Demographic Trends, Low Frequency Fluctuations in the Aggregate Dividend/Price Ratio and the Predictability of Long-Run Stock Market Returns.*

Carlo A. Favero
Bocconi University, IGIER & CEPR

Arie E. Gozluklu
Bocconi University

Andrea Tamoni
Bocconi University

This Version: October, 2009

Abstract

We analyze aggregate long-run stock market return predictability within the dynamic dividend growth model. The crucial assumption of the model for long-run predictability is that of stationarity of the log dividend price ratio. The validity of this assumption has been challenged in the recent literature and its failure has been highlighted as a potential explanation for the mixed evidence on the forecasting performance of the model. We document the existence of a slowly evolving trend in the mean dividend/price determined by demographic variables. Deviations from this slowly evolving long-run component explain transitory (business cycle) movements of aggregate excess stock market returns and increase their out-of-sample predictability . On the basis of this evidence, we exploit the exogeneity and predictability of the demographic variables to simulate the equity risk premium up to 2050.

KEYWORDS: error correction model, long run predictability, equity premium, cointegration, 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, Dondena Seminar at Bocconi University, LBS and EFA 2009. We thank our discussants Philippe Andrade and Michael Halling, participants in Tilburg, Paris, Ancona, Milan, London and Bergen for stimulating discussions. We also thank Massimiliano Marcellino, Gino Favero, Sasson Bar-Yosef, Francesco Billari, Joachim Inkmann, Sami Alpanda, Ralph Koijen for helpful discussions, Andrew Mason for providing us with the data on demographic dividend and Guillaume Vandenbroucke for historical TFP series. 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. The recent empirical literature has replaced the long tradition of the efficient market hypothesis (Fama, 1970) with 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 potential use 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 and Shiller (1988). This model 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 $(p-d)_t$, 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} :

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

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:

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

(1) allows to classify different forecasting regressions 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), based on a constant equilibrium log dividend price, is obtained by augmenting (1) with the hypotheses of constant dividend growth, and constant expected returns. The so-called FED model (Lander et al., 1997), based on 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, a constant relation between the risk premium on long-term bonds and the risk premium on stocks, and a stationary (log) dividend payout ratio ratio. This basic model can be extended (Asness (2003)) by adding the ratio between the historical volatility of stock and bonds. Lettau and Ludvigson (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, does

predict future return on total wealth. In their proposed framework excess consumption proxies p_t^* in that it predicts future discounted returns. Julliard (2004) refines the LL contribution by introducing labour income growth in the empirical model to control for returns on human capital. 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. 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)).

We concentrate on the application of the dynamic dividend growth model for forecasting long-run returns. In this field the 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 use a century of US data to 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. In their paper, breaks are modelled via a purely statistical methods without any explicit relation with economic fundamentals. In a recent working paper Johannes et al.(2008) estimate the process for log dividend price ratio within a particle filtering framework and find evidence on a downward trending and slow-moving dividend price mean.

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 variables explain fluctuations in the dividend yield are supported by evidence based on annual US data. We then exploit stability analysis of 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 non-stationary of the (log) aggregate dividend-price ratio. 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. 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 the Dividend-Price Ratio

We consider a long sample of annual data (1909-2008), to analyze cointegration between dividends and stock prices and stationarity of the (log) dividend-price. 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 lends some support to the recent evidence on non-stationarity (Ribeiro, 2004; LVN, 2007). Differently from LVN we do not use recursive Chow test to identify break points but we analyze 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 (see Appendix C):

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

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

We then apply the trace and maximum eigenvalue tests proposed by Johansen(1988) to identify the number of cointegrating vectors. We then analyze possible structural breaks in the cointegrating relationship by applying the recursive test based on the non zero-eigenvalues suggested in Hansen and Johansen (1999). After an initialization sample for estimation that, as suggested by Warne et al.(2003), is fixed at 35 percent of the full sample, eigenvalues and parameters in the cointegrating relationship are computed recursively by extending by one observation at the time the end point of the estimation sample, t_1, until the full sample is covered.

Figure 2 shows the time path of the recursively calculated log transformed largest non-zero eigenvalues λ_i of the matrix describing the long-run properties of the VAR(2) model together with the 95% confidence bands. We 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 the post-war subsample 1955-2008.

Insert here Table 1

The null of no-cointegration cannot be rejected over the full sample and over the post-war sample.

3 Modelling Low Frequency Fluctuations in the Aggregate 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). The interesting question is now to understand the determinants of the low frequency fluctuations in $(d_t - p_t)$. 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 US. Live births in the US 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 variables. 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 life cycle 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.

3.1 The GMQ Model

GMQ 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 (see appendix). 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

$$c_y^o + q_o c_m^o + q_o q_e c_r^o = w^y + q_o w^m \tag{5}$$

and in even period

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

Moreover, in equilibrium the following resource constraint must be satisfied

$$\begin{aligned} Nc_y^o + nc_m^o + Nc_r^o &= Nw^y + nw^m + D & (7)\\ nc_y^e + Nc_m^e + nc_r^e &= nw^y + Nw^m + D & (8) \end{aligned}$$

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 (thus, we obtain in even period a high MY ratio of $MY = \frac{N}{n} = 1.15$, and in odd period $MY = \frac{n}{N} = 0.87$ (See Figure 3a)). and $w^y = 2, w^m = 3$ to match the ratio (middle to young cohort) of the average annual real income in US. 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 compute the aggregate dividends. 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 would have been violated. 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). 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

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

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

3.2 Putting the GMQ model at work

GMQ model provides a foundation for a long-run relationship between (d$_t$-p$_t$) and demography. GMQ define the empirical conterpat of the 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 report the MY ratio in Figure 3a. Interestingly this variable shows an highly persistent dynamics and a twin peaked behavior with peaks and throughs around 1950, 1980, 2000: the three break points in (d$_t$-p$_t$).

The natural step to put the GMQ model at work is to extend the cointegrating system analyzed in Section 2 to evaluate the empirical performance and parameters' stability of a cointegrating system based on the vector of variables $\mathbf{y}_t' = \begin{bmatrix} d_t & p_t & MY_t \end{bmatrix}$.

Some considerations on the specification of the appropriate system are in order. From the statistical point of view it is important to observe that, as it is evident form the graphical evidence, both $(d_t - p_t)$ and MY_t are trending variables. Johansen(1991) points out that the inclusion of a trend in the cointegrating vector, when appropriate, is important to identify and estimate the cointegrating relationship(s). From the theoretical point of view GMQ explicitly state that 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). On the basis of these arguments we opted for including a including a trend in the cointegrating space. We have experimented with a pure deterministic trend and Total Factor Productivity (TFP).

TFP is a measure of technology accumulation(Kydland & Prescott, 1982); it reflects how efficiently inputs are used in the aggregate production of economy (Comin, 2008). Since stock market is a claim to productive capital to real economy, we include in our specification this variable as an observable empirical proxy for aggregate productivity over time, a state variable which is the main driving force in production based general equilibrium models (Cochrane, 1991; Jermann, 1998, Jermann&Quadrini, 2009). A separate literature points out the importance of technological progress on demography

(Greenwood et al., 2005), as progress of technology relies on abundance of skilled labor who can utilize it to full extent. As the specification with TFP dominated that based on the deterministic trend we report only the results based on TFP trend.

In practice, we model consider the following CVAR specification:

$$\begin{aligned}
\mathbf{y}_t &= \Pi_0 + \Pi_1 \mathbf{y}_{t-1} + \boldsymbol{\alpha}\boldsymbol{\beta}^T \mathbf{y}_{t-1} + \mathbf{v}_t \\
\mathbf{y}_t' &= \begin{bmatrix} d_t & p_t & MY_t & TFP_t \end{bmatrix} \\
\boldsymbol{\beta} &= \begin{bmatrix} 1 & -1 & \beta_3 & \beta_4 \end{bmatrix}
\end{aligned}$$

We report MY_t and TFP_t in Figure 3a-3b. Historical values and predictions up to 2050 are reported. Future projections are made available from Bureau of Census (MY) and Congressional Budget Office (CBO)'s Long-Term Projections for Social Security (TFP, 2009 Update). Using augmented Dickey-Fuller test, we cannot reject the null of a unit root both for MY and TFP.

Insert here Figure 3a - 3b

As in section 2, we apply the Johansen(1991) procedure over the full sample 1909-2008 and the post-war sample (1955-2008). In Table 2 we report the estimation results. 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 different optimal lag-length criteria and the most parsimonious lag selection is reported in the table.

Insert here Table 2

The trace statistics strongly rejects the null hypothesis of no cointegrating relation, and does reject the null of at most one cointegrating vector, both over the full sample (effective sample 1911-2008) and post-war (1955-2008). Hence, we build our VEC model with a single cointegrating vector between p_t, d_t, MY_t and TFP_t that is restricted to be $\begin{pmatrix} -1 & 1 & \beta_3 & \beta_4 \end{pmatrix}$. We report in Table 3 the results of the estimation of the CVAR.

Insert here Table 3

Below, we show point estimates and standard errors for the cointegrating parameters between log dividend-price ratio, MY and TFP.

$$dp_t^{DT} = (d_t - p_t) + \underset{(0.31)}{1.44} \cdot MY_t + \underset{(0.054)}{0.26} \cdot TFP_t + 1.16$$

where dp^{DT} 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.

