

# Consumption Risk and the Cross-Section of Government Bond Returns

Abhay Abhyankar\*, Olga Klinkowska†, Soyeon Lee‡

February 2015

## Abstract

In this paper we provide a consumption-based explanation of risk in nominal US Treasury bond portfolios. We use a consumption-CAPM with Epstein-Zin-Weil recursive preferences. Our model introduces two sources of risk: uncertainty about current consumption (reflected in contemporaneous consumption growth) and uncertainty about prospects of consumption in a long run (reflected in innovations to expectations about future consumption growth). We use a novel approach to estimate pricing factors in our model: we employ a factor-augmented VAR model with common factors, extracted from a large panel of macroeconomic and financial data, as state variables. We find that the important source of risk in US bonds is related to uncertainty in prospects in future consumption and it induces a positive and significant risk premium. We find as well that covariance risk related to innovations in expectations about future consumption growth is greater for long term bond portfolios than for short term bond portfolios, which is consistent with a duration measure of risk and justifies why long term bonds require greater premium than short term bonds. Our model explains well the cross-sectional variation in average excess returns of bonds with different maturities over the period 1975–2011 and compares favorably with competing models.

**JEL classification:** G0, G10, G12

**Keywords:** Epstein–Zin–Weil preferences, consumption risk, asset pricing tests, government bonds, factor analysis

---

\*Xfi Center for Finance and Investment, University of Exeter, Exeter, EX4 4PU, UK, email: a.abhyankar@exeter.ac.uk.

†University of Aberdeen Business School, Edward Wright Building, Dunbar Street, Aberdeen, AB24 3QY, UK, email: o.klinkowska@abdn.ac.uk. Corresponding author.

‡University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK.

# 1 Introduction

We investigate, using a consumption-based capital asset pricing model (C-CAPM) with Epstein–Zin–Weil recursive utility, the cross-section of excess returns on portfolios of US Treasury bonds with varying times to maturity. Specifically, we ask the following questions: what can we learn about bond risk from consumption-based models? Is there a role for consumption risk to play in the explanation of risk premia for nominal bonds with different maturities? More generally, we add to the literature on consumption-based models for pricing bonds that is “surprisingly small, given the vast amount of attention given to consumption-based models of equity pricing” (Campbell (2007)).

The evaluation of risks in nominal government bonds has attracted a considerable attention for quite some time. Campbell et al. (2010), for example, point out that this can be done in many ways. One of them is to measure the covariance of bond returns with a proxy of the marginal utility of the consumers, like the return on market portfolio (as in the classical CAPM) or the aggregate consumption growth (as in the C-CAPM). Indeed early attempts to evaluate the risks of nominal bonds followed this approach (see for example Gultekin and Rogalski (1985)). More recently, Viceira (2012) finds that the consumption beta for bonds is negative, over the 1980s and 1990s, suggesting that nominal bonds help investors hedge aggregate market risks.

Our work is in this spirit and builds on prior works but differs in two important respects. First, we use a consumption CAPM with Epstein–Zin–Weil utility rather than the standard power utility C-CAPM. This allows us to extend a measure of investors’ well-being to include not only contemporaneous consumption growth but also to reflect the future evolution of consumption and the fact that consumers care about a long run risk in consumption as well, in the spirit of Bansal and Yaron (2004) model for equities. We can investigate then whether long run consumption risk can explain positive on average risk premia paid by government bonds. Second, our test assets, which are a set of government bond portfolios with different maturities rather than a single index of government bonds, allow us to study the variation of covariance risk across assets that may be imperfect substitutes. Specifically we investigate whether consumption risk related to uncertainty in prospects in future consumption can explain why bonds with greater maturities require greater excess returns.

Our C-CAPM has two risk factors: consumption growth and innovations to expectations about future consumption growth. While consumption growth is directly measurable, the innovations to expectations about future consumption growth are not and have to be estimated. It is usually done using a Vector Autoregressive (VAR) model<sup>1</sup> where specific state variables are selected that are known to forecast consumption growth well. Our implementation of this methodology is, however, novel. Instead of choosing specific predictor variables we use a set of common factors obtained, following

---

<sup>1</sup>Brunnermeier and Julliard (2008), Campbell and Vuolteenaho (2004), Lustig and Nieuwerburgh (2008).

Stock and Watson (2002a,b), from a large panel of macroeconomic and financial time series. We then estimate a factor-augmented VAR, in the spirit of Bernanke et al. (2005), and extract innovations to expected future consumption growth. This approach has some advantages. First, we can be agnostic in our choice of state variables thus mitigating to some extent concerns about the choice of specific state variables (see for example Chen and Zhao (2009)). Second, there is an evidence (Stock and Watson (2008)) that common factors have good forecasting properties in the presence of structural breaks. We further add to that and demonstrate that extracted factors have good predictive power for consumption growth in in- and out-of-sample tests, which forms the empirical basis for using them as state variables in our VAR model. Third, the pre-estimation of the dynamic factors does not affect the consistency of Ordinary Least Squares (OLS) estimates in the VAR model (Bai and Ng (2008)) which is relevant in our application.

Our main group test assets are bond portfolios that are constructed using US Treasury bonds with times to maturity ranging from over a year to longer than ten years. We use also bond indices with different target maturities and zero coupon bonds as alternatives. The sample period is 1975–2011. We use a covariance risk measure that reflects how bond excess returns covary with consumption growth and innovations to expectation in future consumption growth. We estimate Euler equation with linearized stochastic discount factor (SDF) using Generalized Method of Moments (GMM) in order to study how well our two-factor C-CAPM explains the cross-section of average excess returns on government bonds. This methodology allows us also to estimate the coefficient of relative risk aversion and test the theoretical restrictions of our model. We estimate as well linearized Euler equation with Fama-MacBeth approach. Finally, we compare our measure of long run risk in consumption growth for bonds with a measure of Parker and Julliard (2005) and investigate how it is related to duration, a classical measure of bond risk.

Our main results can be summarized as follows. We find that the risk related to long run uncertainty in consumption plays an important role in pricing US government bonds and induces a risk premium that is positive and significant. We find as well that our two-factor C-CAPM explains well the cross-section of average excess returns on portfolios of US Treasury bonds with differing maturities (around 98.8% of the cross-sectional variation) over the sample period 1975–2011. We also demonstrate that the covariance risk related to innovations to expectations about future consumption growth is greater for portfolios of long term bonds than for portfolios of short term bonds. This provides a consumption risk-based explanation on why long term bonds are paying on average greater returns than short term bonds. We show as well that the covariance risk related to prospects in future consumption is consistent with a classical measure of bond risk – duration and bonds with greater duration have as well greater covariance risk. The estimates of risk aversion parameter implied by the model are lower than for equities. Finally, we find that our model performs well relative to other linear factor models

and prices well a joint portfolio of bonds and equities. Our results are robust to a battery of tests: use of alternate test assets and sample period, alternate measures of consumption growth and estimation methods.

The rest of the paper is organized as follows. Section 2 provides an overview of related research while Section 3 provides the details of our model. Section 4 outlines key features of the applied methodology and Section 5 describes the data. We discuss our empirical results including comparison with other models and tests for robustness in Section 6. Section 7 concludes the paper. The Online Appendix provides the relevant details related to theoretical aspects of our model, results of robustness tests and the description of supplementary data used in the paper.

## 2 Related literature

Expositions of the canonical C-CAPM for equities are now a standard textbook material but applications in the context of bonds are not common; Wolman (2006) is an example of a pedagogic guide to the consumption-based modelling of bonds. We find, as noted earlier, that there is surprisingly little empirical research on the consumption-based explanation of the cross-section of government bond returns.

Gultekin and Rogalski (1985) are possibly the first to study how well Ross’s APT model and the CAPM price the cross-section of constant maturity US government bond portfolios over the 1960–1979 period.<sup>2</sup> They find that average returns on bond portfolios are explained by at least two “priced” factors obtained using factor analysis. Further, using tests for the CAPM available at that time, they find that estimates of the factor risk premia on the market portfolio are all negative but not significantly different from zero. They conclude that “... [their] tests should be viewed simply as the first empirical attempt to properly measure interest-rate risk for bonds using factor-generating models. Our results in terms of the existence of priced risk premia are more favorable to multifactor models than to single-factor models or the CAPM”. In a related study using corporate bond portfolios Chang and Huang (1990) observe that the focus, in the literature,<sup>3</sup> on stocks rather than bonds may be due “ [to] the lack of convincing empirical evidence... show[ing] that covariance risks are priced in bond markets”.

We note here as well that there is a huge related literature on modelling the term structure of interest rates (see Piazzesi (2009) for an excellent survey). This literature assumes that the price of bonds is driven by a continuous time stochastic process, there are continuous trading opportunities and the principle of “no arbitrage” holds. One can then obtain equilibrium term structure models

---

<sup>2</sup>Roll (1971) is an example of early efforts to apply the CAPM to zero coupon bond data.

<sup>3</sup>A more recent example that uses corporate bond data is Gebhardt et al. (2005).

(as in Vasicek (1977) for example) that describe the prices, or equivalently, the yields, of different maturity bonds as functions of one or more state variables. For example, in the Vasicek model the single factor or source of uncertainty is the current level of the short rate. Later models assume that bond prices are driven by multiple state variables. In many multifactor models the factors are latent (unobserved) variables that are identified by data on the yields of different maturity bonds. The factors are further assumed to be “affine” functions of the state variables as in Duffie and Kan (1996). Ang and Piazzesi (2003) link these factors to observed macroeconomic variables. Using the Epstein–Zin–Weil framework, Piazzesi and Schneider (2007) consider a representative agent model with Epstein–Zin–Weil recursive utility and solve it for average yields. Gallmeyer et al. (2007) also demonstrate how the literature on affine models can be linked with a structural equilibrium model of investors’ preferences and opportunities using the Epstein–Zin–Weil utility framework.

In contrast to the vast literature on term structure models, our paper, in the spirit of Gultekin and Rogalski (1985), studies the covariance risk of consumption growth with government bond returns in a linearized factor pricing framework. The role of such linear factor models has been studied for other asset classes including corporate bonds (Gebhardt et al. (2005)), commodity futures (Khan et al. (2007)) as well as for options (Constantinides et al. (2009) and Coval and Shumway (2001)). Finally, linear factor models for bonds potentially have considerable practical applications. For example, in 2008 out of US \$10,349 billion invested in all mutual funds about US \$1,552 billion (15%) was in funds that invested exclusively in US government and related securities<sup>4</sup>. The performance appraisal of these funds is closely linked to a specifications of appropriate asset pricing models that enable the identification of differential performance for investors.<sup>5</sup>

In a recent work Viceira (2012) finds a considerable time variation, persistence and mean reversion in bond market betas over the 1962–2007 period. He reports that CAPM and C-CAPM betas, estimated using a single 5-year maturity bond portfolio, are time dependent. Our work differs from it in two important respects. First, we use a C-CAPM with Epstein–Zin–Weil utility rather than the standard power utility C-CAPM which allows us to study the role of covariance risk related to long run prospects in future consumption in pricing bonds. Secondly, we use, as test assets, bond portfolios with varying times to maturity, which allows us to investigate the role of that risk in the cross section of bond returns. While we recognize the importance of the time variation in covariance risk measures for bonds, we do not undertake this in our work. With a sample relatively short for measuring long term uncertainty of consumption growth and the nature of that risk such an exercise is beyond the scope of this paper.

---

<sup>4</sup>Data from the 2009 Investment Company Institute Fact Book, Investment Company Institute

<sup>5</sup>The literature on the performance of bond mutual funds relies largely on models using *ad hoc* factors like the short rate, term slope curvature, credit spreads, mortgage and liquidity spreads etc (see for example Chen et al. (2010) and references therein).

Our work is also related to the literature that examines whether stock returns are priced by their exposure to consumption risk measured over different horizons. For example, Daniel and Marshall (1997) find that the performance of a standard power utility C-CAPM improves if they use covariances with consumption growth at the two-year horizon. In later works, Parker and Julliard (2005) and Jagannathan and Wang (2007) also find that the power utility C-CAPM explains the cross-section of average stock returns better when risk is measured by the covariances with consumption growth at longer horizons. In this paper, in contrast, the uncertainty of consumption growth over long run is measured as innovations to expectations in future consumption growth over infinite period. Thus e.g. the measure of Parker and Julliard (2005) can be regarded as the truncation of such an infinite series at economically sensible horizons. Given that, we discuss the relation between these measures in our paper.

Finally, our work intersects with the burgeoning literature on long run consumption risk in the C-CAPM using the framework of Epstein–Zin–Weil utility.<sup>6</sup> Malloy et al. (2009), for example, study the role of long run consumption risks for stockholders and non-stockholders using the Consumer Expenditure Survey (CEX) data. They find that a C-CAPM with factors related to current and future consumption growth, as used in this paper, provides a better fit and plausible estimates of the coefficient of risk aversion for households that own stocks. Finally, Boguth and Kuehn (2009) find that consumption growth volatility plays an important role in pricing the cross-section of equity returns.

