THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2007/32

Stochastic Volatility in a Macro-Finance Model of the US Term Structure of Interest Rates 1961-2004.

By

Peter Spencer

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Stochastic Volatility in a Macro-Finance Model of the US Term Structure of Interest Rates 1961-2004.

Peter Spencer*

University of York

*Department of Economics and Related Studies; ps35@york.ac.uk. I am grateful to Karim Abadir, Hans Dewachter, John Hutton, Peter Phillips, Mark Salmon, Peter Smith, Andy Tremayne, Mike Wickens, Zhuoshi Liu and participants in the Econometric Society European Meetings 2006 for useful comments. I also thank the referees of this journal for their many helpful suggestions.

Abstract

This paper generalizes the standard homoscedastic macro-finance model by allowing for stochastic volatility, using the 'square root' specification of the mainstream-finance literature. Empirically, this specification dominates the standard model because it is consistent with the square root volatility found in macroeconomic time series. Thus it establishes an important connection between the stochastic volatility of the mainstream finance model and macroeconomic volatility of the Okun (1971) - Friedman (1977) type. This research opens the way to a richer specification of both macroeconomic and term structure models, incorporating the best features of both macro-finance and mainstream-finance models.

1 Introduction

This paper develops a general affine macro-finance model of the US macroeconomy and the Treasury bond market. As the name suggests the macro-finance approach allows bond yields to reflect macroeconomic variables as well as latent variables representing financial market factors. It is based on the 'central bank model' (CBM) developed by Svensson (1999), Rudebusch (2002), Smets (1999), Kozicki and Tinsley (2005) and others, which represents the behavior of the macroeconomy in terms of the output gap (g_t), inflation (π_t) and the short term interest rate (r_t). The model developed in this paper allows bond yields to reflect changes in macroeconomic volatility related to the underlying rate of inflation.

In turn, the behavior of bond yields helps inform the specification of the macroeconomy, yielding new insights into the operation of monetary policy. In particular, early macro-finance studies showed that although macroeconomic variables provide a good description of the behavior of short rates they do not provide an adequate description of long term yields (Kozicki and Tinsley (2001), Ang and Piazzesi (2003)).

This finding has spawned an important macroeconometric literature which augments the CBM with latent variables, capturing exogenous changes in inflation and interest rates (Kozicki and Tinsley (2005) provides a useful summary). This literature shows that these rates are characterized by a non-stationary common trend (or unit root) that seems to be explained by the underlying rate of inflation. It follows the standard macroeconometric literature in assuming a homoscedastic (fixed) variance structure. This situation is familiar to macroeconomic modelers but poses a potential problem for term structure researchers: it is well-known that asymptotic (long maturity) yields are not properly defined if the interest rate is driven by a random walk (a homoscedastic unit root process).

This theoretical problem was first raised as an empirical issue by Dewachter and Lyrio (2006), but with this notable exception, macro-finance modelers have avoided it by assuming that the underlying inflation variable follows a near-unit root process (Ang and Piazzesi (2003), Rudebusch and Wu (2003), Dewachter, Lyrio, and Maes (2006)). However, because this variable is stationary, it mean-reverts to a constant rather than the variable end-point suggested by unit root macroeconomic models. As Kozicki and Tinsley (2005), Dewachter and Lyrio (2006) and others note, this means that it cannot be interpreted as a long run inflation expectation because it is anchored to a constant that cannot be influenced by monetary policy.

Mainstream finance yield curve research avoids these problems by using heteroscedastic (stochastic volatility) interest rate models based on Cox, Ingersoll, and Ross (1985). I modify their continuous time specification for use with discrete time macroeconomic data, getting a sensible forward rate asymptote without placing constraints on the roots of system. The stochastic trend is estimated using the Extended Kalman Filter, which is also standard in the mainstream finance literature. Model restrictions allow the stochastic trend-volatility term to be interpreted as an inflation trend, consistent with the hypothesis that macroeconomic volatility is influenced by the underlying rate of inflation (Okun (1971), Friedman (1977), Engel (1982))[1]. This specification encompasses the standard macro-finance model, which is decisively rejected by the data.

