Long-Run Factors and Fluctuations in Dividend/Price*

Carlo A. Favero
Bocconi University&IGIER

Arie E. Gozluklu
Bocconi University

Andrea Tamoni
Bocconi University

This Version: May, 2009

Abstract

The dynamic dividend growth model linking the log dividend yield to future expected dividend growth and stock market returns has been extensively used in the literature for forecasting stock returns. The empirical evidence on the performance of the model is mixed as its strength varies with the sample choice. This model is derived on the assumption of *stationary* log dividend-price ratio. The empirical validity of such hypothesis has been challenged in the recent literature (Lettau&Van Nieuwerburgh, 2008) with strong evidence on a time varying mean, due to breaks, in this financial ratio. In this paper, we show that the slowly evolving mean toward which the dividend price ratio is reverting is driven by a demographic factor and a technological trend. We also show that an empirical model using information in long-run factors overperforms virtually all alternative models proposed in the literature within the framework of the dynamic dividend growth model. Finally, we exploit the exogeneity and predictability of the demographic factor to simulate the equity risk premium up to 2050.

KEYWORDS: error correction model, long run predictability, equity premium, cointegration, stock market, demographics.

J.E.L. CLASSIFICATION NUMBERS: G14, G19, C10, C11, C22,C53.

---

*Paper presented at Tilburg University, at the conference "Financial and Real Activity", Paris, at ICEEE 2009 in Ancona and Dondena Seminar at Bocconi University and LSE. We thank our discussant Philippe Andrade and participants both in Tilburg, Paris, Ancona, Milan and London for stimulating discussions. We also thank Massimiliano Marcellino, Gino Favero, Sasson Bar-Yosef, Francesco Billari, Joachim Inkmann, Sami Alpanda, Ralph Koijen for helpful discussions and Andrew Mason for providing us the data on demographic dividend. Carlo A. Favero gratefully acknowledges financial support from Bocconi University.

1 Introduction

Stock market predictability has been an active research area in the past decades. After a long tradition of the efficient market hypothesis (Fama, 1970) that implies that returns are not predictable, the recent empirical literature has moved toward a view of predictability of returns (see, for example, Cochrane, 2007). There is, however, an ongoing debate on the robustness of the predictability evidence and its exploitability from a portfolio allocation perspective (Boudoukh et al., 2008; Goyal&Welch, 2008).

Most of the available evidence on predictability can be framed within the dynamic dividend growth model proposed by Campbell&Shiller (1988). The model of Campbell& Shiller (1988) uses a loglinear approximation to the definition of returns on the stock market. Under the assumption of stationarity of the log of price-dividend ratio pd_t, this variable is expressed as a linear function of the future discounted dividend growth, Δd_{t+j} and of future returns, h^s_{t+j} :

$$pd_t = \overline{pd} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[(\Delta d_{t+j} - \bar{d}) - (h^s_{t+j} - \bar{h})] \tag{1}$$

where $\overline{pd}$, the mean of the price-dividend ratio, $\bar{d}$, the mean of dividend growth rate, $\bar{h}$, the mean of log return and ρ are constants. Once the future variables are expressed in terms of observables (1) can be used to derive an equilibrium price p^*_t as a function of present dividends and future expected dividends and returns; then a forecasting model for logarithmic return is naturally derived by estimating an *Error Correction Model* (ECM) for stock prices:

$$\Delta p_{t+1} = \beta_0 - \beta_1(p_t - p^*_t) + u_t. \tag{2}$$

(2) ensures long-run convergence of stock prices to equilibrium prices allowing for the possibility of short-run disequilibria. This basic relation allows to classify different forecasting regression of stock market returns in terms of different approaches to proxy the future expected variables included in the linearized relations. The classical Gordon growth model (1962) is obtained by augmenting (1) with the hypotheses of constant dividend growth, $E_t\Delta d_{t+j} = g$, and constant expected returns, $E_t h^s_{t+j} = r$. The so-called FED model (Lander et al., 1997) proposing a long-run relation between the price-earning ratio and the long-term bond yield can be understood by substituting out the no-arbitrage restrictions in (1) $E_t h^s_{t+j} = E_t(r_{t+j} + \phi^s_{t+j})$ and then by assuming constant dividend growth, some relation between the risk premium on long-term bonds and the risk premium on stocks, and a stationary (log) earning price ratio. The extension of the FED model proposed by Asness (2003) removes the assumption of proportionality between the stock market risk premium and the bond market risk premium and augments the standard FED model by adding the ratio between the historical volatility of stock and bonds. Lettau and Lud-

vigson (2001, LL henceforth) analyze a linearized version of the consumer intertemporal budget constraint to show that excess consumption with respect to its long-run equilibrium value, a linear combination of labour income and financial wealth, may predict future return on total wealth. If future returns on total wealth are correlated with future stock market return, then excess consumption should forecast future stock market returns. They introduce the well-known cointegrating vector, *cay*, including consumption, assets and income and show empirical evidence strongly supporting their conjecture. In their proposed framework *cay* proxies p_t^* by predicting future discounted returns without concentrating on dividend growth. Julliard (2004) refines the LL contribution by observing that the total return on wealth reflect both returns on financial capital and returns on human capital, therefore the predictive power of excess consumption for stock market returns could be strengthened by controlling for returns on human capital. Labour income growth is proposed as a proxy to control for returns of human capital added to the model on top of *cay*. Ribeiro (2004) also highlights the importance of labour income in predicting future dividends and posits vector error correction model (VECM) for dividend growth and future returns with two cointegrating vectors defined as $(d_t - y_t)$ and $(d_t - p_t)$. Finally, Lamont (1998) argues that the log dividend payout ratio $(d_t - e_t)$ is the most appropriate proxy for future stock market returns and includes it in his specification. The second stage equations (2) based on all these models delivered some degree of predictability, in terms of significance of β_1. However, the degree of predictability varies with the chosen sample and so does the relative performance of different models (see Ang and Bekaert (2007)).

Such mixed evidence of predictability has been recently related to the potential weakness of the fundamental hypothesis of the dynamic dividend growth that log dividend-price ratio is a stationary process (Lettau&Van Nieuwerburgh, 2008, LVN henceforth). LVN show evidence on the breaks in the constant mean $\overline{pd}$ and assert that correcting for the breaks improves predictive power of the dividend yield for stock market excess returns. Interestingly, LVN also give some hints on possible causes for the breaks arising from economic fundamentals due to technology innovations, changes in expected return, etc. but do not explore further the possible effects of fundamentals. Breaks are modelled via a purely statistical methods without any explicit relation with economic fundamentals.

In this paper, we pursue two distinct aims. First, we show that the predictions of the theoretical model by Geanakoplos et al. (2004) that demographic factors, along with a correction for trends, explain fluctuations in the dividend yield is supported by annual US data. We then exploit stability analysis for long-run economic relationships to construct an *equilibrium* dividend-price ratio. Second, we use our measure of disequilibrium obtained as the difference between the actual dividend yield and the equilibrium dividend yield for forecasting market excess returns at different horizons (up to 10 years) and evaluate the forecasting performance of the model based on the corrected dividend-price ratio against different alternative specifications.

The paper is structured as follows. In the next section we provide evidence on the lack of cointegration between log of dividends and stock prices. In section III, we describe the cointegration framework and estimation of cointegration relations. Next we devote a section on forecasting short horizon, followed by a section on forecasting longer horizons up to 10 years and Bayesian model averaging analysis. Then we provide out-of sample evidence. In section V, we introduce different vector error correction (VECM) specifications and simulate the equity premium for the next few decades. The last section concludes.

2 (Non)-Stationarity of Dividend-Price Ratio

In this section, we consider a long sample of annual data (1909-2006), to analyze cointegration between dividends and stock prices and stationarity of the (log) dividend-yield. We report in Figure 1 the time-series of $(d_t - p_t)$.

Insert here Figure 1

The crucial assumption for the validity of the linearized dividend growth model is that this variable is stationary, i.e. that there exists a cointegrating vector with coefficient restricted to $(1, -1)$ between d_t and p_t. The visual inspection of the time series suggest some intuitive support for the recent evidence on non-stationarity (Ribeiro, 2004; LVN, 2007)[1]. Differently from LVN we do not use recursive Chow test to identify break points but we analyze the possibility of breaks and non-stationarity by concentrating on the evidence of cointegration with a (-1,1) vector between d_t and p_t. We follow Warne et al. (2003) to study the non-zero eigenvalues of the matrix describing the long-properties of a bivariate VAR for d_t and p_t used in the Johansen (1991) approach to cointegration analysis.

We consider the following statistical model:

$$\mathbf{y}_t = \sum_{i=1}^{n} \mathbf{A}_i \mathbf{y}_{t-i} + \mathbf{u}_t \tag{3}$$

$$\mathbf{y}_t = \begin{bmatrix} d_t \\ p_t \end{bmatrix}. \tag{4}$$

[1] Researchers in the field have different views on the stationarity of this series with contradictory evidence, but the main point is that model loses its appeal as an approximation as the series deviates from stationarity.

This model can be re-written as follows

$$\begin{aligned} \Delta \mathbf{y}_t &= \mathbf{\Pi}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Pi}_2 \Delta \mathbf{y}_{t-2} + ...\mathbf{\Pi}_{n-1} \Delta \mathbf{y}_{t-n+1} + \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{u}_t \\ &= \sum_{i=1}^{n-1} \mathbf{\Pi}_i \Delta \mathbf{y}_{t-i} + \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{u}_t, \end{aligned} \tag{5}$$

where:

$$\begin{aligned} \mathbf{\Pi}_i &= -\left(I - \sum_{j=1}^{i} \mathbf{A}_j \right), \\ \mathbf{\Pi} &= -\left(I - \sum_{i=1}^{n} \mathbf{A}_i \right). \end{aligned}$$

Clearly the long-run properties of the system are described by the properties of the matrix Π. There are three cases of interest:

1. rank $(\Pi) = 0$. The system is non-stationary, with no cointegration between the variables considered. This is the only case in which non-stationarity is correctly removed simply by taking the first differences of the variables;
2. rank $(\Pi) = 2$, full. The system is stationary;
3. rank $(\Pi) = 1$. The system is non-stationary but there is a cointegrating relationships among the considered variables. In this case $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where α is an (2×1) matrix of weights and $\boldsymbol{\beta}$ is an (2×1) matrix of parameters determining the cointegrating relationships.

Therefore, the rank of Π is crucial in determining the number of cointegrating vectors. The Johansen procedure is based on the fact that the rank of a matrix equals the number of its characteristic roots that differ from zero. The Johansen test for cointegration is based on the estimates of the two characteristic roots of $\mathbf{\Pi}$ matrix. Having obtained estimates for the parameters in the $\mathbf{\Pi}$ matrix, we associate with them estimates for the 2 characteristic roots and we order them as follows $\lambda_1 > \lambda_2$. If the variables are not cointegrated, then the rank of Π is zero and all the characteristic roots equal zero. In this case each of the expression $\ln(1 - \lambda_i)$ equals zero, too. If, instead, the rank of Π is one, and $0 < \lambda_1 < 1$, then $\ln(1 - \lambda_1)$ is negative and $\ln(1 - \lambda_2) = 0$. The Johansen test for cointegration in our bivariate VAR is based on the two following statistics that Johansen derives based on the number of characteristic roots that are different from zero:

$$\begin{aligned}\lambda_{\text{trace}}(k) &= -T\sum_{i=k+1}^{2}\ln\left(1-\widehat{\lambda}_i\right),\\ \lambda_{\max}(k,k+1) &= -T\ln\left(1-\widehat{\lambda}_{k+1}\right),\end{aligned}$$

where T is the number of observations used to estimate the VAR. The first statistic tests the null of at most k cointegrating vectors against a generic alternative. The test should be run in sequence starting from the null of at most zero cointegrating vectors up to the case of at most 2 cointegrating vectors. The second statistic tests the null of at most k cointegrating vectors against the alternative of at most $k+1$ cointegrating vectors. Both statistics are small under the null hypothesis. Critical values are tabulated by Johansen (1991) and they depend on the number of non-stationary components under the null and on the specification of the deterministic component of the VAR.

The main recursive test based on the non zero-eigenvalues is the fluctuation test suggested in Hansen and Johansen (1999). The test starts from estimation of our VAR model over the full sample. After that, we re-estimate the model (the full sample estimates of all coefficients on deterministic variables and lagged first differences are used in order to reduce volatility) and computes recursive eigenvalues and β recursively extending the end point of the recursive sample, t_1, until the full sample is covered, i.e. $t_1 = T_1, T_1+1, ...T$ where the base period is fixed at about 35 percent of the sample, i.e. $T_1 = 0.35 * T$, as suggested in Warne et al. (2003).

Figure 2 shows the time path of the recursively calculated log transformed largest non-zero eigenvalue λ_i from the VAR(2) model together with the 95% confidence bands. We took log transformed eigenvalues to obtain a symmetrical representation of the distribution of λ_i.

$$\xi_i = \log(\lambda_i/(1-\lambda_i))$$

The eigenvalue shows a remarkable amount of variability over the examination period with indication of three break points around 1950, 1980, 2000 and a clear possibility that null of at most zero cointegrating vectors cannot rejected for some relevant part of our sample. Interestingly, this evidence is consistent with that obtained using a different methodology by LVN.

Insert here Figure 2

Table 1 reports the results of the Johansen procedure applied to whole sample, and to two subsamples 1909-1954, 1955-2006.

Insert here Table 1

The null of no-cointegration cannot be rejected over both whole sample and subsamples. Note that validity of the linearized model requires a stronger assumption than cointegration to be satisfied, i.e. the existence of cointegration with restricted cointegrating coefficients.

3 Long-run Factors & the Dividend-Price Ratio

The evidence of instability of the cointegrating relation between log of stock prices and dividends undermines the validity of one of the crucial assumptions of dynamic dividend-growth model (Campbell and Shiller, 1988, Campbell, 1991), i.e. the stationarity of log dividend-price ratio, which is exploited in the loglinear approximation. Geanakoplos, Magill and Quinzii (2004, henceforth, GMQ) offer a potential solution to this problem by considering an overlapping generation model in which the demographic structure mimics the pattern of live births in the U.S. that have featured alternating twenty-year periods of boom and busts. The approach followed by GMQ is part of a strand of literature aimed at explaining stock market fluctuations with demographic factors. In an early paper, Bakshi&Chen (1994) develop two hypotheses; *life-cycle investment hypothesis* which asserts that an investor in early stage of her life allocates more wealth on housing and switches to financial assets at a later stage, and *life cycle risk aversion hypothesis* which posits that an investor's risk aversion increases with age. The authors also test the empirical implications using fraction of people in different age ranges and average age (change in average age) in U.S. estimating an Euler equation. Using post 1945 period, they provide evidence supporting both hypotheses. Starting from this literature, Erb et al. (1996) study the population demographics in international context using population and average age growth and conjecture that it provides information about the risk exposure of a particular economy. On the other hand, Poterba (2001) using age groups finds no robust relationship between demographic structure and asset returns, but hints at the strong link between dividend-price ratio and demographic variables. Goyal (2004) criticizes the use of demographic variables in levels and shows evidence that changes in demographic structure in fact provide support for the traditional lifecycle models. Most of the cited papers concentrate on the slow-moving nature of the demographic variables and their ability to predict long term asset returns (Erb et al., 1996; DellaVigna&Pollet, 2006) and risk premia (Ang&Maddaloni, 2005). Overall the empirical evidence from this literature is mixed.

We first introduce GMQ model that provides a foundation for a long-run relationship between (d_t-p_t) and demography. GMQ study the equilibrium of a cyclical overlapping generations exchange economy to show that the dividend-price ratio should be proportional to the ratio of middle aged to young adults (MY ratio). The authors propose an OLG exchange economy with a single good (income) and three periods; young, middle-aged, retired. Each agent (except retirees) has an endowment, labor income, w= (w^y,w^m,0) and there are two types of financial instruments, riskless bond and risky equity which allows agents to redistribute income over time.

In their simple base model, dividends and wages are deterministic, hence bond and equities are perfect substitutes. GMQ assume that in *odd (even)* periods a large (small) cohort N(n) enters the economy, therefore in every odd (even) period there will be $\{N, n, N\}(\{n, N, n\})$ cohorts living. They conjecture that the life-cycle portfolio behaviour (Bakshi&Chen, 1994) which suggests that agents should borrow when young, invest for retirement when middle-aged, and live off from their investment once they are retired, plays important role in determining equilibrium asset prices.