### 3 Consumption CAPM with Epstein–Zin–Weil preferences

In our theoretical setup we follow Hansen et al. (2005) and Malloy et al. (2009) who consider a representative agent, in an endowment economy, with Epstein–Zin–Weil preferences over the consumption stream. In equilibrium, asset prices are such that it is optimal for the agent to consume her endowment. The preferences are represented by the recursive utility function of Epstein and Zin (1989), Epstein and Zin (1991) and Weil (1989) of the form:

$$V_t = \left\{ (1 - \beta)C_t^{1-\rho} + \beta \left[ E_t(V_{t+1}^{1-\gamma}) \right]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}} \quad (1)$$

where  $V_{t+1}$  is the continuation value,  $\gamma$  is the coefficient of relative risk aversion,  $\frac{1}{\rho}$  represents the elasticity of intertemporal substitution (EIS) and  $\beta$  is the time discount factor. When  $\gamma = \frac{1}{\rho}$ , the expression in eq.(1) collapses to a familiar power utility function. With Epstein–Zin–Weil recursive

---

<sup>6</sup>See Bansal (2007) for an accessible review.

utility the stochastic discount factor (SDF)  $M_{t+1}$  is given by:

$$M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left[ \frac{V_{t+1}}{E_t(V_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}} \right]^{\rho-\gamma} \quad (2)$$

There are two crucial terms in this stochastic discount factor. The first is the consumption growth  $\left( \frac{C_{t+1}}{C_t} \right)$  as in the classical power utility C-CAPM. The second term is related to the continuation value  $V_{t+1}$  and represents future utility. This is a forward looking term linked to a future consumption via the recursion in eq.(1) and is present only when the coefficients of EIS and risk aversion differ. This term is unobservable and for the purpose of empirical implementation and in order to find a closed form solution for the SDF above, we assume that the consumption growth follows some  $MA(\infty)$  process of the general form as specified below:

$$\Delta c_{t+1} = \mu_c + \alpha(L)w_{t+1} = \mu_c + \sum_{s=0}^{\infty} \alpha_s w_{t+1-s} \quad (3)$$

where  $\Delta c_{t+1} \equiv \log\left(\frac{C_{t+1}}{C_t}\right)$ ,  $\alpha(L)$  is a lag polynomial operator defined as  $\alpha(L) = \sum_{s=0}^{\infty} \alpha_s L^s$  and  $w_{t+1}$  is *iid* standard normal process ( $w_{t+1} \sim iid N(0, 1)$ ). Such a specification is quite broad and allows for a wide range of possible models of consumption path over time as noted in Hansen et al. (2005).

We also focus on a special case where the EIS is equal to one since we are interested in studying the cross-section of expected returns. Malloy et al. (2009) argue that higher EIS increases covariances of all asset returns with log SDF by a roughly similar magnitude and they confirm this empirically. As a result in a cross-section the relative differences between expected returns and the cross-sectional price of risk will not be substantially influenced by the value of EIS.

Under the above assumptions, the expression for the log SDF in eq.(2) is given by:<sup>7</sup>

$$m_{t+1} = \log \beta - \Delta c_{t+1} + (1 - \gamma)[\alpha(\beta)\omega_{t+1}] - \frac{(1 - \gamma)^2}{2} \alpha(\beta)^2 \quad (4)$$

where  $m_{t+1} \equiv \log(M_{t+1})$ . Ignoring the log of time discount factor  $\beta$ , there are two components in the above expression that are important in an asset pricing context. The first is the current consumption growth  $\Delta c_{t+1}$  and this is what is captured in the classical power utility C-CAPM. The second term  $\alpha(\beta)\omega_{t+1}$  equals

$$\alpha(\beta)\omega_{t+1} = \left( \sum_{s=0}^{\infty} \alpha_s \beta^s \right) w_{t+1} = \sum_{j=0}^{\infty} \beta^j (E_{t+1} - E_t) \Delta c_{t+1+j}$$

and represents the innovation to expectations about the present value of consumption growth in all future periods. In other words, this term reflects the change in the expectations about future consump-

<sup>7</sup>The detailed derivation of the log SDF is provided in the Online Appendix.

tion growth caused by shock  $\omega_{t+1}$ . For estimation purposes we refer to this forward-looking element as  $\varepsilon_{c,t+1}$ . The last term in eq.(4) is the variance of this innovation. In the above setup, this variance is constant over time. The log SDF, given in eq.(4) can be rewritten as

$$m_{t+1} = \log \beta - \Delta c_{t+1} + (1 - \gamma) \sum_{s=0}^{\infty} \beta^s (E_{t+1} - E_t)(\Delta c_{t+1+s}) - \frac{(1 - \gamma)^2}{2} Var_t \left[ \sum_{s=0}^{\infty} \beta^s (E_{t+1} - E_t)(\Delta c_{t+1+s}) \right] \quad (5)$$

The Euler equation implies that the gross return on any asset  $i$ ,  $R_{t+1}^i$  should be such that  $E(M_{t+1}R_{t+1}^i) = 1$ . The log-linearized version of the Euler equation for our model is the following:

$$E(r_{t+1}^i - r_{t+1}^f) + \frac{1}{2} Var(r_{t+1}^i) - \frac{1}{2} Var(r_{t+1}^f) = -Cov(m_{t+1}, r_{t+1}^i - r_{t+1}^f) \quad (6)$$

$$E(r_{t+1}^i - r_{t+1}^f) + \frac{1}{2} Var(r_{t+1}^i) - \frac{1}{2} Var(r_{t+1}^f) = Cov(\Delta c_{t+1}, r_{t+1}^i - r_{t+1}^f) + (\gamma - 1) Cov \left( \sum_{s=0}^{\infty} \beta^s (E_{t+1} - E_t)(\Delta c_{t+1+s}), r_{t+1}^i - r_{t+1}^f \right)$$

According to the above pricing equation, the expected excess returns are determined by, first, how the excess returns covary with contemporaneous consumption growth and, second, how they covary with innovations in expectations about future consumption growth. This is the main equation in our paper and we evaluate the implications of this equation in empirical applications. We also test the theoretical restriction this equation implies (i.e. the loading on  $Cov(\Delta c_{t+1}, r_{t+1}^i - r_{t+1}^f)$  should be equal to one) and we estimate as well the value of the coefficient of risk aversion  $\gamma$ .

From practical perspective, the forward-looking term  $\varepsilon_{c,t+1}$  is unobservable and needs to be estimated. Our approach here is novel and we estimate the innovations in expectations about future consumption growth from a factor-augmented VAR model for consumption growth.

## 4 Methodology

We proceed towards the evaluation of the empirical performance of Euler equation, given in (6) in the following steps. First we estimate common factors from a large panel of macroeconomic and financial variables and from the extracted factors we select those that have predictive power for consumption growth. They are used in the factor-augmented VAR model as state variables. Next, we estimate the factor-augmented VAR model and extract innovation to expectations about future consumption growth. Finally, we estimate Euler equation and test its implications using two methodologies: linear SDF approach with GMM estimation and Fama-MacBeth cross-sectional regressions.



## 4.1 Estimation of common factors and their selection for VAR model

We estimate common factors from the panel of macroeconomic and financial variables using the methodology of asymptotic principal components with the number of factors determined by the information criteria in Bai and Ng (2002). This is a classical technique in factor analysis and we refer the reader to Bai and Ng (2008) and Stock and Watson (2011) for full details.

In our application we focus on a subset of the common factors that have predictive power for future consumption growth. We select those factors following the procedure detailed in Ludvigson and Ng (2009). The selection is based on statistical measures: Bayesian Information Criterion (*BIC*) and Akaike Information Criterion (*AIC*).

Briefly, let  $r$  be the number of common factors as determined by information criteria in Bai and Ng (2002). Let  $F_{kt}$  denote the set of all possible subsets of common factors, with  $F_{kt}$  denoting the elements of  $F_{kt}$ , where  $k$  indicates the number of factors in each subset  $F_{kt}$  ( $k = 1, 2, \dots, r$ ). For example  $F_{1t}$  has  $r$  possible one-element sets of common factors:  $F_{1t} = \{\{f_{1t}\}, \{f_{2t}\}, \dots, \{f_{rt}\}\}$ , where  $f_{it}$  denotes the  $i^{th}$  common factor. For each  $k$ , the optimal composition of common factors,  $F_{kt}^*$  is determined by *BIC* and *AIC* from the following predictive regressions:

$$\Delta c_{t+1} = \alpha_{k,0} + \alpha'_{k,1} F_{kt} + \xi_{k,t+1} \quad \forall F_{kt} \in F_{kt}, k = 1, 2, \dots, r \quad (7)$$

By construction, all common factors included in  $F_{kt}^*$  are included as well in  $F_{(k+1)t}^*$ . Once the optimal composition of factors  $F_{kt}^*$  is identified for each  $k$ , we select from those optimal sets the one that minimizes both the *BIC* and *AIC* from predictive regressions of the form as in (7). In cases where information criteria do not coincide in the selection, we use the significance of the log-likelihood ratio as a decision tool. Intuitively, this test statistic allows to test whether additional common factor in  $F_{(k+1)t}^*$  in above to  $F_{kt}^*$ , contributes significantly to the predictability of consumption growth, given the forecasting power of to  $F_{kt}^*$ .

## 4.2 Estimation of the factor-augmented VAR

There are two pricing factors in Euler equation (6): consumption growth and innovation to expectations about future consumption growth. While consumption growth is directly measurable, innovation to expectations about future consumption growth is not observable and needs to be estimated. We use a VAR methodology for that purpose. A benefit of using a VAR is that it allows for a time-variation in conditional expectation of consumption growth. We can then extract a forward looking measure of innovations to expectations in future consumption growth. This differs from the approaches in for example, Parker and Julliard (2005) or Jagannathan and Wang (2007), where forward looking

measures related to consumption growth are truncated at specific lag lengths. A cost to this approach, that there is an element of estimation error in the VAR parameter estimates due to mis-specification of state variables used in the VAR, is mitigated with our use of common factors as state variables in the VAR.

We now explain, in brief, the method used to extract the innovations in expectations about future consumption growth,  $\varepsilon_{c,t+1}$ . Let  $\mathbf{Z}_t$  denote a vector which has log consumption growth  $\Delta c_t$  as its first element. The other elements in this vector, denoted as  $\mathbf{x}_t$ , are state variables. Let this vector follow a VAR(1) process:

$$\mathbf{Z}_{t+1} = \mathbf{A}\mathbf{Z}_t + \eta_{t+1} \quad (8)$$

Specifically, we assume the following dynamics for  $\Delta c_t$  and  $\mathbf{x}_t$ :

$$\begin{bmatrix} \Delta c_{t+1} \\ \mathbf{x}_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \mathbf{x}_t \end{bmatrix} + \begin{bmatrix} \eta_{c,t+1} \\ \eta_{x,t+1} \end{bmatrix} \quad (9)$$

Let  $\mathbf{e}_1$  be a vector with the first element equal to 1 and all others equal to zero. Using  $\mathbf{e}_1$  we can e.g. express consumption growth in terms of the elements of the VAR model:  $\Delta c_t = \mathbf{e}'_1 \mathbf{Z}_t$ . Given the joint dynamics of consumption growth and state variables in eq.(9), the innovations to expectations about future consumption growth, defined as  $\varepsilon_{c,t+1} = \sum_{j=0}^{\infty} \beta^j (E_{t+1} - E_t) \Delta c_{t+1+j}$ , are the following:<sup>8</sup>

$$\varepsilon_{c,t+1} = \mathbf{e}'_1 (\mathbf{I} - \beta \mathbf{A})^{-1} \eta_{t+1} \quad (10)$$

In most applications the VAR is estimated using a specific set of state (or predictive) variables. For example, Campbell and Vuolteenaho (2004) use the yield spread, the price-earnings ratio and the small-stock value-spread, to extract shocks or news about changes in expected cash flows and discount rates. These series are then used as pricing factors in a two-factor inter-temporal CAPM to explain the difference between value and growth portfolios. Campbell and Vuolteenaho (2004) note that their results are sensitive to the inclusion of certain specific state variables and this point is further elaborated by Chen and Zhao (2009). In contrast, in this paper, instead of specific predictive variables, we use a set of selected common factors, in the spirit of Bernanke et al. (2005). Ludvigson and Ng (2009) point out that the extracted factors are likely to contribute to the forming of investor's expectations since they reflect a common set of underlying fundamentals and summarize information included in a large panel of macro and financial variables. Our use of common factors in this context is novel and has a number of advantages.<sup>9</sup> First, we can avoid having to choose specific individual

<sup>8</sup>The detailed derivation of (10) is provided in the Online Appendix.

<sup>9</sup>There is a growing literature using a factor analysis in a VAR setting to study the macroeconomic effects of policy interventions or patterns of co-movements in economic activity and as inputs into dynamic stochastic general equilibrium models (see, for example, Bernanke et al. (2005) for a specific application and the surveys by Bai and Ng (2008) and Stock and Watson (2011) for other references).

predictive variables and making our results sensitive to that choice. Second, as Bai and Ng (2008) show, under the assumption that both  $N, T \rightarrow \infty$  while  $\frac{\sqrt{T}}{N} \rightarrow 0$ , the coefficients obtained from ordinary least squares (OLS) estimation of VAR model in eq.(9) are  $\sqrt{T}$ -consistent and asymptotically normal. They also show that the asymptotic variance is such that any inference can proceed as if the factors were observed rather than estimated. In other words, the pre-estimation of the common factors using principal components analysis does not affect the consistency of the OLS estimates or the standard errors in the VAR system. This is of particular relevance in our case since the VAR in eq.(9) is estimated equation-by-equation using OLS. The requirement that  $N$  needs to be large relative to  $T$  is fulfilled in our case since we have  $N = 125$  which is a size similar to that used in previous work (for example, Ludvigson and Ng (2009)). Third, common factors are found to be robust to the structural instability that plagues low-dimensional forecasting regressions (Stock and Watson (2008)) which is relevant in our application as our data span a 37-year period from 1975–2011. The intuition for this result is that such instabilities average out in the construction of common factors if they are sufficiently dissimilar from each other.

### 4.3 Estimation of the Euler equation

#### 4.3.1 Linear SDF approach using GMM estimation

We estimate the Euler equation given in (6) using GMM method. This approach is a standard methodology in asset pricing literature. In our case it directly allows to test a theoretical restriction of the model that the loading on consumption growth risk is equal to one and allows to estimate the parameter of risk aversion from the loading on risk related to innovations in expectations about future consumption growth. The details of this methodology can be found in Cochrane (2005) and below we give a brief summary of this approach.