The research described in this paper was initially motivated by my interest in the asymptotic yield problem raised by early drafts of the Dewachter and Lyrio (2006) paper. I expected the new specification to outperform the standard one in explaining long maturity yields and with this in mind I extended the conventional (10-year maximum) maturity data set to include a 15 year yield, the longest available historically. However, the results are surprising in this respect. The new specification does give a dramatic improvement in fit, but the main reason for this is the importance of the Okun-Friedman heteroscedasticity effect found in the macro data. Once this is allowed for, its more flexible yield curve specification adds very little. This finding suggests that this new macro-finance framework - which uses estimates of macroeconomic volatility to inform the stochastic volatility parameters of the term structure model - should be better at discriminating between rival models than the mainstream finance one (Chen and Scott (1993), Dai and Singleton (2002)), which does not. It further suggests that CBM-based studies of monetary policy should use the heteroscedastic framework rather than the current homoscedastic one, allowing the effects of the stochastic trend on the second as well as the first moments of the system to be analyzed.

The paper is set out as follows. The next section describes the macroeconomic model and its stochastic structure, supported by appendix 1. Section 3 derives the

[1] Ball (1992) offers a theoretical analysis of this phenomenon and the empirical evidence is examined by Brunner and Hess (1993), Holland (1995), Caporale and McKiernan (1997) and others. There is also an extensive literature on the effect of inflation and macroeconomic volatility on the equity risk premium (Brandt and Wang (2003), Lettau, Ludvigson, and Wachter (2006)).

bond pricing model, supported by appendix 2. It discusses the theoretical problems posed by the unit root in the standard specification and shows how these are avoided in the general affine specification. The two respective empirical models are compared in Section 4. Section 5 offers a brief conclusion and suggestions for further research.

2 The general affine macro-model framework

This section specifies the macroeconomic framework. This is 'general affine' or 'exponential-affine' in the sense of (Duffie and Kan (1996), Duffie, Filipovic, and Schachtermayer (2003), to be explained). It consists of a heteroscedastic macroeconomic Vector Autoregression (VAR) augmented by two latent variables, which is specified under the physical (or observed) probability measure $\mathcal{P}$. The yield model is specified under the risk neutral measure $\mathcal{Q}$ in the next section.

2.1 The macroeconomic dynamics

The macro-model is based on the CBM. It represents the behavior of the macroeconomy in terms of the inflation rate (π_t), output gap (g_t) and the 3 month Treasury Bill rate ($r_{1,t}$). These are part of an $n-$vector z_t of macroeconomic variables driven by the difference equation system:

$$z_t = K + \Phi_0 y_t + \Sigma_{l=1}^{L} \Phi_l z_{t-l} + G\eta_t \tag{1}$$

where G is a lower triangular matrix, η_t is an $n-$vector of $i.i.d$ orthogonal errors and y_t is a $k-$vector of latent factors. These follow the first order process:

$$y_t = \theta + \Xi y_{t-1} + \varepsilon_t \tag{2}$$

where ε_t is an $k-$vector of $i.i.d$ orthogonal errors, $\theta = \{\theta_{1,...,}\theta_k\}'$ and $\Xi = Diag\{\xi_{1,...,}\xi_k\}$[2]. It is assumed that z_t is observed without measurement error and that y_t is unobservable. I estimate y_t using the Extended Kalman Filter (Harvey (1989), Duffee and Stanton (2004)) as described in appendix 3 .

The specific model developed in this paper defines $z_t = \{g_t, \pi_t, r_{1,t}\}'$, $y_t = \{y_{\pi^*,t}, y_{\mathbf{r}^*,t}\}'$, $\varepsilon_t = \{\varepsilon_{\pi^*,t}, \varepsilon_{\mathbf{r}^*,t}\}'$, $\xi = \{\xi_{\pi^*}, \xi_{\mathbf{r}^*}\}$ and $\theta = \{\theta_{\pi^*}, \theta_{\mathbf{r}^*}\}'$. In my preferred model, $y_{\pi^*,t}$ is a martingale driving the inflation asymptote: $\pi_t^* = y_{\pi^*,t} + \varphi_{\pi^*}$, where φ_{π^*} is a shift constant. The central tendency $\mathbf{r}^*$ of the real interest rate $\mathbf{r}$ is $y_{\mathbf{r}^*,t}$ plus the constant $\varphi_{\mathbf{r}^*}$. $y_{\mathbf{r}^*,t}$ is assumed to be a mean reverting Gaussian variable ($\xi_2 = \xi_{\mathbf{r}^*}, |\xi_{\mathbf{r}^*}| < 1$), so $\varphi_{\mathbf{r}^*}$ plays an identical role to $\theta_{\mathbf{r}^*}$, which it is convenient to set to zero, making $\varphi_{\mathbf{r}^*}$ the long run mean. The central tendency of the nominal interest rate is thus $r_t^* = y_{\pi^*,t} + y_{\mathbf{r}^*,t} + \varphi_{\pi^*} + \varphi_{\mathbf{r}^*}$, which reverts to the asymptote $r_t^{**} = y_{\pi^*,t} + \varphi_{\pi^*} + \varphi_{\mathbf{r}^*}$. The output gap is also assumed to be a zero-mean reverting variable: $g_t^* = 0$. These equilibrium conditions are enforced by imposing a set of restrictions on (1):