Let $q_o(q_e)$ be the bond price and $\{c_y^o, c_m^o, c_r^o\}(\{c_y^e, c_m^e, c_r^e\})$ the consumption stream in the odd (even) period. The agent born in odd period then faces the following budget constraint

$$c_y^o + q_o c_m^o + q_o q_e c_r^o = w^y + q_o w^m \tag{6}$$

and in even period

$$c_y^e + q_e c_m^e + q_o q_e c_r^e = w^y + q_e w^m \tag{7}$$

Moreover, in equilibrium the following resource constraint must be satisfied

$$N c_y^o + n c_m^o + N c_r^o = N w^y + n w^m + D \tag{8}$$

$$n c_y^e + N c_m^e + n c_r^e = n w^y + N w^m + D \tag{9}$$

where D is the aggregate dividend for the investment in financial markets. If q_o were equal to q_e, the agents would choose to smooth their consumption, i.e. $c_y^i = c_m^i = c_r^i$ for $i = o, e$, but then for values of wages and aggregate dividend calibrated from US data the equilibrium condition above would be violated leading to excess demand either for consumption or saving. To illustrate this point we refer to the calibration provided by GMQ; take $N = 79, n = 69$ as the size (in millions) of Baby Boom (1945-64) and Baby Bust (1965-84) generations[2] and $w^y = 2, w^m = 3$ to match the ratio (middle to young cohort) of the average annual real income in US. Thus we can calculate the total wage in even and odd periods using $Nw^y + nw^m$ for odd periods and $nw^y + Nw^m$ for even periods, and then given the average ratio (0.19) of dividend to wages we can compute the aggregate dividends. Therefore, assuming an annual discount factor of 0.97, which translates to a discount of 0.5 in the model of 20-year periods, if $q_o = q_e = 0.5$ were to hold and agents smooth their consumption, from the budget constraint (eq. 6-7) we obtain $c_y^i = c_m^i = c_r^i = \bar{c} = 2$, but then the resource constraint (eq. 8-9) above

[2]Hence, we obtain in even period a high MY ratio of $MY = \frac{N}{n} = 1.15$, whereas in odd period $MY = \frac{n}{N} = 0.87$ (See Figure 3a).

would have been violated[3]. Therefore, when the MY ratio is small (large), i.e. an odd (even) period, there will be excess demand for consumption (saving) by a large cohort of retirees (middle-aged) and for the market to clear, equilibrium prices of financial assets should adjust, i.e. decrease (increase), so that saving (consumption) is encouraged for the middle-aged. Thus, letting q_t^b be the price of the bond at time t, in a stationary equilibrium, the following holds

$$q_t^b = q_o \text{ when period odd}$$
$$q_t^b = q_e \text{ when period even}$$

together with the condition $q_o < q_e$. Moreover the model predicts a positive correlation between MY and market prices, consequently a negative correlation with the dividend yield.

So, since the bond prices alternate between q_o and q_e, then the price of equity must also alternate between q_t^e and q_t^e as follows

$$\begin{aligned} q_o^{eq} &= Dq_o + Dq_oq_e + Dq_oq_eq_o + ... \\ q_e^{eq} &= Dq_e + Dq_eq_o + Dq_eq_oq_e + ... \end{aligned}$$

which implies

$$\begin{aligned} DP_o &= \frac{D}{q_o^{eq}} = \frac{1 - q_oq_e}{q_oq_e + q_o} \\ DP_e &= \frac{D}{q_e^{eq}} = \frac{1 - q_oq_e}{q_oq_e + q_e} \end{aligned}$$

where DP_o (DP_e) is the dividend price ratio implied by low (high) MY in the model for the odd (even) periods.

In their empirical test of their model, GMQ define MY ratio as the proportion of the number of agents aged 40-49 to the number of agents aged 20-29, which serves as a sufficient statistic for the whole population pyramid.

We find the GMQ model particularly appealing because it provides foundation for using demographic factors to capture the slow evolving mean of dividend-price ratio. Yet, as the authors admit, their deterministic model is not sufficient[4] to explain time series prop-

[3]For instance, an agent from Baby Bust generation would enter in an even period in the model, i.e. (n, N, n) and high MY ratio, and faces the following aggregate resource constraint: $n(c_y^e - w^y) + N(c_m^e - w^m) + nc_r^e - D = 69 \times (2-2) + 79(2-3) + 69 \times 2 - 70 = -11$, where $D = 0.19(\frac{375+365}{2}) = 70$. This leads to excess saving in the economy. For equilibrium conditions to hold, the model implies that asset prices should increase and hence discourage saving in the economy (the experience we observed during 90's in US).

[4]"... in the deterministic model the fluctuations in prices caused by demographic forces alone are not sufficient to generate changes in equity prices of the order of magnitude observed in the US in the postwar period", p.3, Geneakoplos et al., 2004)

erties of dividend-price ratio[5]. In particular, there is empirical evidence (Beaudry&Portier, 2004) in favor of a stochastic trend in this financial ratio. Deterministic GMQ model is mute on this issue, since they "assume that the model has been detrended so that the systematic sources of growth of dividends and wages arising from population growth, capital accumulation and technical progress are factored out." (GMQ, p.6). They rightly address the lack of stochasticity in their model, and extend the model to include business cycle shocks in the form of random shocks to dividend and wages. We take a different route; since stock market is a claim to productive capital to the real economy, we include in our specification total factor productivity (TFP) as an empirical proxy[6] of the historical productivity level which captures the stochastic trend that was left out in the deterministic model. The choice of this series is grounded on the production based general equilibrium models (Cochrane, 1991; Jermann, 1998), where productivity level is modelled as a state variable in the economy that drives equilibrium relations. In particular, under model assumptions (Jermann, 1998), equilibrium dividend-price ratio should be a function of productivity level. Furthermore, shocks to productivity has been attributed as one of the major source of randomness in real business cycle models (RBC, Kydland & Prescott, 1982) and hence inclusion of this variable in our specification circumvents potential problems of misspecification. Therefore, we posit a model that connects the equilibrium log dividend-price ratio with aggregate demand for stock market represented by MY and an observable state variable, i.e. TFP, which controls for joint effects of many factors such as new technologies, economies of scale, managerial skill, changes in the organization of production and suggest the following cointegrating relation:

$$(d_t - p_t) = \beta_0 + \beta_3 MY_t + \beta_4 TFP_t$$

In Figure 3, we report both the historical series of MY and TFP and the out-of-sample projections[7] up to 2050. We notice that the time series behaviour of MY is characterized by slow cyclical movements, whereas TFP resembles a stochastic trend. Table 2a presents summary statistics for (log) annual excess stock market returns with respect to the risk-free rate (equity premium), log dividend-price ratio, TFP, and MY for the whole sample 1909-2006[8]. The last two tables split the whole data set into two subsamples, namely

[5] In fact, our results in robustness section shows that a trivariate model including MY Ratio is misspecified.

[6] We take a TFP series directly from the website of Bureau of Labor Statistics (BLS) for the period 1948-2006. We then extended back the data to the period 1909-1949 by using the original series provided in the classic paper by Solow (1957). We normalized the series from BLS to bring it to the same scale with the Solow data. We compared the results, both with full sample normalized data and post-war original BLS series, and the results are not sensitive to this normalization.

[7] MY projections are taken from Bureau of Census; this projection is quite accurate and we take this variable as exogenous and hence do not provide a confidence band. TFP projections, on the contrary, are obtained from the stochastic simulations using our model and one standard deviation band is provided.

[8] In Table 2b we also provide summary statistics for CRSP dataset spanning from 1926 to 2006.

1909-1954 and 1955-2006, and reports the summary statistics. We consider a sample split in 1954 in the light of the evidence provided by LVN (2008), and of the evidence we reported in the previous section with Eigenvalue analysis, i.e. a first break in $d_t - p_t$ series around the 50's.

In all the tables, the first panel shows the correlation matrix among the relevant variables. The panel below reports the univariate summary statistics of the variables, namely the arithmetic mean, median, mode, standard deviation, minimum, maximum and autocorrelation.

Insert here Table 2a - 2d

At a first glance, our first observation is that the technology and demographic variables have low correlation with equity premium, but relatively higher correlations with log dividend price ratio, these feature is robust to the sample choice. The correlation between TFP and $(d_t - p_t)$ and between MY and $(d_t - p_t)$ is negative, as the intuition and economic modelling suggest, and stable across different subsamples. Importantly, both technology and demography variables are quite persistent like $(d_t - p_t)$. In fact, DF residual based tests for the presence of a unit root in $(d_t - p_t)$, TFP_t, MY_t (not reported but available upon request) do not reject the null hypothesis of a unit root in these series.

It is interesting also to note that MY has a twin peaked behaviour with peaks roughly corresponding to the dates identified by LNV as break points for the mean of the dividend/price ratio. The whole sample correlation between MY and $(d_t - p_t)$ is as high as -0.73. This is a rather striking fact especially because the direct relation between these two variables does not take into account the potential relevance of filtering out trends explicitly cited by GMQ.

Insert here Figure 3a - 3b

Basen on Campbell&Shiller model, the long-run movements in log dividend price ratio can be driven either by movements in expected return or expected dividend growth rates. In other words, structural breaks in this ratio can be attributed either to structural changes in expected returns and/or structural changes in expected dividend growth[9]. Van Binsbergen&Koijen (2009) in a recent paper, extract these two expectation components using Kalman filter approach, and their evidence (Figure4&5) suggests that the breaks, in particular the recent one around 1990's, is caused by a structural change in expected return. This finding coincides with what GMQ model would predict, namely MY ratio is a good proxy for the long-run expected excess return[10]. Our cointegration framework materializes this relationship. In the following graph, we provide graphical evidence on the ability of slow evolving variables MY and TFP to track the movements in the mean

[9] We thank an anonymous referee for this comment.

[10] Expectations are projected foreward in the model for periods longer than a typical business cycle length.

of log dividend-price ratio. We notice that neither TFP nor MY alone is sufficient to capture the evolution of mean dividen-price ratio. In fact, a specification with only TFP, i.e. $\beta_3 = 0$ delivers just a downward-sloping trend, while only MY, i.e. $\beta_4 = 0$, tracks the general tendecy in the financial ratio, but alone is not sufficient to restore the long-run relation. Below, we provide further statistical evidence supporting this claim.

Insert here Figure 4a

We investigate the following conjecture; Once we remove the secular trend captured by MY and TFP, can we associate the residual to business cycle variations in log dividend-price ratio? In other words, we study the relation between the trend in dividend-price extracted using HP filter and the trend obtained by the linear combination of MY and TFP. The figure 4b compares the trend component of dividend-price ratio with a linear combination of MY and TFP where weights are obtained from our cointegration system and 4c compares cycle component of dividend-price ratio with the the error correction vector dp^{TD}. From the figures we can infer that once we condition upon the slow evolving variables, hence capture the trend component, the business cycle component of dividend-price ratio extracted through HP filter coincides fairly well with the error correction vector, dp^{TD}. One argue then that HP filtered dividend-price can be used instead of dp^{TD}. But, our aim is different. We believe that we gain more insight for making inference on relying on fundamentals that drive economy in the long run, rather than solely using a purely statistical method such as a latent regime switching model or a spectral method such as HP filter[11][12].

Insert here Figure4b-c

Beside the graphical evidence, we also report in Table 3 the results of the cointegrating analysis based on the Johansen (1991) procedure. In particular, we report the test based on both $\lambda_{\max}$ and λ_{trace} statistics, critical values are chosen by allowing a linear trend in the data but not in the cointegration relation. The lag length in the VAR specification is chosen on the basis of standard optimal lag-length criteria.

Insert here Table 3

The trace statistics strongly rejects the null of no cointegrating relations, and does not reject the null of at most one cointegrating vector. Therefore, we opt for a specification with a single cointegrating vector between p_t, d_t, TFP_t and MY_t, which is restricted to be

[11] HP filter has been criticized in the litarature, e.g. smoothing parameter choice can affect the results (Canova, 2007).

[12] We use as smoothing parameter λ both 100, which is standard for annual data in macroeconomics literature (Jaimovich& Siu, 2008) and 6.25 following Ravn and Uhlig (2002).

$\begin{pmatrix} -1 & 1 & \beta_3 & \beta_4 \end{pmatrix}$[13]. We estimate our cointegration relation following the maximum likelihood based Johansen procedure and report the results in Table 4a (Sample: 1909-2006). Below, we show the point estimates for the paramaters of the common trend between log dividend-price ratio, MY and TFP.

$$dp^{TD} = (d_t - p_t) + \underset{(5.39)}{0.29} \cdot TFP_t + \underset{(5.19)}{1.554} \cdot MY_t + 1.318$$

where dp^{TD} is the cointegration error from the long-run relation between $(d_t - p_t)$, MY and TFP. The long-run coefficients, β_3 and β_4, describing the impact of TFP_t and MY_t on the price-dividend ratio are both positive and significant. The null hypothesis that the coefficient on p_t is restricted to minus one, and that the coefficient on d_t is restricted to one cannot be rejected at 1% level by the test for the validity of these restrictions on the cointegrating space. This evidence is even more pronounced in the post-war sample 1955-2006 (Table 4b). In particular, the long-run coefficients are stable and much more significant compared to full sample. The restrictions $[-1\ 1]$ on log prices and dividends cannot be rejected at any conventinal level (p-value$= 0.67$). This stronger evidence for the common trend in the post-war period can be due to higher stock market participation (Vissing-Jorgensen, 2002), which implies a stronger link between MY and $(d_t - p_t)$.

Turning to the analysis of the disequilibrium correction (that we report in table 4a[14]), the α coefficients reveal that stock market returns react to disequilibrium while the restriction that α on total factor productivity and dividend growth in our CVAR is zero cannot be rejected.

Insert here Table 4a/b

To facilitate comparison of our cointegration based approach with the evidence based on the statistical analysis of breaks in the mean of $(d_t - p_t)$ provided by LVN, we report in Figure 4d three time series: $(d_t - p_t)$, $\widetilde{dp}_t$ the dividend-price ratio corrected for exogenous breaks in LVN[15] and dp_t^{TD} with an autocorrelation coefficient of 0.66. The graphical

[13] We have also experimented with two cointegrating relationships using an additional demographic variable (See section 3.1). In this case the first cointegrating relation is not different from our chosen specification and the second cointegrating vector could be restricted to a simple linear relationship among the two demographic indicators that is useful only to predict these two variables. Our results should not then be affected of our choice of concentrating to unique cointegrating vector as we never use our CVAR to predict demographic trends that we will consider exogenous and take from the Bureau of Census projections.

[14] See appendix C for model specification.

[15] Following LVN we adopt the following definition:

$$\widetilde{dp}_t = \begin{matrix} dp_t - \overline{dp}_1 & \text{for } t = 1, ..., \tau_1 \\ dp_t - \overline{dp}_2 & \text{for } t = \tau_1 + 1, ..., \tau_2 \\ dp_t - \overline{dp}_3 & \text{for } t = \tau_2 + 1, ..., T \end{matrix}$$

where $\overline{dp}_1$ is the sample mean for 1909-1954, i.e. $\tau_1 = 1954$, $\overline{dp}_2$ is the sample mean for 1955-1994, i.e.

evidence tells us that the cointegration based correction produces very similar results for the break-based correction in LVN (2008).

Insert Figure 4d

We perform stability analysis using the recursively calculated eigenvalues and the Nyblom (1989) Stability test.

Insert here Figure 5a - 5b

Our recursive analysis of the non-zero eigenvalues reveals much more stability compared to baseline case discussed in the first section of this paper, yet there is still some time variation in λ_i. There can be two sources of such time variation: time varying adjustment coefficients, $\boldsymbol{\alpha}$, or time-varying cointegrating parameters, $\boldsymbol{\beta}$. To shed more light on this issue we adopt the test of constancy of the parameters in the cointegrating space proposed by Nyblom (1989). The null hypothesis that the cointegration vectors are constant is tested against the alternative that they are not

$$H_\beta : \beta_{t_1} = \beta_0 \text{ for } t_1 = T_1.....T$$

where we use $\beta_0 = \beta_T$ (Hansen&Johansen, 1999; Warne et al., 2003). In interpreting the results it is important to note that is well known that this test has little power to detect structural change taking place at the end of the sample period (Juselius, 2006). Since we compute the Nyblom statistic for the constancy of β where its asymptotic distribution is unknown theoretically, we approximate by bootstrapping the small sample distribution (we compute 1999 bootstrap samples) using the package SVAR[16] made available by Warne. We estimate the sup-statistics to be 0.4849 (with mean-statistics = 0.2036) for a VEC model of order 1 and allowing for only one cointegration relation with the restrictions specified above. From Figure 5b we can see that the sup-statistics lies in the acceptance region of the bootstrapped distribution, hence the null hypothesis of constancy of β cannot be rejected[17].

3.1 Robustness

We found strong evidence for the model we posit. Nevertheless, one can argue, whether a more parsimonius model can suffice in explaining the fluctuations in log dividend-price ratio. To this end, in this section we test two other model specifications leaving out either TFP (a *demographic model* which would be the counterpart of the deterministic model

$\tau_2 = 1994$, and $\overline{dp}_3$ is the sample mean for 1995-2006.

