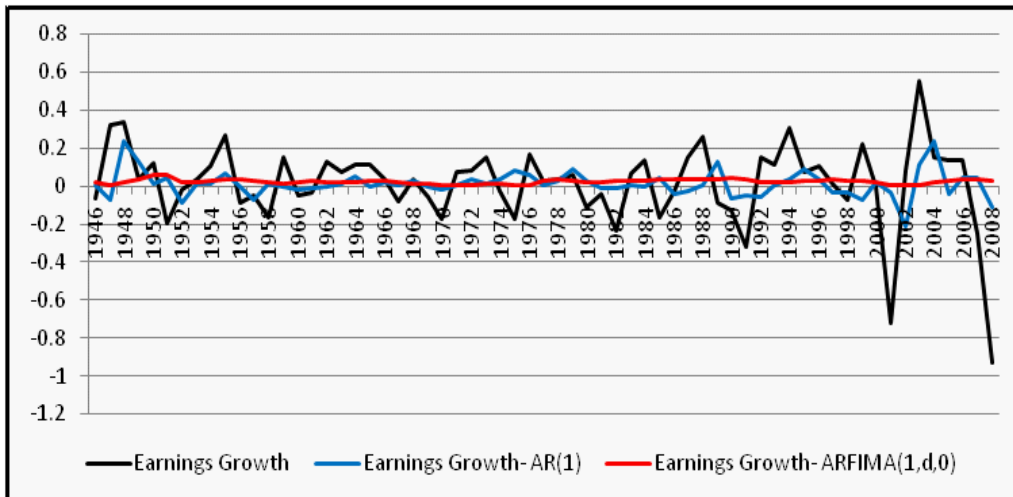
Plot of Dividend Growth for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1946-2008 using earnings growth.



Plot of Dividend Growth for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1926-2008 using earnings growth.

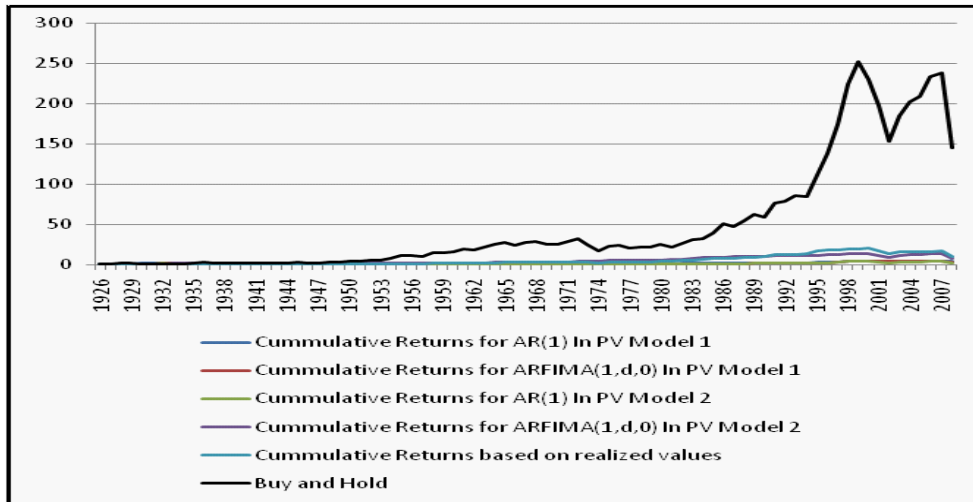


Plot of Earnings Growth for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1926-2008 using earnings growth.