Turning to the analysis of the disequilibrium correction, the α coefficients reveal that stock market returns react to disequilibrium ($\alpha_{11} = 0.304$, t-stat=3.34) while the restriction that α is zero on lagged TFP growth, dividend growth, MY growth cannot be rejected in our cointegrated VAR (CVAR).

We investigated the stability of the cointegrating relationship by using the recursively calculated eigenvalues and the Nyblom (1989) stability test.

Insert here Figure 4a - 4b

Insert here Figure 4c - 4d

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[1] made available by Anders Warne. We estimate the sup-statistics to be 0.4849 (with mean-statistics = 0.2036) for a VEC model of order one and allowing for only one cointegration relation with the restrictions specified above. From Figure 4b 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. We also test for the stability of the cointegration coefficients in the 90's, where most predictive models fail. Recursive parameter estimation of β_3 and β_4 over 1990-2008 suggest that both parameter values remained stable over this period, with a slight kink for β_3 around the turn of the millennium.

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

We further analyze our cointegration-based results by illustrating graphically the ability of slow evolving variables MY and TFP to track the movements in the mean of log dividend-price ratio, $\overline{dp_t}$.

Insert Figure 5a here

We proxy the unobservable $\overline{dp_t}$ using the following long-run relation

$$\overline{dp_t} = \beta_0 + \beta_3 MY_t + \beta_4 TFP_t$$

We note that neither TFP nor MY alone is sufficient to capture the evolution of mean dividend-price ratio, in fact the restrictions $\beta_3 = 0$ and $\beta_4 = 0$ are independently an jointly rejected. To illustrate the point we report in Figure 6a $\overline{dp_t}$ with three specification for its slow moving component: the full cointegrating vector including MY_t and TFP_t, and the two restricted cointegrating vectors obtained by setting in turn $\beta_3 = 0$ and $\beta_4 = 0$. Overall, the graphical evidence from the two restricted vectors shows that MY_t plays a more important role than TFP_t in capturing low-frequency fluctuations in $\overline{dp_t}$.

To further assess the capability of demographics and productivity trend of removing the low frequency component in dividend price we report in Figure 5, the cycle component of $\overline{dp_t}$, obtained by applying an Hodrick-Prescott filter to the original series with the cointegration-based detrended dividend-price, $dp^{DT} = (\overline{dp_t} - \beta_0 - \beta_3 MY_t - \beta_4 TFP_t)$

Insert Figures 5b here

Figure 5 clearly illustrates that evidence in favour of a uniquely identified cycle component.

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 5c three time series: $(d_t - p_t)$, dp_t^{LvN} the dividend-price ratio corrected for exogenous breaks in LVN[2] and dp_t^{DT} . The graphical evidence illustrates how the cointegration based correction matches the break-based correction in LVN (2008).

Insert Figure 5c here

[2]Following LVN we adopt the following definition:

$$dp_t^{LvN} = \begin{array}{ll} dp_t - \overline{dp}_1 & \text{for } t = 1, ..., \tau_1 \\ dp_t - \overline{dp}_2 & \text{for } t = \tau_1 + 1, ..., \tau_2 \\ dp_t - \overline{dp}_3 & \text{for } t = \tau_2 + 1, ..., T \end{array}$$

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. $\tau_2 = 1994$, and $\overline{dp}_3$ is the sample mean for 1995-2008.

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 stock market returns is the only variable that adjusts in presence of disequilibrium. In this section we concentrate on the within sample and out-of-sample predictability of excess returns.

4.1 Within Sample Evidence

Our within sample evidence is constructed by comparing the performance of raw and adjusted dividend-price ratios for predicting excess returns over the sample 1909-2008 and the post-war sample 1955-2008. We split the sample in 1954 in the light of the evidence on breaks discussed in the previous section. 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

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

where $dp_t, dp_t^{LvN}, dp_t^{DT}$ 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)[3]. This correction delivers a stationary time series by attributing the swift decline in dividend-price ratios starting from the 80's to the shifts in corporate payout policies. The procedure is not uncontroversial, in fact Lettau et al. (2006) argue these shifts are unlikely to explain the full decrease in this financial ratio: other financial valuation ratios such as earning-price ratios witness similar declines. The results are shown in Table 4a. We report heteroskedastic and autocorrelated consistent (HAC) covariance matrix estimators using Bartlett kernel weights as described in Newey & West (1987) where the bandwidth has been selected following the procedure described in Newey & West (1994). Alternatively, we also conduct a (wild)bootstrap exercise (Davidson& Flachaire, 2008) to compute p-values. To avoid the critique of focusing predictability tests on only one particular horizon h, we also compute joint tests across horizon within a SUR framework and provide in the last row a χ^2 statistics with associated p-values.

Insert here Table 4a

[3]The series is taken from Prof. Roberts website. The authors suggest 4 new series, we experimented with all series to report only the results with the best performing series.

Results over the full sample (1909-2008) show that dp_t^{DT} is always significant and the pattern of adjusted R^2 suggests that this correction improves in-sample predictability with respect to the row series at all horizons. At the 1-year horizon, adjusted R^2 is at 9 %, to peak at 42% at 5-year horizon and then it slightly declines to reach a level of 27% at the 10-year horizon. When we concentrate on the 1955-2008 subsample, 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. Consistently, with the point made by Lettau et al. (2006), we observe that, even though dp_t^{CFN} performs well over the full sample, it exhibits similar performance to dp_t in the post war sample. On the other hand, dp_t^{LvN} is also shows significant consistently both in full sample and subsamples, but performs worse than dp_t^{DT} both in terms of t-statistics and adjusted R^2.

On the basis of these results, we proceed to compare the performance dp_t^{DT} as a predictor with that of the other financial ratios used in the framework of the dynamic dividend growth model over the sample 1955-2008.

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 4b

We obtain consistent results with the literature. Table 4b suggests that in a univariate model specification one should include cay_t and dp_t^{DT} in all horizons (except 10 years) and both variables have substantial predictive power with in-sample $\bar{\mathrm{R}}^2$ slightly favoring cay_t. To provide further evidence on this issue we consider a forecasting model exploiting simultaneously all the available information.

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

We adopt Bayesian Model Averaging to deal with the problem of potential multicollinearity between regressors. 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) and base our inference on averaging over a set of possible models. In general averaging over all possible models provides provide better predictive power than considering a single model, as 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 5a -5b

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 $H = \{1, ..., 10\}$ 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 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[4] and 50000 draws for the analysis. The sample considered for the analysis spans from 1955-2008, 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^{DT} 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^{DT} is selected in models from 1 to 5 years, while cay_t is favored in relatively longer horizons.

Overall the within sample evidence clearly suggests that the best predicting model for excess return is obtained by using two variables: dp_t^{DT} and cay_t. We find this evidence consistent with the dynamic dividend growth with a time varying mean:

$$(p-d)_t = \overline{(p-d)_t} + \sum_{j=1}^{\infty} \rho^{j-1} E_t[(\Delta d_{t+j} - \bar{d}) - (h_{t+j}^s - \bar{h})] \qquad (9)$$

In fact, with reference to (9) , the demographic variable and the productivity trend

[4] We run the bma_g function provided in Le Sage toolbox: http://www.spatial-econometrics.com. The hyperparameters ν, λ and ϕ are set 4, 0.25 and 3, respectively.

capture the time evolving mean $\overline{(p-d)}_t$, while, as clearly documented by Lettau and Ludvigson(2004) cay_t is a proxy for $\sum_{j=1}^{\infty} \rho^{j-1} E_t[(h^s_{t+j} - \bar{h})]$.Therefore, the combination of these two predictors generates a more precise measure of p^*_t in (2)and a better predictor of excess returns.

4.2 Out-of-Sample Evidence

In this section we follow Goyal and Welch (2008), and analyze the performance of different predictors from the perspective of a real-time investor. We therefore consider out-of-sample evidence.

We run rolling forecasting regressions for the one, three and five years ahead horizon by using as an initialization sample 1955-1981. The forecasting period begins in 1982 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^{DT}_t. 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 6 we report the adjusted $\bar{\mathrm{R}}^2$ and the t-statistics using the full sample 1955-2008. Then we also report mean absolute error (MAE) and root mean square error (RMSE) calculated based on the residuals in the forecasting period, namely 1982-2008. The first column of out-of-sample panel report the out-of-sample R^2 statistics (Campbell&Thomson, 2008) which is computed as

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

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 $T = 2008$. If R^2_{OS} 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^2_{1t} as the squared forecasting error of prevailing mean, and e^2_{2t} as the squared forecasting error of the predictive variables, $d_t = e^2_{1t} - e^2_{2t}$, 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 6

First, we note that the 1-year ahead out-of-sample performance worsens in general with respect to the within-sample performance. However only prediction based on dp_t and dp_t^{LvN} cannot beat those based on the historical mean, while all other predictors maintain a lower MAE and RMSE than the historical mean. In 3-year and 5-year ahead out-of-sample forecast, models including cay_t or dp_t^{DT} clearly outperform forecasts based on the historical mean, with some evidence more strongly in favour of cay_t at the 5-year horizon.

We report in figure 6 the cumulative squared prediction errors of historical mean minus the cumulative squared prediction error of dp_t and dp_t^{DT} .

Insert here Figure 6

We use all the available data from 1909 until 1954 for initial estimation and then we recursively calculate the cumulative squared prediction errors until the sample end, namely 2008. Consistently with the results of the analysis of structural breaks, we note that around 1954, early 1980's and late 90's the financial ratio dp_t predict worse than the historical mean (note the decrease in the cumulative squared prediction error line around the points) , while the corrected dp_t, i.e. dp_t^{DT} performs as well as the historical mean around the 50's and then clearly outperform it afterwards.