Let  $m_{t+1}$  denote a log SDF as given in eq.(5). It can be written in a general form as  $m_{t+1} = a - b'fr_{t+1}$ , where  $fr_{t+1} = \begin{bmatrix} \Delta c_{t+1} & \varepsilon_{c,t+1} \end{bmatrix}'$  and denotes the factors of risk in our model, coefficients  $b = \begin{bmatrix} b_{\Delta c} & b_{\varepsilon_c} \end{bmatrix}'$  are the risk loadings and  $a$  is a scalar chosen based on the normalization of the mean of the log SDF. The risk loadings indicate whether particular pricing factors in a proposed asset pricing model are marginally useful in pricing test assets in the presence of other risk factors. If the test assets are correctly priced by the proposed SDF, then the Euler equation implies that pricing errors will be zero i.e.  $E(m_{t+1}R_{t+1}^{ei}) = 0$ , where  $R^{ei}$  is the excess return on asset  $i$ . Such a form of Euler equation suggests GMM estimation approach using pricing errors as moments. Since the mean of  $m_{t+1}$  cannot be identified from the zero pricing errors, we choose to normalize it to one, i.e.  $E(m_{t+1}) = 1$ , which is a common approach in asset pricing literature. As a result  $m_{t+1} = 1 - b'[fr_{t+1} - \mu_{fr}]$ . This specification implies that the log SDF is a linear function of the demeaned risk factors as in Kan and Robotti (2008)

and:

$$E(m_{t+1}R_{t+1}^{ei}) = E(R_{t+1}^{ei} - R_{t+1}^{ei}(fr_{t+1} - \mu_{fr})\iota b) = 0 \quad (11)$$

The above can be written as

$$\begin{aligned} E(R_{t+1}^{ei}) &= E[R_{t+1}^{ei}(fr_{t+1} - \mu_{fr})\iota]b \\ &= Cov(R_{t+1}^{ei}, fr_{t+1})b \end{aligned} \quad (12)$$

This implies that GMM estimation is equivalent to a cross-sectional regression of average excess returns on the covariances between excess returns and risk factors and it is directly related to Euler equation in eq.(6). This is of special importance in our case as with estimated coefficients  $b$  we can directly test a theoretical restriction of the model in (6) that the risk loading on consumption growth is equal to one, i.e.  $H_0 : b_{\Delta c} = 1$  and we can also estimate the risk aversion parameter as  $\gamma = b_{\varepsilon_c} + 1$  and test its significance.

In terms of GMM estimation we consider the following  $N + K$  moment conditions:

$$g(b, \mu_{fr}) = E \begin{bmatrix} g_{1,t+1}(b, \mu_{fr}) \\ g_{2,t+1}(\mu_{fr}) \end{bmatrix} = E \begin{bmatrix} R_{t+1}^{ei} - R_{t+1}^{ei}(fr_{t+1} - \mu_{fr})\iota b \\ (fr_{t+1} - \mu_{fr}) \end{bmatrix} = 0_{(N+K) \times 1} \quad (13)$$

where  $N$  is the number of test assets and  $K$  is the number of risk factors. The first group of  $N$  moment conditions represents pricing errors – the differences between actual and predicted excess returns on test assets. The second group of  $K$  moment conditions arises from the normalization and a need to estimate the means of risk factors  $\mu_{fr}$ . The sample equivalents of the above moment conditions are the time series averages of relevant expressions:  $g_T(b, \mu_{fr}) = \frac{1}{T} \sum_{t=0}^{T-1} g_{t+1}(b, \mu_{fr})$ .

The GMM objective function is the following:

$$Min_{\{b, \mu_{fr}\}} g_T(b, \mu_{fr})\bar{W}g_T(b, \mu_{fr}) \quad (14)$$

where  $\bar{W}$  is a weighting matrix that determines the importance of specific moment conditions and has the following form:

$$\bar{W} = \begin{bmatrix} W & 0_{N \times 1} & 0_{N \times 1} \\ 0_{1 \times N} & h_1 & 0 \\ 0_{1 \times N} & 0 & h_2 \end{bmatrix}_{(N+K) \times 1} \quad (15)$$

Parker and Julliard (2005) and Malloy et al. (2009) advocate setting  $h_1$  and  $h_2$  sufficiently large numbers so that the estimates of  $\mu_{fr}$  are approximately equal to the time series averages of risk factors  $\frac{1}{T} \sum_{t=0}^{T-1} fr_{t+1}$ . Alternatively Kan and Robotti (2008) impose the restriction  $\hat{\mu}_{fr} = \frac{1}{T} \sum_{t=0}^{T-1} fr_{t+1}$ ,

derive a closed form solution for  $\hat{b}$ , which is equivalent to the one from the GMM minimization using large values of  $h_1$  and  $h_2$ , and specify the asymptotic distribution of the parameter estimates and sample moment conditions. We follow their paper in this respect. We also implement the approach of Parker and Julliard (2005) and Malloy et al. (2009) and verify that this delivers qualitatively similar results.

For the importance of moment conditions related to pricing errors, determined by weighting matrix  $W$ , we use two common, pre-specified matrices. First we use the identity weighting matrix,  $W = I$ . In this case all test assets are treated symmetrically and parameters  $b$  are estimated by minimizing the sum of equally weighted squared pricing errors. Second we use the inverse of the variance matrix of the excess returns on test assets, i.e.  $W = (VarR^e)^{-1}$  as suggested by Kan and Robotti (2008) for the case of de-meaned log SDF. This is a modified version of Hansen-Jagannathan (HJ) matrix which uses the inverse of the second moment of the excess returns. Lettau and Ludvigson (2001b) point out, that parameters can be poorly estimated using the HJ weighting matrix if the size of the available sample  $T$  is small compared to the number of test assets  $N$ . As such, if the parameter estimates from GMM with modified HJ matrix differ greatly from those of GMM with identity matrix, this may be due to the poor finite sample estimate of the asymptotic covariance matrix of the pricing errors.

Finally, we test the null of whether pricing errors from the model are insignificant, using the  $J_T$  test with the usual test statistic:

$$J_T = g_T'(b, \mu_{fr})[Var(g_T(b, \mu_{fr}))]^+ g_T(b, \mu_{fr}) \sim \chi_{N-K}^2 \quad (16)$$

where  $[]^+$  denotes the pseudo-inverse since the variance-covariance matrix of the  $g_T$  is singular. The  $\chi^2$  distribution has degrees of freedom equal to the difference between the number of moments and the number of estimated parameters.

Additionally, we also report some commonly used informal criteria that help assess the goodness-of-fit of the model: root mean square pricing error (RMSE), mean absolute pricing error (MAE),  $R^2$ , adjusted- $R^2$  and a plot of the actual versus the model-predicted excess returns.

### 4.3.2 Fama-MacBeth cross-sectional regressions

We estimate our model using also an alternative methodology: Fama-MacBeth cross-sectional regressions with univariate betas. In the first stage we run the time-series regressions and estimate univariate betas for each risk factor:

$$R_{t+1}^{ei} = \beta_{0,i,\Delta c} + \beta_{i,\Delta c}\Delta c_{t+1} + \xi_{\Delta c,i,t+1} \quad (17)$$

$$R_{t+1}^{ei} = \beta_{0,i,\varepsilon} + \beta_{i,\varepsilon}\varepsilon_{c,t+1} + \xi_{\varepsilon,i,t+1} \quad (18)$$

for  $i = 1, \dots, N$  and  $t = 0, \dots, T - 1$ . In the second stage, we run the cross-sectional regressions of average excess returns on the estimated betas and we estimate the loadings related to risk factors, lambdas  $\lambda$ :

$$\bar{R}_i^e = \lambda_{\Delta c} \hat{\beta}_{i, \Delta c} + \lambda_{\varepsilon} \hat{\beta}_{i, \varepsilon} + \alpha_i \quad (19)$$

Betas, estimated in the first stage represent the riskiness of the assets – the quantities of risk related to specific risk factors and they are directly proportional to covariances of test assets excess returns with risk factors. The factor loadings, estimated in the second stage, are interpreted as risk prices associated with specific risk factors. We test here whether the prices of the two risk factors (i.e. consumption growth risk and the risk related to innovations in expectations about future consumption growth) are significantly different from zero.

Alphas,  $\alpha$ , represent pricing errors and their magnitude determines how well the model prices test assets. We test the null of whether the pricing errors are jointly insignificant. The test statistic is given by  $\alpha'Var(\alpha)^{-1}\alpha$  and asymptotically follows the  $\chi^2_{N-K}$  distribution. We note, in this context, that the null of zero pricing errors may not be rejected, not because of small pricing errors, but because of high sampling error in the estimated betas as underlined by Lettau and Ludvigson (2001b).

We estimate the standard errors of factor loadings and pricing errors under a couple of assumptions. We compute classical iid OLS errors and we also account for the errors-in-variable problem, i.e. that the betas used in the second stage are estimates of the true unknown betas, by computing Shanken-corrected standard errors of as suggested by Shanken (1992). Shanken-correction assumes that asset returns are stationary and conditionally homoskedastic.<sup>10</sup> However we note, as shown in Jagannathan and Wang (1998), that conventional t-statistics in the presence of conditional heteroskedasticity do not necessarily overstate the precision of the standard errors. Finally we estimate as well robust standard errors. For that purpose we employ the GMM framework and estimate betas from the first stage and lambdas from the second stage simultaneously. This is equivalent to the Fama-MacBeth two-stage regressions in the sense that it produces the same estimates of the parameters. The advantage is that it allows to estimate standard errors that not only account for the fact that betas are estimated, but also allow for serial correlation and conditional heteroscedasticity. We follow Newey and West (1987) and estimate heteroscedasticity and autocorrelation consistent standard errors of Newey-West style.

Finally, we report some commonly used informal criteria that help assess the goodness-of-fit of the model: RMSE, MAE,  $R^2$ , adjusted- $R^2$ . To reinforce the conclusions from the Fama-MacBeth estimation results and to address the critique of relying too much on  $R^2$  measure, we estimate the second stage regression given in eq.(19) with constant and test whether the true  $R^2$  of the model is

---

<sup>10</sup>The correction is directly related to the magnitude of each coefficient and inversely related to the variability of the pricing factors. Lettau and Ludvigson (2001b) point out that macro factors are not very volatile and as a result this tends to “blow up” the Shanken correction factor so that the corresponding t-statistics are not significant.

significantly different from zero ( $H_0 : R^2 = 1$ ) and from one ( $H_0 : R^2 = 0$ ). Kan et al. (2009) provide a relevant theory to conduct these tests and we refer reader to the results in their paper for details.

## 5 Data

Our main set of test assets are ten US Treasury bond portfolios from the CRSP Fama Maturity Portfolios Returns Files. These portfolios include bonds and notes with various characteristics: callable, non-callable, non-flower US government notes and bonds, but exclude partially or fully tax-exempt issues. We compute quarterly portfolio returns using monthly holding period returns which are available in the CRSP database and are equal weighted averages of the unadjusted ex-post one-month holding period returns of each bond in the portfolio. Fama Maturity Portfolios (FMP) are sorted according to maturity and in this paper we use FMP with the following maturities: from 13 to 18 months (FMP1), from 19 to 24 months (FMP2), from 25 to 30 months (FMP3), from 31 to 36 months (FMP4), from 37 to 42 months (FMP5), from 43 to 48 months (FMP6), from 49 to 54 months (FMP7), from 55 to 60 months (FMP8) from 61 to 120 months (FMP9) and greater than 120 months (FMP10) from the quote date. Our sample starts from the first quarter of 1975 (the first date from which a complete set of returns for all the FMP is available) through to the fourth quarter of 2011, which gives us 148 quarterly observations. Excess returns on 10 FMP are in excess of 30-day Treasury bills obtained from the CRSP database.

We also use two alternative sets of test assets. The first is the set of seven Fixed Term Indices (FI). These indices were designed to plot a yield curve and each index is represented by a most recently issued bond that is at least six months, but closest to the target maturity date. The target maturities are the following: 1 year (FI1), 2 years (FI2), 5 years (FI5), 7 years (FI7), 10 years (FI10), 20 years (FI20) and 30 years (FI30). The second set of test assets includes five Fama-Bliss Discount Bonds (FB). These represent artificial zero-coupon bonds and were designed to reflect the discount yield term structure. The maturities of the FB discount bonds are the following: 1 year (FB1), 2 years (FB2), 3 years (FB3), 4 year (FB4), 5 years (FB5). Both alternate sets of test assets are used mainly for the purpose of robustness check for our results related to 10 FMP.<sup>11</sup>

We obtain a balanced panel of 125 macroeconomic and financial time series from the Global Insights Basic Economics and the Conference Board's Indicators Databases. These series represent broad categories of economic and financial variables such as: real output and income, employment and labor market, housing market, orders and inventories, money and credit markets, stock market, interest rates and bond yields, exchange rates, prices. We use quarterly data and, following Stock and Watson

---

<sup>11</sup>More details on US bond data used in this paper can be found at the following CRSP webpage: <http://www.crsp.com/products/documentation/supplemental-series>

(2002a) and Stock and Watson (2002b), standardize and transform the data where necessary to ensure stationarity prior to the estimation of common factors. The detailed description of the data as well as the relevant transformations are provided in the Online Appendix.

Finally, we construct the real consumption series using personal consumption expenditure on nondurables and services, deflated by a weighted average of price indices for nondurables and for services. This is then divided by population for the corresponding time period to obtain a real per capita consumption measure. The consumption growth is defined as log change in real per capita quarterly consumption of nondurables and services.<sup>12</sup> The relevant data necessary for the above computations is obtained from the US Department of Commerce, Bureau of Economic Analysis.

We also use, as a robustness check, data on quarterly consumption growth obtained using an alternate procedure. We follow Piazzesi and Schneider (2007) who measure per capita consumption growth as equal to the growth rate of the raw consumption NIPA data minus a constant and assume that population growth is constant. This allows them to mitigate three issues; the large differences in the standard population series available from various data sources, the presence of very large spikes at points where the census data is collected every decade and the related interpolation issues between census years. The details of the consumption data construction are given in Piazzesi and Schneider (2007).

## 6 Empirical results

### 6.1 Summary statistics

In contrast to the stylized facts for equity portfolios the features of the Fama Maturity Portfolios that we use, are less well-known. We report, in table 1, basic summary statistics for excess returns on 10 FMP sorted on maturities, over the period 1975–2011.

We can observe that over the 37-year sample period, average excess returns on bond portfolios increase with the maturity of the constituent bonds: the average excess return on the portfolio of bonds with the shortest maturities (12–18 months) is 1.382% per annum and it increases in a monotonic fashion to 4.433% per annum for the portfolio of the longest maturity bonds (more than 120 months).