[16] Available at Warne's website: http://www.texlips.net/warne/index.html

[17] We also calculated the mean-statistics, the same conclusion holds.

suggested by GQM, where stochasticity arising from TFP is not accounted for) or MY (which we label as *productivity model*), so that we can see the relative importance of two factors we propose for restoring the long-run equilibrium. We report the results in Table 4c and 4d for the post-war sample[18]. The demographic model in line with the theory exhibits some success ($\bar{R}^2 = 11\%$) in explaining price changes and MY enters significantly in the long-run relationship. Yet, the reaction of price to disequilibrium measured by the loading on the cointegrating vector is barely significant (t-stat=2.05). Moreover, the restrictions on log price and log dividend are rejected (p-value=0.00).

Insert here Table 4c

One can argue that the stochastic trend shared by TFP is the main drive of the long-run relationship we have found; this would imply a productivity model similar to the one suggested by Beaudry & Potter (2004) omitting MY ratio. In table 4d, we see that even though TFP significantly (t-stat=3.86) enters the cointegration relation, the restrictions on log prices and dividend are rejected at 10%. But more importantly, the cointegration error has no explanatory power ($\bar{R}^2 = 0.00$) on returns and the reaction of price to disequilibrium is not significant (t-stat=1.06).

All in all, models omitting either the contribution of demographic structure or productivity level to long-run fluctuations in dividend-price ratio cannot provide a complete picture. Large swings in age composition, in particular the portion of society that is relavant for aggregate demand for equity should (as the theory suggests) and does matter iff we account for the productive side of the economy.

Insert here Table 4d

Researchers generally agree upon the role of TFP in restoring the long-run relations in financial markets, yet there is a controversy in the literature on how to construct the right productivity measure. Therefore, we also consider alternative constructions of TFP. Following Beaudry&Portier(2004) we construct two measures of log TFP as

$$
\begin{aligned}
TFP_t &= \log\left(\frac{Y_t}{H_t^{\bar{s}_h} KS_t^{1-\bar{s}_h}}\right) \\
TFP_t^A &= \log\left(\frac{Y_t}{H_t^{\bar{s}_h} (CU_t KS_t)^{1-\bar{s}_h}}\right)
\end{aligned}
$$

where Y_t is the output, H_t is hours, KS_t is the capital services , $\bar{s}_h$ is the average labor share(67.66%) and CU_t is the capacity utilization. All variables are collected from Bureau of Labor Statistics(BLS) and Bureau of Economic Analysis(BEA).

[18] The results for the full sample(1909-2006) is even weaker.

The first series is standard in the literature, while the second one is an adjusted TFP measure that includes capital utilization data to correct for possible variable rate of capital utilization. We obtain consistent results; the cointegrating vector error coefficients do not change significantly, both in terms of magnitude and statistical significance. Moreover, the implications of the model on price changes remain the same[19].

Finally, given the trending behaviour of TFP, we analyze the effect of replacing the stochastic trend with a linear deterministic time trend. The table 4e below suggests that a linear deterministic trend has some success in capturing the long-run swings in dividend-price ratio, but the Johansen cointegration test reveals that there exists a unique cointegration vector only in the case of a stochastic trend. Moreover, the χ^2 statistics testing for the restriction of the coefficients on log dividend and log price fails to reject the model with TFP, whereas it rejects the model with a linear time trend at any conventional confidence levels. This evidence sheds light on the importance of stochastic component in TFP in restoring the long-run relationship.

Insert here Table 4e

To assess the robustness of our cointegrating relationship in identifying the low frequency relation between stock market and demographics, we evaluate the effect of augmenting our baseline relation with an alternative demographic factor. Research in demography has recently concentrated on the economic impact of the *demographic dividend* (Bloom et al., 2003; Mason&Lee, 2005). The demographic dividend depends on a peculiar period in the demographic transition phase of modern population in which the lack of synchronicity between the decline in fertility and the decline in mortality typical of advanced economies has an impact on the age structure of population. In particular a high support ratio is generated, i.e. a high ratio between the share of the population in working age and the share of population economically dependent. Empirical evidence has shown that the explicit consideration of the fluctuations in the support ratio delivers significant results in explaining economic performance (see Bloom et al., 2003). The concept of *Support Ratio (SR)* has been precisely defined by Mason and Lee (2005) as the ratio between the number of effective number of producers, L_t, over the effective number of consumers, N_t (Mason&Lee, 2005). In practice we adopt the following empirical proxy:

$$SR = a2064/(a019 + a65ov)$$

where $a2064$: Share of population between age 20-64, $a019$: Share of population between age 0-19, $a65ov$: Share of population age 65+[20].

[19] We did not report these results, but results are available upon request.

[20] We have checked robustness of our results by shifting the upper limit of the producers to the age of 75. This is consistent with the evidence on the cross-sectional age-wealth profile from Survey of Consumer

Table 4b shows that the restrictions that the coefficient on SR is zero in the cointegrating vector cannot be rejected at 5% confidence level.

Insert here Table 4f

4 Predictability of Stock Market Returns

The long-run analysis of the previous section has shown that there exist a stable cointegrating vector between the dividend-price ratio, total factor productivity and the ratio of the number of agents aged 40-49 to the number of agents aged 20-29. Moreover, the estimated adjustment coefficients α in the CVAR indicates that only that stock market returns adjust in presence of disequilibrium.

In this section we provide more evidence on this issue by concentrating on excess returns to provide within sample and out-of-sample evidence on predictability.

4.1 Within Sample Evidence

Our within sample evidence is constructed by comparing raw and adjusted dividend-price ratios for the sample 1909-2006, 1909-1954 and 1955-2006. We consider a sample split in 1954 in the light of the evidence in provided by LVN. In practice, we consider the following set of regressions where excess returns at different horizons (one to ten years), $r_{m,t+H}-r_{f,t+H}$, are projected on a constant and the relevant measure of the dividend-price ratio

$$\begin{aligned} r_{m,t+H} - r_{f,t+H} &= \gamma_0 + \gamma_1 z_t + \varepsilon_{t+H} \\ z_t &= dp_t, \widetilde{dp}_t, dp_t^{TD}, dp_t^{CFN} \end{aligned}$$

where $dp_t, \widetilde{dp}_t, dp_t^{TD}$ are defined as above and dp_t^{CFN} is the new measure of the cash flow based net payout yield (dividends plus repurchases minus issuances) suggested by Boudoukh et al. (2007)[21]. We included the last series in our analysis, since the authors attribute the swift decline in dividend-price ratios starting from the 80's to the shifts in corporate payout policies. Their suggested measures correct for these shifts and it can therefore provide an alternative explanation for the non-stationarity of dividend-price

Finances, provided in Table 1 of Poterba(2001), which shows that the population share between 64-74 still holds considerable amount of common stocks. Results are available upon request.

[21] The series is taken from Prof. Roberts website. The authors suggest 4 new series, we test all the series and report the results with the best performing series.

ratio. [22] The reuslts reported in Table 5a-5c illustrate the empirical evidence.

Insert here Table 5a - 5c

First we note that over the entire sample (1909-2006) dp_t^{TD} is always significant and the pattern of adjusted R^2 suggests that the correction for non-stationarity improves upon in-sample predictability at all horizons. At 1-year horizon, adjusted R^2 is at 11.1%, it reaches its peak 40.3% at 5-years horizon and remains above 20% until 10 years. From Table 5b, we note that before the first structural break, the log dividend price ratio has forecasting power for excess returns (Newey West- corrected t-statistics in the table are always significant at 95%, except for 1 year). When we restrict our data sample to 1955-2006, we observe that dp_t loses almost all its forecasting power at very short horizons from 1 to 4 years. Instead, once we correct dp_t using the information in demography, we maintain similar forecasting power exhibited in the entire sample, even at short horizons. Consistent with the argument of Lettau et al. (2006), we observe that even though dp_t^{CFN} performs well in the whole sample, it exhibits similar performance to dp_t in the post war data. On the other hand, $\widetilde{dp}_t$ is also shows significant consistently both in full sample and subsamples, but performs worse than dp_t^{TD} both in terms of t-statistics and adjusted R^2.

On the basis of these results, we proceed to compare the performance dp_t^{TD} as a predictor with that of the other financial ratios used in the framework of the dynamic dividend growth model over the sample 1955-2001.

We do so by first considering alternative univariate models based on the different ratios:

$$\begin{aligned} r_{m,t+H} - r_{f,t+H} &= \gamma_0 + \gamma_1 z_t + \varepsilon_{t+H} \\ z_t &= dp_t^{TD}, RREL_t, de_t, term_t, default_t, cay_t, cdy_t, pe_t \end{aligned}$$

where $RREL_t$ is the detrended short term interest rate (Campbell, 1991; Hodrick, 1992), de_t and pe_t are the log dividend earnings ratio and log price earning ratio, respectively (Lamont, 1998). $term_t$ is the long term bond yield (10Y) over 3M treasury bill, $default_t$ is the difference between the BAA and the AAA corporate bond rates, cay_t and cdy_t are cointegration variables introduced by LL (2001, 2005).

Insert here Table 6a - 6b

[22] This is not uncontroversial. Lettau et al. (2006) argue these shifts are unlikely to explain the full decrease in this financial ratio, since other financial valuation ratios such as earning-price ratios witness similar declines

We obtain consistent results with the literature. Table 6 suggests that in a univariate model specification one should include cay_t and dp_t^{TD} in all horizons (except 10 years) and both variables have substantial predictive power with in-sample $\bar{R}^2$ slightly favoring cay_t. Based on the evidence of Table 6, one can also consider other potential candidates for forecasting excess return such as $RREL_t$, $Default_t$ or pe_t .

Therefore, we also consider a forecasting model exploiting simultaneously all the available information.

$$\begin{aligned} r_{m,t+H} - r_{f,t+H} &= \gamma_0 + \gamma_1 \mathbf{x}_t + \varepsilon_{t+H} \\ \mathbf{x}_t &= \left[\begin{array}{ccccccccc} dp_t^{TD} & dp_t^{CFN} & de_t & pe_t & cay_t & cdy_t & RREL_t & term_t & default_t \end{array} \right]^T \end{aligned}$$

To deal with the problem of potential multicollinearity between regressors in the multivariate model we adopt Bayesian Model Averaging. The Bayesian approach allows us to account also for model uncertainty in our linear regression framework. In our analysis we follow Raftery et. al (1997)[23], instead of conditioning on a single selected model, we base our inference on averaging over a set of possible models[24]. Averaging over all possible models provides provide better predictive power than considering a single model, hence the model uncertainty problem is alleviated. Basing inferences on a single "best" model as if the single selected model were the true one underestimates uncertainty about excess returns. The standard Bayesian solution to this problem is

$$\Pr(r_{m,t+H} - r_{f,t+H} | \text{ Data}) = \sum_{i=1}^{K} \Pr(r_{m,t+H} - r_{f,t+H} | \text{ M}_K, \text{Data}) \Pr(\text{M}_K | \text{ Data})$$

where $M = \{\text{M}_1,\text{M}_2....,\text{M}_K\}$ denotes the set of all models considered. This is an average of the posterior distributions under each model weighted by corresponding posterior model probability which we call Bayesian model averaging (BMA). Below we report results

Insert here Table 7a1 -7a2

Insert here Table 7b1-7b2

In the tables we provide the BMA posterior estimates of the coefficients of the regressors (with t-statistics in parentheses) in a multivariate regression for 1,3,5,7,10 years horizon along with the regression R^2 statistics. In a separate table we provide the summary of model selection analysis. We report the two models with highest probability

[23] We run the bma_g function provided in Le Sage toolbox: http://www.spatial-econometrics.com/

[24] A complete Bayesian solution would be averaging over all possible combinations of predictors, but we reduce the set of possible models to a subset of models following Raftery et.al (1997).

and highest number of visits among all the models considered for Bayesian analysis. We also report cumulative probability of each variables, i.e. the probability that a variable appears across all the models considered. We have used flat priors[25] and 50000 draws for the analysis. The sample considered for the analysis spans from 1952-2001[26], the longest sample we have data for each variable. We notice that consistent with the previous section on univariate analysis, both cay_t and dp_t^{TD} are the most selected variables (based on cumulative probability of entering a model visited in BMA analysis) for predicting excess returns. In particular, dp_t^{TD} is selected in models from 1 to 5 years, while cay_t is favored in relatively longer horizons.

4.2 Out-of-Sample Evidence

In this section we follow Goyal and Welch (2008), and we assume that the real-world investor, who does not have access to ex-post information, would have to estimate the prediction equation only with data available strictly before the prediction point, and then make an out-of-sample prediction. Indeed we are not really conducting a true out of sample test since our out-of-sample regressions rely on the very same data points that were used in the in-sample tests to identify the proposed predictors. Therefore we call this a *pseudo* out-of-sample forecast exercise.

We run rolling forecasting regressions for one, three and five years using as an initialization sample 1952-1981, keep the rolling window of 30 data points and make the first forecast in 1982, so the forecasting period includes the anomalous period of late 90's where the sharp increase in stock market index weakens the forecasting power of financial ratios. We select predictors on the basis of our within sample evidence, therefore we focus only on cay_t and dp_t^{TD}.[27] In particular, we consider both univariate and bivariate models and compare the forecasting performance with historical mean benchmark. In the first two columns of Table 8a we report the adjusted $\bar{R}^2$ and the t-statistics using the full sample 1952-2006. Then we also report mean absolute error (MAE) and root mean square error (RMSE) calculated based on the residuals in the forecasting period, namely 1982-2006. The first column of out-of- sample panel report the out-of-sample R^2 statistics

[25] The hyperparameters ν, λ and ϕ are set 4, 0.25 and 3, respectively. See Raftery et al.(1997) for selection of prior distributions.

[26] We also report (Table 7b.1 and Table 7b.2) as robustness check results for the sample that spans the period 1955-2006, where we do not include cdy_t.We estimate our cointegrated vector dp_t^{TD} using only the data points included in the sample. As an additional robustness check, we also include Boudoukh et al. (2007) series for cash-flow based payout yield net of issuances.

[27] We follow Stock and Watson (1993) dynamic least squares (DLS) with 1 lead/lag length to estimate the cointegrating parameters. To have a conservative forecast exercise we reestimate the coefficients of dp_t, TFP_t, and MY_t with data up to the obervation points, whereas for cay_t we use the full sample coefficents (i.e. cay_p (cay post) in Goyal&Welch (2007) terminology).

(Campbell&Thomson, 2008) which is computed as

$$R_{OS}^2 = 1 - \frac{\sum_{t=t_0}^{T} (r_t - \hat{r}_t)^2}{\sum_{t=t_0}^{T} (r_t - \bar{r}_t)^2}$$

where $\hat{r}_t$ is the forecast at $t-1$ and $\bar{r}_t$ is the historical average estimated until $t-1$. In our exercise, $t_0 = 1982$ and $T = 2006$. If R_{OS}^2 is positive, it means that the predictive regression has lower mean square error than the prevailing historical mean. In the last column, we report the Diebold-Mariano (DM) t-test for checking equal-forecast accuracy from two nested models for forecasting h-step ahead excess returns.

$$DM = \sqrt{\frac{(T + 1 - 2 * h + h * (h-1))}{T}} * \left[\frac{\bar{d}}{\widehat{se}(\bar{d})}\right]$$

where we define e_{1t}^2 as the squared forecasting error of prevailing mean, and e_{2t}^2 as the squared forecasting error of the predictive variables, $d_t = e_{1t}^2 - e_{2t}^2$, i.e. the difference between the two forecast errors, $\bar{d} = \frac{1}{T}\sum_{t=t_0}^{T} d_t$ and $\widehat{se}(\bar{d}) = \frac{1}{T}\sum_{\tau=-(h-1)}^{h-1}\sum_{t=|\tau|+1}^{T}(d_t - \bar{d}) * (d_{t-|\tau|} - \bar{d})$. A positive DM t-test statistics indicates that the predictive regression model performs better than the historical mean.

Insert here Table 8a - 8c

First we notice that the 1-year ahead out-of-sample performance deteriorates for the variables considered compared to the in-sample performance. Nonetheless, in the out-of-sample the relative deterioration with respect to prevailing mean becomes evident in case of dp_t and cay_t while the other candidates maintain a lower MAE and RMSE than the one of prevailing historical mean. In particular, the bivariate model including cay_t and dp_t^{TD} performs best based on 3 out of 4 out-of-sample performance criteria. In 3-year ahead out-of-sample forecast, models including cay_t, dp_t^{TD} or both outperform forecasts using prevailing mean. When we move to 5-year ahead out-of-sample forecast the results are consistent with 3-year ahead out-of-sample forecast, but slightly favoring a univariate model including only cay_t over a bivariate alternative.

