

Demographics and the Term Structure of Stock Market Risk

Carlo A. Favero
Bocconi University, IGIER & CEPR

Andrea Tamoni
Bocconi University

First version: February 2010. This version: May 7, 2010

ABSTRACT

The term structure of stock market risk depends on the predictability of stock market returns at different horizons. Intuitive reasoning, formal modeling and empirical evidence show that demographic trends are a slow-moving information variable, whose forecasting power is weak at high frequency but becomes strong at low frequencies, when the effect of the noisy component of stock market fluctuations disappears. We show that the forward solution of the dynamic dividend growth model does progressively eliminate the noise component as the horizon increases. Direct regressions of returns at different horizon on the relevant predictors capture this feature of the model, while VAR based multiperiod iterated forecasts do not, as they are derived from a backward solution of a reduced form empirical model. The combination of direct regression with the use of demographic trends leads us to find a steeply downward sloping term structure of stock market risk.

JEL classification: G17,C53,E44.

Keywords: dynamic dividend growth model, long-run returns predictability, stock market risk, demographics, direct regressions, multiperiod iterated forecasts.

*Carlo A. Favero is at Bocconi University, IGIER & CEPR. Andrea Tamoni is at Bocconi University. Thanks to Francesco Corielli for a number of helpful suggestions. The usual disclaimer applies.

I. Introduction

This paper examines the consequences for the term structure of stock market risk¹ of the impact of demographics on the predictability of stock market returns. In particular we concentrate on the evidence that a slow moving variable determined by demographics has a strongly significant impact on predictability whose strength is positively related to the forecasting horizon.

The best way to introduce our work is to refer the reader to Figures 1(a) and 1(b). Figure 1(a) illustrates, over about one century of US data, the relationship between 1-year real stock market returns and a demographic variable, MY , the ratio of middle-aged to young population. Figure 1(b) relates again demographics and stock market fluctuations, but 20-year real annualized returns are considered instead of 1-year return. The comovement between demographics and stock market returns is negligible for annual returns and remarkable for 20-year returns.

[Insert Figures 1(a), 1(b) about here.]

Intuitive reasoning and formal modeling (see, for example, Geanakoplos, Magill and Quinzii (2004), henceforth GMQ) hints at demography as an important variable to determine the long-run behavior of the stock market, while it is difficult to imagine a relationship between high-frequency fluctuations in stock market prices and a slow-moving trend determined by demographic factors.

The empirical evidence confirms this intuition. In a recent paper Favero, Gozluklu and Tamoni (2010) have shown that, consistently with the prediction of the theoretical model by GMQ, the ratio of middle-aged to young population, MY , captures the slow moving but time-varying mean of the US, dividend-price ratio. A forecasting model for stock market returns over a century of US annual data that uses as predictors the dividend-price ratio and MY performs very well in forecasting stock market returns at long horizons, with a forecasting performance strongly and positively related to the length of the forecasting horizon.

When demographic trends are used to model the slow moving fluctuations in the dividend-price ratio a natural decomposition of this variable into an high volatility “noise” component, and a low-volatility “information” component naturally emerges. The dominance of the “noise” component in short-horizon returns and of the “information” component in long horizon returns implies a positive relation between predictability of returns and forecasting horizon and a negatively sloped term structure of risk.²

¹defined as the conditional variance and covariance per period.

²The use of the term “noise” and “information” has been inspired by our reading of Chapter 3 of Taleb

This paper concentrates on the implication of the relationship between stock market returns and demographics for the slope of the term structure of stock market risk.

While the positive relation between predictability of returns and the forecasting horizon is widely recognized in the literature, the negative slope of the term structure of stock market risk is a controversial issue (see, Campbell and Viceira (2002), Campbell and Viceira (2005), Pastor and Stambaugh (2008) and Pastor and Stambaugh (2009)).

We contribute to the debate by constructing a small “structural” model that allows for an explicit role for demographics in the dynamic dividend growth model. In particular, the dividend-price ratio is modeled as a function of a temporary “noise” component and of a persistent “information” component related to demographics. The term structure of stock market risk is then derived by the simultaneous estimation of a system for stock market returns at different horizon obtained by the forward solution of the model. We show that the forward solution of the dynamic dividend growth model (Campbell and Shiller 1988) does naturally progressively eliminate the noise component as the horizon increases. The explicit comparison of our results with the traditional VAR based methods to derive the term structure of stock market risk shows that the combination of direct regressions methods with the inclusion of the demographic variable in the information set relevant for long-horizon regressions makes the term structure of stock market risk more steeply downward sloping.

The paper is organized as follows. The first section is devoted to illustrate the placement of our contribution in the literature. In the second section a simple structural model linking demographics and the dynamic dividend growth model is introduced to derive the term structure of stock market risk via direct estimation of a system for returns at different horizon. In the third section the term structure of stock market risk derived from the forward-looking solution of the model and estimated via direct regressions is compared with that approximated via the multi-step iterated forecast and the backward solution of a VAR. The last section concludes.

II. Related Literature

This paper adds to a considerable literature on the relation between the predictability of stock market returns and the term structure of stock market risk.

In describing the “verdict of history” on asset returns on a long-sample(1802-1996) of US historical data J.J.Siegel((Siegel 2007),pp.32), pointed out that “...stocks are riskier than fixed-income investment over short-term holding periods. But once the holding period in-

(2001).

creases to between 15 and 20 years, the standard deviation of average annual returns,..., become lower than the standard deviation of average bond and bill returns...”.

This statement on unconditional second moments has been strengthened by Campbell-Viceira(2002, 2005) who exploited the predictability of returns by estimating VAR models for returns and predictors and by using VAR-based multiperiod iterated forecasts to find that the conditional variance of stock return does not grow in proportion with the investment horizon but it grows more slowly. As a consequence the term-structure of stock market risk is downward sloping and the findings by Siegel on the property of the unconditional distribution of stocks returns are extended and strengthened when the conditional distribution of returns is used to measure stock market risk.

However, the downward sloping term-structure of stock market risk, has been recently questioned by Pastor-Stambaugh(2008,2009) who show that, allowing for coefficient uncertainty and imperfect predictors in a Campbell-Viceira type of VAR, the conditional variance of stock returns does increase with the horizon and it can even exceed the unconditional variance and the current variance.

The VAR-based approach to measure the term structure of stock market risk uses the (log) dividend-price ratio as the predictor for stock market returns at different horizon. This specification has its foundation in the dynamic dividend growth (DDG) model proposed by Campbell and Shiller (1988). In fact, the DDG model predicts that (log) dividends and prices share a common stochastic trend and that deviations of (log) prices from the common trend in (log) dividends summarize expectations of either stock market returns, or dividend growth or some combination of the two. The empirical investigation of the DDG has established a number of relevant results.

First, the log dividend-price ratio, dp_t , does not have important long-horizon forecasting power for future discounted dividend-growth (Campbell (1991), Cochrane (1991), Campbell, Lo and Mackinlay (1997), Cochrane (2001) and Cochrane (2008a)). Second, dp_t is a very persistent time-series and forecasts stock market returns and excess returns over horizons of many years (Fama and French (1988), Campbell and Shiller (1988), Cochrane (2001, Ch. 20), and Cochrane (2008a)). Third, the very high persistence of dp_t has led some researchers to question the evidence of its forecasting power for returns, especially at short-horizon. Careful statistical analysis that takes full account of the persistence in dp_t provides little evidence in favour of predictability of stock-market returns and excess returns based on the log-dividend-price ratio (Nelson and Kim (1993); Stambaugh (1999); Ang and Bekaert (2007); Valkanov (2003); Goyal and Welch (2003) and Welch and Goyal (2008)). Structural breaks have also been found in the relation between dp_t and future returns (Neely and Weller (2000) and Weller(2000) and Paye and Timmermann (2006), Rapach and Wohar (2006)).

A recent strand of the empirical literature has related the contradictory evidence on the dynamic dividend growth model to the potential weakness of its fundamental hypothesis that log dividend-price ratio is a stationary process (Lettau and Nieuwerburgh (2008), LVN henceforth). LVN use a century of US data to show evidence on the breaks in the constant mean \overline{dp} . The potential time-variation of the linearization point creates a link between demographics and the DDG model.