$$\Phi_0 = (I - \Sigma_{l=1}^{L}\Phi_l)R; \quad K = \Phi_0\varphi; \tag{3}$$

$$\text{where}: \quad \varphi' = \{\varphi_{\pi^*}, \varphi_{\mathbf{r}^*}\}; \quad R = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

to give the equilibrium solution $z_t^* = (I - \Sigma_{l=1}^{L}\Phi_l)^{-1}\Phi_0(y_t + \varphi) = R(y_t + \varphi)$.

This system can be consolidated by defining $x_t = \{y_t', z_t'\}'; v_t = \{\varepsilon_t', \eta_t'\}'$ and combining (1) and (2) to get an $L-$th order difference system described in appendix 1 as (23). The yield model employs the state space form (Harvey (1989)), obtained by arranging this as first order difference system describing the dynamics of the state

[2] In this paper, $Diag\{\delta\}$ represents a matrix with the elements of the row vector δ in the main diagonal and zeros elsewhere. 0_a is the $(a \times 1) \times 1$ zero vector; 1_a is the $(a \times 1) \times 1$ summation vector; $0_{a,b}$ the $(a \times b)$ zero matrix; and I_a the a^2 identity matrix.

vector:

$$X_t = \Theta + \Phi X_{t-1} + W_t \tag{4}$$

where $X_t = \{y_t', z_t', ..., z_{t-l}'\}'$ is the state vector, $W_t = C.\{\varepsilon_t', \eta_t', 0_{1,N-k-n}\}'$ and Θ, Φ & C are defined in appendix 1. X_t has dimension $N = k + nl$.

The macroeconomic data were provided by Datastream and are shown in chart 1. π_t is the annual CPI inflation rate and $r_{1,t}$ the 3 month Treasury Bill rate. The output gap series g_t is the quarterly OECD measure, derived from a Hodrick-Prescott filter. The yield data were taken from McCulloch and Kwon (1991), updated by the New York Federal Reserve Bank[3]. These have been extensively used in the empirical literature on the yield curve. To represent this curve I use 1,2,3,5,7,10 and 15 year maturities. These yield data are available on a monthly basis, but the macroeconomic data dictated a quarterly time frame (1961Q4-2004Q1, a total of 170 periods). The quarterly yield data are shown in chart 2. The 15 year yield is shown at the back of the chart, while the shorter maturity yields are shown at the front.

These inflation and interest rates all exhibit a high degree of persistence, which could be the effect of slow mean reversion, unit roots or of structural breaks. Table 1 shows the means, standard deviations and first order autocorrelation coefficients of these data, as well as KPSS and ADF test results. The ADF tests show that the null hypothesis of non-stationarity for these variables cannot be rejected at the 5% level. The KPSS (Kwiatowski, Phillips, Schmidt, and Shin (1992)) statistics for inflation and 1-10 year interest rates are only significant at the 10% level, suggesting that the null hypothesis of stationarity may just be acceptable. However, the KPSS statistic for the 15 year rate is almost significant at the 5% level. Moreover, Fama (2006) argues persuasively that the long upswing and downswing in rates evident in the

[3] I am grateful to Tony Rodrigues of the New York Fed for supplying a copy of this yield dataset.

charts was the result of a succession of permanent shocks that were on balance positive until 1981 and negative thereafter. In this paper, I follow Kozicki and Tinsley (2005), Dewachter and Lyrio (2006) and Fama (2006) in analyzing a macroeconometric model characterized by a unit root.