5 Equity Premium Projections

Long-run horizon forecast for MY_t and TFP_t, the two exogenous factors explaining low frequency fluctuation in the dividend/price ratio, are readily available. In fact, the Bureau of Census(BoC) and CBO provide on their website projections up to 2050 for MY_t and TFP_t. We can then feed these forecasts in our CVAR model 1 to produce projections for stock market equity premia over the period 2009-2050. We augment our VEC specification with an autoregressive process for nominal risk free rate and using the simulation output from our model, we construct the equity premium first for 1990-2008 and then for 2009-2050, 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.

We first validate the model by using it to form (pseudo) out-sample equity premium forecasts, that can be assessed against realized excess returns in our sample. We conduct the pseudo out-of-sample exercise by estimating the model with data up to 1990, and then by solving it forward stochastically to obtain out-of-sample forecasts until 2008. We report in figure 8 of the mean equity premia (with one standard deviation band) generated from the model along with the actual historical equity premium and in-sample fit of the models. We compare the fit of the model with a baseline specification, a bivariate VAR including only p_t and d_t.

Insert here Figure 7

The forecast from the VEC model, using information from demography and productivity, capture the general tendency of data (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) but they miss large deviations from the mean. Root mean square error test (RMSE) confirms the improvement of the CVAR forecasts with respect to those based on the bivariate VAR (RMSE$_{dp_t}$ =19.81, RMSE$_{dp_t^{DY}}$ =17.64)

Insert here Figure 8

In light of this strong predictability evidence, we also provide a comparison of our model predictions with respect to historical mean for the next few decades. Our simulation (Figure 8a) predicts a rapid stock market recovery for the next two years followed by a sudden reversion to historical mean with cyclical declines in the premium around 2030's. In its current form, the model does not foresee a dramatic market meltdown, a "doomsday" scenario, due to a collective exit from the stock market by retired the baby boomers. GQM model relies on the cyclicality of young and middle aged cohorts, and the projection of MY up to 2050 does not suggest any meltdown scenario.

6 The GMQ Model and Our Empirical Specification: Some Further Considerations and Robustness Analysis.

We have mapped the GMQ model into the dynamic dividend growth model by showing that the demographic variable singled out by GMQ , together with a productivity-related, helps to explain the time varying mean of the aggregate (log) dividend-price ratio in the following specification:

$$(p-d)_t = \overline{(p-d)}_t + \sum_{j=1}^{\infty} \rho^{j-1} E_t[(\Delta d_{t+j} - \bar{d}) - (h^s_{t+j} - \bar{h})] \tag{11}$$

We have then considered the predictive power for stock market returns and excess returns of deviation of observed log price dividend ratio $(p-d)_t$ from its slowly evolving mean, $\overline{(p-d)_t}$, to find some clear and stable evidence for predictability. On the basis of this evidence we exploited the exogeneity and predictability of the drivers of the low frequency fluctuations in the dividend price ratio to provide Equity Premium projections up to 2050.

Before drawing conclusions, we consider three further issues.

First, we have extended the dynamic dividend growth model to include a slowly evolving mean $\overline{(p-d)_t}$ and used the prediction of the GMQ model to model it by using two observable variables MY_t and TFP_t. Consistently with this choice we have interpreted the statistical evidence in favour of a model including $(p-d)_t$, MY_t , TFP_t and cay_t as the best model to predict excess returns by attributing to cay_t the role of predictor $\sum_{j=1}^{\infty} \rho^{j-1} E_t[(h_{t+j}^s - \bar{h})]$. However, there is a possible alternative interpretation of our results that maintains the standard dynamic dividend growth model with constant $\overline{(p-d)}$ and rationalizes our evidence by attributing to MY_t and TFP_t the status of significant predictors of future long-horizon dividend growth and future stock market returns. Within this framework the evidence for a very slow mean-reversion in the dividend price is attributed to the very slow mean reversion of the determinants of fluctuations around a constant mean rather than to a slowly evolving mean. We provide evidence on this issues by comparing the forecasting performance for future stock market returns, future dividend growth and future GDP growth of the three variables:$(p-d)_t$, $\overline{(p-d)_t}$, and $\left[(p-d)_t - \overline{(p-d)_t}\right]$ .We report in Table 7 results for the 3-year, 5-year, 10-year horizon. These results illustrate that.$\left[(p-d)_t - \overline{(p-d)_t}\right]$ uniformly dominates .the other two variables as a predictor of stock market returns at all different horizons. The performance of all three variables in predicting real activity and real dividend growth is generally clearly inferior to that in predicting stock market returns, however the evidence in favour of $\left[(p-d)_t - \overline{(p-d)_t}\right]$ as the best predictor is confirmed. Overall the evidence lends support to the interpretation of demographic trends as explanatory variables for the low frequency fluctuations in the time-varying mean of the dividend/price.

Second, in the GMQ model bond and stock are perfect substitutes, therefore the evaluation of the performance of MY_t and TFP_t in forecasting yields to maturity of long-term bonds seems a natural extension of our empirical investigation. In fact, the debate on the so-called FED model (Lander et al., 1997) of the stock market , based on a long-run relation between the price-earning ratio and the long-term bond yield, brings some interesting evidence on this issue. The FED model is based on the equalization, up to a constant, between long-run stock and bond market returns This feature is shared by the GMQ framework, and it requires a constant relation between the risk premium on long-term bonds and the risk premium on stocks. It has been shown that, although the FED model performs well in period where the stock and bond market risk premia are strongly correlated, some measure of the fluctuations in their relative premium is necessary to model periods in which volatilities in the two markets have been different (see, for example, Asness (2003)). As a consequence, to put MY_t and TFP_t at work to explain the bond yields, some modelling of the relative bond/stock risk premia is also in order. We consider this as an interesting extension that is on our agenda for future work but it is beyond the scope of this paper.

Third, there are a number of different potential measures for demographic trends. We have therefore conducted robustness analysis of our cointegration results to the introduction of different measures of demographic structure of the population and productivity trends. The results, discussed in Appendix B, are supportive our preferred specification.

7 Conclusions

The significance of the dividend-price ratio in forecasting stock market returns has been recently questioned on the basis of mixed empirical evidence. We concentrate on the possibility that the lack of modelling of a slowly evolving component in the mean dividend/price ratio might explain the available evidence. In particular we have related, theoretically and empirically, the low-frequency fluctuations in the aggregate dividend/price to demographic trends. We have shown that incorporating demographic information along with an aggregate productivity trend provides an explanation for time variation in the mean of dividend-price ratio. We then use deviations of the dividend-price ratio from the proposed equilibrium relation (shared trend between stock market, demography and productivity) to predict business cycle variations of stock market returns. Eventual reversion to the long-run evolving mean guarantees return predictability and a detrended dividend yield improves out-of sample predictions with respect to traditional models for stock market annual excess returns at different horizons. Exploiting the exogeneity and the predictability of long-run anchors, we have also provided projections for equity risk premia up to 2050. Our simulations point to some, albeit not dramatic, decline of the equity risk premium for the next 10 years preceded by a sharp stock-market rally over the next two 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, Fall 2003, 30(1), pp. 11-24.

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

[7] Beaudry, Paul and Franck Portier, 2006, Stock Prices, News and Economic Fluctuations, American Economic Review, 96, 1293-1307.

[8] 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.

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

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

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

[16] Cochrane, John H., 2007, The Dog that Did Not Bark: A Defense of Return Predictability, Review of Financial Studies, 20, 5.

[17] Comin, Diego, 2008, Total Factor Productivity, The New Palgrave Dictionary of Economics. 2^{nd}ed., Edited by Steven Durlauf and Lawrence Blume, Palgrave Macmillan.

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

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

[20] Davidson, Russell and Emmanuel Flachaire, 2008, The Wild bootstrap, Tamed at Last, Journal of Econometrics,146, 1, 162-169.

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

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

[23] 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.

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

[25] 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.

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

[27] Greenwood, Jeremy, Seshadri, Ananth and Guillaume Vandenbroucke, 2005, The Baby Boom and Baby Bust, American Economic Review, 95,1, 183-207.

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

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

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

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

[32] Jermann, Urban and Vincenzo Quadrini, 2007, Stock Market Boom and the Productivity Gains of the 1990s, Journal of Monetary Economics, 54, 2.

[33] Jermann, Urban, 1998, Asset Pricing in Production Economies, Journal of Monetary Economics, 257-275.

[34] Johannes, Michael, Korteweg, Arthur and Nicholas Polson, 2008, Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns, working paper.

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

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

[41] 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.

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

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

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

[45] 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.

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

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

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

[49] 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.

[50] Newey, Whitney K. and Kenneth D. West, 1987, A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 3, 703-08.

[51] Newey, Whitney K. and Kenneth D. West, 1994, Automatic Lag Selection in Covariance Matrix Estimation, Review of Economic Studies, 61, 631-653.

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

[53] Pastor, Lubos and Robert F. Stambaugh, 2009, Are Stocks Really Less Volatile in the Long Run?, SSRN working paper.

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

[55] 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.