In a related paper Favero, Gozluklu and Tamoni(2010) have rationalized the structural breaks found by LVN in the dividend-price process with demographic trends. They point out that theoretical model by GMQ predicts that a specific demographic variable, MY , the ratio of middle-aged to young population, explains fluctuations in the dividend yield.

GMQ consider an overlapping generation model in which the demographic structure mimics the pattern of live births in the US, that have featured alternating twenty-year periods of boom and busts. They conjecture that the life-cycle portfolio behavior (Bakshi and Chen 1994), which suggests that agents should borrow when young, invest for retirement when middle-aged, and live off their investment once they are retired, plays an important role in determining equilibrium asset prices. Consumption smoothing by the agents, given the assumed demographic structure requires that when the MY ratio is small (large), 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. As the dividend-price ratio is negatively related to fluctuations in prices, the model predicts a negative relation between this variable and MY . When the GMQ model is taken to the data via the conjecture that fluctuations in MY could capture a slowly evolving mean in dp_t within the dynamic dividend growth model, strong evidence is found in favour of using this variable together with the dividend-price ratio in long-run forecasting regressions for stock market returns. Interestingly, the fluctuations in MY match very well the break-points identified by LVN in the fifties and the nineties. This paper differs from Favero, Gozluklu and Tamoni(2010) by concentrating on the implications of the relation between demographics and the dividend-price ratio for the term structure of stock market risk. We propose an empirical strategy potentially capable of identifying separately the importance of demographic variables for high-frequency and low-frequency fluctuations in asset prices. Investigations conducted in the literature on the interaction between asset prices and demographic variables have traditionally concentrated either on high-frequency or low frequency fluctuations but have never considered an empirical framework based on the dynamic dividend growth model, capable of accommodating both of them, with a different role (see Erb, Harvey and Viskanta (1997), Poterba (2001), Goyal (2004), Ang and Maddaloni (2005) and DellaVigna

and Pollet (2005)).

III. The Dynamic Dividend Growth Model, Demographics and the Term Structure of Stock Market Risk.

The objective of this paper is to propose an alternative model to measure the term structure of stock market risk.

Consider the continuously compounded stock market return from time t to time $t + 1$, \mathbf{r}_{t+1} . Define μ_t , the conditional expected log return given information up to time t , as follows:

$$\mathbf{r}_{t+1} = \mu_t + \mathbf{u}_{t+1}$$

where \mathbf{u}_{t+1} is the unexpected log return. Define the k -period cumulative return from period $t + 1$ through period $t + k$, as follows:

$$\mathbf{r}_{t,t+k} = \sum_{i=1}^k \mathbf{r}_{t+i}$$

The term structure of risk is defined as the conditional variance of cumulative returns, given the investor's information set, scaled by the investment horizon

$$\Sigma_r(k) \equiv \frac{1}{k} Var(\mathbf{r}_{t,t+k} | D_t) \quad (1)$$

where $D_t \equiv \sigma\{z_k : k \leq t\}$ consists of the full histories of returns as well as predictors that investors use in forecasting returns.

In the light of the results of the empirical investigations on the DDG model and on the evidence of the relation between demographics and the dividend price ratio dp_t , we consider the following small "structural" model:

$$r_{t+1}^s = \Delta d_{t+1} - \rho [dp_{t+1} - \bar{dp}_{t+1}] + [dp_t - \bar{dp}_t] \quad (2)$$

$$\Delta d_{t+1} = \varepsilon_{1,t+1} \quad (3)$$

$$dp_{t+1} = \varphi_{22} dp_t + \varphi_{23} MY_{t+1} + \varepsilon_{2,t+1} \quad (4)$$

$$\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \sim \begin{bmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \begin{matrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{matrix} \end{bmatrix}$$

Equation (2) defines real returns using the log-linearized approximation proposed by

Campbell and Shiller (1988) and has no error attached to it³. The specification of this equation differs from the standard linearization only in that the equilibrium long-run mean around which the dividend price ratio is linearized is not constant. In fact, the dividend-price ratio itself depends on the age structure of the population, MY, a slowly evolving highly predictable variable (the Bureau of Census makes available through its web page projections of this variable up to 2050). Such a modification is justified by the theoretical model of Geanakoplos et al. (2004) and by the empirical evidence provided in Favero, Gozluklu and Tamoni (2010). MY constitutes the information component of the dividend price ratio and there is no uncertainty attached to it: we take it as an exogenous variable whose path for the relevant future is known. However, the dividend-price is also affected by some short-term idiosyncratic noise $\varepsilon_{2,t+1}$. In the process generating dividends we take $|\varphi_{22}| < 1$: dividend-price are mean reverting toward their long-run trend determined by the information variable and the effect of the noise shock on the process is only temporary. In fact, our empirical results show that the speed of mean reversion of the dividend-price ratio toward its long-run mean determined by demographic trends is much higher than that of the dividend-price process itself. This reduced persistence in the financial ratio is important for at least two reasons. First it makes inference less problematic since there is little doubt on the stationarity of dividend-price ratio around a demographic trend. Second the reduced half-life of the price ratios shocks makes innovations to expected returns more likely to be linked to many plausible economic risk factors, such as those linked to business cycles. In the model specification there is also a second shock, the innovation to real dividend growth $\varepsilon_{1,t+1}$. This is a simple parameterization for the dividend growth process, that is fully consistent with the evidence of very little predictability of dividend growth⁴.

By solving eq. (2) forward we obtain:

$$\sum_{j=1}^m \rho^{j-1} (r_{t+j}^s) = [dp_t - \bar{dp}_t] + \sum_{j=1}^m \rho^{j-1} (\Delta d_{t+j}) - \rho^m [dp_{t+m} - \bar{dp}_{t+m}] \quad (5)$$

Eq. (5) clearly shows that the model implies the predictability of **long-run** returns. In fact the equation states that the deviations of the dividend/price ratio from its equilibrium value should have predictive power for m -period ahead stock market returns (and/or dividend

$${}^3r_{t+1}^s = \ln \left(\frac{P_{t+1} + D_{t+1}}{P_t} \right),$$

$$dp_t = \ln \left(\frac{D_t}{P_t} \right)$$

⁴Following the approach of Lacerda and Santa-Clara (2010) we also tried a specification where agents forecast the dividend growth rate from the average of past dividend growth rates, i.e. $E_t[\Delta d_{t+k}] = \bar{g}_t$. Results are unaffected and available upon requests.

growth), for sufficiently large m so that the transversality condition can be applied and the last term in the equation becomes of negligible importance. To bring the model to the data, we assume that the relevant linearization value for computing returns from time t to time $t+m$ is the conditional expectation of the dividend-yield for time $t+m$, given the information available at time t . We then have

$$\begin{aligned}
\sum_{j=1}^m \rho^{j-1} (r_{t+j}^s) &= dp_t - \left[\varphi_{22}^m dp_t + \sum_{j=1}^m \varphi_{22}^{j-1} \varphi_{23} MY_{t+m+1-j} \right] + u_{t+m} \quad (6) \\
&= (1 - \varphi_{22}^m) dp_t - \sum_{j=1}^m \varphi_{22}^{j-1} \varphi_{23} MY_{t+m+1-j} + u_{t+m} \\
u_{t+m} &= \sum_{j=1}^m \rho^{j-1} (\varepsilon_{1,t+j}) - \rho^m \sum_{j=1}^m \varphi_{22}^{j-1} \varepsilon_{2,t+m+1-j}
\end{aligned}$$

Note that the relevance of the noisy component in the distribution of m -period returns natural decreases with the horizon: as the horizon gets longer the mean-reversion of the dividend-yield process around the information variable makes the informative content of this variable dominant. The speed at which the noise disappears depends on the speed of mean reversion of the dividend process and on the discount parameter ρ . However, even for values of ρ close to unity, the mean reversion in dividend-prices is sufficient to cause a cancellation of the noise $\varepsilon_{2,t}$. The second component of the noise in m -period returns is the uncertainty in the dividend process that dies out much more slowly than the effect of the noise $\varepsilon_{2,t+j}$ and it becomes persistent when ρ approaches the unit value. Eq. (6) implies that the fit of direct predictive regressions projecting returns at different horizon on the information available at time t should improve with the horizon. It also predicts that the residuals of such predictive regressions have a moving-average component that should be taken care of in estimation. This is a well-known result (see for example, Valkanov (2003)). Interestingly, the model also predicts that the coefficient on the dividend-yield in the projections of long-horizon returns on this variable should be increasing with the horizon.