2.2 The stochastic structure

The standard macro-finance model assumes that the volatility structure is homoscedastic and Gaussian: $W_t \sim N(0_N, \Omega)$, while mainstream finance models usually assume that volatility is stochastic, driven by square root processes in one or more of the state variables[4]. Dai and Singleton (2000) derive 'admissibility' conditions to ensure that these state variables remain non-negative and the variances are well-defined. They classify an admissible model with N state variables and m independent square root factors conditioning volatility is classed as $\mathbb{A}_m(N)$. Thus the standard macro-finance model (which is homoscedastic) is classified as $\mathbb{A}_0(N)$ and the mainstream specification (with a single stochastic volatility term) as $\mathbb{A}_1(N)$.

This paper develops a model that encompasses the $\mathbb{A}_0(N)$ and $\mathbb{A}_1(N)$ specifications. These models all generate affine yield curves because the probability distributions underpinning them are all 'exponential-affine' in the sense of Duffie, Filipovic, and Schachtermayer (2003). They define a process as exponential-affine under any measure $\mathcal{M}$ if the conditional Moment Generating Function ($L^{\mathcal{M}}[\nu,\, X_t] = E^{\mathcal{M}}[\exp[\nu' X_{t+1}] \mid X_t]$) for X_{t+1}, is an exponential-affine (loglinear) function of X_t. E^M denotes the expectation under the measure $\mathcal{M}$ while E & V denote the mean & variance under the state price density and ν is a vector of Laplace coefficients. The Moment Generating Function (MGF) is the Laplace Transform of the density

[4] Preliminary tests showed no significant evidence of Autoregressive Conditional Heteroscedasticity (ARCH) in this data set.

of X_{t+1}. For example $\mathbb{A}_0(N)$ assumes that $\varepsilon_{1,t}$ is normally distributed with mean zero and standard deviation δ_{01} and we use the formula for the MGF of a normal variable:

$$E[\exp[\nu.y_{\pi^*,t+1}|y_{\pi^*,t}]] = \exp[\nu(\mu + \xi_{\pi^*}y_{\pi^*,t}) + \frac{1}{2}\nu^2\delta_{01}]. \tag{5}$$

In $\mathbb{A}_1(N)$, this latent variable also drives volatility through a square root process similar to the diffusion for the spot interest rate in CIR (1985). They show that in discrete time, this has a non-central conditional χ^2 distribution. If we normalize the time interval $(s-t)$ in their equation (18) as unity and replace r by y_{π^*} :

$$y_{\pi^*,t+1} \sim \chi^2[2cy_{\pi^*,t+1}; 2c\mu; 2\xi_{\pi^*}cy_{\pi^*,t}] \tag{6}$$

where $2c$ is the scale factor, $2\xi_{\pi^*}cy_{\pi^*,t}$ is the non-centrality parameter and $2c\mu$ shows the degrees of freedom.

This process is exponential-affine because its conditional Moment Generating Function (MGF) is a log-linear function of $y_{\pi^*,t}$:

$$E[\exp[\nu.y_{\pi^*,t+1}|y_{\pi^*,t}]] = \exp[\frac{\nu\xi_{\pi^*}y_{\pi^*,t}}{1-\nu/c} - c\mu\ln[1-\frac{\nu}{c}]] \tag{7}$$

(provided that: $\nu < c$, Johnson and Kotz (1970)). Differentiating (7) w.r.t. ν once, twice and then setting ν to zero gives the conditional mean and variance:

$$E[y_{\pi^*,t+1}|y_{\pi^*,t}] = \mu + \xi_{\pi^*}y_{\pi^*,t}; \quad V[y_{\pi^*,t+1}|y_{\pi^*,t}] = \delta_{01} + \delta_{11}y_{\pi^*,t}; \tag{8}$$
$$\text{where :} \quad \delta_{01} = \mu/c, \delta_{11} = 2\xi_{\pi^*}/c.$$

In the limiting case of a unit root, the degree of freedom parameter is zero. This model is studied by Seigel (1979) and his basic results are reported in Chapter 29 of Johnson and Kotz (1970). Important results have also been obtained for this case by Gourieroux and Jasiak (2002). In this limit: $\xi_{\pi^*} = 1$; $\delta_{01} = \mu = 0$; $c = 2/\delta_{11}$, and (8) simplifies to:

$$E[y_{\pi^*,t+1}|y_{\pi^*,t}] = y_{\pi^*,t}; \quad V[y_{\pi^*,t+1}|y_{\pi^*,t}] = \delta_{11}y_{\pi^*,t}. \tag{9}$$