In the figures below we plot the cumulative squared prediction errors of prevailing mean minus the cumulative squared prediction error of dp_t and dp_t^{TD} where a positive line means that the predictive regression improves upon historical mean (the zero line is drawn in the figure to graphically detect performance).

Insert here Figure 6a-6b

In figure 6a, we use all the available data from 1909 until 1954 for initial estimation and then we recursively calculate the cumulative squared prediction errors until sample end, namely 2006. Consistent with the breaking point analysis, we notice that around the breaking points of 1954 and early 1980's and late 90's the financial ratio dp_t predict worse than prevailing mean (note the decrease in the cumulative squared prediction error line around the points) , while the corrected dp_t, i.e. dp_t^{TD} performs as well as the historical mean around the 50's and then improves upon the benchmark, in particular during last stock market bubble. Figure 6b repeats the same exercise using a larger initial estimation period, namely 1909-1967, we notice that we we exclude the very recent data points, dp_t still performs well compared to the historical benchmark, consistent with the literature which favors this financial ratio as a major predictive regressor, but once we also include the data points around the millennium, this financial ratio loses its forecasting power (as evident in the figure), whilst dp_t^{TD} even improves its performance upon the historical benchmark, thanks to the correction mechanism driven by fundamentals which are immune to temporary bubbles.

5 Equity Premium Projections

One of the interesting aspects of the demographics variable is that long-forecasts for these variables are readily available. In fact, the Bureau of Census provides on its website projected data up to 2050. Having now shown that the CVAR model introduced in section 1 provides forecasts for stock market returns that are at least comparable to those produced by the best available models suggested in previous literature based on financial ratios, we go back to it and use it to produce projections for stock market equity premia over the period 2007-2050. For there is no validation for future projection performance of our model[28], we first use our model to form (pseudo) out-sample equity premium forecasts, which can then be validated against realized excess returns in our sample. To start the pseudo out-of-sample exercise, we split our sample into two, the first part of the sample (up to 1990) is used for estimation, and we solve forward the model stochastically as explained in appendix C to obtain out-of-sample forecasts until 2006. Below we report the figures of the mean equity premia (with one standard deviation band) generated from the three models along with the actual historical equity premium and in-sample fit of the models.

Insert here Figure 7a/b/c

We observe in figure 7a that the forecast from the first model, using information from demography and TFP, is able to capture a portion of the last run-up in stock markets around the millenium. Naturally, the model does not capture the whole phenomenon as some may call it a stock market bubble or irrational exuberance (Shiller, 2005), since we

[28] We thank an anonymous referee for this comment.

start with a model that does not allow for price bubbles. Yet, it captures the general tendency (one standard deviations around the mean predictions provide the upper and lower bounds for the actual data we observe historically in the past two decades) since we modelled explicitly the smooth movement in the mean of the financial ratio. The second model (LL) (figure 7b) performs surprisingly well in sample predictions as previous results suggested, but once we move to out-of-sample it only forecasts the historical mean and does not provide additional information. Finally, the combined model (figure 7c) is also successful in capturing the major tendencies in the data, since it relies upon information extracted from both models[29].

Insert here figure 8

In light of this strong predictability evidence, we also provide a comparison of alternative models in projecting equity premium for the next few decades. Our simulation (Figure 8) confirms the evidence in favour of the often quoted claim that the end of the baby boomers generation will cause a reduction in the equity premium. The model based on demographics and TFP shows a reduction in the equity premium between 2010 and 2020 which is promptly reverted in the following years. This results are robust to the inclusion of excess consumption in the model, while the model without demographics information predicts a much flatter profile for the equity premium.

6 Conclusions

The intuition that demographic information should be incorporated in long-run stock return has long been considered in the theoretical and empirical finance literature. Yet, there is still controversy on the channels through which demographic factors affect stock markets fluctuations. In this paper, we follow the idea that demographic factors, along with a stochastic trend captured by TFP, are important long-run anchors for slowly evolving mean of stock prices. We show that incorporating demographic information along with a technological trend provides an explanation for the breaks in the dividend-price ratio and the linear combination of these variables capture a slowly evolving mean toward which the dividend-price ratio is reverting. The deviations of dividend-price ratio from the identified long-run evolving mean are powerful within-sample and pseudo out-of sample predictors for stock market excess returns at different horizons. We find that an empirical model based on long-run factors, namely MY and TFP along with a demand factor as captured by excess consumption in the sense of Lettau and Ludvigson (2004), outperforms all alternative models proposed in the empirical literature within

[29] Although these results provide strong empirical evidence in favor of our predictor, we should stress the fact that it is only a pseudo and not a real out-of-sample forecast.

the framework of the dynamic dividend growth model. On the basis of these results we exploit the exogeneity and the predictability of demographic factors to simulate the equity risk premium up to 2050. Our results points to some, albeit not dramatic, decline of the equity risk premium for the next 10 years.

References

[1] Abel, Andrew B., 2003, The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security, Econometrica, 71, 2, 551-578.

[2] Abel, Andrew B., 2001, Will Bequests Attenuate the Predicted Meltdown in Stock Prices When Baby Boomers Retire?, Review of Economics and Statistics, 83, 4, 589-595.

[3] Ang, Andrew and Geert Bekaert, 2007, Stock Return Predictability: Is It There? The Review of Financial Studies, 20, 651–707.

[4] Ang, Andrew and Angela Maddaloni, 2005, Do Demographic Changes Affect Risk Premiums? Evidence from International Data, Journal of Business, 78, 341-380.

[5] Asness, Clifford, 2003, Fight the Fed Model: the Relationship between Future Returns and Stock and Bond Market Yields, Journal of Portfolio Management.

[6] Bakshi, Gurdip S., and Zhiwu Chen, 1994, Baby Boom, Population Aging, and Capital Markets, Journal of Business, 67, 2, 165-202.

[7] Bansal, Ravi and Amir Yaron, 2004, Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles, Journal of Finance, 59,4,1481-1509.

[8] Beaudry, Paul and Franck Portier, 2003, Stock Prices, News and Economic Fluctuations, CEPR Discussion Papers 3844, C.E.P.R. Discussion Papers.

[9] Beaudry, Paul and Franck Portier, 2004, Stock Prices, News and Economic Fluctuations, working paper.

[10] Bloom, David E., David Canning, and Jaypee Sevilla, 2003, The Demographic Dividend. A new Perspective on the Economic Consequences of Population Change, Rand Corporation, Santa Monica.

[11] Boudoukh, Jacob, Richardson, Matthew and Robert F. Whitelaw, 2008, The Myth of Long-Horizon Predictability, The Review of Financial Studies, 21, 4, 1577-1605.

[12] Boudoukh, Jacob, Michaely, Roni, Richardson, Matthew and Michael Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, Journal of Finance, forthcoming.

[13] Brooks, Robin J.,2002, Asset-Market Effects of the Baby Boom and Social-Security Reform, American Economic Review, 92, 2, 402-406.

[14] Brooks,Robin J., 2000, What Will Happen to Financial Markets When The Baby Boomers Retire?, Computing in Economics and Finance, 92, Society for Computational Economics.

[15] Campbell, John Y., and Samuel B. Thomson, 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, The Review of Financial Studies, 21, 1509-1531.

[16] Campbell, John. Y., 2001, A Comment on James M. Poterba's "Demographic Structure and Asset Returns", The Review of Economics and Statistics, 83, 4, 585-588.

[17] Campbell, John Y., and Robert Shiller, 1988, Stock Prices,Earnings, and Expected Dividends, Journal of Finance, 43, 661-676.

[18] Canova, F., 2007, Methods for Applied Macroeconomic Research. Princeton University Press.

[19] Cochrane, John H., 2007, The Dog that Did Not Bark: A Defense of Return Predictability, Review of Financial Studies.

[20] Cochrane, John H., 2001, Asset Pricing, Princeton University Press.

[21] Cochrane, John H., 1997, Where is the Market Going? Uncertain Facts and Novel Theories, Federal Reserve Bank of Chicago - Economic Perspectives, 21(6), 3–37.

[22] Cooper, Ilan and Priestley, Richard, 2008, Time Varying Risk Premia and the Output Gap, Review of Financial Studies, forthcoming.

[23] Dalla Vigna S. and J.Pollet, 2007, Demographics and Industry Returns, American Economic Review, Vol 97, pp. 1167-1702.

[24] Erb, Claude B., Campbell R. Harvey, and Tadas E. Viskanta, 1996, Demographics and International Investment, Financial Analysts Journal (July/August), 14-28.

[25] Fama, Eugene, 1970, Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance, 25, 383–417.

[26] Fama, Eugene and Kenneth R. French, 1988a, Dividend Yields and Expected Stock Returns, Journal of Financial Economics, 22, 3-26.

[27] Geanakoplos, John, Magill, Michael and Martine Quinzii, 2004, Demography and the Long Run Behavior of the Stock Market, Brookings Papers on Economic Activities, 1: 241-325.

[28] Gordon, Myron J., 1962, The Investment, Financing and the Valuation of the Corporation, Homewood, Ill.: R.D. Irwin.

[29] Goyal, Amit, and Ivo Welch, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. The Review of Financial Studies, 21-4, 1455-1508.

[30] Goyal, Amit, 2004, Demographics, Stock Market Flows, and Stock Returns, Journal of Financial and Quantitative Analysis, 39, 1, 115-142.

[31] Hansen, Henrik and Soren Johansen, 1999, Some Tests for Parameter Constancy in cointegrated VAR-models, The Econometrics Journal, 2, 306-333.

[32] Hodrick, Robert, and Edward C. Prescott, 1997, Postwar U.S. Business Cycles: An Empirical Investigation, Journal of Money, Credit, and Banking, 1-16.

[33] Hodrick, Robert, 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, Review of Financial Studies, 5, 357–386.

[34] Jaimovich, Nir and Henry E. Siu, 2008, The Young, the Old, and the Restless: Demographics and Business Cycle Volatility, NBER working paper, 14063.

[35] Johansen, Soren, 1988, Statistical analysis of cointegrating vectors, Journal of Economic Dynamics and Control,12, 231–254.

[36] Johansen, Soren, 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica, 56, 1551–1580.

[37] Julliard, Christian, 2004, Labor Income Risk and Asset Returns, Job Market Paper, Princeton University.

[38] Juselius, Katarina, 2006, The Cointegrated VAR Model: Methodology and Applications, Oxford University Press.

[39] Kandel, Shmuel, and Robert Stambaugh, 1989, Modeling expected stock returns for long and short horizons, Working paper 42-88, Rodney L. White Center, Wharton School, University of Pennsylvania.

[40] Koop, Gary, 2003, Bayesian Econometrics, Chapter 11, Wiley-Interscience.

[41] Kydland, Finn E. and Edward C. Prescott, 1982, Time to Build and Aggregate Fluctuations, Econometrica, 50, 1345-1371.

[42] Lamont, Owen, 1998, Earnings and Expected returns, Journal of Finance, 53, 5, 1563-1587.

[43] Lander, Joel, Athanasios Orphanides, and Martha Douvogiannis, 1997, Earnings Forecasts and the Predictability of Stock Returns: Evidence from trading the S&P, Journal of Portfolio Management, 23(4), 24-35.

[44] Lettau, Martin, Ludvigson, Sydney and Jessica Wachter, 2006, The Declining Equity Premium: What Role does Macroeconomic Risk Play?, The Review of Financial Studies, forthcoming.

[45] Lettau, Martin, and Sydney Ludvigson, 2005, Expected Returns and Expected Dividend Growth, Journal of Financial Economics, 76, 583-626.

[46] Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, Aggregate Wealth and Expected Stock Returns, Journal of Finance, 56, 3, 815-849.

[47] Lettau, Martin, and Sydney Ludvigson, 2004, Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," American Economic Review, 94, 1, 279-299.

[48] Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, Review of Financial Studies, 21, 4, 1607-1652.

[49] Lewellen, Jonathan, 2004, Predicting Returns with Financial Ratios. Journal of Financial Economics 74:209–35.

[50] Macunovich, Diane J., 2002, Birth Quake, Chicago University Press.

[51] Mason, Andrew, and Ronald Lee, 2005, Reform and Support Systems for the Elderly in Developing Countries: Capturing the Second Demographic Dividend, *Genus*, 2, 11-35.

[52] Nyblom, Jukka, 1989, Testing for the Constancy of Parameters over Time. Journal of the American Statistical Association 84, 223-230.

[53] Poterba, James M., 2001, Demographic Structure and Asset Returns, The Review of Economics and Statistics, 83, 4, 565-584.

[54] Raftery, Adrian E., David Madigan, and Jennifer A. Hoeting, 1997, Bayesian model averaging for linear regression models, Journal of the American Statistical Association, 92, 179–191.

[55] Ravn, Morten O. and Harald Uhlig, 2002. On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. Review of Economics and Statistics 84, 371-376.

[56] Ribeiro, Ruy M., 2004, Predictable Dividends and Returns: Identifying the Effect of Future Dividends on Stock Prices, Wharton School,University of Pennsylvan.

[57] Shiller, Robert J., 2005, Irrational Exuberance, second edition, Princeton University Press.

[58] Solow, Robert, 1957, Technical Change and the Aggregate Production Function, Review of Economics and Statistics, 39, 312-320.

[59] Van Binsbergen, Jules H. and Ralph S.J. Koijen, 2009, Predictive Regressions: A Present-Value Approach, SSRN working paper.

[60] Vissing-Jorgensen, A., 2002, Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy*, 110, 825-853.

[61] Warne, Anders, Bruggeman, Annick and Paola Donati, 2003, Is the Demand for Euro Area M3 Stable?, ECB Working Paper Series No. 255.

[62] Yoo, Peter S., 1997, Population Growth and Asset Prices, Federal Reserve Bank of St. Louis working paper no.1997-016A.

[63] Yoo, Peter S., 1994a, Age Dependent Portfolio selection, Federal Reserve Bank of St. Louis working paper no. 94-003A.

APPENDIX A: TABLES

L-Max		Trace		$H_0 = r$
Test Statistics	95%CV	Test Statistics	95%CV	r =
Panel A: Whole Sample (1909-2006)				
9.73	14.26	10.44	15.49	0
0.71	3.84	0.71	3.84	1
Panel B: Subsample (1909-1954)				
14.44*	14.26	14.81	15.49	0
0.37	3.84	0.37	3.84	1
Panel C: Subsample (1951-2006)				
4.89	14.26	4.93	15.49	0
0.04	3.84	0.04	3.84	1

Table 1. Johansen Cointegration Test using log dividend and log price series. We report both L-Max and Trace test statistics with 95% critical values. The null hypothesis is that there are r cointegration relations.