We measure the term structure of stock market risk by estimating the following “structural” system of eleven equations⁵:

⁵Our “structural” estimation is similar to that by Van Binsbergen and Koijen (2009) with two main differences: equations at all relevant horizons are simultaneously estimated and all variables included in the model are observable.

$$\begin{aligned}
\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s) &= \delta_{0,m} + \frac{1}{\sqrt{m}} (1 - \varphi_{22}^m) dp_t - \frac{\varphi_{23}}{\sqrt{m}} \left(\sum_{j=1}^m \varphi_{22}^{j-1} MY_{t+m+1-j} \right) + u_{t+m} \quad (7) \\
m &= 1, \dots, 10 \\
dp_{t+1} &= \varphi_{20} + \varphi_{22} dp_t + \varphi_{23} MY_{t+1} + \varepsilon_{2,t+1}
\end{aligned}$$

The specification of (7) slightly differs from the model in that we use as a dependent variable the unweighted annualized period-returns ($\rho = 1$). This is because the objective of our exercise is to compare the term structure of stock market risk obtained by direct regression and by iterative multi-step iterated VAR based forecasts. To assess the potential cost of the approximation introduced by using $\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s)$ instead of $\sum_{j=1}^m \rho^{j-1} (r_{t+j}^s)$ an unrestricted version of (7) is also estimated to perform a test of the validity of the relevant restrictions:

$$\begin{aligned}
\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s) &= \delta_{0,m} + \frac{\delta_{1m}}{\sqrt{m}} dp_t + \frac{\delta_{2m}}{\sqrt{m}} \left(\sum_{j=1}^m \varphi_{22}^{j-1} MY_{t+m+1-j} \right) + u_{t+m} \quad (8) \\
m &= 1, \dots, 10 \\
dp_{t+1} &= \varphi_{20} + \varphi_{22} dp_t + \varphi_{23} MY_{t+1} + \varepsilon_{2,t+1}
\end{aligned}$$

Note that (7) and (8) are both specified with $\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s)$ as the dependent variable to obtain directly the conditional annualized standard error of returns from the standard error of the regression.

We estimate the model on a dataset of annual observations for the period 1910-2008. The data are from Welch and Goyal (2008)⁶, who provide detailed descriptions of the data and their sources. Stock returns are measured as continuously compounded returns on the S&P 500 index, including dividends. To compute real returns we calculate inflation rate from the CPI (all urban consumers). The predictor for the equity premium is the dividend-price ratio, computed as the difference between the log of dividends paid on the S&P 500 index and log of stock prices (S&P 500 index), where dividends are measured using a one-year moving sum.

The results of the estimation are reported in Table 1. The GMM estimation uses the OLS normal conditions to estimate the restricted model parameters φ_{22} and φ_{23} . The Table

⁶The data are available at www.bus.emory.edu/AGoyal/Research.html.

shows an highly significant effect of MY both in the equation for dp_t and in all ten predictive regressions. The performance of the restricted model, that estimates only two parameters in addition to eleven constants, is very similar in term of adjusted R^2 and standard error of the equations to that of the unrestricted model that estimates twenty more parameters and the restrictions are accepted by the relevant chi-square test. Put differently, the normal conditions for the short- and long-horizon moments are satisfied at the model-implied long-horizon predictability coefficients.

[Insert Table 1 about here.]

The term structure of stock market risk described by the estimation of the structural system of direct regression is steeply downward sloping as it can be read directly off the standard errors of regressions reported in Table 1.

The estimates of the parameters φ_{22} and φ_{23} show that demographics are clearly significant in explaining the dividend-price ratio and that the dividend-price ratio is clearly mean reverting around a mean determined by MY. Figure 2 brings more evidence on this issue by reporting dp_t along with the time-varying linearization point used in the model and the breaks identified by LVN.

[Insert Figure 2 about here.]

Finally the term structure of stock market risk described by the estimation of the structural system of direct regression is steeply downward sloping as it can be read directly off the standard errors of regressions reported in Table 1.

IV. Understanding our Empirical Results on the Term Structure of Stock Market Risk

Our empirical results on the term structure of stock market risk differ rather importantly from those derived in the literature based on VAR model. To illustrate the point we consider first a simple representation of the VAR adopted by Campbell-Viceira by estimating a VAR for continuously compounded total stock market returns, r_t^s , and the log dividend price, dp_t :

$$(z_t - E_z) = \Phi_1 (z_{t-1} - E_z) + \nu_t$$

$$\nu_t \sim \mathcal{N}(0, \Sigma_\nu)$$

where

$$\begin{aligned}
z_t &= \begin{bmatrix} r_t^s \\ dp_t \end{bmatrix}, E_z = \begin{bmatrix} E_{r^s} \\ E_{d-p} \end{bmatrix} \\
\Phi_1 &= \begin{bmatrix} 0 & \varphi_{1,2} \\ 0 & \varphi_{2,2} \end{bmatrix} \\
\begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} &\sim \begin{bmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \begin{matrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{matrix} \end{bmatrix}
\end{aligned}$$

The bivariate model for returns and the predictor features a restricted dynamics such that only the lagged predictor is significant to determine current returns ($\varphi_{1,1} = 0$) and the predictor is itself a strongly exogenous variable ($\varphi_{2,1} = 0$).

Given the VAR representation and the assumption of constant Σ_ν

$$\begin{aligned}
Var_t[(z_{t+1} + \dots + z_{t+k}) | D_t] &= \Sigma_\nu + (I + \Phi_1)\Sigma_\nu(I + \Phi_1)' + & (9) \\
& (I + \Phi_1 + \Phi_1^2)\Sigma_\nu(I + \Phi_1 + \Phi_1^2)' + \dots \\
& +(I + \Phi_1 + \dots + \Phi_1^{k-1})\Sigma_\nu(I + \Phi_1 + \dots + \Phi_1^{k-1})'
\end{aligned}$$

from which we can derive:

$$\begin{aligned}
\Sigma_r(k) &= \frac{1}{k} \sum_{i=0}^{k-1} D_i \Sigma D_i' \\
D_i &= I + \Phi_1 \Xi_{i-1} \quad i > 0 \\
\Xi_i &= \Xi_{i-1} + \Phi_1^i \quad i > 0 \\
D_0 &\equiv I, \quad \Xi_0 \equiv I
\end{aligned}$$

Note that, under the chosen specification of the matrix Φ_1 we can write the generic term $D_i \Sigma D_i'$, as follows:

$$\begin{aligned}
D_i \Sigma D_i' &= \left(\begin{array}{c|c} M_{11} & M_{12} \\ \hline M'_{12} & M_{22} \end{array} \right) & (10) \\
M_{11} &= \Sigma_{1,1} + \Phi_{1,2} \Xi_{i-1}^{(22)} \Sigma'_{1,2} + \Sigma_{1,2} \Xi_{i-1}^{(22)'} \Phi'_{1,2} + \Phi_{1,2} \Xi_{i-1}^{(22)} \Sigma_{2,2} \Xi_{i-1}^{(22)'} \Phi'_{1,2} \\
M'_{12} &= \Xi_i^{(22)} \Sigma'_{1,2} + \Xi_i^{(22)} \Sigma_{2,2} \Xi_{i-1}^{(22)'} \Phi'_{1,2} \\
M_{22} &= \Xi_i^{(22)} \Sigma_{2,2} \Xi_i^{(22)'}
\end{aligned}$$

where we have used the fact that

$$\begin{aligned}\Xi_i &= \sum_{j=0}^i \Phi_1^j \\ &= \left(\begin{array}{c|c} 0 & \phi_{1,2} \sum_{j=0}^{i-1} \phi_{2,2}^j \\ \hline 0 & \sum_{j=0}^i \phi_{2,2}^j \end{array} \right)\end{aligned}$$

and

$$\begin{aligned}D_i &= I + \Phi_1 \Xi_{i-1} \\ &= \left(\begin{array}{c|c} I & \phi_{1,2} \sum_{j=0}^{i-1} \phi_{2,2}^j \\ \hline 0 & \sum_{j=0}^i \phi_{2,2}^j \end{array} \right)\end{aligned}$$