This process is a martingale: the expectation of any future value is equal to the current value. However, unlike the random walk model, the error variance is also proportional to this value. These models can all be represented as: $y_{\pi^*,t+1} = \theta_{\pi^*} + \xi_{\pi^*}y_{\pi^*,t} + \varepsilon_{\pi^*,t+1}$. To be consistent with (2) we set the intercept θ_{π^*} equal to μ for the models (5) and (8) and to zero in (9).

In $\mathbb{A}_1(N)$, this stochastic trend also conditions the volatility of the other variables. It is ordered as $x_{1,t} = y_{\pi^*,t}$, the first variable in the $(k+n)$ vector x_t. The other contemporaneous variables are put into an $(k+n-1)$ vector $x_{2,t}$, so that: $x_t = \{x_{1,t}, x'_{2,t}\}'$ and conformably: $v_t = \{w_{1,t}, v'_{2,t}\}'$ and $w_t = \{w_{1,t}, w'_{2,t}\}'$, where $w_{1,t} = \varepsilon_{\pi^*,t}$. Similarly, writing $X_t = \{x_{1,t}, X'_{2,t}\}'$ and partitioning W_t, Θ, Φ, C conformably (see appendix 1), (4) becomes:

$$\begin{bmatrix} x_{1,t+1} \\ X_{2,t+1} \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \Theta_2 \end{bmatrix} + \begin{bmatrix} \xi_1 & 0'_{N-1} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ X_{2,t} \end{bmatrix} + \begin{bmatrix} w_{1,t+1} \\ W_{2,t+1} \end{bmatrix} \tag{10}$$

where $\theta_1 = \theta_{\pi^*}$, $\xi_1 = \xi_{\pi^*}$ and $w_{1,t+1} = \varepsilon_{\pi^*,t+1}$. In this paper subscripts 1 and 2 denote partitions of N (or $k+n$) dimensional vectors and matrices into 1 and $N-1$ (or $k+n-1$). The stochastic structure for (10) is described in appendix 1. The distribution of $x_{2,t}$ and $X_{2,t}$ conditional upon $x_{1,t-1}$ is assumed to be Gaussian. The conditional covariance of $X_{2,t}$ is $\Sigma_0 + \Sigma_1 x_{1,t-1}$, where: $\Sigma_i = C_{22}\Delta_i C'_{22}$ and $\Delta_i =$

$Diag\{\{\delta_{i2}, ..., \delta_{i(k+n)}\}, 0'_{N-k-n}\}$; $i = 0, 1$. C_{22} is a lower triangular and $\Delta_i; i = 0, 1$ are deficient diagonal $(N-1)^2$ matrices.

This model is the discrete time analogue of the 'general affine' $\mathbb{A}_1(N)$ model developed by Duffie and Kan (1996), which generalizes the CIR model by adding a translation or shift constant to a model variable like an interest rate or stochastic trend when defining the volatility term in the rate diffusion. However, it is convenient to allow for this shift by using φ_{π^*} in (3) instead, keeping y_{π^*} a CIR process but making the inflation asymptote π^* a Duffie and Kan process. Importantly, the Duffie and Kan $\mathbb{A}_1(N)$ model encompasses $\mathbb{A}_0(N)$. Similarly, the discrete time $\mathbb{A}_0(N)$ model is a special case of my $\mathbb{A}_1(N)$ model, which can be obtained by taking the limit as c tends to infinity. Setting δ_{11} and Σ_1 to zero then makes $x_{1,t}$ and $x_{2,t}$ homoscedastic. This also renders the distribution of $y_{\pi^*,t+1}$ Gaussian, allowing this to be defined in the $\mathbb{A}_0(N)$ model as another zero-mean reverting latent variable $(\theta_1 = 0)$[5]. These models are 'admissible' in the sense (of Dai and Singleton (2000) and (2002)) that they ensure that the variance structure remains non-negative definite[6]. They can be used independently of the yield specification, as for example by (Kozicki and Tinsley (2005)). However my interest is to use them jointly with bond market data, using the macro-yield framework developed in the next section.