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

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

TABLES

L-Max		Trace		$H_0 = r$
Test Statistic	95% CV	Test Statistic	95% CV	r =
Panel A : Whole Sample (1911-2008), Lag in VAR =1				
12.12	14.26	12.15	15.49	0
0.03	3.84	0.03	3.84	1
Panel B : Post-war Sample (1955-2008), Lag in VAR =2				
4.71	14.26	4.86	15.49	0
0.15	3.84	0.15	3.84	1

Table 1. Johansen Cointegration Test. Series: log S&P 500 dividend and log S&P 500 index price.

L-Max		Trace		$H_0 = r$
Test Statistic	95% CV	Test Statistic	95% CV	r =
Panel A : Whole Sample (1911 -2008), Lag in VAR= 1				
26.47*	27.58	53.80**	47.86	0
16.12	21.13	27.32	29.80	1
10.97	14.26	11.20	15.49	2
0.23	3.84	0.22	3.84	3
Panel B : Post-war Sample (1955-2008), Lag in VAR= 2				
24.73	27.58	51.02**	47.86	0
19.90	21.13	26.30	29.80	1
6.35	14.26	6.40	15.49	2
0.05	3.84	0.05	3.84	3

Table 2. Johansen Cointegration Test. Series: log S&P 500 dividend and log S&P 500 index price, total factor productivity index (TFP) and middle-young ratio (MY).

A constant is included in the cointegration relation. We report both L-Max and Trace test statistics: The columns labeled "Test Statistics" give the value of the test and "95% CV" gives the 95 percent confidence interval. The null hypothesis is that there are r cointegration relations. The lag length in the VAR model is chosen according to optimal information criteria, i.e. sequential LR test, Akaike (AIC), Schwarz (SIC), Hannan-Quinn (HQ) information criterion.

Cointegration Relation (Sample: 1911-2008)				
Cointegrating vector	p_t	d_t	TFP_t	MY_t
coefficients	-1	1	0.26	1.44
(t-stat)			(4.80)	(4.66)
Error correction	Δp_t	Δd_t	ΔTFP_t	ΔMY_t
dp_{t-1}^{TD}	0.304	-0.078	0.035	0.003
(t-stat)	(3.34)	(-1.73)	(1.46)	(0.39)
Δp_{t-1}	0.239	0.341	0.076	-0.001
(t-stat)	(2.08)	(5.99)	(2.53)	(-0.10)
Δd_{t-1}	-0.343	0.185	-0.037	0.016
(t-stat)	(-2.14)	(2.32)	(-0.87)	(1.32)
ΔTFP_{t-1}	-0.118	-0.064	0.023	-0.062
(t-stat)	(-0.30)	(-0.33)	(0.22)	(-2.04)
ΔMY_{t-1}	1.331	-0.400	-0.284	0.793
(t-stat)	(1.57)	(-0.95)	(-1.27)	(12.14)
constant	0.046	0.021	0.031	0.002
(t-stat)	(2.02)	(1.84)	(5.11)	(1.20)
$\bar{R}^2$	0.15	0.41	0.03	0.63

Table 3. Johansen VECM Estimation. Series: log S&P 500 dividend and log S&P 500 index price, total factor productivity index (TFP) and middle-young ratio (MY). A constant (c =1.16) is included in the cointegration relation. The table reports estimated coefficients from cointegrated first order vector autoregression, where the coefficients on log price and log dividend are restricted to be -1,1, respectively. t-statistics are reported in parentheses. The lag length (n=1) is selected by using optimal information criteria, i.e. sequential LR test, Akaike (AIC), Schwarz (SIC), Hannan-Quinn (HQ) information criterion.

Horizon / Predictors ($\bar{R}^2$)	dp_t	$\bar{R}^2$	dp_t^{LvN}	$\bar{R}^2$	dp_t^{CFN}	$\bar{R}^2$	dp_t^{DT}	$\bar{R}^2$
coefficient (1 year)	0.073	0.02	0.167	0.03	0.671	0.23	0.258	0.09
(t-stat) / (w.b. p-value)	(1.69)	(0.1061)	(2.75)	(0.0235)	(6.02)	(0.0010)	(3.48)	(0.0010)
coefficient (2 years)	0.158	0.06	0.450	0.12	1.286	0.39	0.591	0.25
(t-stat) / (w.b. p-value)	(2.30)	(0.0415)	(3.94)	(0.0005)	(4.79)	(0.0005)	(6.96)	(0.0000)
coefficient (3 years)	0.201	0.07	0.519	0.12	1.526	0.39	0.721	0.28
(t-stat) / (w.b. p-value)	(1.48)	(0.0240)	(2.55)	(0.0065)	(4.84)	(0.0000)	(3.89)	(0.0000)
coefficient (4 years)	0.277	0.10	0.663	0.15	1.593	0.33	0.939	0.36
(t-stat) / (w.b. p-value)	(1.94)	(0.0090)	(2.44)	(0.0065)	(6.67)	(0.0000)	(5.13)	(0.0000)
coefficient (5 years)	0.356	0.14	0.752	0.17	1.579	0.28	1.097	0.42
(t-stat) / (w.b. p-value)	(2.57)	(0.0030)	(2.52)	(0.0020)	(5.48)	(0.0000)	(6.61)	(0.0000)
coefficient (6 years)	0.403	0.17	0.714	0.14	1.517	0.25	1.098	0.40
(t-stat) / (w.b. p-value)	(2.63)	(0.0040)	(2.09)	(0.0045)	(3.20)	(0.0100)	(6.93)	(0.0000)
coefficient (7 years)	0.460	0.19	0.719	0.13	1.699	0.31	1.127	0.39
(t-stat) / (w.b. p-value)	(2.61)	(0.0035)	(2.24)	(0.0005)	(3.70)	(0.0040)	(8.43)	(0.0000)
coefficient (8 years)	0.538	0.22	0.849	0.16	1.917	0.35	1.204	0.39
(t-stat) / (w.b. p-value)	(2.45)	(0.0075)	(2.58)	(0.0006)	(3.88)	(0.0030)	(8.44)	(0.0000)
coefficient (9 years)	0.592	0.22	0.835	0.14	2.127	0.39	1.160	0.33
(t-stat) / (w.b. p-value)	(2.54)	(0.0210)	(2.23)	(0.0005)	(3.62)	(0.0085)	(6.43)	(0.0000)
coefficient (10 years)	0.642	0.21	0.788	0.11	2.402	0.42	1.113	0.27
(t-stat) / (w.b. p-value)	(2.42)	(0.0325)	(1.72)	(0.0015)	(3.67)	(0.0050)	(4.93)	(0.0000)
χ^2	3.17		2.53		10.18		8.55	
(p-value)	(0.0005)		(0.0052)		(0.0000)		(0.0000)	

Table 4a. Univariate Predictive Regressions. Series: log dividend price ratio (dp_t), log dividend price ratio corrected for the breaks in the mean (dp_t^{LvN},LvN, 2008), cash-flow based net payout yield (dp_t^{CFN},Boudoukh et al., 2007), de-trended log dividend price ratio, dp_t^{DT}. This table reports the results of h-period ahead regressions of returns on the S&P 500 index in excess of 3-month Treasury Bill rate. We report Newey-West (1987,1994) HAC consistent t-statistics with optimal selected lags and adjusted $\bar{R}^2$. The sample is annual and spans the period 1909-2008 (1926-2003 for dp_t^{CFN}). In the last two rows we also report χ^2 and p-value for the joint significance of the regression coefficients across different horizons (SUR estimation).

Horizon / Predictors ($\bar{R}^2$)	$RREL_t$	$\bar{R}^2$	$TERM_t$	$\bar{R}^2$	$Default_t$	$\bar{R}^2$	cay_t	$\bar{R}^2$	dp_t^{DT}	$\bar{R}^2$
coefficient (1 year)	-1.137	-0.01	2.521	0.02	3.595	-0.01	2.570	0.13	0.489	0.18
(t-stat) / (w.b. p-value)	(-0.65)	(0.5663)	(2.69)	(0.1111)	(0.84)	(0.3042)	(2.49)	(0.0015)	(4.50)	(0.0035)
coefficient (2 years)	-3.690	0.03	4.672	0.06	5.102	-0.01	4.841	0.26	0.938	0.40
(t-stat) / (w.b. p-value)	(-2.41)	(0.0005)	(3.20)	(0.0010)	(0.81)	(0.2816)	(2.64)	(0.0010)	(6.02)	(0.0010)
coefficient (3 years)	-2.518	-0.00	4.757	0.05	1.653	-0.02	6.280	0.31	0.974	0.37
(t-stat) / (w.b. p-value)	(-2.40)	(0.0000)	(2.96)	(0.0000)	(0.19)	(0.7064)	(2.90)	(0.0035)	(4.60)	(0.0020)
coefficient (4 years)	-3.843	0.01	6.776	0.10	5.828	-0.01	7.852	0.33	1.066	0.35
(t-stat) / (w.b. p-value)	(-2.45)	(0.0045)	(3.21)	(0.0000)	(0.47)	(0.3597)	(3.69)	(0.0000)	(4.76)	(0.0000)
coefficient (5 years)	-4.238	0.01	8.403	0.13	11.530	0.01	9.792	0.35	1.205	0.36
(t-stat) / (w.b. p-value)	(-2.58)	(0.0005)	(2.17)	(0.0225)	(0.97)	(0.0350)	(4.27)	(0.0000)	(4.93)	(0.0000)
coefficient (6 years)	-3.724	0.00	8.054	0.09	14.761	0.02	11.425	0.37	1.213	0.33
(t-stat) / (w.b. p-value)	(-1.47)	(0.0260)	(1.27)	(0.0875)	(1.33)	(0.0020)	(3.53)	(0.0010)	(4.17)	(0.0000)
coefficient (7 years)	-3.834	-0.00	8.131	0.08	18.058	0.04	11.610	0.34	1.227	0.30
(t-stat) / (w.b. p-value)	(-0.87)	(0.1581)	(1.08)	(0.1191)	(1.53)	(0.0000)	(3.52)	(0.0000)	(5.67)	(0.0000)
coefficient (8 years)	-6.314	0.03	12.053	0.16	25.474	0.08	12.273	0.32	1.256	0.25
(t-stat) / (w.b. p-value)	(-1.45)	(0.0090)	(1.61)	(0.0285)	(1.98)	(0.0000)	(4.30)	(0.0000)	(3.99)	(0.0000)
coefficient (9 years)	-5.274	0.01	9.054	0.06	32.731	0.12	14.746	0.36	1.337	0.21
(t-stat) / (w.b. p-value)	(-1.17)	(0.0135)	(1.30)	(0.0035)	(2.28)	(0.0000)	(4.51)	(0.0000)	(4.71)	(0.0000)
coefficient (10 years)	-4.796	-0.00	10.053	0.08	36.340	0.14	17.575	0.36	1.253	0.14
(t-stat) / (w.b. p-value)	(-1.21)	(0.0000)	(1.20)	(0.0000)	(2.50)	(0.0000)	(4.98)	(0.0000)	(1.88)	(0.0000)
χ^2	1.02		1.70		1.92		7.70		4.34	
(p-value)	(0.4291)		(0.0789)		(0.0409)		(0.0000)		(0.0000)	