Eq. (10) implies that, in our simple bivariate example, the term structure of stock market risk takes the form

$$\sigma_r^2(k) = \sigma_1^2 + 2\varphi_{1,2}\sigma_{1,2}\psi_1(k) + \varphi_{1,2}^2\sigma_{2,2}^2\psi_2(k) \quad (11)$$

where

$$\begin{aligned}\psi_1(k) &= \frac{1}{k} \sum_{l=0}^{k-2} \sum_{i=0}^l \varphi_{2,2}^i \quad k > 1 \\ \psi_2(k) &= \frac{1}{k} \sum_{l=0}^{k-2} \left(\sum_{i=0}^l \varphi_{2,2}^i \right)^2 \quad k > 1 \\ \psi_1(1) &= \psi_2(1) = 0\end{aligned}$$

So the total stock market risk can be decomposed in three components: i.i.d uncertainty, mean reversion, uncertainty about future predictors. Without predictability ($\varphi_{1,2} = 0$) the entire term structure is flat at the level σ_1^2 . This is the classical situation where portfolio choice is independent of the investment horizon. The possible downward slope of the term structure of risk depends on the second term, and it is therefore crucially affected by predictability and a negative correlation between the innovations in dividend price ratio and in stock market returns ($\sigma_{1,2}$), the third term is always positive and increasing with the horizon when the autoregressive coefficient in the dividend yield process is positive. Note that the slope of the term structure of risk depends on the contemporaneous correlation between innovation in the returns and in the predictors and on the persistence of the predictors.

[Insert Table 2 about here.]

Table 2 summarizes the results of the estimation of the system. The estimation results

confirm the noisy nature of 1-year stock market returns and the high persistence of the dividend-price ratio. The covariance structure of the innovations is such that the unexpected log excess stock returns are highly negatively correlated with the innovations in the log dividend price ratio. Figure 3 plots the term structure of risk resulting from the estimation of the restricted VAR and its decomposition. The evidence of a downward sloping curve with risk halving from the one-year to the thirty year horizon replicates the results in Campbell and Viceira (2002), based on the estimation of a larger model including bond and stock excess returns, the nominal and real risk free rate together with the dividend-yield and the yield spread as predictors.

[Insert Figure 3 about here.]

Pastor and Stambaugh(2008, 2009) illustrate how the possibility of “imperfect predictors” in the predictive system adds to uncertainty and affects the entire term structure of risk. We illustrate such a possibility in our simple set-up by introducing a predictive relationship linking stock market returns to an unobserved variable μ_t , that in turns is only imperfectly related to the observed dividend-price. In this case, the relevant empirical model can be written as follows:

$$\begin{aligned}
 (r_t^s - E_{r^s}) &= (\mu_{t-1} - E_{r^s}) + u_{1,t} \\
 (dp_t - E_{dp}) &= \varphi_{22} (dp_{t-1} - E_{dp}) + u_{2,t} \\
 (\mu_t - E_{r^s}) &= \varphi_{32} (dp_t - E_{dp}) + u_{3,t} \\
 \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix} &\sim \begin{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{matrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{matrix} \end{bmatrix}
 \end{aligned} \tag{12}$$

form which the following VAR representation is derived:

$$\begin{aligned}
 (z_t - E_z) &= \Phi_1 (z_{t-1} - E_z) + \nu_t \\
 \nu_t &\sim \mathcal{N}(0, \Sigma_\nu)
 \end{aligned} \tag{13}$$

where

$$z_t = \begin{bmatrix} r_t^s \\ dp_t \\ \mu_t \end{bmatrix}, E_z = \begin{bmatrix} E_{r^s} \\ E_{dp} \\ E_{r^s} \end{bmatrix}$$

$$\Phi_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \varphi_{22} & 0 \\ 0 & \varphi_{32}\varphi_{22} & 0 \end{bmatrix}$$

$$\begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix} \sim \begin{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, & \sigma_1^2 & \sigma_{12} & \varphi_{32}\sigma_{12} + \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \varphi_{32}\sigma_2^2 + \sigma_{23} \\ \varphi_{32}\sigma_{12} + \sigma_{13} & \varphi_{32}\sigma_2^2 + \sigma_{23} & \varphi_{32}^2\sigma_2^2 + \sigma_3^2 + 2\varphi_{32}\sigma_{23} \end{bmatrix}$$

In our simple example the term structure of stock market risk takes now the form

$$\sigma_r^2(k) = \sigma_1^2 + 2\sigma_{13} + \sigma_3^2 + 2\varphi_{3,2}(\sigma_{12} + \sigma_{23})\psi_1(k) + \varphi_{2,2}^2\sigma_{2,2}^2\psi_2(k)$$

where

$$\psi_1(k) = \frac{1}{k} \sum_{l=0}^{k-2} \sum_{i=0}^l \beta_{22}^i \quad k > 1$$

$$\psi_2(k) = \frac{1}{k} \sum_{l=0}^{k-2} \left(\sum_{i=0}^l \beta_{22}^i \right)^2 \quad k > 1$$

$$\psi_1(1) = \psi_2(1) = 0$$

Note that the specification of the relevant VAR to project the term structure of risk involves an unobservable variable. VAR estimation needs that to deal with this problem that it is best solved within a Bayesian framework. Such a framework in turn generates naturally another source of volatility, namely parameters uncertainty. In the Pastor-Stambaugh(2008, 2009) framework the term-structure of risk can be decomposed in five components: the original three in eq. (11) plus other two, one reflecting uncertainty around the mean of the process generating returns and one reflecting parameters' uncertainty.

Table 3 shows instead the results from the estimation of the three-variate predictive system that it is specified following Pastor-Stambaugh (2008,2009). Within this Bayesian framework the prior beliefs on the correlation between innovations in the equation for returns and innovations in the equation for expected returns (i.e. ρ_{ν_1, ν_3} , the Stambaugh Correlation) substantially affects estimates of expected returns as well as various inferences about predictability. In our estimation we impose the belief $\rho_{\nu_1, \nu_3} < 0$ following the evidence of

Campbell (1991) and of Van Binsbergen and Koijen (2009). Interestingly, Robertson and Wright (2009) show that for a plausible range of ARMA parameters the Stambaugh Correlation is bounded away from zero and very close to (minus) unity. Therefore we follow Pastor and Stambaugh and specify an informative prior on ρ_{ν_1, ν_3} that the implied prior on ρ_{ν_1, ν_3}^2 has 99.9% of its mass above 0.5, with a mean of about 0.77⁷. The Table 3 reports the R^2 in the regression of r_{t+1} on $E[r_{t+1}|D_t]$ for the predictive system and shows that this R^2 is higher than the R^2 in Table 2 because $dp_t \in D_t$ and therefore the estimates of the expected returns from the predictive system are at least as precise as the estimates from the predictive regression and VAR. Moreover the R^2 (not reported) from a regression of μ_t on dp_t larger than 0.5 receive very little posterior probability suggesting that the predictor is not perfectly correlated with the latent expected return and therefore the predictive system has superior ability in extracting information and forming the proxy for the true unobservable μ_t .

[Insert Table 3 about here.]

[Insert Figure 4 about here.]