3 The general affine yield curve framework

This section is supported by appendix 2 and shows how exponential-affine MGFs can be used to model the yield curve consistently with the macro models of the previous section. It is based on the fact (Duffie, Filipovic, and Schachtermayer (2003)) that the MGF of a distribution that is exponential-affine under the risk neutral measure

[5] This parameter is equivalent to φ_1 in (3) in this case.

[6] In $\mathbb{A}_1(N)$ the variable driving volatility $x_{1,t}$ has a non-central χ^2 distribution and is non-negative, keeping the variance structure $\Sigma_0 + \Sigma_1 x_{1,t-1}$ for $X_{2,t}$ non-negative definite.

$\mathcal{Q}$ generates an exponential-affine discount bond price model:

$$P_{\tau,t} = \exp[-\gamma_\tau - \Psi'_\tau \; X_t]; \;\; \tau = 1, ..., M. \tag{11}$$

The natural logarithm of this price is denoted by $p_{\tau,t}$ and is linear in X_t. Reversing sign and dividing by maturity τ gives the discount yields: $r_{\tau,t} = -p_{\tau,t}/\tau = \alpha_\tau + \beta'_\tau X_t$; where: $\alpha_\tau = \gamma_\tau/\tau$; $\beta_\tau = \Psi_\tau/\tau$. The slope coefficients β_τ are known as 'factor loadings.' Stacking the yield equations for $\tau = 4, 8, 12, 20, 28, 40$ & 60 quarters and adding an error vector e_t gives the multivariate regression model used in section 4 for the (7×1) vector r_t :

$$r_t = \alpha + B'X_t + e_t = \alpha + B'_0 y_t + \Sigma_{l=1}^{L} B'_l z_{t+1-l} + e_t \tag{12}$$

$$\text{where : } e_t \sim N(0, \bar{P}); \bar{P} = Diag\{\rho_1, \rho_2, ..., \rho_7\}.$$

and e_t is an i.i.d.error vector.[7] This is the affine yield curve framework used in this paper. The remainder of this section discusses the measure $\mathcal{Q}$ used for asset pricing and the 'essentially affine' yield models $\mathbb{E}\mathbb{A}_0(N)$ and $\mathbb{E}\mathbb{A}_0(N)$ corresponding to the macromodels $\mathbb{A}_0(N)$ and $\mathbb{A}_0(N)$ of the previous section.

3.1 The risk neutral measure Q

Assets are priced under the risk neutral measure, which adjusts the state probabilities in such a way that they all have the same expected return. These adjustments depend upon 'price of risk' parameters that show the effect of model variables on the risk premia. For the yield model to be affine these prices must also be affine in the state variables. So for example, the variable $\lambda_{1,t}$ shows the price of risk associated with

[7] The usual conventional in macro-finance models is that this represents measurement error.

the stochastic trend, which plays an important role in this analysis:

$$\lambda_{1,t} = \lambda_{10} + \lambda_{11} x_{1,t} + \Lambda_{12} X_{2,t}. \tag{13}$$

If this is zero, then an 'asymptotic' or 'end-point' portfolio that is constructed so that it is only exposed to shocks in $x_{1,t}$ has a zero risk premium and is expected to earn the spot rate. If it is constant ($\lambda_{1,t} = \lambda_{10}$), then variations in this risk premium depend only upon variations in volatility, such as those induced by $x_{1,t}$ in $\mathbb{A}_1(N)$. This parameter plays the key role in that model. If λ_{11} is also non-zero then the trend can influence the risk premia thorough variations in the price of risk, even if volatility is fixed as in $\mathbb{A}_0(N)$, so λ_{11} plays the key role in that model. Appendix 2 shows how the prices of risk associated with the other variables are adjusted, following Duffee (2002). After this modification, the models $\mathbb{A}_0(N)$ and $\mathbb{A}_1(N)$ are classified as the 'essentially affine' models $\mathbb{EA}_0(N)$ and $\mathbb{EA}_1(N)$, respectively represented by the empirical models M0 and M1 in section 4. Following Dewachter and Lyrio (2006), both models incorporate the restrictions: $\xi_1 = 1; \Lambda_{12} = 0_{N-1}$.