The volatility of excess returns also increases with maturity but at a greater pace: from 2.5% for the shortest-maturity portfolio up to 11.889% for the longest-maturity portfolio. This greater pace is reflected in the annualized Sharpe ratios, that are falling with maturity: from 0.552 for the shortest-maturity portfolio up to 0.372 for the longest-maturity portfolio. It is worth noting that the magnitude of Sharpe ratios for bond portfolios is similar to the Sharpe ratio of broad equity portfolio, represented

---

<sup>12</sup>We consider expenditure on nondurables and services, following a large literature on consumption-based models (see, for example, Lettau and Ludvigson (2001a)).



by S&P500 index, which is approximately 0.43 for the same period. We can also see from table 1 that there is little autocorrelation in quarterly excess returns for all the bond portfolios and in vast majority it is not significant. Our results are similar to those reported by Pilotte and Sterbenzr (2006) using similar data.

## 6.2 Selection on common factors for VAR model and estimation of the VAR

As discussed in earlier sections, the number of common factors well representing a large panel of data is determined, using the information criteria in Bai and Ng (2002). In our case we find that the common variation in the panel of 125 time series is well captured by eight common factors, which we denote as  $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{5t}, \hat{f}_{6t}, \hat{f}_{7t}, \hat{f}_{8t}$ . In table 2 we report the results related to the selection of the common factors that have a strong predictive power for consumption growth. According to BIC, the optimal selection of factors for VAR model includes the following common factors:  $\hat{F}_{5t}^* = \{\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{7t}\}$  however the AIC points at  $\hat{F}_{7t}^* = \{\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{5t}, \hat{f}_{7t}, \hat{f}_{8t}\}$  as the optimal selection. The LL-ratio indicates that the marginal contribution of  $\hat{f}_{5t}$  and  $\hat{f}_{8t}$  to the predictability of consumption growth is not significant so the set of common factors that are carried forward and included in the VAR model as state variables, is  $\hat{F}_{5t}^*$ .

A point of interest in including these factors in the VAR is whether they have any economic interpretation. In general, interpretation of the factors as representing specific types of macroeconomic or financial series is inappropriate since the construction of each one is affected to some degree by all the variables in our large dataset. As a result none of the common factors would correspond exactly to a precise economic concept like output or unemployment especially when such series are naturally correlated. With this caveat, but with a view to get some intuition of what the factors might represent, we follow Stock and Watson (2002b) and Ludvigson and Ng (2009) in characterizing the factors as they relate to the 125 variables in our panel dataset.<sup>13</sup> We depict in figure 1 the marginal  $R^2$  from regressions of each of the 125 individual series from our panel data onto each of selected five factors  $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}$  and  $\hat{f}_{7t}$ , one at a time, using the full sample of data. Figure 1 displays the  $R^2$  statistics as bars in the chart separately for each of the four factors we use. We can observe from this figure that the first factor  $\hat{f}_{1t}$  loads heavily measures related to real output, industrial production, employment and manufacturing orders while displaying little correlation with stock indices and prices. The second factor  $\hat{f}_{2t}$  on the other hand, appears to load most heavily on prices and measures of inflation but displays little correlation with macroeconomic measures. The third factor  $\hat{f}_{3t}$  appears to be largely

<sup>13</sup>While interesting, this analysis takes us away from the main theme of our paper and we refer the reader to Ludvigson and Ng (2009) who have a detailed and an interesting analysis of this issue when using common factors in empirical analysis.

related to several interest spreads while the fourth factor  $\hat{f}_{4t}$  reflects mainly housing market measures and term structure of interest rates. Finally the last factor,  $\hat{f}_{7t}$  correlates mainly to foreign exchange rates.

Finally, we report in table 3 the estimation results of the factor-augmented VAR model with consumption growth and the five common factors as state variables. We can observe from this table that the selected common factors indeed possess strong predictive power for consumption growth, even in the presence of lagged consumption growth. They jointly explain more than 40% of the variation in consumption growth, which is an impressive amount. It turns out that such a good quality of the predictive model for consumption growth accounts partially for an explanation of the reasons that drive our asset pricing tests as will be discussed in later sections.

### 6.3 Asset pricing results

Table 4 reports the GMM estimation results of the linear SDF representation of the model using 10 Fama Maturity Portfolios as test assets. We can observe that when the identity matrix is used to weight moment conditions, the estimate of loading on risk factor related to innovations to expectations about future consumption growth  $\hat{b}_\varepsilon$  equals 50.703 and is statistically significant which indicates that this risk is an important source of variation in the cross-section of excess returns on bond portfolios with different maturities. On the other hand the estimate of the loading related to contemporaneous consumption growth risk  $\hat{b}_{\Delta c}$  is equal to  $-66.742$ , but this value is not significant in statistical terms implying that this risk factor does not seem to play an important role in pricing the cross-section of government bonds in our sample. However, high sampling error related to the estimate of  $\hat{b}_{\Delta c}$  makes it statistically indifferent from one so the theoretical restriction of the model cannot be rejected in our data. The estimate of risk aversion parameter is equal to 51.703 and is statistically different from zero. This value is quite high but most of the consumption-based asset pricing models produce high estimates of risk aversion. Finally, the J statistic rejects the null of insignificant pricing errors in our model. We suspect that the reason of that is high sampling variability of moment conditions. On the other hand, the model explains well the cross-section of bond excess return on 10 Fama Maturity Portfolios – 98.8% of the cross-sectional variation in excess returns is explained by the variation in the two risk factors. Also the RMSE is relatively low, 0.084% on annual basis comparing to the magnitude of annualized excess returns that is between 1.382% and 4.433%. These results are robust to the use of an alternate weighting matrix as specified in panel B of this table. The striking differences for informal statistics like  $R^2$  or RMSE is quite common in these applications, when weighting matrix is different that the identity matrix.

We report in table 5 the estimates of risk aversion when other bond portfolios are used as test

assets. When Fixed Term Indices or Fama-Bliss Discount Bonds are used, the risk aversion estimates are even lower, respectively 35.633 and 49.278. For comparison we report as well the risk aversion estimates for equity portfolios or for the mixture of 10 Fama Maturity Portfolios and different equity portfolios. In these cases the risk aversion estimates are higher, they have a magnitude of 64 – 123 and are mostly significant.

In table 6 we report the results of the Fama-MacBeth regressions for 10 Fama Maturity Portfolios as test assets. We find that the estimate of the price of risk related to innovations to expectations about future consumption growth ( $\lambda_\varepsilon$ ) is positive and statistically significant, which indicates that this source of risk is priced for bond portfolios and the factor risk premium equals 0.239. On the other hand, the price of contemporaneous consumption growth risk is negative ( $-0.138$ ), though not significantly different from zero. This makes this risk source unimportant for pricing bond portfolios which further confirms the conclusions from the linear SDF approach in the previous section. This is also consistent with the well-documented poor performance of the standard C-CAPM for equity portfolios (Hansen and Singleton (1982)). The tests of the hypothesis on joint insignificance of pricing errors reject this hypothesis, again as in case of the linear SDF approach. In panel B of table 6 we include the estimation results from the variation of Fama-MacBeth cross-sectional regression in which there is a constant included in the second stage. As we can observe from the results, the constant is not statistically significant, as predicted by the theory and the other results are qualitatively the same as in panel A where no constant is included in the second stage. Additionally, from the test of whether the  $R^2$  is significantly different from zero and one, we can conclude that the true value of  $R^2$  is not significantly different from one, but we reject the null that is equal to zero. These further confirm that high values of  $R^2$  are not an effect of model misspecification or sampling errors.

Finally, we picture the performance of the model in figure 2 where we plot the mean excess returns predicted by the model vs the actual ones. The closer the bond portfolios lie to the 45 degree line from the origin, the better model performance and the lower the pricing errors. It is clear from the plot that our two-factor model provides a close fit to the data.

## 6.4 Discussion of the results

In this section we discuss what underlines our results and give some intuition on the importance of the risk factor related to innovations in expectations in future consumption growth  $\varepsilon_{c,t+1}$ . We think there are a couple of reasons that drive the results and allow our model to explain well the cross-section of excess returns on bond portfolios.

First, we take a closer look at the measure of two risks in our model, expressed by the relevant covariances  $cov(R_{t+1}^{ei}, \Delta c_{t+1})$  and  $cov(R_{t+1}^{ei}, \varepsilon_{c,t+1})$ . We can observe from figure 3 that there is little

cross-sectional variation in covariances of excess returns with consumption growth and the relation between them and average excess returns on 10 Fama Maturity Portfolios is almost flat. In contrast the spread in covariances of excess returns with  $\varepsilon_{c,t+1}$  is much greater. We can observe from the figure as well that these covariances are clearly increasing with with average excess returns. Thus the portfolio of bonds with long maturity requires a greater return than a portfolio of short-maturity bonds as a compensation for a greater risk related to uncertainty of consumption in all future periods and the relevant expectations of this consumption stream. More importantly, we show in table 7 that the findings pictured in figure 3 are statistically important rather than a matter of sampling error. From Panel A of table 7 we can see that the covariances of excess returns with consumption growth, although individually significant, are small and there is a weak evidence that they are jointly significant or statistically different from each other (the highest covariance is -0.085 vs the lowest -0.319). In striking contrast, the covariances of excess returns with innovations to expectations in future consumption growth, reported in Panel B of table 7, are much larger, varying from 0.514 to 1.783. In comparison to covariances with consumption growth the spread between them is more than five times greater. They are not only individually significant but also jointly significant and, what's more important, they are significantly different from each other in statistical terms. This implies that the covariance risk related to uncertainty in consumption growth over all future periods is significantly different for portfolios of bonds with different maturities: short-maturity bonds have low covariance risk so are less risky and as a result pay low risk premium, the covariance risk increases with maturity and long-term bonds have high covariance risk so they are less risky and as a result they pay high risk premium. These findings are further confirmed by the fact that innovations to expectations in future consumption growth correlate more tightly with excess returns on all bond portfolios than the consumption growth. We can observe from table 7 that the magnitude of correlations is roughly 0.50 vs 0.16 in absolute terms. Such a tighter co-movement of innovations to expectations in future consumption growth and excess bond returns further confirms the importance of long run risk in consumption growth in pricing bond portfolios.

We now turn back our attention to the predictability of consumption growth by common factors, given that the measure of  $\varepsilon_c$  is derived from factor-augmented VAR model. As have already pointed, the selected set of common factors  $\{\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{7t}\}$  have an extraordinary forecasting power for consumption growth, even in the presence of lagged consumption growth. They explain jointly around 40% of the variation in consumption growth. More important is, however, that these common factors retain their forecasting power in an out-of-sample prediction. Briefly, the aim of the out-of-sample prediction is to mimic the real-time forecasting and evaluate the power of candidate predictors in situation when the forecast for the time  $t + 1$  can be made using the data available only up to time  $t$ . As a result any model used for a prediction can be also estimated with data up to time  $t$ , in

contrast to a classical predictability exercise where the models are estimated and evaluated with a full sample. In our case it means a recursive estimation of common factors at each time we make a prediction for consumption growth. We compare the performance of the factor-augmented model with a benchmark model that does not include common factors. We select two benchmark models: a historical mean and an AR(1) model for consumption growth and we evaluate the relative performance of the factor-augmented model for three different evaluation periods as specified in table 8.<sup>14</sup> We can observe from this table that the factor-augmented forecasting model performs on average better than the benchmark model (no matter which benchmark or evaluation period we select). It always produces lower forecasting errors as can be observed e.g. from comparison of standard metrics like RMSE and MSE. We also employ a set of standard statistical tools to compare the forecasting power of two nested models: the MSE-t test of Diebold and Mariano (1995) and West (1996) (to test equal MSE of two models), the MSE-F test of McCracken (2004) (to test for equal forecasting accuracy of two models), the ENC-T test of Harvey et al. (1998) and the ENC-NEW test of Clark and McCracken (2001) (both to test for forecast encompassing). We find that all the four test statistics are greater than the critical values when constant benchmark is used, indicating that in each case we reject the null of identical forecasting power of the two models. The results are slightly weaker when AR(1) is used as a benchmark. This out-of-sample forecasting exercise further confirms that the strong forecasting power of common factors for consumption growth is indeed preserved in a real-like forecasting exercise.

An interesting and important question is how our measure of long run risk in consumption growth is related to other forward looking measures used in the literature. Our approach is an alternative to e.g. Parker and Julliard (2005) who express a long run risk in consumption as a consumption growth over longer periods (with an optimal period of around 3 years). We show that, in spite of different measurements, we capture similar aspects of the long run risk in consumption for bond portfolios. In figure 4 we plot the covariances of consumption growth over horizons of 1 up to 20 quarters with the returns on a bond factor for three groups of bonds: Fama Maturity Portfolios, Fixed Term Indices and Fama-Bliss Discount Bonds. The bond factor represents a difference between the returns of the highest and lowest maturities bonds and is constructed in a similar spirit as e.g. size or book-to-market factors for stock portfolios. We can observe in figure 4 that the covariances increase with a horizon over which consumption growth is measured with a peak of around 15 quarters when they are of a similar magnitude as the relevant covariances of innovations to expectations in future consumption growth  $\varepsilon_c$  with the bond factor returns. Such a pattern is consistent for all three groups of bonds that we consider. We conclude that innovations to expectations about future consumption growth are similar, in terms of covariance risk measurement, to consumption growth over almost 4 years. In graph

---

<sup>14</sup>The out-of-sample forecasting is becoming a standard approach e.g. for predicting stock returns or equity premia. The details of out-of-sample forecasting methodology can be found e.g. in Diebold and Mariano (1995), Harvey et al. (1998), Clark and McCracken (2001), McCracken (2007).

From figure 4 we plot the covariances of consumption growth over different horizons with our forward looking measure of consumption risk  $\varepsilon_c$  and observe the peak for around 15 quarters, which further confirms our conclusions. Interestingly Parker and Julliard (2005) observe an equivalent pattern for size and book-to-market factors as shown in the figure 2 of their article. However, in contrast to our findings, they demonstrate that the "optimal" horizon for consumption growth is 11 quarters when equity portfolios are tested. What distinguishes our forward looking measure of consumption risk from a simple consumption growth over longer periods is the correlation with bond returns. We can observe in figure 5 that the correlation coefficients between innovations to expectations about future consumption growth and bond factor returns for different groups of bonds are not only significant in statistical terms but also of a magnitude of 0.3–0.4, which is much greater than the magnitude of relevant correlations with consumption growth over different horizons. These, on the other hand, are pretty stable over different horizons and not greater than 0.07 in absolute terms.<sup>15</sup> Our measure of long run consumption risk co-moves then more tightly with bond returns on different maturities, which explains why in the cross-section this type of risk has a positive price. We believe, this is another driver of our good empirical results when pricing bond portfolios with different maturities.