Table 4b. Univariate Predictive Regressions. Series: detrended short rate ($RREL_t$), long rate(10Y) minus short rate (3mTB), ($TERM_t$),BBA minus AAA corporate bond rate, ($Default_t$), consumption-wealth ratio, cay_t, de-trended log dividend price ratio, dp_t^{DT}. This table reports the results of h-period ahead regressions of returns on the S&P 500 index in excess of 3-month Treasury Bill rate. We report Newey-West (1987,1994) HAC consistent t-statistics with optimal selected lags, and adjusted $\bar{R}^2$. We also report wild bootstrap p-values in parentheses. The sample is annual and covers the post-war period 1955-2008. In the last two rows we also report χ^2 and p-value for the joint significance of the regression coefficients across different horizons (SUR estimation).

Bayesian Model Averaging (BMA) Posterior Estimates									
Model: $r_{m,H}-r_{f,H-1}$	$const$	dp^{TD}_{t-1}	de_{t-1}	pe_{t-1}	cay_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$	R^2
coefficient (1 year)	0.030	0.336	-0.002	-0.001	1.846	-0.030	0.443	0.076	0.25
(t-stat)	(1.20)	(2.38)	(-0.02)	(-0.03)	(1.33)	(-0.02)	(0.28)	(0.02)	
coefficient (2 years)	0.036	0.725	-0.001	0.008	5.199	-0.287	1.641	0.001	0.57
(t-stat)	(2.07)	(4.43)	(-0.01)	(0.12)	(2.78)	(-0.16)	(0.98)	(-0.00)	
coefficient (3 years)	0.112	0.585	0.001	0.004	8.329	-0.027	0.621	- 0.237	0.56
(t-stat)	(4.29)	(3.24)	(0.01)	(0.06)	(4.11)	(-0.02)	(0.33)	(-0.04)	
coefficient (4 years)	0.142	0.590	-0.001	0.000	9.777	-0.015	2.416	- 0.071	0.61
(t-stat)	(4.36)	(3.00)	(-0.01)	(0.00)	(4.38)	(-0.01)	(1.22)	(-0.01)	
coefficient (5 years)	0.155	0.768	0.004	-0.000	9.529	0.133	4.852	0.199	0.64
(t-stat)	(3.87)	(3.58)	(0.03)	(-0.00)	(3.78)	(0.05)	(2.21)	(0.03)	
coefficient (6 years)	0.294	0.431	0.016	-0.018	13.630	-0.005	0.897	0.731	0.56
(t-stat)	(5.71)	(1.74)	(0.08)	(-0.20)	(4.95)	(0.01)	(0.33)	(0.10)	
coefficient (7 years)	0.436	0.381	0.018	-0.057	12.990	-0.030	1.060	1.312	0.49
(t-stat)	(5.15)	(1.35)	(0.08)	(-0.56)	(4.21)	(-0.01)	(0.34)	(0.15)	
coefficient (8 years)	0.466	0.358	0.044	-0.069	12.132	0.022	4.270	4.541	0.51
(t-stat)	(3.85)	(1.17)	(0.17)	(-0.57)	(3.61)	(0.04)	(1.24)	(0.48)	
coefficient (9 years)	0.749	0.030	0.019	-0.176	18.023	0.020	0.315	11.940	0.55
(t-stat)	(3.04)	(0.08)	(0.07)	(-1.36)	(5.54)	(0.00)	(0.09)	(1.23)	
coefficient (10 years)	1.582	-0.027	0.016	-0.453	19.085	-0.057	1.083	5.656	0.59
(t-stat)	(4.29)	(-0.06)	(0.05)	(-3.38)	(5.66)	(-0.02)	(0.30)	(0.54)	

Table 5a.

Bayesian Model Averaging (BMA) Model Selection									
Model: $r_{m,H}-r_{f,H-1}$	dp^{TD}_{t-1}	de_{t-1}	pe_{t-1}	cay_{t-1}	$RREL_{t-1}$	$Term_{t-1}$	$Default_{t-1}$	Prob.	Visit
Best Model (1 year)	√	–	–	–	–	–	–	34.01	2314
Cum. Probability	0.74	0.05	0.06	0.42	0.06	0.17	0.05		
Best Model (2 years)	√	–	–	√	–	–	–	37.90	2179
Cum. Probability	0.99	0.05	0.11	0.86	0.12	0.41	0.05		
Best Model (3 years)	√	–	–	√	–	–	–	58.26	1839
Cum. Probability	0.94	0.05	0.08	0.99	0.06	0.19	0.06		
Best Model (4 years)	√	–	–	√	–	√	–	37.95	1687
Cum. Probability	0.90	0.05	0.06	0.99	0.07	0.50	0.05		
Best Model (5 years)	√	–	–	√	–	√	–	56.49	1637
Cum. Probability	0.96	0.06	0.06	0.98	0.09	0.76	0.06		
Best Model (6 years)	√	–	–	√	–	–	–	36.43	2410
Cum. Probability	0.65	0.08	0.14	0.99	0.06	0.19	0.09		
Best Model (7 years)	–	–	–	√	–	–	–	19.35	4372
Cum. Probability	0.51	0.08	0.27	0.95	0.06	0.19	0.12		
Best Model (8 years)	–	–	–	√	–	–	–	14.23	3383
Cum. Probability	0.42	0.13	0.27	0.87	0.08	0.48	0.25		
Best Model (9 years)	–	–	–	√	–	–	–	28.09	1841
Cum. Probability	0.10	0.08	0.51	1.00	0.07	0.09	0.47		
Best Model (10 years)	–	–	√	√	–	–	–	44.45	2066
Cum. Probability	0.09	0.08	0.91	0.99	0.07	0.19	0.25		

Table 5b

Table 5a reports BMA posterior estimates of the coefficients of the regressors (with t-statistics in parentheses) in a multivariate regression for H=$\{1, .., 10\}$ years horizon along with the regression R^2 statistics. Table 5b. reports Bayesian Model Selection. We report the model with the highest probability along with the number of visits among all the models considered for Bayesian analysis. "$\sqrt{}$" denotes the variables included in the "best" model. We also report the the probability that a variable appears across all possible models ($2^n, n$: number of variables). We use flat priors and 50000 draws. The sample period is 1955-2008.

	In-Sample				Out-of-Sample			
Panel A (H=1 year)	R^2	t-stat	MAE	RMSE	R^2_{OS}	MAE	RMSE	DM
dp_t	3.03	1.64	12.92	16.17	-11.22	14.54	18.60	-17.43
dp_t^{IvN}	6.36	2.64	11.93	16.20	- 5.25	13.58	18.09	- 8.09
cay_t	16.05	3.43	12.74	16.04	1.07	13.03	17.54	0.46
dp_t^{DT}	18.28	4.15	11.10	16.04	0.60	12.22	17.58	0.23
Historical Mean	–	–	12.92	16.70	–	13.40	17.63	–
Panel B (H= 3 years)	R^2	t-stat	MAE	RMSE	R^2_{OS}	MAE	RMSE	DM
dp_t	6.34	1.99	18.26	24.87	-88.56	33.98	43.59	- 1.32
dp_t^{IvN}	10.76	1.70	17.91	24.81	-25.89	28.40	35.61	- 4.41
cay_t	40.58	4.25	16.11	20.23	37.46	21.04	25.10	2.52
dp_t^{DT}	36.72	5.20	16.31	21.80	17.24	23.00	28.88	1.37
Historical Mean	–	–	19.38	26.65	–	25.26	31.74	–
Panel C (H= 5 years)	R^2	t-stat	MAE	RMSE	R^2_{OS}	MAE	RMSE	DM
dp_t	9.66	4.20	20.85	28.78	-73.61	49.47	60.30	-1.40
dp_t^{IvN}	4.98	1.54	22.63	31.56	-45.68	43.04	52.31	-2.11
cay_t	40.58	6.11	18.48	23.64	54.21	25.43	29.33	5.93
dp_t^{DT}	36.34	5.71	18.78	24.24	35.56	30.92	34.79	2.05
Historical Mean	–	–	25.05	32.25	–	38.39	43.34	–

Table 6. Out-of Sample Tests. We report statistics on H-year ahead forecast errors for stock returns. The sample starts in 1955 and we construct first forecast in 1982. 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_{OS} compares the forecast error from forecasts based on the historical mean with the forecast from predictive regressions.