As appendix 2 explains, variations in $\lambda_{1,t}$ are a nuisance in the $\mathbb{EA}_1(N)$ model M1 and the conventional assumption is that $\lambda_{1,t} = \lambda_{10}$ for that model. On the other hand, it is important to allow $\lambda_{1,t}$ to reflect variations in $x_{1,t}$ in M0, so in this case $\lambda_{1,t} = \lambda_{10} + \lambda_{11} x_{1,t}$. This means that M1 does not encompass M0. However, an encompassing specification that nests both models can be constructed by relaxing the usual macro-finance assumption that bond market participants use the true value of ξ_1 in pricing. This gives my 'baseline' model M2, which uses a new parameter $\xi_1^{\mathcal{B}}$ to describe their estimate of the speed of adjustment of $x_{1,t}$ under $\mathcal{P}$, which may differ from the parameter ξ_1 defined in the macro-model[8]. This modification is a technical

[8] A shift in the parameters from (4) to (14) could occur either because of the risk adjustment

one, designed to allow encompassing tests, but could detect an error in the market estimate of ξ_1. Appendix 2 derives the MGF of M2 under measure $\mathcal{Q}$ and shows how it can be used as a 'model generating function' to derive the other empirical models. It can for instance be used as a moment generating function to give the dynamics of this system under $\mathcal{Q}$:

$$\begin{bmatrix} x_{1,t+1} \\ X_{2,t+1} \end{bmatrix} = \begin{bmatrix} \theta_1^{\mathcal{Q}} \\ \Theta_2^{\mathcal{Q}} \end{bmatrix} + \begin{bmatrix} \xi_1^{\mathcal{Q}} & 0'_{N-1} \\ \Phi_{21}^{\mathcal{Q}} & \Phi_{22}^{\mathcal{Q}} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ X_{2,t} \end{bmatrix} + \begin{bmatrix} w_{1,t+1}^{\mathcal{Q}} \\ W_{2,t+1}^{\mathcal{Q}} \end{bmatrix} \tag{14}$$

where the time t expectations of the error terms are zero under $\mathcal{Q}$. Table 2 shows parameter values for models M0-M2, where:

$$H_0 = \Sigma_0\Lambda_{20} + \delta_{01}\lambda_{10}C_{21}; H_1 = \Sigma_0\Lambda_{20}; \Upsilon_0 = \Lambda_{21} + \delta_{01}\lambda_{11}C_{21}; \Upsilon_1 = \Lambda_{21} + \Sigma_1\Lambda_{20}. \tag{15}$$

These reduced form parameters show the effects of the coefficients λ_{10} and λ_{11} modelling the price of risk associated with $x_{1,t}$. The vectors Λ_{20} & Λ_{21} and the matrix Λ_{22} are defined in (37) and model the price of risk associated with $x_{2,t}$. Obviously, if all of these parameters are set to zero, the parameters revert to those shown (for $\mathcal{P}$) in the first column. The parameter values for model M2 are shown in the second column. Those for M1 in the next column assume the market uses the true value of $\xi_1^{\mathcal{B}} = \xi_1$. The parameters for M0 in the final column are standard and appendix 2 shows that they can be derived from those for M2 by setting δ_{11} & Σ_1 to zero and taking the limit as c tends to infinity. In this case we replace $\xi_1^{\mathcal{B}}$ by $\xi_1^{\mathcal{Q}_0} = \xi_1 - \delta_{01}$ λ_{11}, which is the Duffee (2002) risk-adjusted parameter.

implied by (36) or because the bond market does not use the true values of the parameters in (4). However, it is not possible to distinguish these two effects without imposing restrictions on (36).

3.2 The $EA_1(N)$ yield specifications (M1 and M2)

The coefficients of (11) are partitioned Ψ_τ conformably with (10) as $\{\psi_{1,\tau}, \Psi'_{2,\tau}\}'$. They are recursive in maturity. Since $-p_{\tau,1} = r_{1,t}$ its coefficients have the starting values: $\gamma_1 = \psi_{1,1} = 0$; and $\Psi_{2,1} = J_{2,r}$, where $J_{2,r}$ is a selection vector such that $J'_{2,r}X_{2,t} = r_{1,t}$. It is also recursive in the sense that $\Psi_{2,\tau}$ does not depend upon $\psi_{1,\tau-1}$ (or $\gamma_{\tau-1}$) :