Finally, we address the questions of whether the duration of bonds, a classical measure of bond risk, is in any way related to the consumption-based measures of risk used in this paper. For that reason we use Fixed Term Indices and Fama-Bliss Discount Bonds because their durations are simply the relevant target maturities. We plot in figure 6 the covariance risk measures vs duration for the two groups of bonds. We find that covariance risk related to consumption growth is actually negatively related to duration, which clearly contradicts the theory. On the other hand the covariance risk related to long run uncertainty in consumption growth is increasing with duration. We find that not only the covariances are individually significant, but they are also jointly significant and statistically different from each other as can be viewed from the detailed estimation results included in the Online Appendix. These findings confirm that our consumption-based risk measure is in line with a traditional measure of risk for bonds which is duration.

## 6.5 Comparisons with other models

We now present, in table 9, the results of the comparison showing how some prominent models perform in pricing the cross-section of government bond returns. We consider the following linear factor models: the classical CAPM with market risk premium, the Fama–French three-factor model which adds size and book-to-market factors and the five-factor model of Fama and French (1993) additionally augmented over the three-factor mode with two bond market factors: term premium and default pre-

---

<sup>15</sup>The data used to create figures 4 and 5 are available from authors upon request. We do not report them here to conserve the space.

mium. We measure term premium (UTS) as the yield spread between 20-year and 1-year US Treasury bonds and default premium (UPR) as the yield spread between BAA-rated and AAA-rated corporate bonds. We also compare the performance of our model with the linearized versions of the power utility C-CAPM, the Parker and Julliard (2005) C-CAPM with ultimate consumption growth over 11 and 15 quarters and the Epstein and Zin (1991) C-CAPM with market risk and consumption growth as risk factors.

We find that the classical CAPM model perform very poorly in pricing government bonds – it has a negative  $R^2$  and very high RMSE. But when augmented with size and book-to-market factors, or additionally with term and default premia, it improves significantly on these statistics over the simple CAPM. The inclusion of two bond market factors over the three-factor model decreases the HJ distance but also introduces more sampling error for this statistics. Surprisingly in our sample only the three-factor model produces the HJ distance that is not significantly different from zero. The results in table 9 suggest also that the two bond market factors are not statistically significant when pricing government bond portfolios. When we compare our model with other consumption-based models we find that it produces the highest  $R^2$  of 99% and the lowest RMSE of 0.021% per quarter, though the differences are not of a great magnitude. The values of HJ distance are similar for all consumption-based model as well. It is also interesting to observe that the Parker and Julliard (2005) model with ultimate consumption growth seems to be sensitive to a choice of period over which consumption growth is taken. Just by increasing the growth period from 11 to 15 quarters, the model improves significantly on all the statistics e.g. the  $R^2$  increases from 70% to 95%.

Overall, we conclude that our consumption-based model compares reasonably well with other related models in pricing the cross-section of average government bond returns.

## 6.6 Tests for robustness

As a main robustness check for our model, we investigate whether our model prices well the cross-section of return on two groups of assets: stocks and bonds. While there literature on pricing the cross-section of equities is huge, it has become a standard now to evaluate empirical performance of asset pricing models for a mixture of bond and stock portfolios. We employ this as well and assess the performance of our model for the following four groups of test assets: 10 Fama Maturity Portfolios mixed with (i) market portfolio (ii) 10 size portfolios, (iii) 10 book-to-market portfolios and (iv) 25 Fama-French portfolios. The visualization of the performance is presented in figure 7, along with the  $R^2$  values marked in relevant graphs. We find that our model also prices well the equity portfolios and has slightly more explanation power for joint bond and size portfolios (98.9%) than for joint bond and book-to-market portfolios (95.8%). We can observe as well from this figure that excess returns on

bond and stocks account for a complementary picture with bond excess returns up to 1.5% per quarter and stock excess returns of 2%–4% per quarter.

We also perform several other tests to assess the robustness of our results. Here we describe only the conclusions of these tests and provide full details in the Online Appendix.

First, we use two alternate sets of test assets. The first includes Fixed Term Indices and the second set comprises of Fama-Bliss Discount Bonds. Our results for these two groups of test assets are qualitatively similar to those obtained using the Fama Maturity portfolio returns. Specifically, the innovations to expectations about future consumption growth in an important risk factor when pricing the cross-section of quarterly returns on these assets.

Next, we estimate the model using a different consumption measure in extracting the innovations to expectations about future consumption growth following Piazzesi and Schneider (2007) and test the model for 10 Fama Maturity Portfolios. The results, detailed in the Online Appendix, again support the main results in our paper obtained using the classical measure of consumption growth.

We also stress test our model using a sub-sample of our data. Our full sample period, 1975–2011 is dictated by the availability of non-missing data on the Fama Maturity Portfolios. We therefore test the model using data over the sub-period 1982–2011 – a period marked by a large decline in volatility of major macroeconomic variables like GDP growth or inflation rate, known as the Great Moderation. As a result, we can test the model during a period characterized by a common macroeconomic environment while at the same time avoid exacerbating the problem of small sample size given the span of data points available to us. We find, again, that there is strong support for our main conclusions obtained using the full sample of data.

Finally we report the relevant estimation results for the model in which we include the third pricing factor: the variance of expected future consumption growth shock,  $Var(\varepsilon_{c,t})$  estimated as a square of innovations to expectation about future consumption growth. Bansal and Yaron (2004) assign a crucial role for a conditional volatility of consumption growth and we take that on board in our paper as well. The conclusion from these results is that the volatility factor is not statistically important in pricing the cross-section of government bonds during our sample period.

## 7 Summary and conclusion

In this paper we use a consumption-based asset pricing model with Epstein–Zin–Weil recursive preferences to explain the differences between the average excess returns on portfolios of nominal US Treasury bonds with different times to maturity, ranging from one year to more than ten years. In other words we investigate whether consumption-based asset pricing models provide any explanation of risk premia in bond markets. Our consumption-CAPM has two risk factors: contemporaneous con-



sumption growth, as in case of classical power utility CAPM, and innovations to expectations about future consumption growth, a forward looking measure of uncertainty in long run prospects in future consumption.

We use a VAR model to estimate the forward looking measure of uncertainty in consumption. A novelty of our application of VAR is that, instead of specific predictive variables, we use common factors obtained from a large panel of macroeconomic and financial time series. We show in this paper that selected common factors possess very good forecasting abilities for consumption growth and, along with lagged consumption growth, are able to explain more than 40% of the variation in consumption growth. What's more, we find that such a good forecasting power of common factors is preserved in an out-of-sample prediction as well. This further justifies the use of common factor in VAR as the quality of prediction of consumption growth determines the quality of measure of innovations in expectations about future consumption growth.

Our main finding is that consumption-based models, like ours, can provide a coherent story of the riskiness of government bonds. We find that, for our sample, the risk related to long run prospects in future consumption plays a crucial role in pricing nominal US government bonds. Risk premium related to that source or risk is found to be positive and statistically significant. Further, we explain why long term government bonds pay on average greater returns than short term government bonds: because they are more risky. And the risk that is rewardable, is the one related to long run prospects in future consumption growth. Specifically, we measure that risk by covariances of portfolios returns with innovations in expectations about future consumption growth and we demonstrate that long term bond portfolios have significantly greater covariances, thus risk, than short term bond portfolios. Our two-factor model explains around 98.8% of the cross-sectional variation in average excess returns of portfolios of bonds with different maturities, over the period 1975–2011. We show as well that covariance risk related to innovations in expectations about future consumption growth, is consistent with a standard measure of risk in bonds – duration. Bonds with greater duration have as well greater covariance risk related to long run risk in prospects in future consumption. We find also, that the estimates of the coefficient or relative risk aversion are lower when bond portfolios are used as test assets, rather than equity portfolios, though the values are still high comparing to the values implied by economic theory. Our model is also robust when pricing jointly bond and equity portfolios. Finally, we find that our model does well, in terms of pricing the cross-section of government bond excess returns, when compared to several competing CAPM and consumption-based CAPM models and our conclusions remain robust to a variety of checks for robustness. Overall the results imply that investors must be rewarded to hold government bond portfolios which are risky rather than safe assets and which may not be useful in hedging against risk related to long run prospects in consumption growth.

## **Acknowledgements**

We would like to thank David Babbel, Angela Black, Laurence Copeland, Antonio Diez de los Rios, Kabir Dutta, Javier Gil-Bazo, Lynda Khalaf, Chung-Ming Kuan, Patrick Minford, Francisco Penaranda, Jesper Rangvid, Enrique Sentana and seminar participants at the Universities of Aarhus, Aberdeen, Autonoma de Barcelona, Cardiff, Carlos III de Madrid, Essex, National Central University (Taiwan), National Taiwan University, Pompeu Fabra, Reading and the participants at the 2009 Warsaw International Economic Meeting, 2009 Econometric Society European Meeting Barcelona, 2009 ASSET Istanbul, XVII Foro Finanzas Madrid, XXXIV SAEe Valencia, 5th PhD Meeting of RES London for helpful discussions and comments.

## References

- Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50 (4), 745–787.
- Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models? *Econometrica* 70 (1), 191–221.
- Bai, J., Ng, S., 2008. Large dimensional factor analysis. *Foundations and Trends in Econometrics* 3 (2), 89–163.
- Bansal, R., 2007. Long-run risks and financial markets. *St. Louis Federal Reserve Bank Review* 89 (4), 287–300.
- Bansal, R., Yaron, A., 2004. Risks for the long-run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59 (4), 1481–1509.
- Bernanke, B., Boivin, J., Elias, P., 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics* 120 (1), 387–422.
- Boguth, O., Kuehn, L., 2009. Consumption volatility risk. Working Paper, Sauder School of Business University of British Columbia.
- Brunnermeier, M., Julliard, C., 2008. Money illusion and housing frenzies. *Review of Financial Studies* 21 (1), 135–180.
- Campbell, J., 2007. Comments on *Equilibrium Yield Curves* by M. Piazzesi and M. Schneider. *NBER Macroeconomics Annual* 21, 459–468.
- Campbell, J., Sunderam, A., Viceira, L., 2010. Inflation bets or deflation hedges? the changing risks of nominal bonds. Working Paper, Harvard University, 157–179.
- Campbell, J., Vuolteenaho, T., 2004. Bad beta, good beta. *American Economic Review* 94 (5), 1249–1275.
- Chang, E., Huang, R., 1990. Time-varying return and the risk in the corporate bond market. *Journal of Financial and Quantitative Analysis* 25 (3), 323–340.
- Chen, L., Zhao, X., 2009. Return decomposition. *Review of Financial Studies* 22 (12), 5213–5249.
- Chen, Y., Ferson, W., Peters, H., 2010. Measuring the timing ability and performance of bond mutual funds. *Journal of Financial Economics* 98 (1), 72–89.
- Clark, T., McCracken, M., 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105 (1), 85–110.
- Cochrane, J., 2005. *Asset Pricing*. Princeton University Press.
- Constantinides, G., Jackwerth, J., Savov, A., 2009. The puzzle of index option returns. Working Paper, University of Chicago.
- Coval, J., Shumway, T., 2001. Expected option returns. *Journal of Finance* 56 (3), 983–1009.
- Daniel, K., Marshall, D., 1997. The equity premium puzzle and risk-free rate puzzles at long horizons. *Macroeconomic Dynamics* 1 (2), 452–484.
- Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13 (3), 253–263.

- Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. *Mathematical Finance* 6 (4), 379–406.
- Epstein, L., Zin, S., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica* 57 (4), 937–969.
- Epstein, L., Zin, S., 1991. Substitution, risk aversion and the temporal behavior of consumption and asset returns: an empirical investigation. *Journal of Political Economy* 99 (6), 263–286.
- Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1), 3–56.
- Gallmeyer, M., Hollifield, B., Palomino, F., Zin, S., 2007. Arbitrage-free bond pricing with dynamic macroeconomic models. *Federal Reserve Bank of St. Louis Review*, 305–326.
- Gebhardt, W., Hvidkjaer, S., Swaminathan, B., 2005. The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics* 75 (1), 85–114.
- Gultekin, N., Rogalski, R., 1985. Government bond returns, measurement of interest rate risk, and the arbitrage pricing theory. *Journal of Finance* 40 (1), 43–61.
- Hansen, L., Heaton, C., Li, N., 2005. Consumption strikes back? measuring long run risk. NBER Working Paper No. 11476.
- Hansen, L., Singleton, K., 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50 (5), 1269–1286.
- Harvey, D., Leybourne, S., Newbold, P., 1998. Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16, 254–259.
- Jagannathan, R., Wang, Z., 1998. An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *Journal of Finance* 53 (4), 1285–1309.
- Jagannathan, R., Wang, Z., 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. *Journal of Finance* 62 (4), 1623–1661.
- Kan, R., Robotti, C., 2008. Specification tests of asset pricing models using excess returns. *Journal of Empirical Finance* 15 (5), 816–838.
- Kan, R., Robotti, C., Shanken, J., 2009. Pricing model performance and the two-pass cross-sectional regression methodology. Working Paper, Federal Reserve Bank of Atlanta.
- Khan, S., Khokher, Z., Simin, T., 2007. Scarcity and risk premiums in commodity futures. Working Paper, University of Western Ontario.
- Lettau, M., Ludvigson, S., 2001a. Consumption, aggregate wealth and expected stock returns. *Journal of Finance* 56 (3), 815–849.
- Lettau, M., Ludvigson, S., 2001b. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109 (6), 1238–1287.
- Ludvigson, S., Ng, S., 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22 (12), 5027–5067.
- Lustig, H., Nieuwerburgh, S. V., 2008. The returns on human capital: Good news on wall street is bad news on main street. *Review of Financial Studies* 21 (5), 2097–2137.
- Malloy, C., Moskowitz, T., Vissing-Jorgensen, A., 2009. Long-run stockholder consumption risk and asset returns. *Journal of Finance* 64 (6), 2427–2479.