Figure 4 plots the conditional variance $Var[r_{t,t+k}|D_t]$ and its components for the three-variate predictive system with the dividend price as an observable predictor. It is interesting to note that the three components of $Var[r_{t,t+k}|\mu_t, \Theta, D_t]$ ⁸, namely the i.i.d. (top right panel), the mean reverting (mid left panel) and the uncertainty about future values of μ_t (mid right panel) are fairly similar to the one we compute under the VAR approach (see Figure 3). Therefore the sum of these three contribution, namely $Var[r_{t,t+k}|\mu_t, \Theta, D_t]$ (see continues line in top-right panel of Figure 4 and top right panel in Figure 3) is almost identical under the two alternative approaches. Nevertheless Pastor and Stambaugh (2009) show that other two important blocks affect the conditional variance $Var[r_{t,t+k}|D_t]$: one is the predictor imperfection that reflects the uncertainty about the current conditional expected returns and the other is the estimation risk that reflects uncertainty about parameters in Φ_1 . In particular, when we add the predictor imperfection component to the $Var[r_{t,t+k}|\mu_t, \phi, D_t]$ we obtain the dashed line in top-right panel of Figure 3 which shows no evidence for a downward sloping term structure of stock-market risk.

Why are our results so different from those based on traditional VAR analysis? There are two reasons: the inclusion of MY_t in the information set and the derivation of the term structure of risk via direct regression.

⁷As noted in PS 2008 this prior reflects the belief that at least half of the variance of market returns is due to discount rate news.

⁸We indicate with Θ the full set of parameters.

The direct regression of returns at different horizons on the relevant predictors in our model deliver the following term structure of stock market risk:

$$\begin{aligned}\sigma_r^2(k) &= \psi_1(k)\sigma_1^2 + \psi_2(k)\sigma_2^2 \\ \psi_1(k) &= \frac{1}{k} \sum_{j=1}^k \rho^{2(j-1)} \\ \psi_2(k) &= \frac{\rho^{2k}}{k} \sum_{j=1}^k \varphi_{22}^{2(j-1)}\end{aligned}\tag{14}$$

which is downward sloping as the effect of the noisy component of the dividend-price dies out as the horizon m increases.

Consider now the case in which a VAR is fitted to the data generated by eqs. (2) -(4):

$$\begin{aligned}r_{t+1}^s &= \varphi_{10} + \varphi_{12}dp_t + \varphi_{13}MY_{t+1} + \varepsilon_{1,t+1} - \rho\varepsilon_{2t+1} \\ dp_{t+1} &= \varphi_{20} + \varphi_{22}dp_t + \varphi_{23}MY_{t+1} + \varepsilon_{2,t+1}\end{aligned}$$

As noted by Cochrane (2008b), deliciously, the regression and “structural” model match almost perfectly ⁹.

However the derivation of the term structure of stock market risk by backward projection of the VAR on the information available at time t will deliver a different shape from that obtained by the direct regression:

$$\begin{aligned}\sigma_r^2(k) &= (\sigma_1^2 + \rho^2\sigma_2^2) - 2\varphi_{1,2}\rho\sigma_2^2\psi_1(k) + \varphi_{1,2}^2\sigma_2^2\psi_2(k) \\ \psi_1(k) &= \frac{1}{k} \sum_{l=0}^{k-2} \sum_{i=0}^l \varphi_{22}^i \quad k > 1 \\ \psi_2(k) &= \frac{1}{k} \sum_{l=0}^{k-2} \left(\sum_{i=0}^l \varphi_{22}^i \right)^2 \quad k > 1 \\ \psi_1(1) &= \psi_2(1) = 0\end{aligned}\tag{15}$$

⁹Recall that MY is an exogenous variable and therefore we can use MY_{t+1} in the VAR.

as long-run returns are obtained by aggregating high-frequency returns and the relevant model is solved backward, the effect of the noisy component does not decrease with the forecasting horizon. Direct comparison of (14) with (15) shows that the VAR based backward-looking term structure of risk delivers biased estimates of the DGP based forward looking term structure of stock market risk.

The results of the VAR estimation are reported in Table 4.

[Insert Table 4 about here.]

Note that the inclusion of MY as an exogenous variable in the VAR causes an improvement in the fit of both equations, reduces the persistence of the VAR (by reducing with respect to the benchmark the estimates of both φ_{12} and φ_{22}), but it does not affect significantly the correlation between the VAR innovations that stays negative at -0.85.

Figure 5 summarizes the main points of our paper by allowing a comparison between the term structure of risk derived by direct estimation of a model including MY_t , with the term structure of risk based on the recursive iteration of two VARs, one with and the other without MY_t . The results show clearly that MY_t plays an important role in determining the conditional mean of the system but also that the use of direct estimation of a forward-looking model rather than iterative recursive multi-step forecast is a source of a major shift in the measured term structure of stock market risk. Consistently with the prediction of the dynamic dividend growth model, such shift is far from being a parallel one. Importantly the measure of the term structure of risk based on the direct regression is very little affected by the “imperfect predictors” problem pointed out by Pastor-Stambaugh. In fact, the existence of imperfect predictors would change the interpretation but not the shape of the term structure based on direct regression reported in Figure 5. Only parameter uncertainty could be an issue, but this is an issue of a certainly limited relevance as the of the term structure of risk derived from our structural system is based on the estimation of very few parameters, all them very well determined. Figure 6 illustrates our initial point on “noise” and “information” by reporting actual and predicted returns at the 1-year and 10-year horizons. We consider predictions based on two different models: the three-variables VAR with one imperfect predictor, and the “structural” model with MY . Noise blurs the pictures at the short-end but at the long-end the “information” generated by using the predictor determined by demographics in the structural model strongly dominates. In fact the mean square error (MSE) and mean absolute error (MAE) for the predictive system are equal to 0.0023 and 0.0409, respectively, whereas the MSE and MAE for the direct regression are equal to 0.0012 and 0.0280.

[Insert Figures 5, 6 about here.]

V. Conclusions

We started this paper by arguing that the emergence of relative importance of “information” versus “noise” in the determination of stock market returns at different horizon could generate a downward sloping term structure of stock market risk. We have shown that this is indeed the case when i) a demographic variable is used to capture the slow-moving information component in the dividend-price ratio and in stock market returns and ii) direct regressions based on the structural estimation of a forward-looking specification consistent with the dynamic dividend growth model is adopted. We have also shown that the use of backward-looking iterated multi-step forecasts to derive the term structure of risk leads to an underestimation of the importance of the emergence of “information” as the horizon increases.

Table 1: System Estimation (1910-2009)

$dp_{t+1} = \varphi_{20} + \varphi_{22}dp_t + \varphi_{23}MY_{t+1} + \varepsilon_{2t+1}$											
UM:	$\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s)$	$= \delta_{0m} + \frac{\delta_{1m}}{\sqrt{m}} dp_t + \frac{\delta_{2m}}{\sqrt{m}} \left(\sum_{j=1}^m \varphi_{22}^{j-1} MY_{t+j} \right) + u_{t+m}$									$m = 1, \dots, 10$
RM:	$\frac{1}{\sqrt{m}} \sum_{j=1}^m (r_{t+j}^s)$	$= \delta_{0m} + \frac{1}{\sqrt{m}} (1 - \varphi_{22}^m) dp_t - \frac{\varphi_{23}}{\sqrt{m}} \left(\sum_{j=1}^m \varphi_{22}^{j-1} MY_{t+j} \right) + u_{t+m}$									
horizon m in years											
UM		1	2	3	4	5	6	7	8	9	10
δ_{1m}		0.18	0.38	0.48	0.61	0.70	0.73	0.78	0.86	0.89	0.89
(t -stat)		(3.87)	(5.57)	(5.61)	(6.96)	(7.99)	(7.17)	(7.25)	(8.73)	(7.63)	(6.15)
δ_{2m}		0.41	0.52	0.56	0.64	0.69	0.70	0.74	0.78	0.81	0.83
(t -stat)		(3.23)	(3.69)	(3.85)	(4.28)	(4.55)	(4.69)	(4.78)	(4.93)	(4.94)	(5.01)
φ_{22}		0.61									
(t -stat)		(9.21)									
φ_{23}		-0.83									
(t -stat)		(-3.79)									
RM											
φ_{22}		0.76									
(t -stat)		(19.31)									
φ_{23}		-0.53									
(t -stat)		(4.41)									
χ_{12}^2		13.45	χ_{20}^2	17.19							
		(0.34)		(0.64)							
σ_{DepVar}		0.195	0.198	0.187	0.185	0.181	0.174	0.172	0.173	0.171	0.168
$\sigma_{u_{t+m}}$	UM	0.188	0.179	0.164	0.152	0.140	0.131	0.125	0.118	0.112	0.109
$\sigma_{u_{t+m}}$	RM	0.189	0.179	0.164	0.152	0.141	0.133	0.127	0.120	0.115	0.112
$adjR^2$	UM	—	0.18	0.24	0.33	0.41	0.44	0.48	0.54	0.57	0.58
$adjR^2$	RM	0.06	0.18	0.23	0.32	0.40	0.42	0.46	0.52	0.54	0.55