Correlation Matrix (Sample 1909-2006)					
	$r_{m,t} - r_{f,t}$	dp_t	TFP	SR	MY
$r_{m,t}-r_{f,t}$	1.00	-0.24	0.01	0.08	0.10
dp_t	-0.24	1.00	-0.75	-0.21	-0.73
TFP	0.01	-0.75	1.00	0.34	0.47
SR	0.08	-0.21	0.34	1.00	0.17
MY	0.10	-0.73	0.47	0.17	1.00
Univariate Summary Statistics					
Mean	0.057	-3.217	2.278	1.290	0.793
Median	0.093	-3.133	2.143	1.294	0.752
Std	0.179	0.453	0.937	0.133	0.172
Min	-0.530	-4.450	0.983	1.048	0.550
Max	0.423	-2.287	3.976	1.509	1.149
Autocorrelation	0.081	0.880	0.972	0.974	0.967

Table 2a. Summary Statistics (whole sample, 1909-2006, using S&P500 data from Robert Shiller's website)

Correlation Matrix (CRSP: Sample 1926-2006)					
	$r_{m,t}-r_{f,t}$	dp_t	TFP	SR	MY
$r_{m,t}-r_{f,t}$	1.00	-0.07	0.08	0.06	0.03
dp_t	-0.07	1.00	-0.72	-0.09	-0.69
TFP	0.08	-0.72	1.00	0.17	0.29
SR	0.06	-0.09	0.17	1.00	0.01
MY	0.03	-0.69	0.29	0.01	1.00
Univariate Summary Statistics					
Mean	0.096	-3.299	2.525	1.312	0.827
Median	0.134	-3.210	2.724	1.357	0.780
Std	0.194	0.424	0.841	0.137	0.169
Min	-0.586	-4.499	1.197	1.048	0.550
Max	0.454	-2.627	3.976	1.509	1.149
Autocorrelation	0.095	0.915	0.964	0.977	0.973

Table 2b. Summary Statistics (whole sample using CRSP data, 1926-2006)

Correlation Matrix (Sample 1909-1954)					
	$r_{m,t}-r_{f,t}$	dp_t	TFP	SR	MY
$r_{m,t}-r_{f,t}$	1.00	-0.69	0.26	0.09	0.24
dp_t	-0.69	1.00	-0.04	-0.05	-0.05
TFP	0.26	-0.04	1.00	0.76	0.88
SR	0.09	-0.05	0.76	1.00	0.66
MY	0.24	-0.05	0.88	0.66	1.00
Univariate Summary Statistics					
Mean	0.063	-2.891	1.394	1.296	0.720
Median	0.085	-2.900	1.246	1.275	0.735
Std	0.208	0.235	0.325	0.107	0.082
Min	-0.530	-3.323	0.983	1.155	0.563
Max	0.423	-2.288	2.033	1.468	0.915
Autocorrelation	0.107	0.398	0.929	0.968	0.883

Table 2c. Summary Statistics (first subsample 1909-1954, using S&P500 data from Robert Shiller's website)

Correlation Matrix (Sample 1955-2006)					
	$r_{m,t} - r_{f,t}$	dp_t	TFP	SR	MY
$r_{m,t} - r_{f,t}$	1.00	-0.18	-0.03	0.08	0.10
dp_t	-0.18	1.00	-0.54	-0.42	-0.82
TFP	-0.03	-0.54	1.00	0.85	0.11
SR	0.08	-0.42	0.85	1.00	0.09
MY	0.10	-0.82	0.11	0.09	1.00
Univariate Summary Statistics					
Mean	0.052	-3.507	3.060	1.285	0.858
Median	0.093	-3.435	3.106	1.329	0.899
Std	0.151	0.402	0.500	0.154	0.202
Min	-0.365	-4.450	2.106	1.048	0.550
Max	0.285	-2.925	3.976	1.509	1.149
Autocorrelation	-0.025	0.907	0.925	0.973	0.976

Table 2d. Summary Statistics (first subsample 1955-2006, using S&P500 data from Robert Shiller's website)

L-Max		Trace		$H_0 = r$
Test Statistics	95%CV	Test Statistics	95%CV	r =
Panel A: Whole Sample (1909-2006)				
27.86*	27.58	55.10*	47.86	0
17.18	21.13	27.24	29.78	1
9.86	14.26	10.06	15.49	2
0.20	3.84	0.20	3.84	3

Table 3. Johansen Cointegration Test. We use the general model including nominal log dividends, log prices, TFP, MY.

Cointegrating Eq:	dp^{TD}_{t-1}	t-stat	χ^2	Prob.
p_{t-1}	-1		6.49	0.01
d_{t-1}	1			
TFP_{t-1}	0.290	(5.39)		
MY_{t-1}	1.554	(5.19)		
constant	1.318			
Error Correction	Δp_t	Δd_t	ΔTFP_t	ΔMY_t
dp^{TD}_{t-1}	0.315	-0.070	0.042	0.002
	(3.78)	(-1.59)	(1.81)	(0.30)
Δp_{t-1}	0.245	0.347	0.083	-0.001
	(2.27)	(6.07)	(2.78)	(-0.15)
Δd_{t-1}	-0.319	0.186	-0.036	0.017
	(-2.09)	(2.30)	(-0.84)	(1.38)
ΔTFP_{t-1}	-0.252	-0.072	-0.005	-0.061
	(-0.66)	(-0.36)	(-0.05)	(-1.97)
ΔMY_{t-1}	1.077	-0.411	-0.306	0.790
	(1.32)	(-0.95)	(-1.36)	(11.87)
Adj. R^2	0.17	0.40	0.04	0.63

Table 4a. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1909-2006. t-statistics are reported in parentheses.

Cointegrating Eq.	dp_t^{TD}	t-stat	χ^2	$Prob$
p_{t-1}	-1		0.17	0.67
d_{t-1}	1			
TFP_{t-1}	0.435	(12.48)		
MY_{t-1}	1.363	(14.41)		
$constant$	1.007			
Dependent variable	Δp_t	Δd_t	ΔTFP_t	ΔMY_t
dp_{t-1}^{TD}	0.841	0.071	0.038	0.011
	(5.09)	(1.75)	(0.61)	(0.53)
Δp_{t-1}	0.305	0.104	0.138	0.018
	(2.34)	(3.27)	(2.86)	(1.12)
Δd_{t-1}	-1.076	0.485	-0.456	-0.049
	(-2.08)	(3.84)	(-2.38)	(-0.78)
ΔTFP_{t-1}	0.025	0.184	0.064	-0.047
	(0.07)	(2.03)	(0.46)	(-1.04)
ΔMY_{t-1}	0.016	-0.287	-0.341	0.817
	(0.02)	(-1.74)	(-1.36)	(9.88)
Adj. R^2	0.45	0.53	0.26	0.76

Table 4b.The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1955-2006. t-statistics are reported in parentheses.

Cointegrating Eq.	dp_t^{D}	t-stat	χ^2	$Prob$
p_{t-1}	-1		23.39	0.00
d_{t-1}	1			
MY_{t-1}	2.181	(4.39)		
$constant$	1.624			
Dependent variable	Δp_t	Δd_t	ΔMY_t	
dp_{t-1}^{D}	0.158	-0.014	-0.013	
	(2.05)	(-0.88)	(-1.64)	
Δp_{t-1}	-0.019	-0.019	0.007	
	(-0.14)	(2.71)	(0.55)	
Δd_{t-1}	0.317	0.632	-0.037	
	(0.62)	(5.91)	(-0.73)	
ΔMY_{t-1}	1.428	-0.180	0.872	
	(1.95)	(-1.17)	(11.92)	
Adj. R^2	0.11	0.44	0.75	

Table 4c. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1955-2006. t-statistics are reported in parentheses.

Cointegrating Eq.	dp_t^T	t-stat	χ^2	$Prob$
p_{t-1}	-1		3.721	0.05
d_{t-1}	1			
TFP_{t-1}	0.679	(3.86)		
$constant$	1.436			
Dependent variable	Δp_t	Δd_t	ΔTFP_t	
dp_{t-1}^T	0.075	0.025	-0.028	
	(1.06)	(1.86)	(-1.38)	
Δp_{t-1}	0.066	0.077	0.098	
	(0.47)	(2.89)	(2.44)	
Δd_{t-1}	0.185	0.504	-0.289	
	(0.31)	(4.32)	(-1.65)	
ΔTFP_{t-1}	-0.545	0.210	-0.005	
	(-1.15)	(2.33)	(-0.03)	
Adj. R^2	0.00	0.49	0.19	

Table 4d. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1955-2006. t-statistics are reported in parentheses.

Cointegrating Eq.	dp_t^{TD}	t-stat	χ^2	$Prob$
p_{t-1}	-1		11.06	0.00
d_{t-1}	1			
$Trend_{t-1}$	0.015	(12.23)		
MY_{t-1}	1.307	(14.35)		
$constant$	1.007			
Dependent variable	Δp_t	Δd_t	ΔMY_t	
dp_{t-1}^{TD}	0.836	0.035	0.022	
	(5.68)	(0.91)	(1.23)	
Δp_{t-1}	0.288	0.096	0.021	
	(2.33)	(2.97)	(1.33)	
Δd_{t-1}	-0.771	0.584	-0.068	
	(-1.70)	(4.93)	(-1.20)	
ΔMY_{t-1}	-0.974	-0.313	0.781	
	(-1.31)	(-1.61)	(-8.34)	
Adj. R^2	0.42	0.44	0.74	

Table 4e. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1955-2006. t-statistics are reported in parentheses.

Cointegrating Eq:	dp_{t-1}^{TD}	t-stat	χ^2	Prob.	
p_{t-1}	-1		5.40	0.07	
d_{t-1}	1				
TFP_{t-1}	0.196	(4.09)			
MY_{t-1}	2.083	(7.49)			
SR_{t-1}	0				
constant	1.114				
Error Correction	Δp_t	Δd_t	ΔTFP_t	ΔMY_t	ΔSR_t
dp_{t-1}^{TD}	0.297	-0.122	0.045	0.002	-0.006
	(2.95)	(-2.38)	(1.68)	(0.24)	(-1.28)
Δp_{t-1}	0.254	0.307	0.085	-0.001	0.002
	(2.11)	(5.02)	(2.67)	(-0.07)	(0.36)
Δd_{t-1}	-0.272	0.171	-0.032	0.018	-0.013
	(-1.74)	(2.14)	(-0.77)	(1.46)	(-1.90)
ΔTFP_{t-1}	-0.256	-0.060	-0.022	-0.057	0.046
	(-0.65)	(-0.30)	(-0.21)	(-1.84)	(2.67)
ΔMY_{t-1}	1.502	-0.680	-0.372	0.825	0.019
	(1.61)	(-1.43)	(-1.50)	(11.14)	(0.47)
ΔSR_{t-1}	1.534	-1.338	-0.097	0.102	0.841
	(0.92)	(-1.57)	(-0.22)	(0.77)	(11.35)
Adj. R^2	0.13	0.42	0.05	0.63	0.77

Table 4f. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coeffiencts on log price and log dividend are restricted to be -1,1, respectively. χ^2 along with probability is the LR test statistics for binding restrictions. The sample is annual and spans the period 1909-2006. t-statistics are reported in parentheses.

$z_t =$	Horizon h (in years)				
	1	2	3	4	5
dp_t	0.073	0.158	0.201	0.276	0.355
	(1.60)	(2.28)	(2.36)	(2.87)	(3.31)
	[2.18]	[6.08]	[7.27]	[10.5]	[14.4]
$\tilde{dp}_t$	0.171	0.454	0.526	0.673	0.767
	(2.20)	(4.08)	(2.94)	(3.03)	(3.21)
	[2.97]	[12.6]	[12.7]	[16.0]	[17.8]
dp_t^{TD}	0.259	0.568	0.688	0.885	1.025
	(3.80)	(5.78)	(4.77)	(5.32)	(6.36)
	[11.1]	[25.9]	[28.6]	[35.5]	[40.3]
dp^{CFN}	0.671	1.286	1.526	1.593	1.579
	(5.93)	(5.88)	(6.31)	(7.08)	(6.05)
	[22.7]	[38.7]	[39.5]	[33.2]	[28.9]

$z_t =$	Horizon h (in years)				
	6	7	8	9	10
dp_t	0.403	0.460	0.538	0.592	0.642
	(3.26)	(3.36)	(3.35)	(3.11)	(2.71)
	[16.8]	[19.4]	[22.0]	[22.2]	[20.8]
$\tilde{dp}_t$	0.736	0.748	0.882	0.842	0.879
	(2.93)	(3.13)	(4.10)	(3.12)	(3.71)
	[15.4]	[14.5]	[18.1]	[12.9]	[16.1]
dp_t^{TD}	1.028	1.017	1.051	0.986	0.951
	(6.56)	(7.00)	(7.53)	(6.12)	(5.36)
	[37.7]	[34.2]	[33.0]	[26.9]	[22.3]
dp^{CFN}	1.517	1.698	1.917	2.127	2.402
	(3.59)	(3.46)	(3.92)	(3.63)	(3.56)
	[24.8]	[30.9]	[35.4]	[38.9]	[42.2]

Table 5a. This table reports the results of h-period regressions of returns on the S&P500 index in excess of a 3-month Treasury bill rate. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses) and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1909-2006(1926-2003 for dp^{CFN}).

$z_t =$	Horizon h (in years)				
	1	2	3	4	5
	0.129	0.462	0.604	0.898	1.022
dp_t, $\tilde{dp}_t$	(1.10)	(3.45)	(2.26)	(2.78)	(3.21)
	[0.18]	[10.2]	[12.8]	[22.5]	[26.3]
	0.184	0.446	0.610	0.885	1.068
dp_t^{TD}	(2.50)	(5.80)	(4.19)	(4.96)	(6.10)
	[5.35]	[17.2]	[22.6]	[34.7]	[42.3]

$z_t =$	Horizon h (in years)				
	6	7	8	9	10
	0.892	0.814	0.850	0.758	0.819
dp_t, $\tilde{dp}_t$	(2.99)	(3.07)	(3.78)	(3.22)	(3.09)
	[20.5]	[17.1]	[18.6]	[15.1]	[15.3]
	1.046	1.059	1.148	1.075	1.104
dp_t^{TD}	(6.11)	(7.20)	(8.38)	(5.64)	(4.24)
	[40.0]	[40.4]	[46.0]	[42.5]	[41.3]

Table 5b. This table reports the results of h-period regressions of returns on the S&P500 index in excess of a 3-month Treasury bill rate. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses) and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1909-1954.

	Horizon h (in years)				
$z_t =$	1	2	3	4	5
dp_t	0.090	0.153	0.174	0.211	0.268
	(1.53)	(1.66)	(1.71)	(2.27)	(3.63)
	[3.03]	[5.77]	[6.34]	[7.46]	[9.66]
$\tilde{dp}_t$	0.235	0.449	0.419	0.398	0.387
	(2.57)	(2.42)	(1.76)	(1.54)	(1.37)
	[6.36]	[15.1]	[10.8]	[7.22]	[4.98]
dp_t^{TD}	0.583	1.042	1.086	1.129	1.245
	(5.25)	(6.67)	(5.22)	(4.70)	(4.82)
	[32.3]	[53.7]	[47.8]	[40.4]	[38.4]
dp^{CFN}	0.689	0.942	0.940	0.936	1.268
	(3.16)	(2.27)	(1.92)	(1.65)	(1.69)
	[9.60]	[9.17]	[6.71]	[4.36]	[7.21]

	Horizon h (in years)				
$z_t =$	6	7	8	9	10
dp_t	0.323	0.396	0.479	0.581	0.667
	(5.35)	(5.60)	(5.48)	(5.43)	(4.02)
	[12.2]	[16.2]	[18.6]	[19.9]	[18.9]
$\tilde{dp}_t$	0.374	0.463	0.665	0.845	0.910
	(1.11)	(1.33)	(2.17)	(2.55)	(2.34)
	[3.77]	[5.88]	[12.0]	[16.9]	[17.2]
dp_t^{TD}	1.312	1.197	1.016	0.952	0.847
	(5.10)	(5.11)	(4.12)	(2.40)	(1.64)
	[35.0]	[24.2]	[13.1]	[8.68]	[5.22]
dp^{CFN}	1.656	2.13	2.459	3.202	3.685
	(1.59)	(2.32)	(2.81)	(3.62)	(4.41)
	[11.5]	[19.0]	[22.6]	[34.3]	[41.2]

Table 5c. This table reports the results of h-period regressions of returns on the S&P500 index in excess of a 3-month Treasury bill rate. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses) and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1955-2006 (1955-2003 for dp^{CFN}).

$z_t =$	Horizon h (in years) 1	2	3	4	5
dp_t^{TD}	0.544	0.978	0.957	0.993	1.217
	(5.04)	(5.32)	(3.74)	(3.49)	(3.76)
	[29.2]	[47.7]	[36.9]	[27.7]	[29.4]
de_t	0.003	0.097	0.130	0.207	0.335
	(0.04)	(0.55)	(0.55)	(0.90)	(1.49)
	[-2.04]	[-1.17]	[-0.99]	[-0.12]	[2.01]
pe_t	-0.050	-0.084	-0.037	0.001	-0.004
	(-0.95)	(-1.06)	(-0.40)	(0.00)	(-0.02)
	[-0.74]	[-0.00]	[-2.01]	[-2.44]	[-2.50]
$TERM_t$	0.339	1.036	4.219	7.168	6.606
	(0.23)	(0.44)	(1.72)	(1.88)	(1.22)
	[-2.2]	[1.93]	[3.07]	[11.1]	[5.81]
$DEFAULT_t$	4.922	4.152	2.060	5.529	12.149
	(1.24)	(0.81)	(0.30)	(0.63)	(1.44)
	[0.42]	[-1.19]	[-1.96]	[-1.26]	[1.28]
$RREL_t$	-3.180	-4.307	-1.819	-3.701	-4.677
	(-2.61)	(-2.65)	(-1.53)	(-2.19)	(-2.98)
	[5.56]	[5.60]	[-1.05]	[1.80]	[2.40]
cay_t	5.706	10.110	11.092	12.326	14.211
	(3.53)	(4.70)	(4.12)	(4.36)	(4.73)
	[25.6]	[45.6]	[47.3]	[42.1]	[40.1]
cdy_t	0.242	5.727	5.086	5.727	6.832
	(0.13)	(3.05)	(1.85)	(1.70)	(1.74)
	[-2.21]	[11.7]	[7.90]	[7.41]	[7.41]

Table 6a. This table reports the results of h-period regressions of returns on the S&P500 index in excess of a 3-month Treasury bill rate. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses) and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1955-2001.