- McCracken, M., 2004. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105 (1), 85–110.
- McCracken, M., 2007. Asymptotics for out of sample tests of granger causality. *Journal of Econometrics* 140 (2), 719–752.
- Newey, W., West, K., 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3), 703–708.
- Parker, J., Julliard, C., 2005. Consumption risk and the cross-section of expected returns. *Journal of Political Economy* 113 (1), 185–222.
- Piazzesi, M., 2009. Affine term structure models. In: *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia and L.P. Hansen. North-Holland, Amsterdam, pp. 691–766.
- Piazzesi, M., Schneider, M., 2007. Equilibrium yield curves. *NBER Macroeconomics Annual* 21, 389–442.
- Pilotte, E., Sterbenz, F., 2006. Sharpe and treynor ratios on treasury bonds. *Journal of Business* 71 (1), 149–180.
- Roll, R., 1971. Investment diversification and bond maturity. *Journal of Finance* 26 (1), 51–66.
- Shanken, J., 1992. On the estimation of beta pricing models. *Review of Financial Studies* 5 (1), 1–34.
- Stock, J., Watson, M., 2002a. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97 (460), 1167–1179.
- Stock, J., Watson, M., 2002b. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20 (2), 147–162.
- Stock, J., Watson, M., 2008. Forecasting in dynamic factor models subject to structural instability. In: *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor D.F. Hendry*. Oxford University Press, pp. 173–205.
- Stock, J., Watson, M., 2011. Dynamic factor models. In: *Oxford Handbook of Economic Forecasting*, ed. by M.P. Clements and D.F. Hendry. Oxford University Press, pp. 35–60.
- Vasicek, O., 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5 (2), 177–188.
- Viceira, L., 2012. Bond risk, bond return volatility, and the term structure of interest rates. *International Journal of Forecasting* 28 (1), 97–117.
- Weil, P., 1989. The equity premium puzzle and the risk-free puzzle. *Journal of Monetary Economics* 24 (3), 401–421.
- West, K., 1996. Asymptotic inference about predictive ability. *Econometrica* 64 (5), 1067–1084.
- Wolman, A., 2006. Bond price premiums. *Federal Reserve Bank of Richmond Economic Quarterly* 94 (2), 317–336.

## Tables

Table 1: Summary statistics of excess returns for Fama Maturity Portfolios

	FMP1	FMP2	FMP3	FMP4	FMP5	FMP6	FMP7	FMP8	FMP9	FMP10
Mean	1.382	1.567	1.866	2.132	2.398	2.509	2.608	2.719	3.117	4.433
St dev	2.500	3.208	3.987	4.538	4.961	5.511	5.902	6.515	7.680	11.889
Sharpe Ratio	0.552	0.488	0.468	0.469	0.483	0.455	0.442	0.417	0.405	0.372
Skewness	1.438	1.031	0.921	0.769	0.534	0.567	0.455	0.621	0.481	0.687
Kurtosis	10.123	6.846	6.141	4.625	2.588	2.318	1.611	2.525	1.515	1.606
Minimum	-18.215	-22.065	-27.849	-31.405	-30.174	-32.351	-33.414	-37.828	-44.577	-64.325
Maximum	31.150	36.293	44.200	46.738	44.444	48.775	48.687	60.582	59.905	86.356
$\rho_1$	-0.132	-0.118	-0.118	-0.112	-0.090	-0.076	-0.072	-0.096	-0.037	-0.044
$\rho_2$	0.034	0.006	0.013	0.011	0.019	-0.015	-0.010	-0.017	-0.022	-0.069
$\rho_3$	0.144	0.162*	0.132	0.133	0.146	0.151	0.144	0.103	0.106	0.079
$\rho_4$	0.010	0.036	0.040	0.035	0.050	0.040	0.053	0.041	0.032	-0.005

Notes: Table 1 presents summary statistics for excess returns (annualized % excess returns) on Fama Maturity Portfolios over the quarterly 30-day Treasury bill rates. The quarterly holding period returns on the Fama Maturity Portfolios are computed using monthly returns obtained from the CRSP US Treasury Database. Quarterly T-bill rates are obtained from the CRSP US Treasury Database. The Fama Maturity Portfolios consist of non-callable, non-flower notes and bonds and are defined by six-month maturity intervals. In this table, FMP1 represents excess returns on a portfolio of bonds that mature from 13 to 18 months from the quote date, FMP2 includes bonds with maturity from 19 to 24 months from the quote date. The other portfolios are: FMP3 – maturity from 25 to 30 months, FMP4 – maturity from 31 to 36 months, FMP5 – maturity from 37 to 42 months, FMP6 – maturity from 43 to 48 months, FMP7 – maturity from 49 to 54 months, FMP8 – maturity from 55 to 60 months, FMP9 – maturity from 61 to 120 months, and FMP10 – maturity greater than 120 months. The data span the period 1975Q1–2011Q4. Also reported in this table are autocorrelations up to 4 quarters: those marked with asterisk are significant, exceeding the 95% confidence interval of  $\pm 1.96/\sqrt{T} \approx \pm 0.161$ .

Table 2: Selection of common factors for VAR model

	LL-Ratio	$R^2$	$\bar{R}^2$	BIC	AIC
$F_{1t}^* = \{f_{4t}\}$	0.000	0.125	0.120	1.313	1.297
$F_{2t}^* = \{f_{1t}, f_{4t}\}$	22.349*	0.214	0.206	1.231	1.199
$F_{3t}^* = \{f_{1t}, f_{3t}, f_{4t}\}$	21.328*	0.291	0.281	1.153	1.105
$F_{4t}^* = \{f_{1t}, f_{2t}, f_{3t}, f_{4t}\}$	19.839*	0.356	0.343	1.083	1.019
$F_{5t}^* = \{f_{1t}, f_{2t}, f_{3t}, f_{4t}, f_{7t}\}$	7.644*	0.379	0.364	1.072	0.992
$F_{6t}^* = \{f_{1t}, f_{2t}, f_{3t}, f_{4t}, f_{5t}, f_{7t}\}$	4.639	0.393	0.375	1.076	0.979
$F_{7t}^* = \{f_{1t}, f_{2t}, f_{3t}, f_{4t}, f_{5t}, f_{7t}, f_{8t}\}$	3.993	0.404	0.383	1.082	0.969
$F_{8t}^* = \{f_{1t}, f_{2t}, f_{3t}, f_{4t}, f_{5t}, f_{6t}, f_{7t}, f_{8t}\}$	0.150	0.405	0.381	1.107	0.978

Notes: Table 2 presents the relevant steps of the selection procedure that aims to select an optimal set of common factors with predictive power for consumption growth for a VAR model.  $F_{kt}^*$  denotes an optimal set of  $k$  common factors out of eight possible, in terms of highest predictive power for consumption growth. For each of such optimal sets (i.e. for  $k = 1, 2, \dots, 8$ ) we report  $R^2$ , adjusted- $R^2$  ( $\bar{R}^2$ ), BIC and AIC information criteria. The log-likelihood ratio (LL-Ratio) reflects the marginal contribution of  $F_{(k+1)t}^*$  with respect to  $F_{kt}^*$  to the predictability of consumption growth. Asterisks \* indicate statistical significance at 1% significance level. The data span the period 1975Q1–2011Q4.

Table 3: Estimation of factor-augmented VAR model

							R <sup>2</sup>	BIC
	$\Delta c_{t-1}$	$\hat{f}_{1,t-1}$	$\hat{f}_{2,t-1}$	$\hat{f}_{3,t-1}$	$\hat{f}_{4,t-1}$	$\hat{f}_{7,t-1}$	$\bar{R}^2$	AIC
$\Delta c_t$	0.241	0.078	0.111	0.113	0.154	-0.051	0.411	1.019
se OLS	(0.072)**	(0.033)**	(0.026)**	(0.027)**	(0.027)**	(0.027)**	0.397	0.938
se BOOT	[0.073]**	[0.033]**	[0.027]**	[0.026]**	[0.027]**	[0.026]**		
$\hat{f}_{1t}$	0.213	0.624	0.153	0.291	0.265	0.027	0.680	1.829
se OLS	(0.109)**	(0.050)**	(0.040)**	(0.041)**	(0.041)**	(0.041)	0.672	1.749
se BOOT	[0.106]**	[0.048]**	[0.040]**	[0.041]**	[0.042]**	[0.043]		
$\hat{f}_{2,t}$	-0.076	0.202	-0.311	0.035	0.200	0.112	0.185	2.766
se OLS	(0.174)	(0.080)**	(0.064)**	(0.065)	(0.065)**	(0.066)**	0.164	2.686
se BOOT	[0.177]	[0.079]**	[0.063]**	[0.066]	[0.064]**	[0.067]**		
$\hat{f}_{3,t}$	0.245	-0.441	0.036	0.668	-0.056	-0.028	0.630	1.975
se OLS	(0.117)**	(0.054)**	(0.043)	(0.044)**	(0.044)*	(0.044)	0.621	1.894
se BOOT	[0.113]**	[0.051]**	[0.045]	[0.043]**	[0.044]*	[0.043]		
$\hat{f}_{4,t}$	0.238	0.231	0.069	-0.182	0.269	-0.166	0.246	2.688
se OLS	(0.167)*	(0.077)**	(0.062)	(0.063)**	(0.063)**	(0.063)**	0.227	2.607
se BOOT	[0.169]*	[0.073]**	[0.062]	[0.060]**	[0.061]**	[0.062]**		
$\hat{f}_{7,t}$	-0.512	0.211	-0.279	0.194	0.153	0.185	0.233	2.663
se OLS	(0.165)**	(0.076)**	(0.061)**	(0.062)**	(0.062)**	(0.062)**	0.214	2.582
se BOOT	[0.162]**	[0.076]**	[0.060]**	[0.062]**	[0.062]**	[0.059]**		

Notes: Table 3 presents the estimation results of a factor-augmented VAR(1) model, for consumption growth and five optimal common factors, of the following form:  $Z_{t+1} = AZ_t + w_{t+1}$ , where  $Z_t = [\Delta c_t \hat{f}_{1t} \hat{f}_{2t} \hat{f}_{3t} \hat{f}_{4t} \hat{f}_{7t}]$ ,  $\Delta c_t$  is real per capita consumption growth and  $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{7t}$  denote estimated common factors, selected for the VAR model. The series of consumption growth  $\Delta c_t$  have been multiplied by 100 for convenience. Each row in the table corresponds to a relevant, single equation in the VAR(1) model. We report in the table the coefficient estimates for matrix  $A$ , two types of standard errors for coefficient estimates: OLS standard errors in parentheses and bootstrapped standard errors in square brackets,  $R^2$  and adjusted- $R^2$  ( $\bar{R}^2$ ), BIC and AIC information criteria. Asterisks indicate statistical significance of the coefficient estimates at \*10% and \*\* 5% levels.



Table 4: GMM estimation of Euler equation with linear SDF for Fama Maturity Portfolios

Panel A: W=I	$\hat{b}_{\Delta c}$	$\hat{b}_{\varepsilon}$	$H_0: b_{\Delta c}=1$	$\hat{\gamma}$	$J_T$ stat	p-value		
estimates	-66.742	50.703		51.703			RMSE	0.021
pv HAC(0)	[0.291]	[0.084]	[0.288]	[0.080]	26.110	[0.001]	MAE	0.018
pv HAC(6)	[0.287]	[0.100]	[0.284]	[0.096]	39.688	[0.000]	$R^2$	0.988
pv HAC(auto)	[0.288]	[0.100]	[0.285]	[0.097]	19.045	[0.014]	$\bar{R}^2$	0.986
Panel B:								
W=Var( $R^e$ ) <sup>-1</sup>	$\hat{b}_{\Delta c}$	$\hat{b}_{\varepsilon}$	$H_0: b_{\Delta c}=1$	$\hat{\gamma}$	$J_T$ stat	p-value		
estimates	50.755	50.393		51.393			RMSE	0.252
pv HAC(0)	[0.217]	[0.022]	[0.222]	[0.020]	21.918	[0.005]	MAE	0.240
pv HAC(6)	[0.257]	[0.051]	[0.261]	[0.049]	35.714	[0.000]	$R^2$	-0.489
pv HAC(auto)	[0.221]	[0.027]	[0.225]	[0.025]	18.817	[0.015]	$\bar{R}^2$	-0.862

Notes: Table 4 presents the results of the GMM estimation of Euler equation with linear SDF with weighting matrix  $W = I$  (in Panel A) and  $W = Var(R^e)^{-1}$  (in Panel B) for 10 Fama Maturity Portfolios. We report the estimates of coefficients  $b$  along with p-values for individual significance, related to the following types of standard errors: Newey-West with 0 and 6 lags and with auto-lag selection (HAC). We also report p-values related to testing the theoretical restriction of  $H_0 : b_{\Delta c} = 1$  and the estimate of risk aversion parameter  $\hat{\gamma}$  along with the p-values related to its statistical significance. We present as well the  $J_T$  statistics, used for testing a joint zero pricing errors hypothesis along with relevant p-values. Finally, we report the goodness-of-fit statistics like  $R^2$ , adjusted- $R^2$  ( $\bar{R}^2$ ), RMSE and MAE (in % per quarter). All the p-values are reported in square brackets. The data span the period 1975Q1-2011Q4.

Table 5: Risk aversion estimates for different groups of test assets

	10FMP	7FI	5FB	10Size	10BM	25FF	10FMB +10Size	10FMB +10BM	10FMB +25FF
$\hat{\gamma}$	51.703	35.633	49.278	103.004	64.438	123.441	84.617	84.931	91.397
pv HAC(0)	[0.080]	[0.167]	[0.148]	[0.075]	[0.171]	[0.024]	[0.008]	[0.009]	[0.005]
pv HAC(6)	[0.096]	[0.183]	[0.193]	[0.104]	[0.175]	[0.055]	[0.032]	[0.036]	[0.028]
pv HAC(auto)	[0.097]	[0.178]	[0.177]	[0.075]	[0.162]	[0.024]	[0.032]	[0.031]	[0.008]

Notes: Table 5 presents the estimates of the risk aversion parameter from GMM estimation of Euler equation with linear SDF with identity weighting matrix for different groups of test assets. These are the following: 10 Fama Maturity Portfolios (10FMP), 7 Fixed Term Indices (7FI), 5 Fama-Bliss Discount Bonds (5FB), 10 equity size portfolios (10Size), 10 equity book-to-market portfolios (10BM), 25 Fama-French equity portfolios (25FF). We report as well the p-values (in square brackets) of statistical significance of estimated parameters, related to the following types of standard errors: Newey-West with 0 and 6 lags and with auto-lag selection (HAC). The data span the period 1975Q1–2011Q4.