Long Horizon Regressions	Real Stock Market Returns			Real Dividend Growth			Real GNP Growth		
	3 Years	5 Years	10 Years	3 Years	5 Years	10 Years	3 Years	5 Years	10 Years
dp_t	0.176	0.319	0.612	-0.107	-0.095	0.002	0.029	0.048	0.091
(t-stat)	(1.21)	(2.00)	(2.13)	(-1.02)	(0.79)	(0.01)	(0.57)	(0.72)	(1.36)
$\left[\bar{R}^2\right]$	[0.05]	[0.10]	[0.16]	[0.04]	[0.02]	[-0.01]	[0.00]	[0.02]	[0.03]
$\overline{dp_t}$	0.000	-0.000	-0.003	0.001	0.002	0.001	-0.062	-0.134	-0.280
(t-stat)	(0.05)	(-0.08)	(-0.68)	(0.58)	(0.41)	(0.37)	(-0.65)	(-0.99)	(-1.25)
$\left[\bar{R}^2\right]$	[-0.01]	[-0.01]	[0.01]	[0.03]	[0.02]	[0.00]	[0.02]	[0.10]	[0.30]
$dp_t - \overline{dp_t}$	0.715	1.109	1.243	-0.095	0.087	0.454	0.026	-0.013	-0.115
(t-stat)	(5.01)	(8.74)	(6.15)	(-0.77)	(0.63)	(2.33)	(0.44)	(-0.13)	(-0.36)
$\left[\bar{R}^2\right]$	[0.26]	[0.40]	[0.29]	[0.00]	[-0.00]	[0.12]	[-0.01]	[-0.01]	[0.01]

Table 7. Forecasting Regressions, dependent variables are reported in the first row, regressors are reported in the first colum. We report Newey and West HAC consistent t-statistics and adjusted R^2 for each model. Annual sample 1909-2008 (1929-2008 for real GDP growth).

APPENDIX B: FIGURES

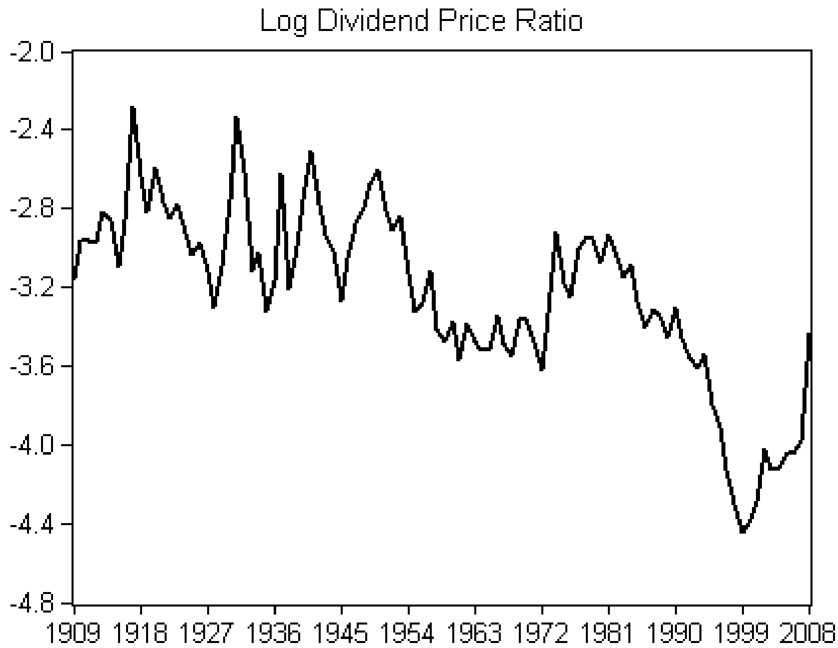

Figure 1. The time series of log dividend price ratio ($d_t - p_t$). Annual data from 1909 to 2008.

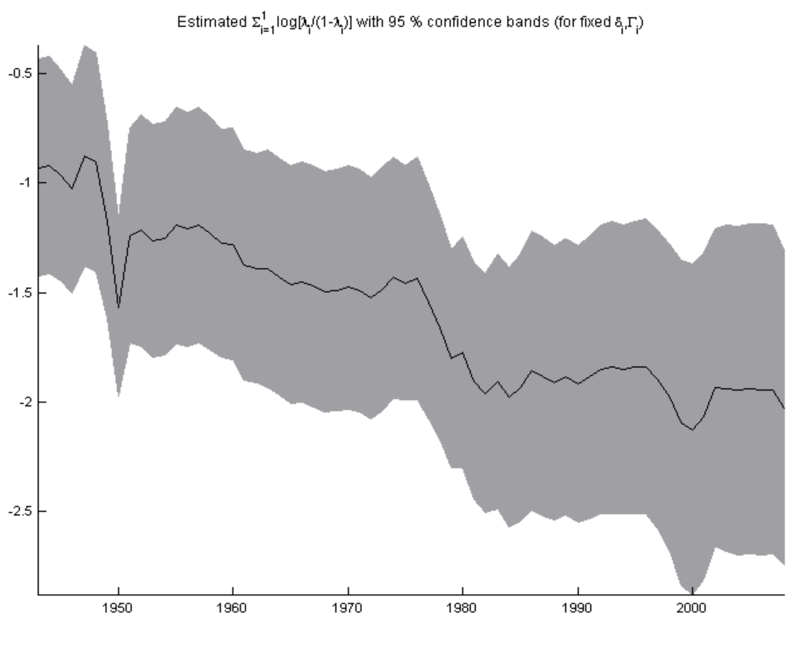

Figure 2. Recursive Eigenvalue Test using log nominal prices and log nominal dividends.

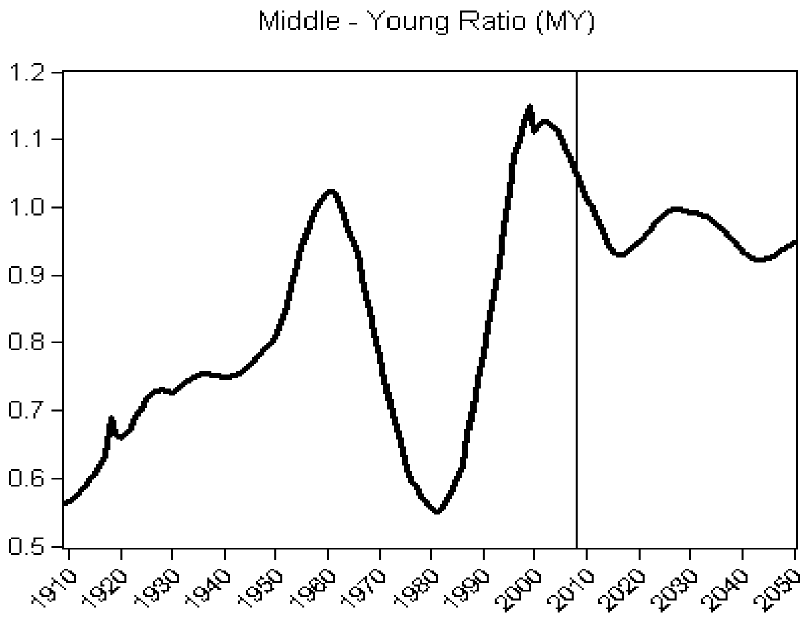

Figure 3a. Middle-Young (MY) ratio and projections provided by Bureau of Census for the period 2009-2050.

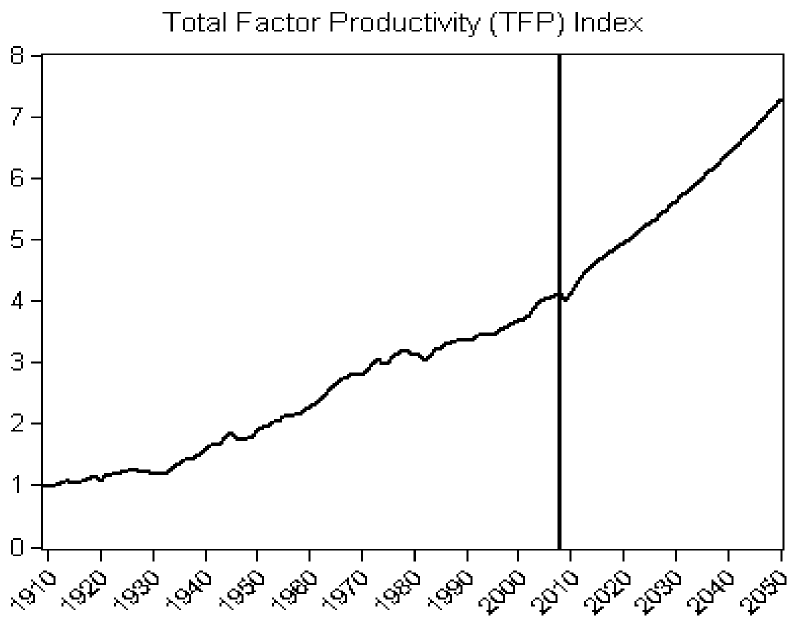

Figure 3b. Total Factor Productivity (TFP) index normalized to 1 at the beginning of our sample and projections provided by Congressional Budget Office (CBO) for the period 2009-2050.