Table 1: This table compares the univariate OLS long-horizon regression coefficients, to the GMM estimates that impose the restrictions suggested by the present-value model with demographics. The estimation is by GMM, where the moments are the OLS normal conditions. Standard errors are by Newey-West with optimal bandwidth selection. The first-stage weighting matrix is the identity matrix. σ_{DepVar} is the annualized unconditional standard deviation. $\sigma_{u_{t+m}}$ is the annualized conditional standard deviation of the compounded (over m periods) returns, i.e. our measure of stock market risk. The effective sample period is 1910-2009.

Table 2: A simple bivariate VAR (1910-2009)

$$\begin{aligned} (r_{t+1}^s - E_{r^s}) &= \varphi_{12} (dp_t - E_{dp}) + \nu_{1,t+1} \\ (dp_{t+1} - E_{dp}) &= \varphi_{22} (dp_t - E_{dp}) + \nu_{2,t+1} \end{aligned}$$

φ_{12} (<i>t-stat</i>)	φ_{22} (<i>t-stat</i>)	χ_2^2 $\varphi_{11}=0, \varphi_{21}=0$	σ_1	σ_2	$\frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$	$adj R_{r_{t+1}^s}^2$	$adj R_{dp_{t+1}}^2$
0.067	0.892	5.673	0.194	0.219	-0.856	0.02	0.78
(1.70)	(18.80)	(0.06)					

Table 2: The table reports coefficient estimates (with t-statistics in parentheses) and the R^2 statistic for each equation. We also report the standard deviations and correlations of residuals.

Table 3: A three-variate VAR with imperfect predictors (1910-2009)

$$\begin{aligned} (r_{t+1}^s - E_{r^s}) &= (\mu_t - E_{r^s}) + \nu_{1,t+1} \\ (dp_{t+1} - E_{dp}) &= \varphi_{22} (dp_t - E_{dp}) + \nu_{2,t+1} \\ (\mu_{t+1} - E_{r^s}) &= \varphi_{33} (\mu_t - E_{r^s}) + \nu_{3,t+1} \end{aligned}$$

φ_{22} (<i>low-upp</i>)	φ_{33} (<i>low-upp</i>)	σ_1	σ_2	σ_3
0.847	0.939	0.187	0.285	0.171
[0.757 0.934]	[0.811 0.993]	[0.166 0.214]	[0.251 0.327]	[0.088 0.335]
Pred R^2		$\frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$	$\frac{\sigma_{13}}{\sigma_{11}\sigma_{33}}$	$\frac{\sigma_{23}}{\sigma_{22}\sigma_{33}}$
0.04		-0.631	-0.655	0.451
		[-0.732 -0.502]	[-0.841 -0.325]	[0.205 0.623]

Table 3: This table shows the posterior median and [0.025 0.975] quantile obtained with the predictive system described in Pastor and Stambaugh (2008). The sample period is 1910-2009.

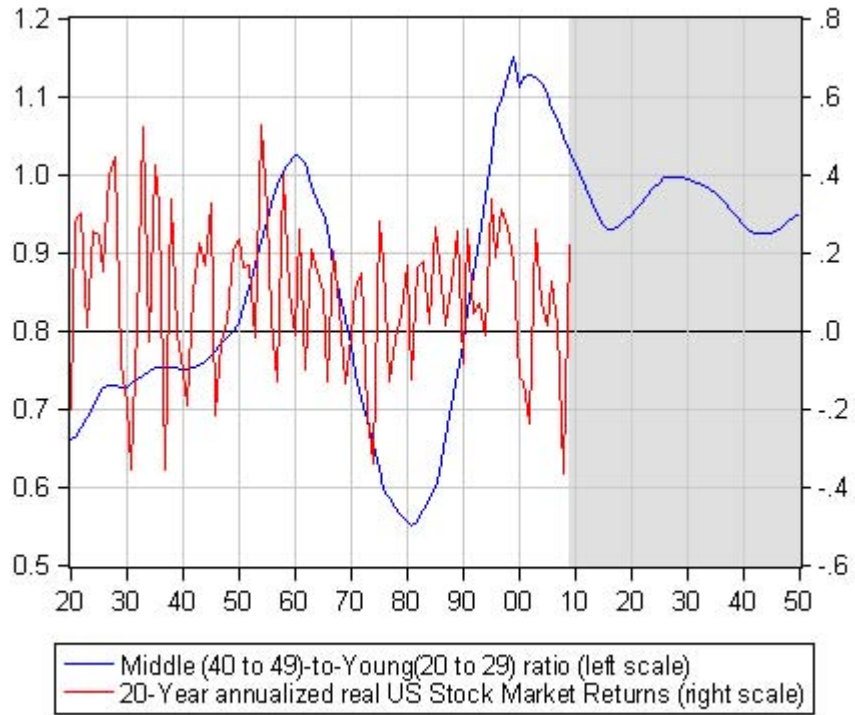
Table 4: A bi-variate VAR with MY (1910-2009)

$$r_{t+1}^s = \varphi_{10} + \varphi_{12}dp_t + \varphi_{13}MY_{t+1} + \nu_{1,t+1}$$

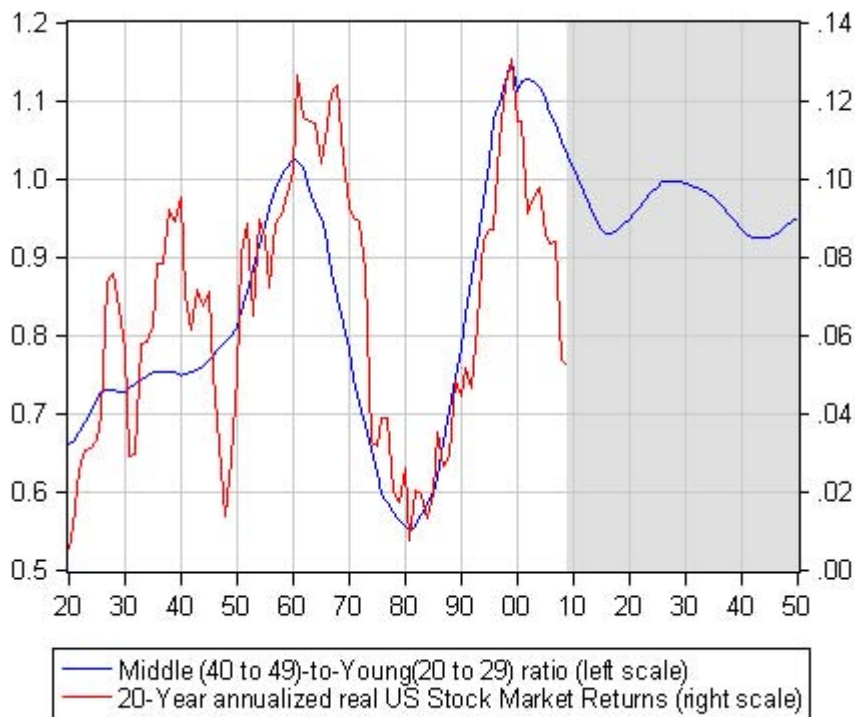
$$dp_{t+1} = \varphi_{20} + \varphi_{22}dp_t + \varphi_{23}MY_{t+1} + \nu_{2,t+1}$$

φ_{12} (<i>t-stat</i>)	φ_{13} (<i>t-stat</i>)	φ_{22} (<i>t-stat</i>)	φ_{23} (<i>t-stat</i>)	χ_2^2 $\varphi_{11}=0, \varphi_{21}=0$	σ_1	σ_2	$\frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$	$adj R_{r_{t+1}^s}^2$	$adj R_{dp_{t+1}}^2$
0.179	0.410	0.728	-0.603	4.74	0.188	0.207	-0.846	0.07	0.80
(3.07)	(2.67)	(11.33)	(-3.56)	(0.09)					

Table 4: The table reports coefficient estimates (with t-statistics in parentheses) and the R^2 statistic for each equation. We also report the standard deviations and correlations of residuals.