Table 6: Fama-MacBeth regressions for Fama Maturity Portfolios

Panel A		$\lambda_{\Delta_c}$	$\lambda_\varepsilon$	$\chi^2$	p-value			
estimates		-0.138	0.239			RMSE	0.021	
pv OLS		[0.304]	[0.095]	38.397	[0.000]	MAE	0.018	
pv Shanken		[0.316]	[0.110]	33.526	[0.000]	$R^2$	0.988	
pv HAC(0)		[0.291]	[0.077]	26.110	[0.001]	$\bar{R}^2$	0.986	
pv HAC(6)		[0.293]	[0.057]	39.688	[0.000]			
pv HAC(auto)		[0.293]	[0.057]	13.959	[0.082]			
Panel B		<i>const</i>	$\lambda_{\Delta_c}$	$\lambda_\varepsilon$	$\chi^2$	p-value		
estimates		0.026	-0.119	0.237		RMSE	0.020	
pv OLS		[0.436]	[0.273]	[0.114]	22.211	[0.002]	MAE	0.017
pv Shanken		[0.439]	[0.285]	[0.128]	19.672	[0.006]	$R^2$	0.990
pv HAC(0)		[0.429]	[0.266]	[0.097]	16.592	[0.020]	$\bar{R}^2$	0.987
pv HAC(6)		[0.419]	[0.277]	[0.069]	25.302	[0.000]	$H_0:R^2=1$	[0.185]
pv HAC(auto)		[0.419]	[0.277]	[0.069]	11.519	[0.117]	$H_0:R^2=0$	[0.027]

Notes: Table 6 presents the estimation results of cross-sectional regression using the second stage of Fama-MacBeth methodology without constant (in Panel A) and with constant (in Panel B) for 10 Fama Maturity Portfolios. We report the estimates of factor risk prices  $\lambda$  along with p-values for individual significance, related to the following types of standard errors: OLS, Shanken-corrected, Newey-West with 0 and 6 lags and with auto-lag selection (HAC). The values of  $\lambda$  are multiplied by 100 for convenience. We also report the  $\chi^2$  statistics, used for testing a joint zero pricing errors hypothesis with relevant p-values. Finally, we report the goodness-of-fit statistics like  $R^2$ , adjusted- $R^2$  ( $\bar{R}^2$ ), RMSE and MAE (in % per quarter) and p-value related to testing the following nulls:  $H_0:R^2=1$  and  $H_0:R^2=0$ . All the p-values are reported in square brackets. The data span the period 1975Q1–2011Q4.

Table 7: Covariances of excess returns on Fama Maturity Portfolios with risk factors

Panel A	FMP1	FMP2	FMP3	FMP4	FMP5	FMP6	FMP7	FMP8	FMP9	FMP10	
$cov(R^e, \Delta c)$	-0.085	-0.119	-0.149	-0.176	-0.185	-0.209	-0.234	-0.266	-0.253	-0.319	
pv HAC(0)	[0.168]	[0.132]	[0.129]	[0.109]	[0.098]	[0.096]	[0.080]	[0.088]	[0.117]	[0.154]	
pv HAC(6)	[0.096]	[0.063]	[0.057]	[0.041]	[0.035]	[0.035]	[0.024]	[0.026]	[0.040]	[0.053]	
pv HAC(auto)	[0.051]	[0.025]	[0.020]	[0.012]	[0.012]	[0.010]	[0.006]	[0.004]	[0.013]	[0.012]	
	Wald stat (joint eq of cov)				p-value		Wald stat (joint sign of cov)				p-value
pv HAC(0)					14.302				14.464		[0.152]
pv HAC(6)					15.843				16.109		[0.096]
pv HAC(auto)					13.887				14.251		[0.161]
$corr(R^e, \Delta c)$	-0.150	-0.163	-0.164	-0.170	-0.164	-0.166	-0.174	-0.179	-0.144	-0.117	
Panel B	FMP1	FMP2	FMP3	FMP4	FMP5	FMP6	FMP7	FMP8	FMP9	FMP10	
$cov(R^e, \varepsilon_c)$	0.514	0.632	0.752	0.825	0.850	0.934	0.939	1.049	1.227	1.783	
pv HAC(0)	[0.003]	[0.002]	[0.003]	[0.002]	[0.002]	[0.001]	[0.002]	[0.003]	[0.001]	[0.000]	
pv HAC(6)	[0.034]	[0.034]	[0.035]	[0.034]	[0.030]	[0.030]	[0.030]	[0.033]	[0.025]	[0.016]	
pv HAC(auto)	[0.045]	[0.045]	[0.046]	[0.045]	[0.041]	[0.041]	[0.040]	[0.042]	[0.034]	[0.022]	
	Wald stat (joint eq of cov)				p-value		Wald stat (joint sign of cov)				p-value
pv HAC(0)					30.783				31.954		[0.000]
pv HAC(6)					24.114				24.171		[0.007]
pv HAC(auto)					21.484				21.575		[0.017]
$corr(R^e, \varepsilon_c)$	0.598	0.573	0.548	0.529	0.498	0.493	0.462	0.468	0.464	0.436	

Notes: Table 7 presents the estimates of the covariances of excess returns on 10 Fama Maturity Portfolios with consumption growth  $cov(R^e, \Delta c)$  (in Panel A) and with innovations to expectations in future consumption growth  $cov(R^e, \varepsilon_c)$  (in Panel B) along with p-values for individual significance, related to the following types of standard errors: OLS, Shanken-corrected, Newey-West with 0 and 6 lags and with auto-lag selection (HAC). Covariances are multiplied by 10,000 for convenience. We also report the Wald statistics with relevant p-values to test joint equality of covariances (joint eq of cov) and joint significance of covariances (joint sign of cov). Finally, we report as well the correlation coefficients. All the p-values are reported in square brackets. The data span the period 1975Q1–2011Q4.

Table 8: Predictability of consumption growth in an out-of-sample exercise

Panel A: benchmark model $\Delta c_{t+1} = \alpha_0 + \xi_{t+1}$		MSE-F	MSE-t	ENC-NEW	ENC-t	$\frac{MSE_u}{MSE_r}$	$\frac{RMSE_u}{RMSE_r}$
Evaluation period:	test stat.	80.613	2.665	88.788	4.073	0.647	0.804
1975Q1–2011Q4	95% CV	(0.459)	(0.081)	(4.542)	(1.465)		
	99% CV	(3.783)	(0.677)	(7.495)	(2.154)		
Evaluation period:	test stat.	50.066	2.535	47.477	3.362	0.683	0.826
1985Q1–2011Q4	95% CV	(1.396)	(0.335)	(3.407)	(1.466)		
	99% CV	(4.350)	(0.989)	(5.633)	(2.168)		
Evaluation period:	test stat.	56.933	2.765	48.578	3.123	0.544	0.737
1995Q1–2011Q4	95% CV	(2.238)	(0.780)	(2.537)	(1.550)		
	99% CV	(4.452)	(1.371)	(3.971)	(2.209)		
Panel B: benchmark model $\Delta c_{t+1} = \alpha_0 + \alpha_1 \Delta c_t + \zeta_{t+1}$		MSE-F	MSE-t	ENC-NEW	ENC-t	$\frac{MSE_u}{MSE_r}$	$\frac{RMSE_u}{RMSE_r}$
Evaluation period:	test stat.	34.370	1.634	54.062	3.461	0.811	0.900
1975Q1–2011Q4	95% CV	(0.459)	(0.081)	(4.542)	(1.465)		
	99% CV	(3.783)	(0.677)	(7.495)	(2.154)		
Evaluation period:	test stat.	6.576	0.564	18.017	2.675	0.942	0.970
1985Q1–2011Q4	95% CV	(1.396)	(0.335)	(3.407)	(1.466)		
	99% CV	(4.350)	(0.989)	(5.633)	(2.168)		
Evaluation period:	test stat.	7.850	0.784	14.872	2.276	0.896	0.946
1995Q1–2011Q4	95% CV	(2.238)	(0.780)	(2.537)	(1.550)		
	99% CV	(4.452)	(1.371)	(3.971)	(2.209)		

Notes: Table 8 presents the estimation results on the evidence of out-of-sample predictability of consumption growth by common factors  $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{7t}$ . We compare a performance of the factor-augmented (unrestricted) forecasting model for consumption growth relative to the benchmark (restricted) model which is either a simple constant model (Panel A) or an AR(1) model (Panel B). For each forecast of  $\Delta c_{t+1}$  with factor-augmented model, the factors  $\hat{f}_{1t}, \hat{f}_{2t}, \hat{f}_{3t}, \hat{f}_{4t}, \hat{f}_{7t}$  are recursively estimated using data up to time  $t$ . The relative model evaluation is conducted for different sample periods, specified in the first column of each panel. We report the values of four test statistics along with their 95% and 99% asymptotic critical values (in parentheses). The test statistics are the following: the MSE-t test of Diebold and Mariano (1995) and West (1996), the MSE-F test of McCracken (2004), the ENC-T test of Harvey et al. (1998) and the ENC-NEW test of Clark and McCracken (2001). They are used to test equal mean square errors of the two models (MSE-t), equal forecasting accuracy (MSE-F) and for forecast encompassing (ENC-T and ENC-NEW). Asymptotic critical values are obtained from Clark and McCracken (2001) and McCracken (2007). We report as well in the table the ratios of forecasting errors of the factor-augmented model vs the benchmark model ( $MSE_u/MSE_r, RMSE_u/RMSE_r$ ). The data span the period 1960Q1–2011Q4.

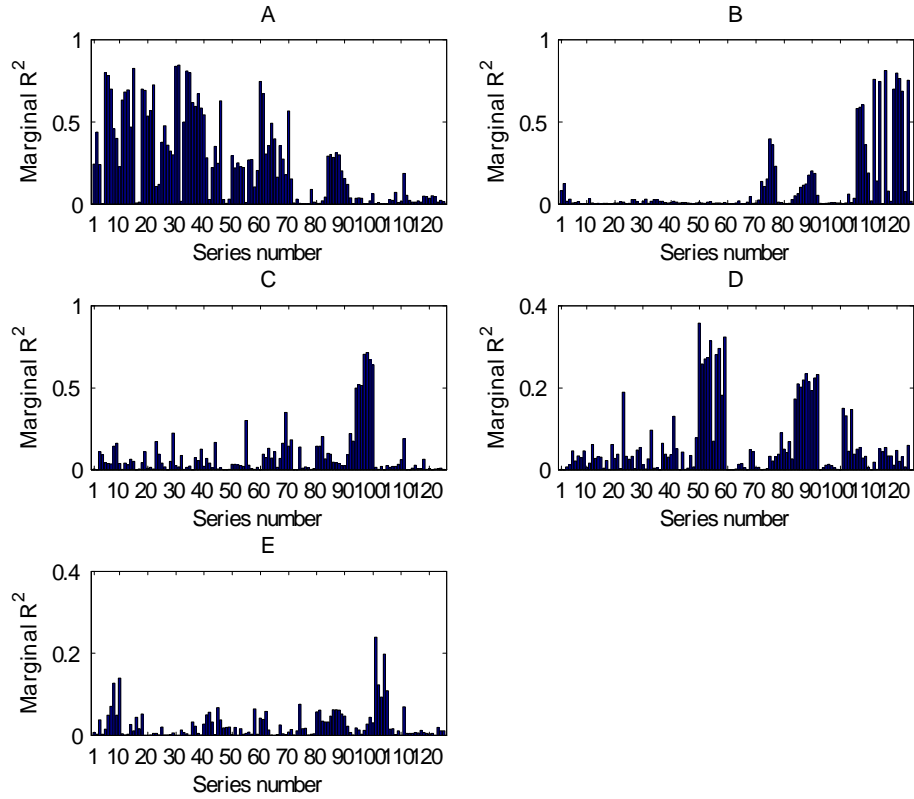
Table 9: Model comparison for Fama Maturity Portfolios

MODEL	$R_m - R_f$	SMB	HML	UTS	UPR	$R^2$ $\bar{R}^2$	RMSE MAE	HJ-dist
CAPM	33.260					-7.628	0.607	0.386
pv HAC(6)	[0.267]					-8.586	0.475	[0.022]
FF-3f	14.952	-23.056	12.665			0.982	0.027	0.337
pv HAC(6)	[0.073]	[0.173]	[0.206]			0.975	0.018	[0.234]
FF-5f	-10.821	33.265	-14.860	164.705	329.848	0.991	0.019	0.287
pv HAC(6)	[0.395]	[0.380]	[0.395]	[0.320]	[0.272]	0.982	0.014	[0.096]
MODEL	$\Delta c$	$R_m$	$\Delta c_S$	$\varepsilon_c$				
C-CAPM	-306.862					0.885	0.069	0.382
pv HAC(6)	[0.099]					0.873	0.050	[0.035]
PJ2005 (S=11)			77.385			0.703	0.093	0.333
pv HAC(6)			[0.197]			0.670	0.077	[0.056]
PJ2005 (S=15)			48.319			0.947	0.033	0.319
pv HAC(6)			[0.107]			0.941	0.029	[0.074]
EZ model	-306.961	1.723				0.889	0.068	0.380
pv HAC(6)	[0.101]	[0.426]				0.862	0.047	[0.029]
Our model	-66.742			50.703		0.988	0.021	0.380
pv HAC(6)	[0.287]			[0.100]		0.986	0.018	[0.040]

Notes: Table 9 presents the results of the comparison of the performance of our model with different asset pricing models: CAPM with market risk premium ( $R_m - R_f$ ) as a risk factor, Fama-French three-factor and five-factor models (FF-3f with two additional factors: size SMB and value HML, and FF-5f respectively with two more risk factors: term premium UTS and default premium UPR), C-CAPM with consumption growth  $\Delta c$ , Parker and Julliard (2005) model with ultimate consumption growth  $\Delta c_S$  over 11 and 15 quarters (PJ2005 (S=11) and PJ2005 (S=15) respectively), Epstein-Zin model with market return and consumption growth as risk factors. For comparison we also include our model. Test assets are 10 Fama Maturity Portfolios. We report in this table the estimates of coefficients  $b$  along their p-values (in square brackets) of individual significance related to Newey-West standard error with 6 lags, estimated by GMM with identity weighting matrix. We also show the goodness-of-fit statistics:  $R^2$ , adjusted- $R^2$  ( $\bar{R}^2$ ), RMSE and MAE (in % per quarter). Finally, we report the values of HJ-distance with the associated p-values (in square brackets). The data span the period 1975Q1–2011Q4.