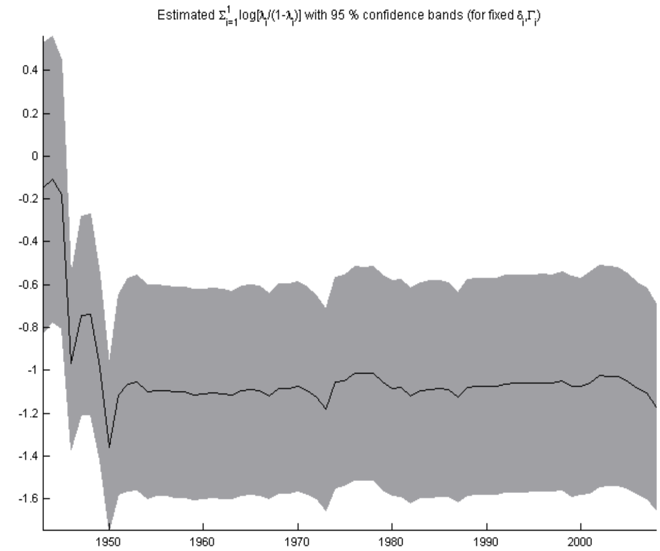

Figure 4a. Recursive Eigenvalue test. We include nominal log dividends, log prices, total factor productivity (TFP) and middle-young ratio(MY).

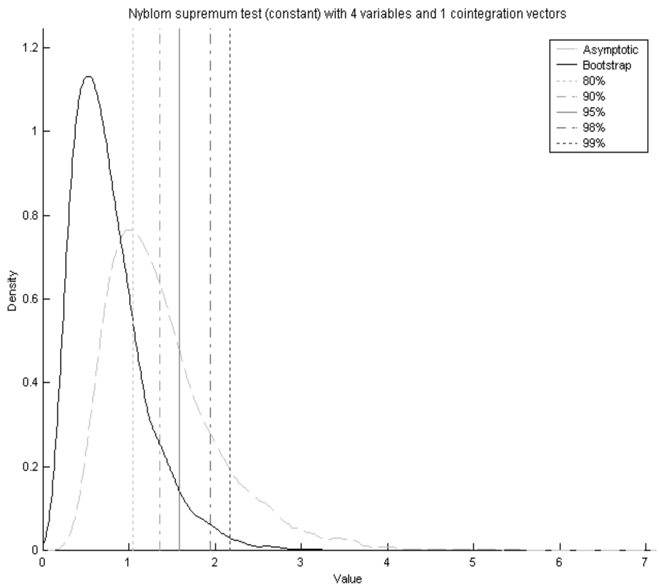

Figure 4b. Nyblom Bootstrap Test for a our model. The sup-statistics is 0.4849 (with mean-statistics = 0.2036) for a vector error correction(VEC) model of order one allowing for only one cointegration relation.

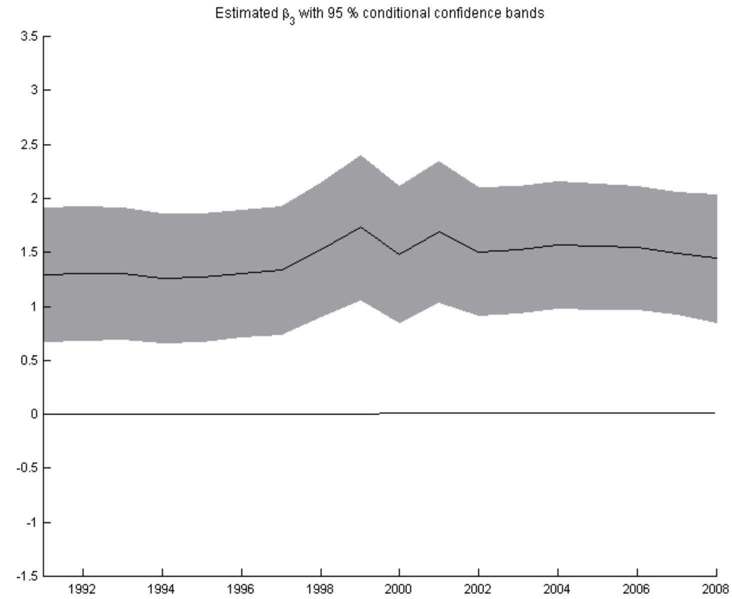

Figure 4c. Parameter stability. Recursive parameter estimation of β_3 in the vector error correction(VEC) model.

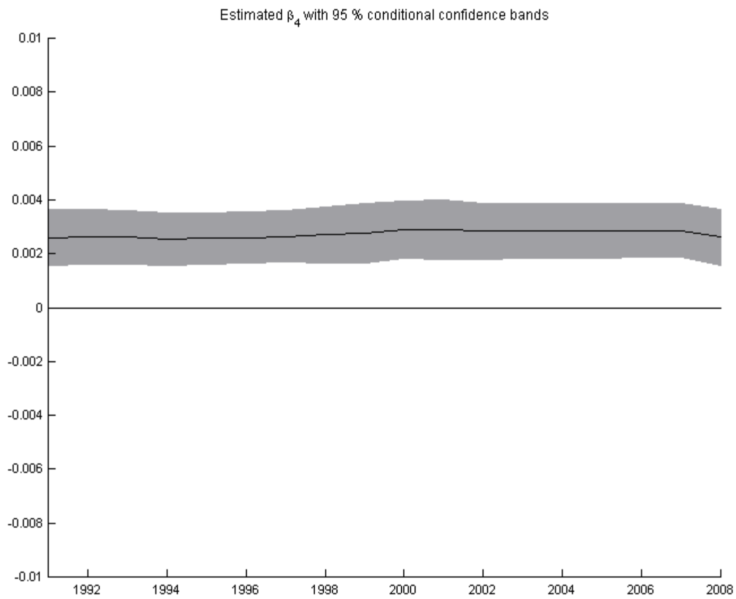

Figure 4d. Parameter stability. Recursive parameter estimation of β_4 in the vector error correction(VEC) model.

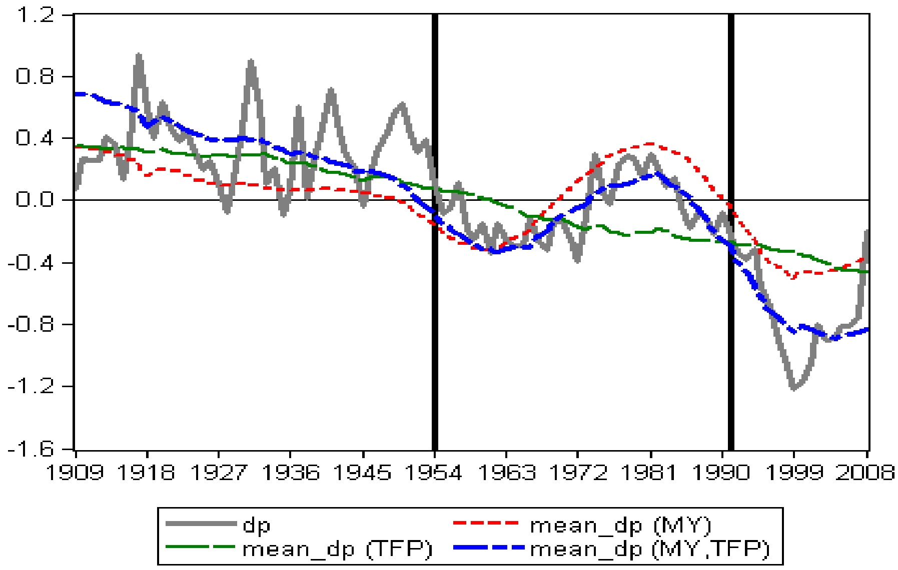

Figure 5a. Log of dividend-price ratio, time varying mean driven by MY, mean_dp(MY), driven by TFP, mean_dp(TFP), and driven by both MY and TFP, mean_dp(MY, TFP). The sample period is 1909-2008.

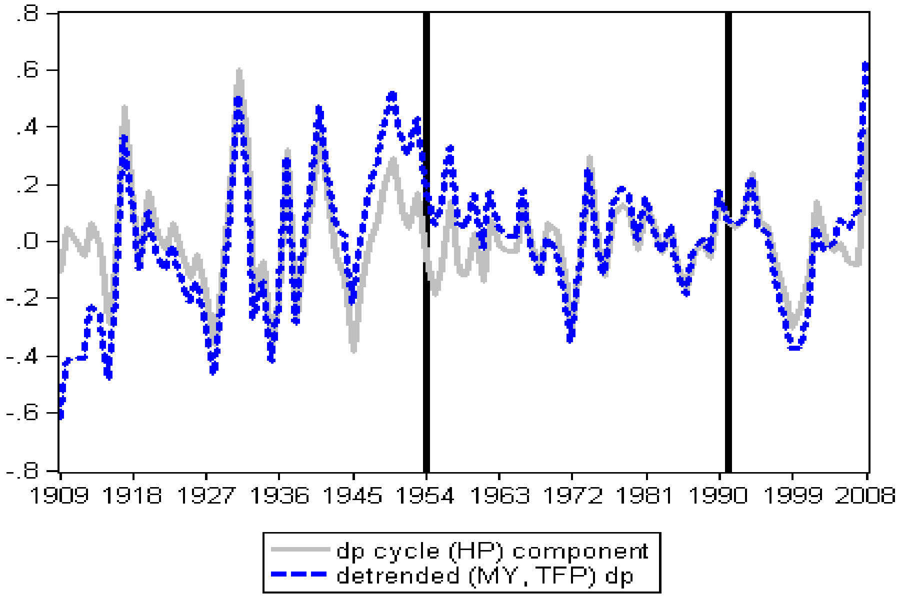

Figure 5b. Log dividend-price cycle component obtained using HP filter and detrended log dividend-price ratio, dp_t^{DT}.