$z_t =$	Horizon h (in years)				
	6	7	8	9	10
dp_t^{TD}	1.346	1.166	1.037	0.958	0.820
	(3.42)	(3.58)	(3.36)	(2.30)	(1.67)
	[28.9]	[18.8]	[11.2]	[7.59]	[3.88]
de_t	0.447	0.389	0.135	-0.205	-0.557
	(1.36)	(0.87)	(0.28)	(-0.47)	(-1.42)
	[3.87]	[1.47]	[-2.01]	[-1.71]	[2.11]
pe_t	-0.004	-0.038	-0.073	-0.343	-0.532
	(-0.02)	(-0.15)	(-0.25)	(-1.27)	(-1.99)
	[-2.50]	[-2.42]	[-2.15]	[4.96]	[12.82]
$TERM_t$	7.148	8.105	6.252	6.883	8.383
	(1.31)	(1.48)	(1.16)	(1.28)	(1.38)
	[5.26]	[6.36]	[1.84]	[1.59]	[2.60]
$DEFAULT_t$	16.300	19.039	22.734	29.906	37.051
	(1.95)	(2.25)	(2.55)	(2.46)	(2.75)
	[3.06]	[4.37]	[5.97]	[10.3]	[14.9]
$RREL_t$	-4.523	-6.113	-6.808	-5.454	-6.200
	(-1.86)	(-2.34)	(-1.89)	(-1.45)	(-1.94)
	[1.14]	[3.48]	[3.80]	[0.86]	[1.37]
cay_t	17.549	18.122	19.359	20.849	21.326
	(5.25)	(4.60)	(4.28)	(4.09)	(4.37)
	[49.0]	[45.1]	[40.3]	[39.1]	[35.7]
cdy_t	8.673	6.934	6.871	6.908	7.942
	(1.59)	(1.10)	(1.23)	(1.06)	(1.19)
	[10.02]	[4.64]	[3.11]	[2.28]	[3.16]

Table 6b. This table reports the results of h-period regressions of returns on the S&P500 index in excess of a 3-month Treasury bill rate. For each regression, the table reports OLS estimates of the regressors, Newey-West (1987) corrected t-statistics (in parentheses) and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1955-2001.

Bayesian Model Averaging (BMA) Posterior Estimates (Sample:1952-2001)										
$r_{m,H}-r_{f,H-1}$	$const$	dp_{t-1}^{TD}	dp_{t-1}^{CFN}	de_{t-1}	pe_{t-1}	cay_{t-1}	cdy_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$
H:1 Year (R^2= 0.45)										
coefficient	0.057	0.519	0.092	-0.007	0.010	1.096	-0.009	-1.173	0.077	1.260
t-statistics	(3.27)	(4.65)	(0.36)	(-0.06)	(0.15)	(0.78)	(-0.01)	(-0.84)	(0.05)	(0.30)
H:3 Years (R^2= 0.65)										
coefficient	0.167	0.945	0.021	-0.011	-0.003	3.752	0.208	-0.087	3.780	0.734
t-statistics	(7.84)	(6.33)	(0.07)	(-0.08)	(-0.04)	(2.14)	(0.12)	(-0.05)	(1.95)	(0.14)
H:5 Years (R^2= 0.66)										
coefficient	0.250	1.507	0.036	0.007	0.025	1.066	0.183	-0.140	11.660	21.409
t-statistics	(8.72)	(7.35)	(0.08)	(0.04)	(0.22)	(0.43)	(0.08)	(-0.06)	(4.52)	(3.13)
H:7 Years (R^2= 0.66)										
coefficient	0.314	0.500	0.858	0.809	-0.376	10.893	0.148	-0.143	3.390	4.677
t-statistics	(7.58)	(1.87)	(1.49)	(2.16)	(-1.90)	(3.63)	(0.05)	(-0.05)	(1.00)	(0.50)
H:10 Years (R^2= 0.64)										
coefficient	0.438	0.035	2.562	0.396	-0.367	11.002	0.287	-0.217	0.364	1.306
t-statistics	(8.22)	(0.11)	(3.49)	(0.77)	(-1.52)	(2.62)	(0.08)	(-0.06)	(0.08)	(0.12)

Table 7a1. Bayesian Posterior Estimates. We report the BMA posterior estimates of the coefficients of the regressors (with t-statistics in parentheses) in a multivariate regression for for 1,3,5,7,10 years horizon along with the regression R^2 statistics.

Bayesian Model Averaging (BMA) Model Selection (Sample:1952-2001)											
Model: $r_{m,H}-r_{f,H-1}$	dp^{TD}_{t-1}	dp^{CFN}_{t-1}	de_{t-1}	pe_{t-1}	cay_{t-1}	cdy_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$	Prob.	Visit
H:1 Year											
Model 1	1	0	0	0	0	0	0	0	0	16.13	4265
Model 2	1	0	0	0	0	0	1	0	0	14.77	2000
Cum. Probability	0.99	0.20	0.07	0.11	0.36	0.05	0.38	0.07	0.17		
H:3 Years											
Model 1	1	0	0	0	1	0	0	1	0	23.15	1951
Model 2	1	0	0	0	1	0	0	0	0	21.28	1673
Cum. Probability	1.00	0.08	0.08	0.07	0.71	0.10	0.07	0.66	0.11		
H:5 Years											
Model 1	1	0	0	0	0	0	0	1	1	44.46	1461
Model 2	1	0	0	0	1	0	0	1	1	12.19	1944
Cum. Probability	1.00	0.09	0.07	0.16	0.23	0.08	0.08	0.97	0.91		
H:7 Years											
Model 1	0	0	1	1	1	0	0	0	0	17.37	726
Model 2	0	1	1	1	1	0	0	0	0	9.24	787
Cum. Probability	0.53	0.54	0.65	0.56	0.85	0.07	0.08	0.39	0.23		
H:10 Years											
Model 1	0	1	0	0	1	0	0	0	0	19.90	1804
Model 2	0	1	1	1	1	0	0	0	0	14.84	392
Cum. Probability	0.12	0.88	0.36	0.52	0.82	0.08	0.08	0.09	0.11		

Table 7a2. Model selection analysis. We report the two models with highest probability and highest number of visits among all the models considered for Bayesian analysis. 1's in the cells denote that the variable is included in the model, whereas 0's indicate that those variables no not enter the model. We report cumulative probability of each variables, i.e. the probability that a variable appears across all the models considered and two models with highest probability. We have used flat priors and 50000 draws for the analysis. The sample considered for the analysis spans from 1952-2001.

Bayesian Model Averaging (BMA) Posterior Estimates (Sample:1955-2006)								
$r_{m,H} - r_{f,H-1}$	$const$	dp_{t-1}^{TD}	de_{t-1}	pe_{t-1}	cay_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$
H:1 Year (R^2=0.37)								
coefficient	0.011	0.564	-0.006	0.004	0.414	-0.579	0.195	0.298
t-statistics	(1.24)	(4.71)	(-0.07)	(0.08)	(0.28)	(-0.43)	(0.14)	(0.07)
H:3 Years (R^2=0.67)								
coefficient	0.063	0.807	-0.005	0.007	6.174	-0.078	3.075	0.029
t-statistics	(2.86)	(4.95)	(-0.04)	(0.11)	(3.35)	(-0.05)	(1.80)	(0.01)
H:5 Years (R^2=0.65)								
coefficient	-0.254	1.094	-0.000	0.104	6.172	-0.210	5.629	10.613
t-statistics	(1.70)	(4.78)	(0.00)	(1.07)	(2.32)	(-0.09)	(2.26)	(1.45)
H:7 Years (R^2=0.54)								
coefficient	0.157	0.392	0.030	0.014	12.383	-0.145	3.174	9.051
t-statistics	(3.98)	(1.33)	(0.13)	(0.09)	(4.232)	(-0.05)	(0.98)	(1.05)
H:10 Years (R^2=0.55)								
coefficient	1.162	-0.004	0.028	-0.327	18.174	-0.022	1.377	14.662
t-statistics	(2.81)	(-0.01)	(0.06)	(-2.07)	(5.11)	(-0.01)	(0.34)	(1.41)

Table 7b1. We report the BMA posterior estimates of the coefficients of the regressors (with t-statistics in parentheses) in a multivariate regression for for 1,3,5,7,10 years horizon along with the regression R^2 statistics. The sample period is 1955-2006.

Model: $r_{m,H}-r_{f,H-1}$	dp_{t-1}^{TD}	de_{t-1}	pe_{t-1}	cay_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$	Prob.	Visit
H:1 Year									
Model 1	1	0	0	0	0	0	0	44.17	2778
Model 2	1	0	0	0	1	0	0	14.06	3131
Cum. Probability	0.99	0.07	0.08	0.17	0.23	0.11	0.08		
H:3 Years									
Model 1	1	0	0	1	0	1	0	46.14	2195
Model 2	1	0	0	1	0	0	0	25.49	2862
Cum. Probability	1.00	0.06	0.10	0.94	0.06	0.66	0.05		
H:5 Years									
Model 1	1	0	0	1	0	0	0	25.08	1590
Model 2	1	0	1	0	0	1	1	24.44	484
Cum. Probability	0.96	0.07	0.39	0.65	0.09	0.62	0.45		
H:7 Years									
Model 1	0	0	0	1	0	0	0	32.85	2376
Model 2	0	0	0	1	0	0	1	8.85	2873
Cum. Probability	0.39	0.11	0.16	0.85	0.07	0.32	0.39		
H:10 Years									
Model 1	0	0	1	1	0	0	0	29.55	1629
Model 2	0	0	0	1	0	0	1	17.98	1385
Cum. Probability	0.10	0.12	0.67	0.99	0.07	0.20	0.50		

Table 7b2. Model selection analysis. We report the two models with highest probability and highest number of visits among all the models considered for Bayesian analysis. 1's in the cells denote that the variable is included in the model, whereas 0's indicate that those variables no not enter the model. We report cumulative probability of each variables, i.e. the probability that a variable appears across all the models considered and two models with highest probability. We have used flat priors and 50000 draws for the analysis. The sample considered for the analysis spans from 1955-2006.

$z_t =$	In-Sample				Out-Of-Sample			
	R^2	t-stat	MAE	RMSE	R^2_{OS}	MAE	RMSE	DM
dp_t	2.33	1.41	11.08	12.71	-15.83	12.56	14.90	-21.68
$\tilde{dp}_t$	6.30	2.04	9.87	11.79	12.82	10.85	12.92	14.86
dpp_t^{TD}	21.50	4.17	9.71	11.29	26.84	9.98	11.84	12.79
cay	21.00	4.02	10.30	12.76	3.35	10.95	13.61	1.15
cay,dpp_t^{TD}	32.08	3.33,3.34	9.87	11.69	11.54	10.43	13.02	3.31
Historical Mean	-	-	11.07	13.10	-	11.70	13.84	-

Table 8a. This table presents statistics on 1-year ahead forecast errors (in-sample and out-of-sample) for stock returns. The sample starts in 1952 and we construct first forecast in 1982. All numbers are in percent. RMSE is the root mean square error, MAE is the mean absolute error. DM is the Diebold and Mariano (1995) t-statistic for difference in MSE of the unconditional forecast and the conditional forecast. The out-of-sample R^2 compares the forecast error of the historical mean with the forecast from predictive regressions.

$z_t =$	In-Sample				Out-Of-Sample			
	R^2	t-stat	MAE	RMSE	R^2_{OS}	MAE	RMSE	DM
dp_t	6.65	1.66	16.79	23.09	-58.04	28.37	36.99	-2.28
$\tilde{dp}_t$	5.44	1.68	15.77	22.77	2.65	23.43	29.03	1.75
dpp_t^{TD}	28.89	4.28	15.42	19.73	29.70	20.76	24.67	3.81
cay	35.58	4.13	13.23	16.47	57.55	16.58	19.17	4.00
cay,dpp_t^{TD}	49.49	4.38,3.99	10.92	14.22	61.95	15.53	18.15	4.00
Historical Mean	-	-	17.97	24.65	-	23.80	29.42	-

Table 8b. This table presents statistics on 3-year ahead forecast errors (in-sample and out-of-sample) for stock returns. The sample starts in 1952 and we construct first forecast in 1982. All numbers are in percent. RMSE is the root mean square error, MAE is the mean absolute error. DM is the Diebold and Mariano (1995) t-statistic for difference in MSE of the unconditional forecast and the conditional forecast. The out-of-sample R^2 compares the forecast error of the historical mean with the forecast from predictive regressions.

$z_t =$	In-Sample R^2	t-stat	MAE	RMSE	Out-Of-Sample R^2_{OS}	MAE	RMSE	DM
dp_t	14.29	3.85	19.29	27.89	-38.34	40.85	48.68	-1.80
$\tilde{dp}_t$	2.08	1.12	21.83	30.97	-20.06	37.01	45.35	-1.12
dpp_t^{TD}	33.77	5.47	19.88	26.60	30.02	29.77	34.63	3.40
cay	36.38	5.10	15.40	20.84	57.64	22.68	26.94	4.52
cay,dpp_t^{TD}	52.70	4.57,4.55	14.66	20.64	59.16	20.67	26.45	3.26
Historical Mean	-	-	23.18	30.77	-	37.19	41.39	-

Table 8c. This table presents statistics on 5-year ahead forecast errors (in-sample and out-of-sample) for stock returns. The sample starts in 1952 and we construct first forecast in 1982. All numbers are in percent. RMSE is the root mean square error, MAE is the mean absolute error. DM is the Diebold and Mariano (1995) t-statistic for difference in MSE of the unconditional forecast and the conditional forecast. The out-of-sample R^2 compares the forecast error of the historical mean with the forecast from predictive regressions.

APPENDIX B: FIGURES

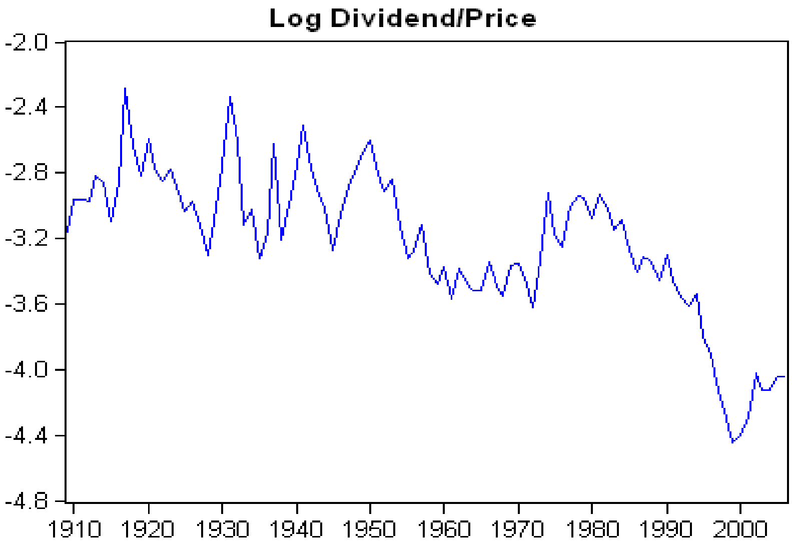

Figure 1. The time series of log dividend price ratio $(d_t - p_t)$. Annual data from 1909 to 2006.

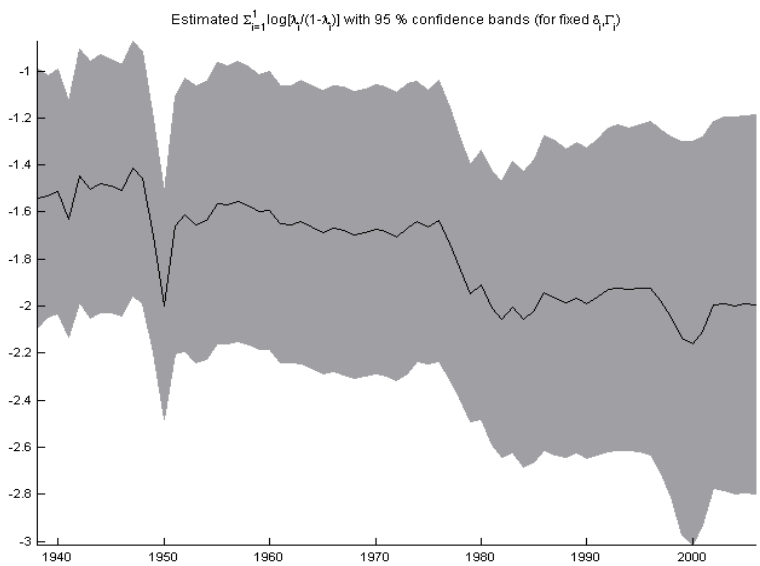

Figure 2. Recursive Eigenvalue Test using log nominal prices and log nominal dividends.