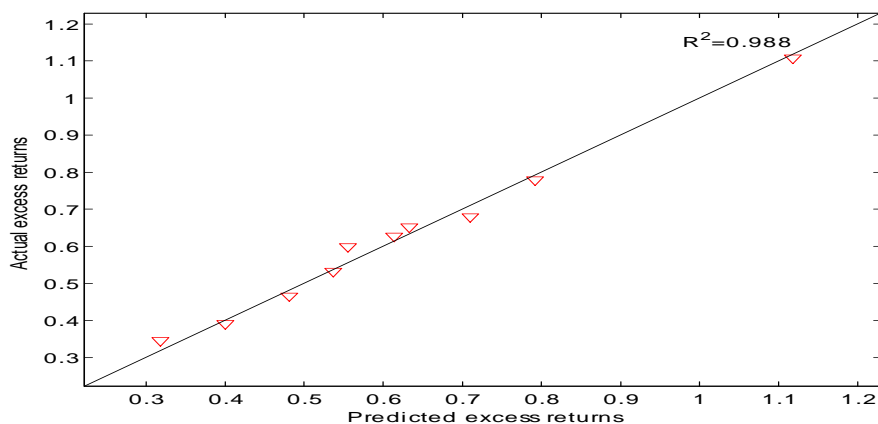
# Figures

Figure 1: Marginal  $R^2$  for selected common factors



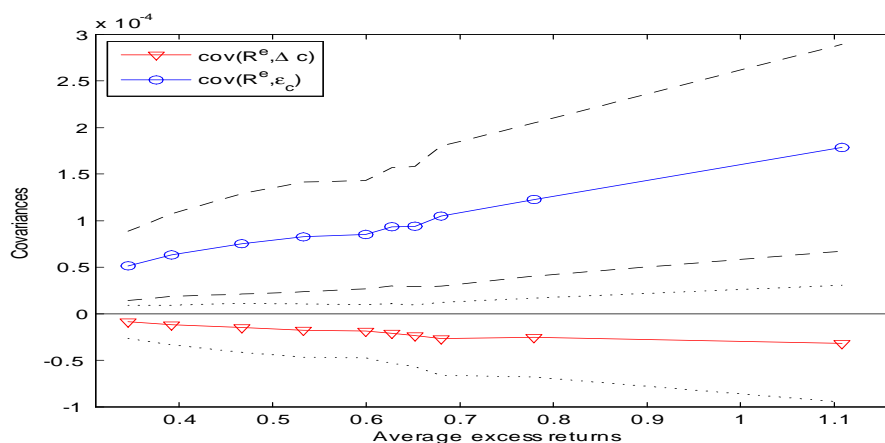
Notes: Figure 1 plots marginal  $R^2$  (on the y-axes) from regressing each of macroeconomic or financial series, denoted by a series number (on the x-axes) on each of the following common factors:  $\hat{f}_1$ ,  $\hat{f}_2$ ,  $\hat{f}_3$ ,  $\hat{f}_4$ ,  $\hat{f}_7$  (A for  $\hat{f}_1$ , B for  $\hat{f}_2$ , C for  $\hat{f}_3$ , D for  $\hat{f}_4$ , E for  $\hat{f}_7$ ). The series are grouped into the following categories: Real Output and Income (1–19), Employment and Labor Market (20–49), Housing Market (50–59), Orders and Inventories (60–71), Money and Credit Markets (72–79), Stock Market (80–83), Interest Rates and Bond Yields (84–100), Exchange Rates (101–105), Prices (106–125). The series numbers for specific series are given in the Online Appendix. The data span period 1960Q1–2011Q4.

Figure 2: Predicted vs actual excess returns on Fama Maturity Portfolios



Notes: Figure 2 plots average excess returns predicted by the model against actual average excess returns on 10 Fama Maturity Portfolios (in % per quarter) over the period 1975Q1–2011Q4. We report as well the goodness-of-fit measure  $R^2$ .

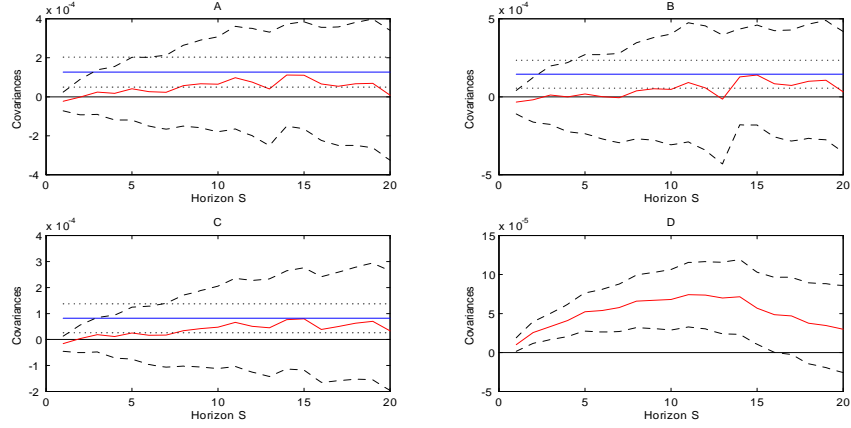
Figure 3: Covariances vs actual excess returns for Fama Maturity Portfolios



Notes: Figure 3 plots the covariances of excess returns on 10 Fama Maturity Portfolios with consumption growth  $Cov(R_{t+1}^e, \Delta c_{t+1})$  and with innovations to expectations in future consumption growth  $Cov(R_{t+1}^e, \epsilon_{c,t+1})$  against the actual average excess returns on these portfolios. We plot as well two standard error bands for covariances, calculated using Newey-West standard errors with 6 lags. The data span the period 1975Q1–2011Q4.

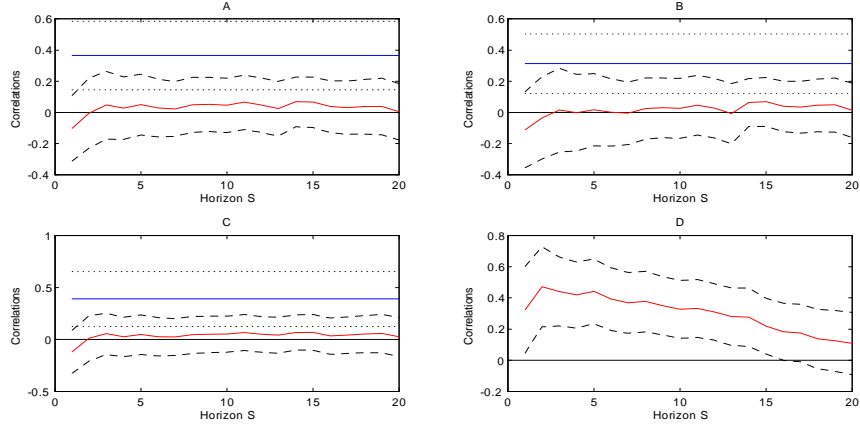


Figure 4: Covariances with consumption growth over different horizons.



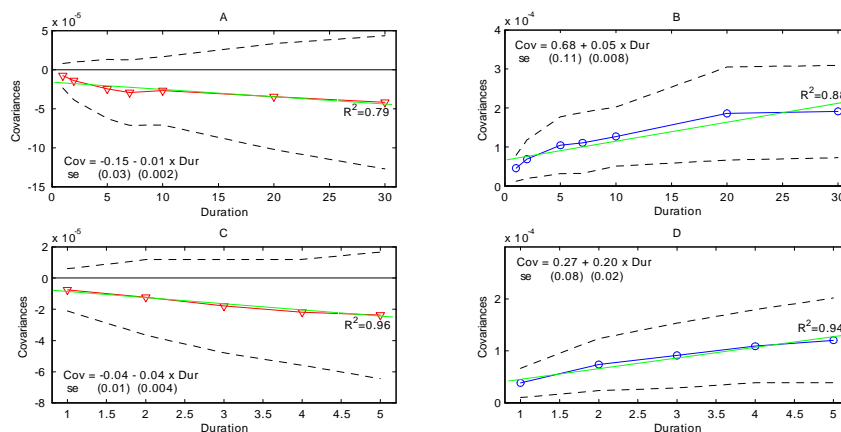
Notes: Figure 4 plots the covariances of consumption growth over  $S$  quarters and of innovations to expectations in future consumption growth with returns on High-Minus-Low (HML) bond factors for three groups of bond portfolios, along with two standard error bands. Respectively  $Cov(\ln C_{t+S}/C_t, R_{HML,t+1})$  are plotted in red line with relevant bands in dashed lines and  $Cov(\varepsilon_{c,t+1}, R_{HML,t+1})$  is plotted in blue horizontal line with relevant bands in dotted lines. The three groups of bond portfolios are: Fama Maturity Portfolios (A), Fixed Term Indices (B) and Fama-Bliss Discount Bonds (C). The returns on HML bond factors  $R_{HML}$  are computed as follows:  $R_{FMP10} - R_{FMP1}$  for Fama Maturity Portfolios,  $R_{FI30} - R_{FI1}$  for Fixed Term Indices and  $R_{FB5} - R_{FB1}$  for Fama-Bliss Discount Bonds. In graph D we plot the following covariances  $Cov(\ln C_{t+S}/C_t, \varepsilon_{c,t+1})$ . Two standard error bands are calculated using Newey-West standard errors with 6 lags. The data span the period 1975Q1–2011Q4.

Figure 5: Correlations with consumption growth over different horizons.



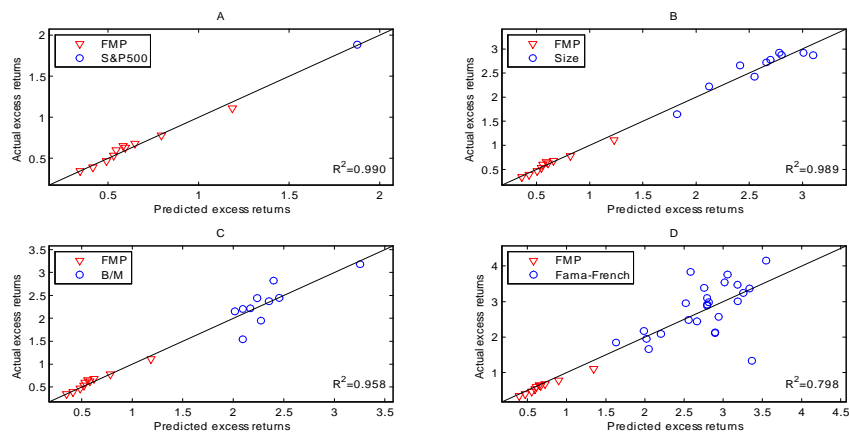
Notes: Figure 5 plots the correlations of consumption growth over  $S$  quarters and of innovations to expectations in future consumption growth with returns on High-Minus-Low (HML) bond factors for three groups of bond portfolios, along with two standard error bands. Respectively  $Corr(\ln C_{t+S}/C_t, R_{HML,t+1})$  are plotted in red line with relevant bands in dashed lines and  $Corr(\varepsilon_{c,t+1}, R_{HML,t+1})$  is plotted in blue horizontal line with relevant bands in dotted lines. The three groups of bond portfolios are: Fama Maturity Portfolios (A), Fixed Term Indices (B) and Fama-Bliss Discount Bonds (C). The returns on HML bond factors  $R_{HML}$  are computed as follows:  $R_{FMP10} - R_{FMP1}$  for Fama Maturity Portfolios,  $R_{FI30} - R_{FI1}$  for Fixed Term Indices and  $R_{FB5} - R_{FB1}$  for Fama-Bliss Discount Bonds. In graph D we plot the following correlations  $Corr(\ln C_{t+S}/C_t, \varepsilon_{c,t+1})$ . Two standard error bands are calculated using Newey-West standard errors with 6 lags. The data span the period 1975Q1–2011Q4.

Figure 6: Consumption-based risk measures vs duration



Notes: Figure 6 plots the covariances of excess returns on 7 Fixed Term Indices with consumption growth ( $Cov(R_{t+1}^e, \Delta c_{t+1})$ , graph A) and with innovations to expectations in future consumption growth ( $Cov(R_{t+1}^e, \varepsilon_{c,t+1})$ , graph B) and the covariances of excess returns on 5 Fama-Bliss Discount Bonds with consumption growth (graph C) and with innovations to expectations in future consumption growth (graph D) against duration. We plot as well two standard error bands for covariances, calculated using Newey-West standard errors with 6 lags, the linear trend line (in green) along with the relevant estimated equation of the form  $Cov_i = \alpha_0 + \alpha_1 D + \varepsilon_i$ , OLS standard errors (in parentheses) and goodness-of-fit-measure  $R^2$ . The data span the period 1975Q1–2011Q4.

Figure 7: Predicted vs actual excess returns on Fama Maturity Portfolios combined with equity portfolios



Notes: Figure 7 plots average excess returns predicted by the model against actual average excess returns (in % per quarter) for the following groups of assets: 10 Fama Maturity Portfolios + S&P500 Index (A), 10 Fama Maturity Portfolios + 10 Size portfolios (B), 10 Fama Maturity Portfolios + 10 Boon-to-Market portfolios (C), 10 Fama Maturity Portfolios + 25 Fama-French portfolios (D). In each figure bond portfolios are denoted as red triangles, while equity portfolios are denoted as blue circles. We report as well the goodness-of-fit measure  $R^2$ . The data span the period 1975Q1–2011Q4.