(a) 1-year real US stock market returns and demographic trends.



(b) 20-year real US stock market returns and demographic trends.

Figure 1: Stock market returns and demography.

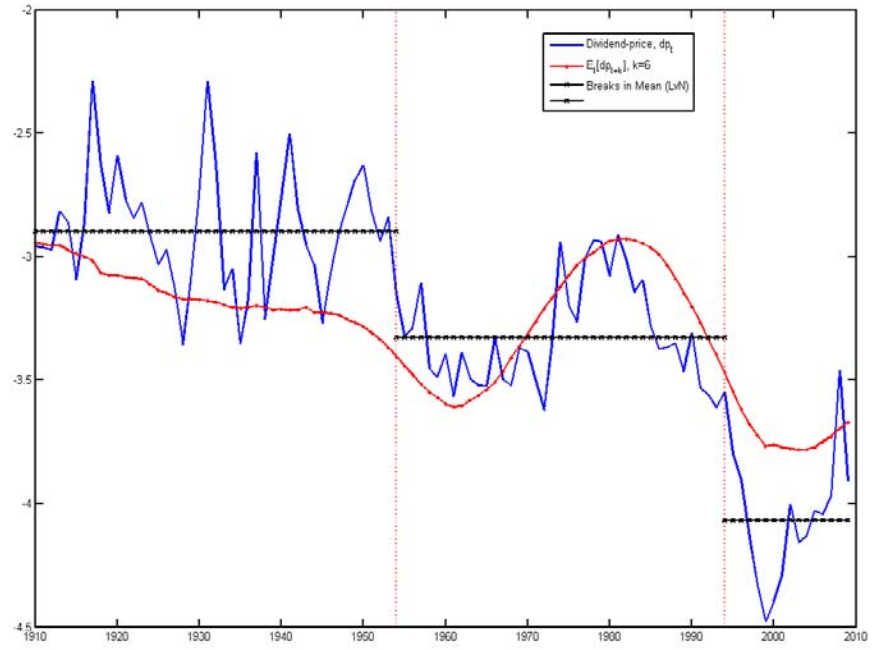


Figure 2: dp_t along with the time-varying linearization point used in our model and the breaks identified by LVN.

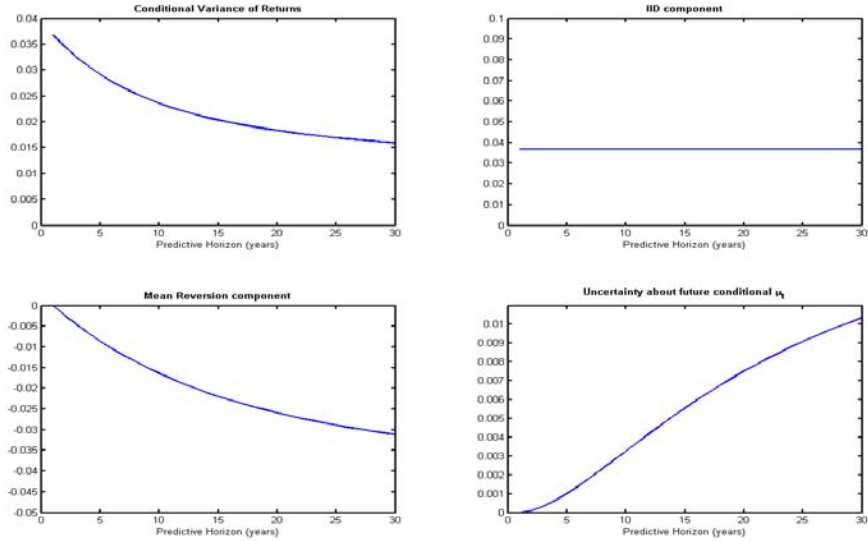


Figure 3: The term structure of stock market risk from a bi-variate VAR.

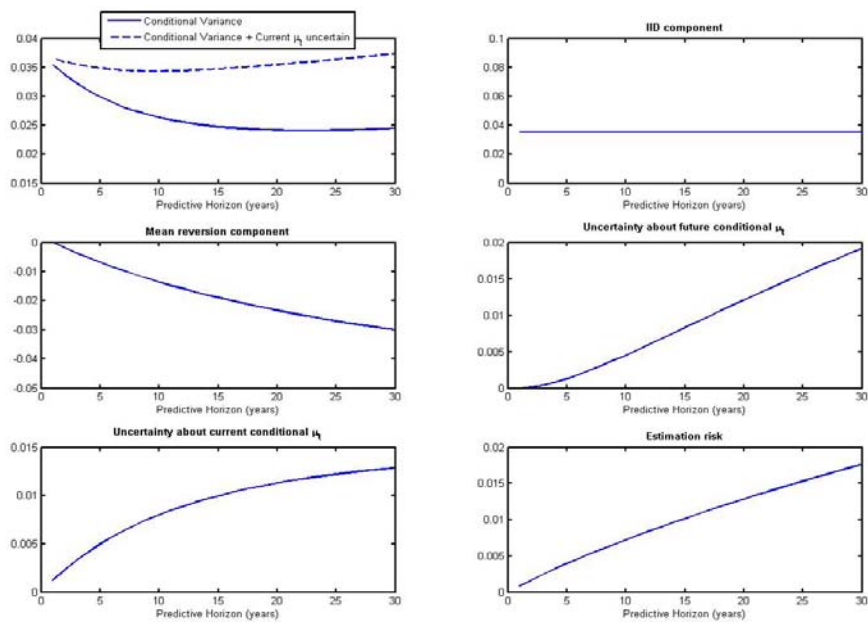


Figure 4: The term structure of US stock market risk from a three-variate VAR with an unobservable component.

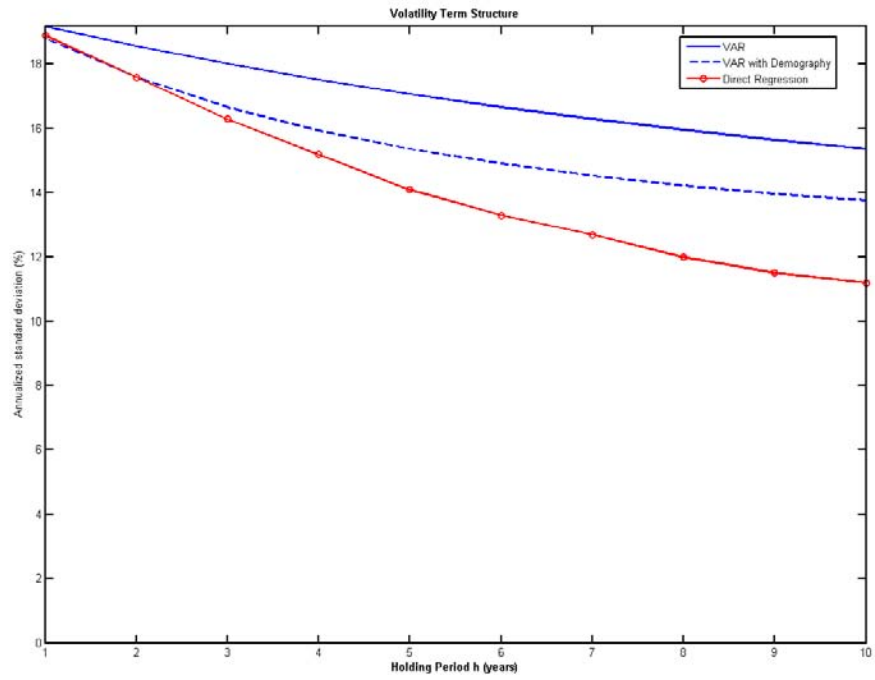


Figure 5: Three alternative measures of the TS of stock market risk.