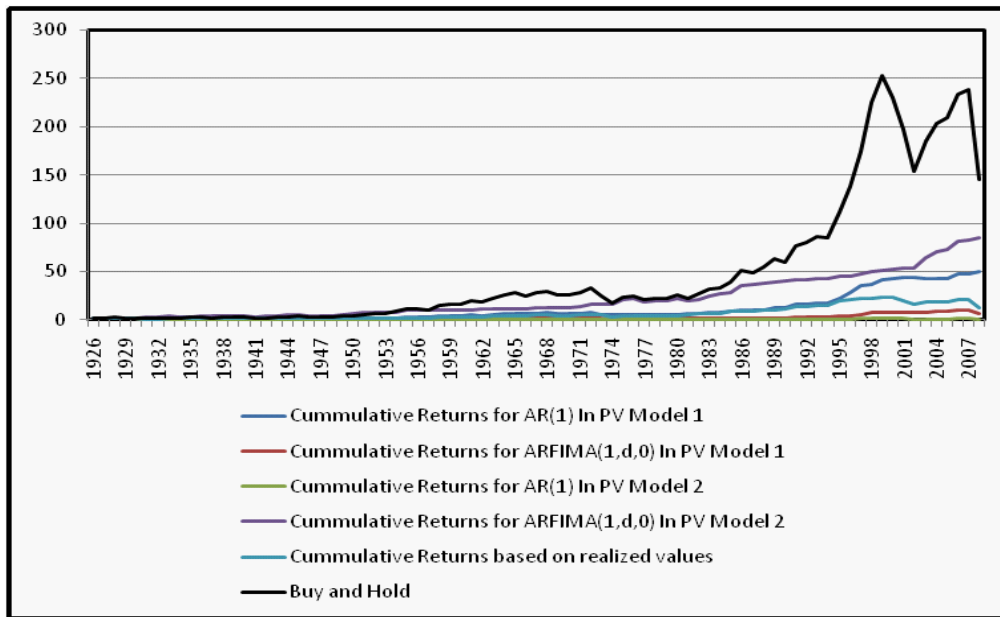


Plot of Earnings Growth for AR(1) and ARFIMA(1,d,0) and Realized Returns for 1946-2008 using earnings growth.

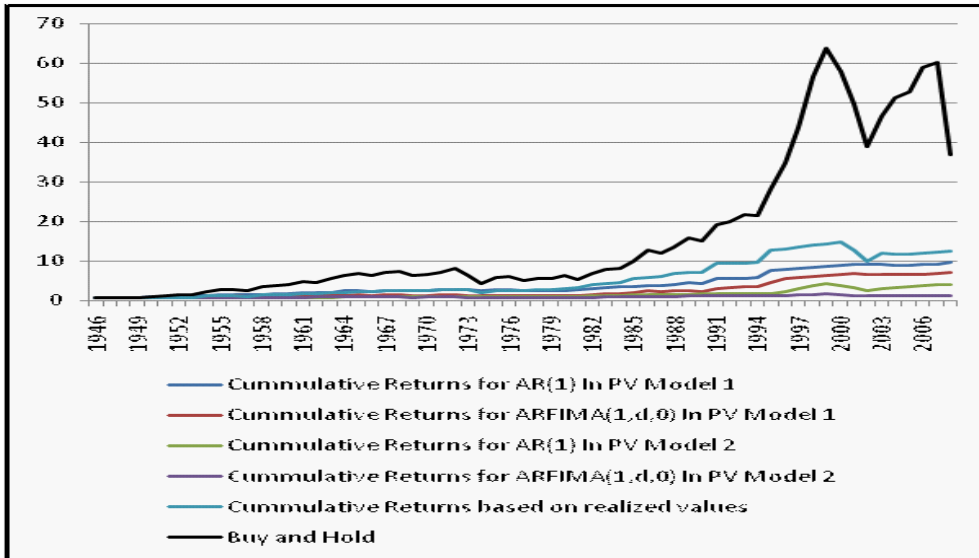
F Plot of Trading Returns



Cummulated Returns based on the the price dividend ratio for the sample 1926-2008.



Cummulative Returns under different strategies based on the price earnings ratio 1926-2008



Cumulated Returns for the period 1946-2008 using the price dividend ratio.