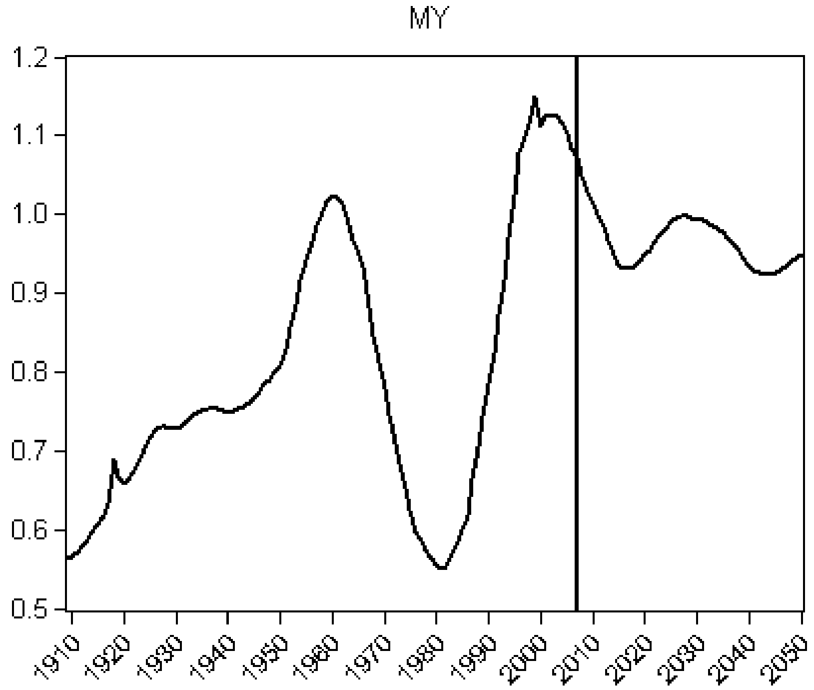

Figure 3a. Middle/young (MY) ratio from 1909 to 2006 and Bureau of Census projections from 2007-2050.

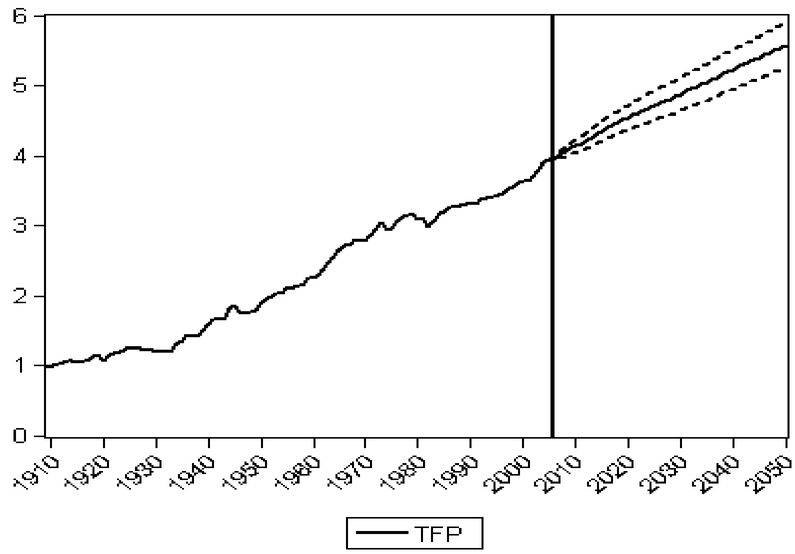

Figure 3b. Total Factor Productivity (TFP) normalized to 1 at the beginning of our sample and projections out-of-sample (with one standard deviation band) obtained from stochastic simulation of VECM model for the period 2007-2050.

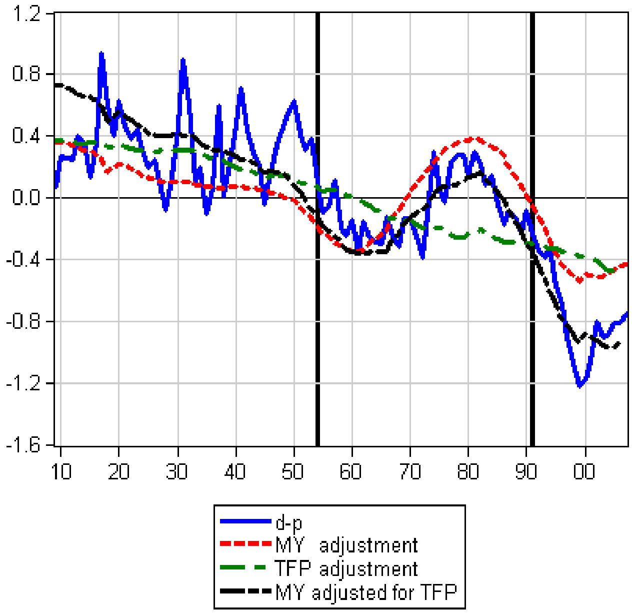

Figure 4a. Log of dividend-price ratio, MY correction, TFP correction and MY&TFP correction with the coefficients given by the cointegrating vector. All variables demeaned.

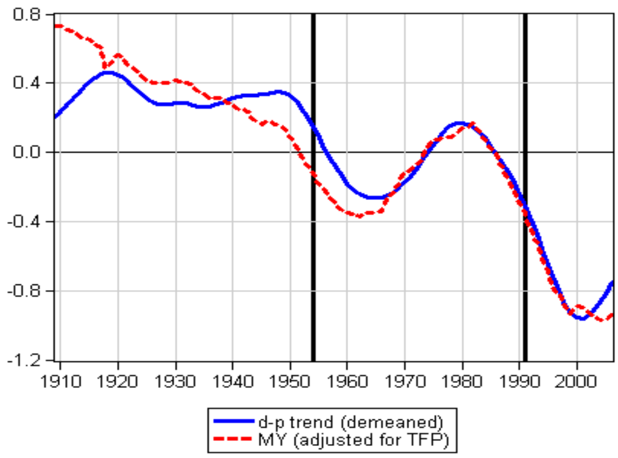

Figure 4b. dividend-price trend component obtained using HP filter (smoothing parameter 100) and linear combination of MY and TFP.

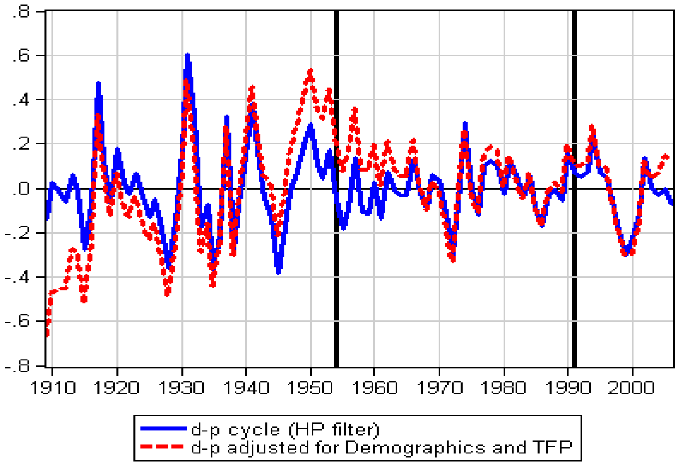

Figure 4c. dividend price cycle component obtained using HP filter and the cointegration vector, log dividends, log prices, MY and TFP.

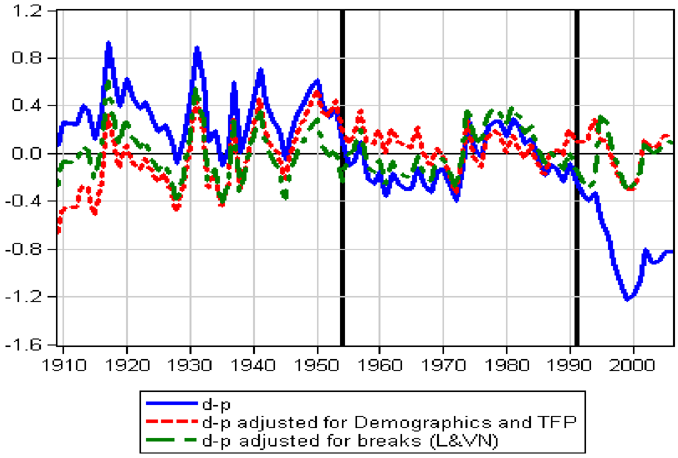

Figure 4d. log dividend price ratio, log dividend price ratio adjusted for exogenous breaks (LVN, 2007) and log dividend price ratio adjusted for demography and TFP.

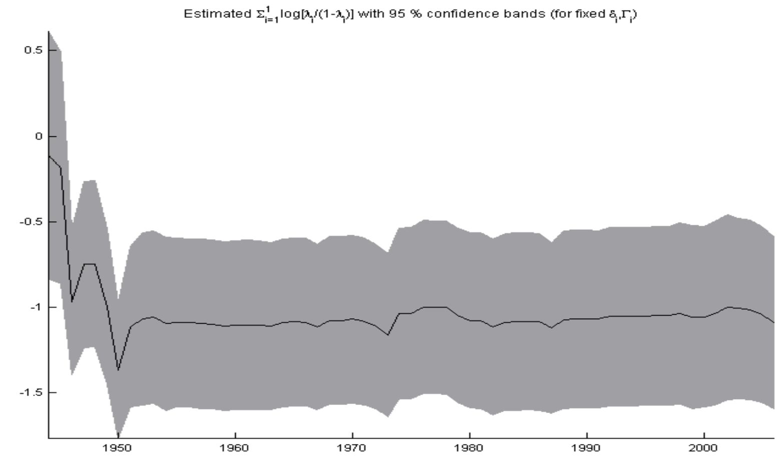

Figure 5a. Recursive Eigenvalue test using the general model. We include nominal log dividends, log prices, TFP, MY and SR.

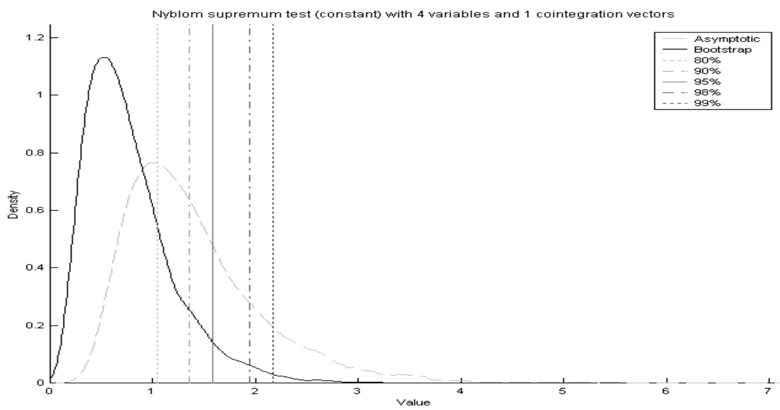

Figure 5b. Nyblom Bootstrap Test for a our model. The sup-statistics is 0.4849 (with mean-statistics = 0.2036) for a VEC model of order 1 allowing for only one cointegration relation

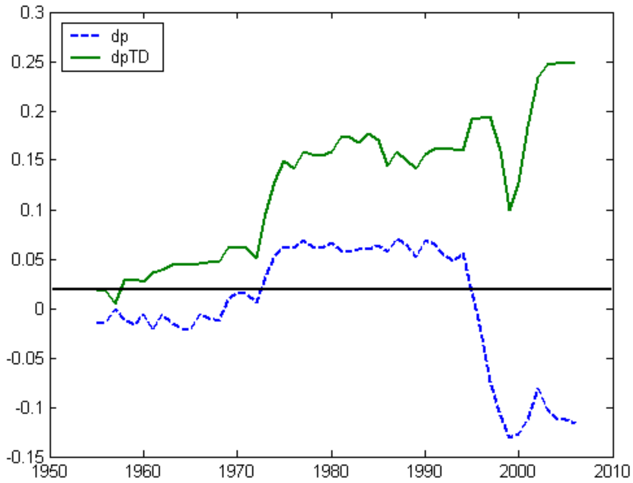

Figure 6a. Out-of sample performance for annual predictive regression. Difference between cumulative squared forecast errors based on a linear regression incuding just a constant and a linear regression including the predictive variable (dp^{TD} or dp). The units are in percent. First forecast in 1955.

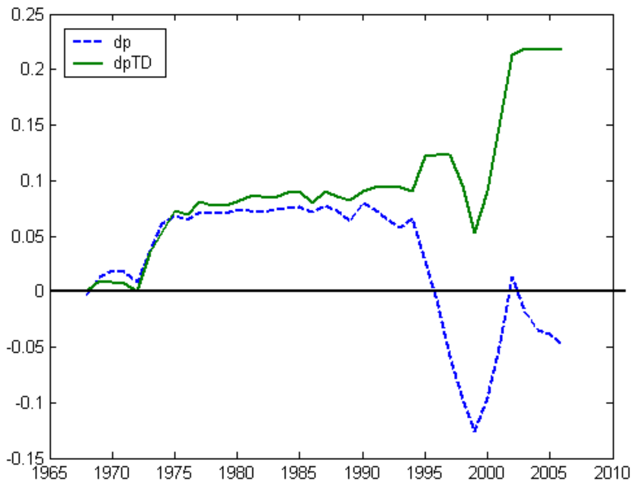

Figure 6b. Out-of sample performance for annual predictive regression. Difference between cumulative squared forecast errors based on a linear regression incuding just a constant and a linear regression including the predictive variable (dp^{TD} or dp). The units are in percent. First forecast in 1968.

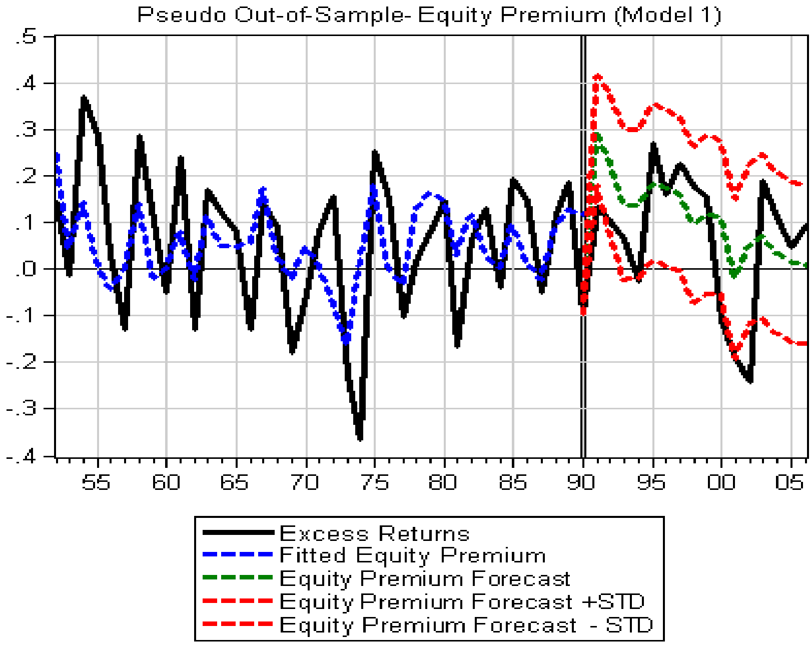

Figure 7a. Pseudo Out-of-Sample Forecast of Equity Premium using the specification with p_t, d_t, TFP_t as endogenous variables and MY_t as exogenous variable.

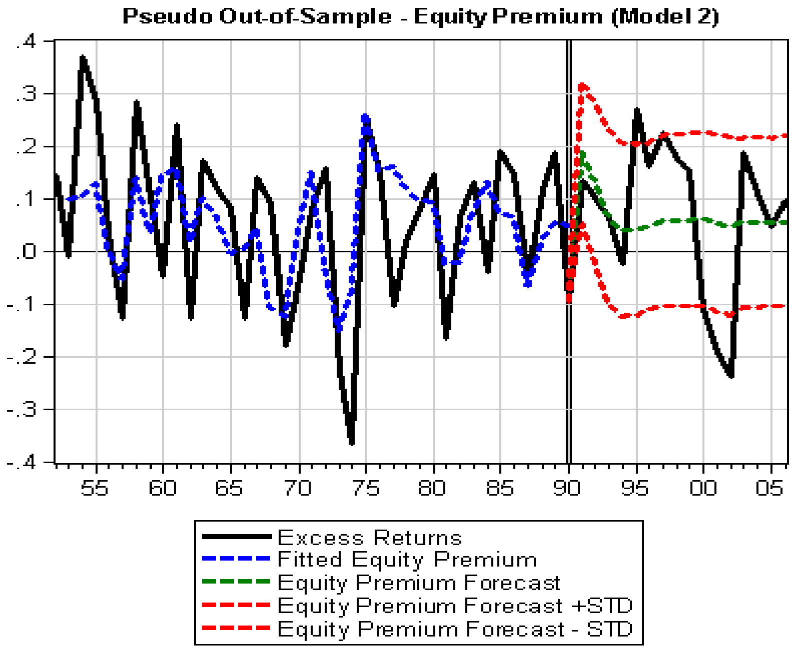

Figure 7b. Pseudo Out-of-Sample Forecast of Equity Premium using the specification with c_t, a_t, y_t as endogenous variables and without any exogenous variables.