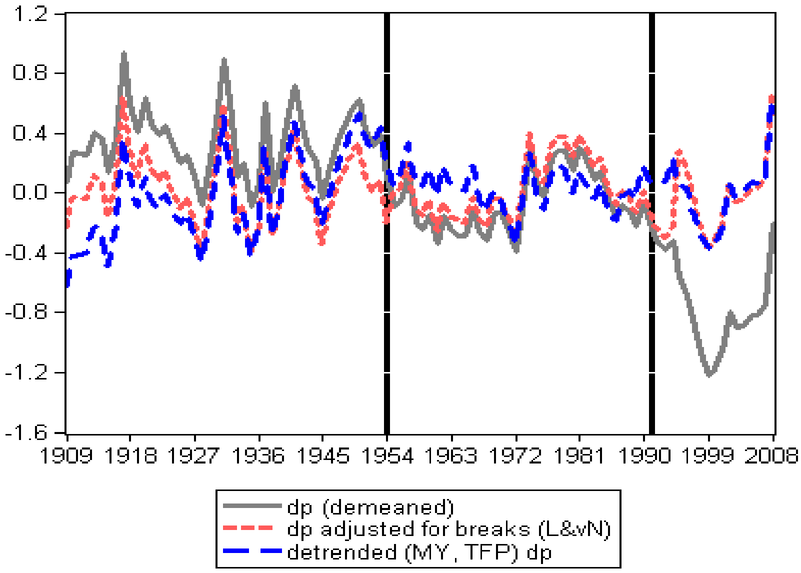

Figure 5c. Log dividend-price ratio, log dividend price ratio adjusted for exogenous breaks, dp_t^{LvN} (LvN, 2007) and detrended log dividend price ratio, dp_t^{DT}.

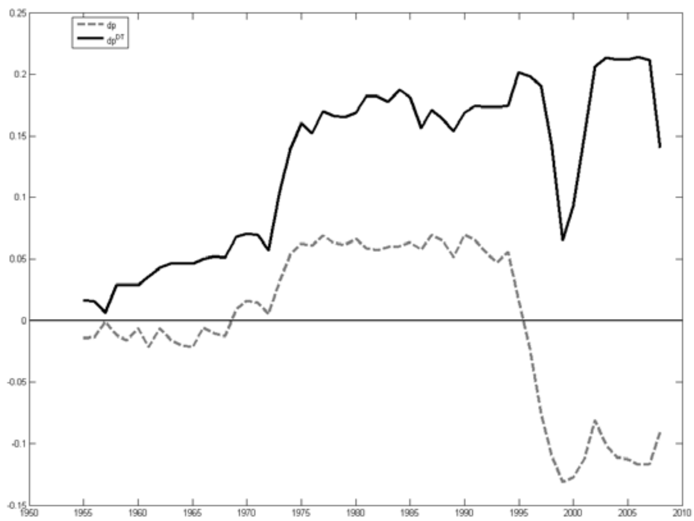

Figure 6. Out-of sample performance for annual predictive regression. Difference between cumulative squared forecast errors based on a linear regression incuding just a constant and a linear regression including the predictive variable (dp^{DT} or dp). The units are in percent. First forecast in 1955.

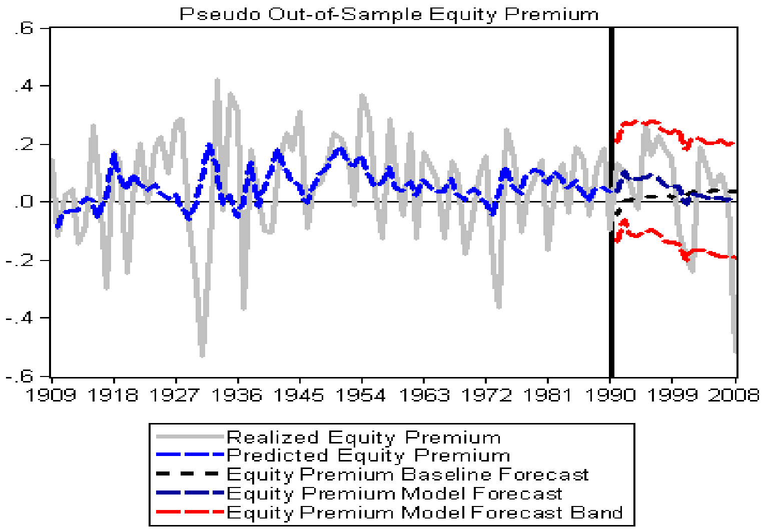

Figure 7. Pseudo Out-of-Sample Forecast of Equity Premium using the vector error correction (VEC) model with p_t, d_t as endogenous variables and MY_t ,TFP_t as exogenous variables. We also plot the predictions from a bivariate VEC model including only p_t and d_t.

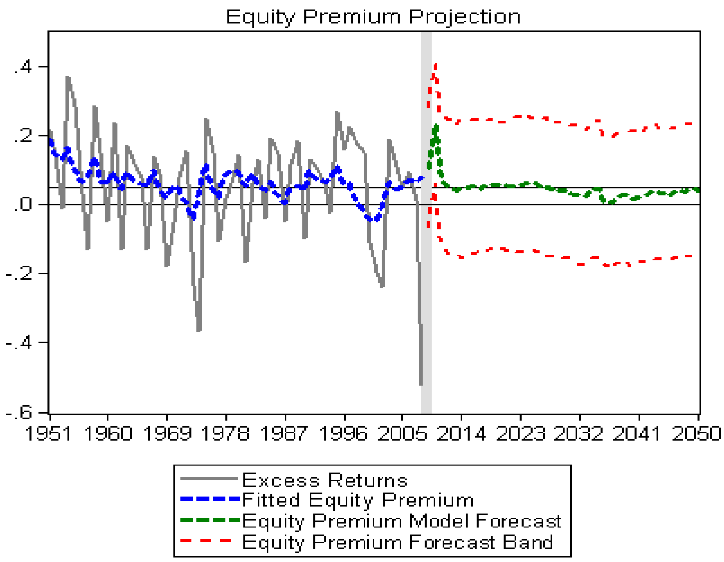

Figure 8 Long-run Equity Premium Projections. We estimate the VEC model in the full sample 1909-2008 and plot the fitted equity premium. We solve the model through stochastic simulations (1000 repetitions) for the period 2009-2050 and plot the average equity premium together with one standard deviation forecast bands. The horizontal line at 0.051 indicates the historical average equity premium.

APPENDIX A: The Statistical Model

We consider the following statistical model:

$$\mathbf{y}_t = \sum_{i=1}^{n} \mathbf{A}_i \mathbf{y}_{t-i} + \mathbf{u}_t \tag{12}$$

$$\mathbf{y}_t \text{ is an } mx1 \text{ vector of variables} \tag{13}$$

This model can be re-written as follows

$$\begin{aligned} \Delta \mathbf{y}_t &= \mathbf{\Pi}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Pi}_2 \Delta \mathbf{y}_{t-2} + ...\mathbf{\Pi}_{n-1} \Delta \mathbf{y}_{t-n+1} + \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{u}_t \\ &= \sum_{i=1}^{n-1} \mathbf{\Pi}_i \Delta \mathbf{y}_{t-i} + \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{u}_t, \end{aligned} \tag{14}$$

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 $(\Pi) = m$, full. The system is stationary;
3. rank $(\Pi) = k < m$. The system is non-stationary but there are k cointegrating relationships among the considered variables. In this case $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$, where α is an $(m \times k)$ matrix of weights and $\boldsymbol{\beta}$ is an $(k \times m)$ 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 m characteristic roots and we order them as follows $\lambda_1 > \lambda_2 > ...\lambda_m$. 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 $\ln(1-\lambda_2) = \ln(1-\lambda_3) = ... = \ln(1-\lambda_m) = 0$. 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:

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

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 m 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.

APPENDIX B: Robustness analysis for the cointegrating evidence

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

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

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).

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. 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+[5].

SR did not attract a significant coefficient when we augmented our cointegrating specification with this variable.

[5] 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 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.

APPENDIX C: Below, we describe thetime-series used in our empirical investigation.

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

Stock Prices: S&P 500 index yearly prices from 1909 to 2008 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 2008.

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[6] 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 1934-2008. 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

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

[6] 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.

(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 (1934-2008); 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 1921-2008.

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 (1934-2008) and available from 1920 to 2008.

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 to 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 (1952-2008, 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 to 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) level , which measures the technology accumulation over time. Shocks to this variable has been considered as 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 1948-2008, 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 collected the data for the same period also from (U.S. Bureau of the Census, 1975, Series W6) , the results remain the same. As a robustness check, 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/

Congressional Budget Office: http://www.cbo.gov/doc.cfm?index=10457&zzz=39352

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/