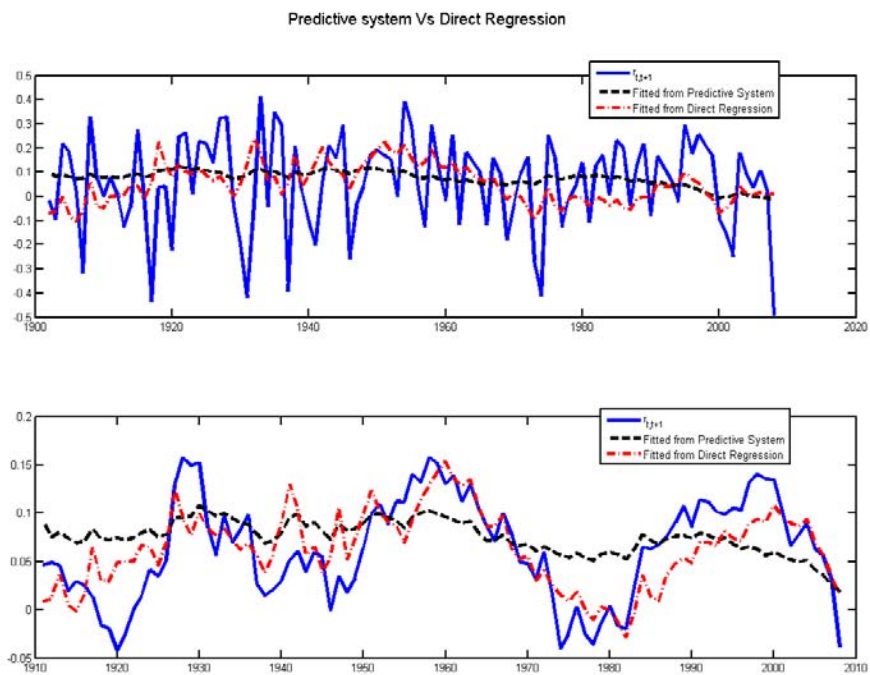


Figure 6: Noise and Information reconsidered.

References

- Ang, Andrew and Angela Maddaloni**, “Do Demographic Changes Affect Risk Premiums? Evidence from International Data,” *Journal of Business*, 2005, 78 (1), 341–380.
- and **Geert Bekaert**, “Stock Return Predictability: Is It There?,” *Review of Financial Studies*, 2007, 20 (3), 651–707.
- Bakshi, Gurdip S and Zhiwu Chen**, “Baby Boom, Population Aging, and Capital Markets,” *Journal of Business*, April 1994, 67 (2), 165–202.
- Binsbergen, Jules H. Van and Ralph S. Koijen**, “Predictive Regressions: A Present-Value Approach,” *SSRN eLibrary*, 2009.
- Campbell, John Y.**, “A Variance Decomposition for Stock Returns,” *Economic Journal*, 1991, 101 (405), 157–79.
- and **Luis M. Viceira**, “Strategic Asset Allocation,” 2002.
- Campbell, John Y and Luis M Viceira**, “The Term Structure of the Risk-Return Tradeoff,” CEPR Discussion Papers 4914, C.E.P.R. Discussion Papers February 2005.
- Campbell, John Y. and Robert J. Shiller**, “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies*, 1988, 1 (3), 195–228.
- , **Andrew W. Lo, and A. Craig Mackinlay**, *The Econometrics of Financial Markets*, Princeton University Press, June 1997.
- Cochrane, John H.**, “Volatility tests and efficient markets : A review essay,” *Journal of Monetary Economics*, June 1991, 27 (3), 463–485.
- , *Asset Pricing.*, Princeton University Press, Princeton, NJ, 2001.
- , “The Dog That Did Not Bark: A Defense of Return Predictability,” *Review of Financial Studies*, July 2008, 21 (4), 1533–1575.
- , “State-Space vs. VAR Models for Stock Returns,” *SSRN eLibrary*, 2008.
- DellaVigna, Stefano and Joshua M. Pollet**, “Attention, Demographics, and the Stock Market,” *SSRN eLibrary*, 2005.
- Erb, Claude B., Campbell R. Harvey, and Tadas E. Viskanta**, “Demographics and International Investment,” *Financial Analysts Journal*, 1997.

- Fama, Eugene F. and Kenneth R. French**, “Dividend yields and expected stock returns,” *Journal of Financial Economics*, October 1988, *22* (1), 3–25.
- Favero, Carlo A., Arie E. Gozluklu, and Andrea Tamoni**, “Demographic Trends, the Dividend-Price Ratio and the Predictability of Long-Run Stock Market Returns,” *Journal of Financial and Quantitative Analysis*, *forthcoming*, 2010, pp. –.
- Geanakoplos, John, Michael Magill, and Martine Quinzii**, “Demography and the Long-Run Predictability of the Stock Market,” *Brookings Papers on Economic Activity*, 2004, *35* (2004-1), 241–326.
- Goyal, Amit**, “Demographics, Stock Market Flows, and Stock Returns,” *Journal of Financial and Quantitative Analysis*, March 2004, *39* (01), 115–142.
- and **Ivo Welch**, “Predicting the Equity Premium with Dividend Ratios,” *Management Science*, 2003, *49* (5), 639–654.
- Lacerda, Filipe and Pedro Santa-Clara**, “Forecasting Dividend Growth to Better Predict Returns,” Technical Report February 2010.
- Lettau, Martin and Stijn Van Nieuwerburgh**, “Reconciling the Return Predictability Evidence,” *Review of Financial Studies*, July 2008, *21* (4), 1607–1652.
- Neely, Christopher J. and Paul Weller**, “Predictability in International Asset Returns: A Reexamination,” *The Journal of Financial and Quantitative Analysis*, 2000, *35* (4), 601–620.
- Nelson, Charles R. and Myung J. Kim**, “Predictable Stock Returns: The Role of Small Sample Bias,” *The Journal of Finance*, 1993, *48* (2), 641–661.
- Pastor, Lubos and Robert F Stambaugh**, “Predictive Systems: Living with Imperfect Predictors,” *Journal of Finance*, *forthcoming*, February 2008, pp. –.
- and — , “Are Stocks Really Less Volatile in the Long Run?,” CEPR Discussion Papers 7199, C.E.P.R. Discussion Papers March 2009.
- Paye, Bradley S. and Allan Timmermann**, “Instability of return prediction models,” *Journal of Empirical Finance*, 2006, *13* (3), 274 – 315.
- Poterba, James M.**, “Demographic Structure And Asset Returns,” *The Review of Economics and Statistics*, November 2001, *83* (4), 565–584.

- Rapach, David E. and Mark E. Wohar**, “In-sample vs. out-of-sample tests of stock return predictability in the context of data mining,” *Journal of Empirical Finance*, March 2006, *13* (2), 231–247.
- Robertson, Donald and Stephen Wright**, “The Limits to Stock Return Predictability,” Technical Report 2009.
- Siegel, Jeremy J.**, *Stocks for the Long Run*, McGraw-Hill Companies, November 2007.
- Stambaugh, Robert F.**, “Predictive regressions,” *Journal of Financial Economics*, December 1999, *54* (3), 375–421.
- Taleb, Nassim Nicholas**, *Fooled by Randomness. The Hidden Role of Chance in the Markets and in Life.*, Texere XLC, June 2001.
- Valkanov, Rossen**, “Long Horizon Regressions: Theoretical Results and Applications,” *Journal of Financial Economics*, 2003, *68*, 201–232.
- Welch, Ivo and Amit Goyal**, “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies*, July 2008, *21* (4), 1455–1508.