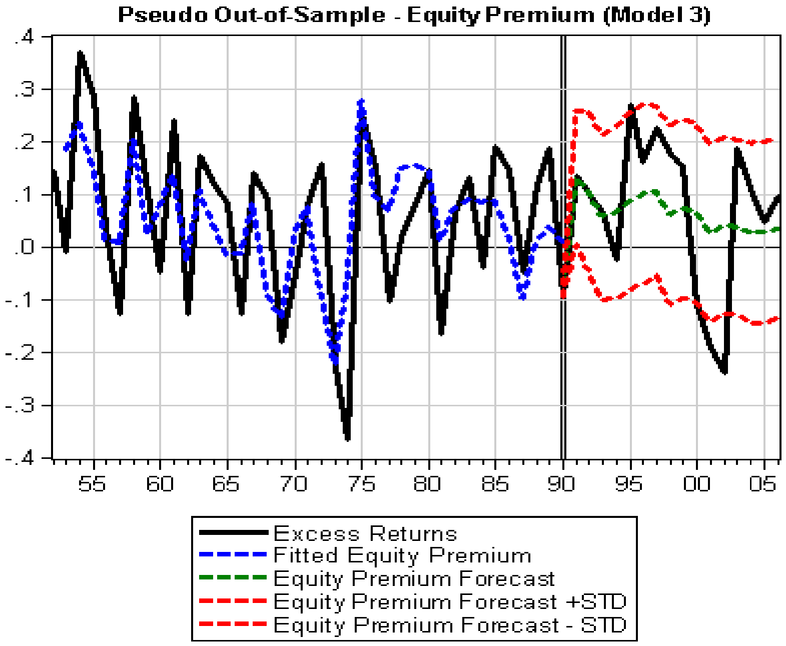

Figure 7c. Pseudo Out-of-Sample Forecast of Equity Premium using the specification with $p_t, d_t, TFP_t, c_t, a_t$ and y_t as endogenous variables and MY_t as exogenous variable.

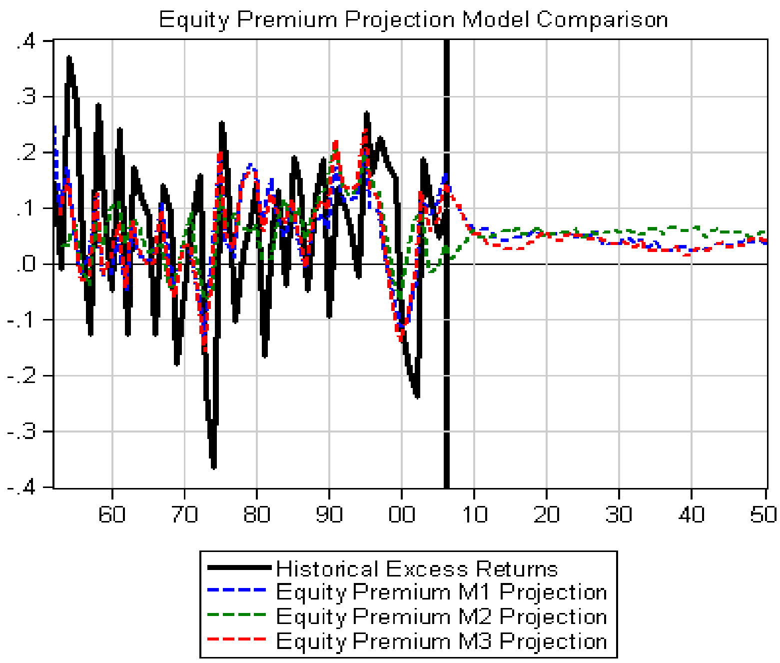

Figure 8. Model Comparison. We graph from 1952 to 2006 the fitted values from three alternave model we consider in this section and from 2007 to 2050 we also graph model predictionss.

APPENDIX C

In order to produce simulations, we take directly the projections from Bureau of Census for our exogenous variable MY_t and we project our endogenous variables by solving a model through stochastic simulations[30], i.e. the model solution generates a distribution of outcomes for the endogenous variables in every period. Through the projected variables, both exogenous and endogenous, we construct the predictive regressors needed for equity premium forecast.

In particular we focus on three models where we augment our VEC specification with an autoregressive process for nominal risk free rate[31].

The first VEC model is already introduced in section 3 and we repeat it here for reader's convenience

$$\begin{pmatrix} \Delta p_t \\ \Delta d_t \\ \Delta TFP_t \end{pmatrix} = \Pi_0 + \Pi_1 \begin{pmatrix} \Delta p_{t-1} \\ \Delta d_{t-1} \\ \Delta TFP_{t-1} \\ \Delta MY_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{pmatrix} \begin{pmatrix} -1 & 1 & \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} p_{t-1} \\ d_{t-1} \\ TFP_{t-1} \\ MY_{t-1} \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}$$

where MY_t is taken as exogenous and it assumed to take the value generated by the Bureau of census predictions over the relevant period. Moreover, we assume constant total factor productivity growth, and hence set $\alpha_{31} = 0$. Using the simulation output from our model, we construct the equity premium first for 1990-2006 and then for 2007-2050, i.e.

$$equity\ premium_t = \log\left(\frac{\tilde{P}_t + \tilde{D}_t}{\tilde{P}_{t-1}}\right) - \tilde{r}_{f,t} \tag{10}$$

where $\tilde{P}_t, \tilde{D}_t, \tilde{r}_{f,t}$ are simulated series from the model.

In the second VEC model, we use the cointegrated system suggested in Lettau&Ludvigson (2001)[32], namely

$$\begin{pmatrix} \Delta c_t \\ \Delta a_t \\ \Delta y_t \end{pmatrix} = \bar{\Pi}_0 + \bar{\Pi}_1 \begin{pmatrix} \Delta c_{t-1} \\ \Delta a_{t-1} \\ \Delta y_{t-1} \end{pmatrix} + \begin{pmatrix} \bar{\alpha}_{11} \\ \bar{\alpha}_{21} \\ \bar{\alpha}_{31} \end{pmatrix} \begin{pmatrix} 1 & \bar{\beta}_2 & \bar{\beta}_3 \end{pmatrix} \begin{pmatrix} c_{t-1} \\ a_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \bar{v}_{1t} \\ \bar{v}_{2t} \\ \bar{v}_{3t} \end{pmatrix}$$

where we augment the model with an autoregressive process for the nominal risk-free rate and a predictive regression for equity premium, i.e. $equity\ premium_t = f(cay_{t-1})$.

[30] In fact the coefficients in our equations are estimated, rather than fixed at known values. One way to reflect this uncertainty about our coefficients in the results from our model is by using stochastic simulation.

[31] We opt for an autoregressive model AR(1) given our sample evidence.

[32] For the estimation of the model, we restricted the insignificant coeffients to zero (consistent with the evidence in LL, 2005), to keep the parameter space small given our short annual sample.

In particular, we assume that the functional relation is linear, i.e.

$$equity\ premium_t = \gamma_0 + \gamma_1 * \widetilde{cay}_{t-1} + \varepsilon_t$$

Notice the difference in forecasting the equity premia in both models. In the former, we simulate the dividend, price and risk free rate processes from the model and the equity premium accounts (in a highly non linear way) for the uncertainty in all of these random variables while in the latter we simulate the equity premium process in a univariate regression where $\gamma = \begin{bmatrix} \gamma_0 & \gamma_1 \end{bmatrix}$ is estimated in the sample 1952-2006 (1990) and the regressor, cointegrating vector $\widetilde{cay}_t$, is reconstructed with the simulated series from the second model for 2007-2050 (1990-2006 for pseudo out-of-sample).

Finally, we combine the two VEC models, where $p_t, d_t, TFP_t, c_t, a_t, y_t$ enter as endogenous variables, MY_t as exogenous variable in the model. We augment the model again with an autoregressive process for the nominal risk-free rate and we reconstruct the equity premium according to equation (6). Given the high number of parameters to be estimated in the model, we set the following restrictions:

- the cointegrating vector cay_t only affects Δa_t
- we assume constant income growth Δy_t and constant ΔTFP_t
- as in model 2, we let only Δa_{t-1} and Δy_{t-1} affect Δc_t[33]
- we model ex-dividend return $\Delta p_t = \delta_0 + \delta_1 * \Delta p_{t-1} + \delta_2 * dp_{t-1}^{TD} + \delta_3 * cay_{t-1} + \varepsilon_t$

To calculate statistics in order to describe the distributions of our endogenous variables, namely p_t, d_t and TFP_t in the first model, c_t, a_t and y_t in the second model and p_t, d_t, TFP_t, c_t, a_t and y_t in the last model, we used a Monte Carlo approach[34], where the model is solved many times with random numbers drawn from a normal distribution with variance covariance resembling the estimation period in sample variance and substituted for the unknown errors at each repetition and then calculating statistics, namely the mean and standard deviation, over all the different outcomes. This method provides only approximate results. However, as the number of repetitions is increased, we would expect the results to approach their true values. We set the number of repetitions to be performed during the stochastic simulation to 10000 and the forecast sample is from 2007 to 2050 (1990-2006). Since our main aim is to simulate future returns, we focus on the long-run price dynamics among other endogenous variables.

[33] We also tried different specifications but results do not change.

[34] We also solved the model bootstrapping the unknown error from estimated in sample residuals, but the results do not change.

APPENDIX D:

In Appendix C, we describe our data construction and provide the links to the data sources. We report results with annual data, but we also cross-check the results using quarterly data. We opt for annual frequency for several reasons; first, the demography variables move slowly and do not change much in quarterly frequency (in that case we interpolate the series), second we are mainly concerned with long term prediction (up to 10 years!) and thus we correct for overlapping data. This way, we also remove the seasonality effects of the data, mainly the dividends. But these advantages come with the trade-off of few data points, which might be particular concern for estimation. Below, we describe the main series we have constructed;

First, the dependent variable, the excess return over the risk free rate:

Stock Prices: S&P 500 index yearly prices from 1909 to 2006 are from Robert Shiller's website, but we took the last month's observation for each year. Alternatively, we also use CRSP annual end-of-year data for value-weighted market (NYSE+AMEX+NASDAQ) index (cum dividend) from 1926 to 2006.

Stock Returns: For S&P 500 index, to construct the continuously compounded return r_t, we take the ex-dividend price P_t add dividend D_t[35] over P_{t-1} and take the natural logarithm of the ratio. On the other hand, for CRSP value-weighted market return, we directly download the cum-dividend market return (*retd*) add 1 and take the natural logarithm to construct the continuously compounded market return.

Risk-free Rate: We download secondary market 3-Month Treasury Bill rate from St.Louis (FRED) from 1934-2006. The risk-free rate for the period 1920 to 1933 is from New York City from NBER's Macrohistory data base. Since there was no risk-free short-term debt prior to the 1920's, we estimate it following Goyal&Welch (2007). We obtain commercial paper rates for New York City from NBER's Macrohistory data base. These are available for the period 1871 to 1970. We estimate a regression for the period 1920 to 1971, which yielded

$$T - billRate = -0.004 + 0.886 \times \; CommercialPaperRate.$$

Therefore, we instrument the risk-free rate for the period 1909 to 1919 with the predicted regression equation.

Hence we build our dependent variable which is the equity premium $(r_{m.t} - r_{f,t})$, i.e., the rate of return on the stock market minus the prevailing short-term interest rate in the year $t-1$ to t.

Second, we construct the independent variables commonly used in the long horizon stock market prediction literature; namely

[35] In Robert Shiller's database, Prices are beginning of period, i.e. January prices, whereas dividends are distributed at the end of the period. In the last section, we simulated our models with december prices.

Log Dividend-Price Ratio (dp_t): is the difference between the log of dividends and the log of prices. For S&P 500 index, i.e. data taken from Robert Shiller's website, we take the natural logarithm of D_t over P_t, in the case of CRSP data we construct dividends D_t by substracting vwretx$_t$ from $vwretd_t$ and multiplying it by $vwindx_{t-1}$.Then dp_t is constructed by taking the natural logarithm of D_t over $P_t(vwindx_t)$. This variable is one of the best candidates for long horizon stock market prediction and is extensively used in the literature (Rozeff (1984), Shiller (1984), Campbell (1987), Campbell and Shiller (1988), Campbell and Shiller (1989), Fama and French (1988a), Hodrick (1992), Barberis (2000), Campbell and Viceira (2002), Campbell and Yogo (2003), Lewellen (2004). See Cochrane (1997) for a survey on dividend price ratio prediction literature).

Log Dividend-Earnings (payout) ratio: Both annual dividend and earning series are taken from Robert Shiller's website. The variable is constructed by taking the natural logarithm of D_t over E_t (Lamont,1998).

Log Earnings Price ratio: Both annual price and earning series are taken from Robert Shiller's website. The variable is constructed by taking the natural logarithm of E_t over P_t (Lamont,1998).

RREL: This variable, the stochastically detrended riskless rate, is constructed using monthly 3-Month Treasury Bill yield data from NBER Macrohistory Data Base (from 1920 to 1933) and 3-Month Treasury Bill: Secondary Market Rate from FRED St. Loius (1934-2006); i.e. we define RREL for month t, $RREL_t$ is r_t minus the average of r_t from months $t-12$ to $t-1$. Yearly $RREL_t$ is the last observation at the end of the year (Campbell,1991; Hodrick,1992). The data is available from 1921-2006.

TERM: is the difference between the long-term government bond yield (10year) from Robert Shiller's Website and 3-Month T-Bill yield from NBER Macrohistory Data Base (from 1920 to 1933) and 3-Month Treasury Bill: Secondary Market Rate from FRED St. Loius (1934-2006) and available from 1920 to 2006.

DEFAULT: is the difference between the BAA and the AAA corporate bond rates. Both series are collected from St.Louis (FRED) and available from 1919 to 2008.

Consumption, wealth, income ratio (cay): is suggested in Lettau and Ludvigson (2001). Data for its construction is available from Sydney Ludvigson's website at annual frequency from 1948 2001. Lettau-Ludvigson estimate is described in equation (4) in their paper, where two lags are used in annual estimation (k = 2). This variable is named as $cayp(post)$ by Goyal&Welch (2008), which they claim contains look-ahead bias, we also consider their variable $caya(ante)$ that eliminates the bias, but report the results using $cayp$, since this gives us a more conservative benchmark. We also use their updated quarterly cay (1952-2006, last quarter as annual observation) for BMA analysis.

Consumption, dividend, income ratio (cdy): is suggested in LL (2005). Data for its construction is available from Sydney Ludvigson's website at annual frequency from 1948 2001. Lettau-Ludvigson estimate is described in equation (4) in their paper, where two lags are used in annual estimation (k = 2).

In addition to the independent variables commonly used in the literature, we also use demography and technology variables in a cointegration framework to explain the long run movement of prices driven by fundamentals.

Demography Variables

The U.S annual population estimates series are collected from U.S Census Bureau and the sample covers estimates from 1900-2050.

Technology Variable

Among other candidates such as Industrial production, number of patents or a variable extracted from a large dataset using principal component, we first focus on a single technology variable, total factor productivity (TFP), which is typically the main source of randomness in standard Real Business Cycle models (RBC, Kydland & Prescott, 1982).

Total Factor Productivity (TFP): We take the net multifactor productivity index (annual) for Private Business Sector (excluding Government Enterprises) from 1948-2006, a series available on the website of Bureau of Labor Statistics (BLS). In order to have a longer time series, we merged this series with the TFP data from 1909 to 1949 provided in the original paper by Solow (1957). We normalized the series from BLS to bring it to the same scale with Solow data. We also constructed TFP series following Beaudry and Portier (2004) and obtained consistent results (available upon request).

DATA SOURCES

Robert Shiller's Website

http://www.econ.yale.edu/~shiller/

NBER Macrohistory Data Base

http://www.nber.org/databases/macrohistory/contents/chapter13.html.

Martin Lettau's Website

http://faculty.haas.berkeley.edu/lettau/

WRDS

http://wrds.wharton.upenn.edu/

US Census Bureau

http://www.census.gov/popest/archives/

Andrew Mason's Website

http://www2.hawaii.edu/~amason/

Bureau of Labor Statistics Webpage

http://www.bls.gov/data/

FRED

http://research.stlouisfed.org/fred2/

Michael R. Roberts' Website

http://finance.wharton.upenn.edu/~mrrobert/