$$\begin{aligned}\Psi_{2,\tau} &= (\Phi_{22}^{\mathcal{Q}})'\Psi_{2,\tau-1} + J_{2,r} \\ &= (I - (\Phi_{22}^{\mathcal{Q}})')^{-1}(I - ((\Phi_{22}^{\mathcal{Q}})')^{\tau})J_{2,r}\end{aligned} \tag{16}$$

where $\Phi_{22}^{\mathcal{Q}} = \Phi_{22} - \Lambda_{22}$ (defined in table 2) adjusts Φ_{22} to allow for the effect of $X_{2,t}$ on the associated prices of risk. I assume that the roots of this sub-system are stable under $\mathcal{Q}$, so this has the asymptote:

$$\Psi_2^* = lim_{\tau\to\infty}\Psi_{2,\tau} = (I - (\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r} \tag{17}$$

Dividing by τ and taking the limit as this goes to infinity gives the limit for β_2 shown in table 3. This Gaussian sub-structure is common to all models, but the structure of the remaining coefficients depends critically upon the model specification, particularly if the system is non-stationary.

Importantly, as Dai and Singleton (2002) and others note, non-stationarity under $\mathcal{Q}$ is not a problem in the EA_1 specification. Indeed their results, like previous mainstream $\mathbb{EA}_1(N)$ estimates (Chen and Scott (1993)) suggest that there is a root which is significantly larger than unity under $\mathcal{Q}$. The volatility of $x_{1,t+1}$ is linear in $x_{1,t}$ and so the equation determining its price coefficient includes a non-linear Jensen effect. For M2:

$$\psi_{1,\tau} = \frac{\xi_1^{\mathcal{B}}[\psi_{1,\tau-1} + \lambda_{10} + \Psi'_{2,\tau-1}C_{21}]}{1 + [\psi_{1,\tau-1} + \lambda_{10} + \Psi'_{2,\tau-1}C_{21}]/c} - \frac{\xi_1^{\mathcal{B}}\lambda_{10}}{1 + \lambda_{10}/c} - \Psi'_{2,\tau-1}\Upsilon_1 - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_1\Psi_{2,\tau-1}. \tag{18}$$

For a regular solution:

$$\lambda_{10} + c > 0; \quad [\psi_{1,\tau-1} + \lambda_{10} + \Psi'_{2,\tau-1}C_{21}] + c > 0. \tag{19}$$

As Campbell, Lo, and MacKinlay (1996) note in a similar heteroscedastic yield curve model[9], the price parameter ψ_1^* is determined by a quadratic rather than a linear equation and is well-defined (with $\beta_1^* = 0$) even if $\left|\xi_1^{\mathcal{Q}_1}\right| \geq 1$[10]. For the intercept:

$$\Delta\gamma_\tau = (\Theta_2 - C_{21}\mu - H_1)'\Psi_{2,\tau-1} - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_0\Psi_{2,\tau-1} + c\mu\ln[\frac{c + \psi_{1,\tau-1} + \lambda_{10} + \Psi'_{2,\tau-1}C_{21}}{c + \lambda_{10}}]. \tag{20}$$

Unit roots are not a problem in $\mathbb{EA}_1(N)$. Indeed, they simplify the model structure, giving model M1 with $\xi_1^{\mathcal{B}} = \xi_1 = 1$. Substituting $\mu = 0$ into (20) 'switches off' the logarithmic term and makes γ_τ and hence the asymptotic forward rate f^* independent of ψ_1^*, as shown in appendix 2. In other words, because the volatility of $x_{1,t}$ is proportional to $x_{1,t-1}$, the associated Jensen effects are found in (18), but not in (20). This also simplifies the risk premium, derived in appendix 2 as (43).

[9] Their model uses a Gaussian approximation to the Cox, Ingersoll, and Ross (1985) process describing the spot rate, due originally to Sun (1992).

[10] Substituting (17) into (18) gives ψ_1^* as the solution to: $\psi_1^* = \frac{\xi_1(\psi_1^* + \lambda_{10} + \Psi_2^{*\prime}C_{21})}{1 + (\psi_1^* + \lambda_{10} + \Psi_2^{*\prime}C_{21})/c}$- $\frac{\xi_1\lambda_{10}}{1+\lambda_{10}/c}$- $\Psi_2^{*\prime}\Upsilon_1$- $\frac{1}{2}\Psi_2^{*\prime}\Sigma_1\Psi_2^*$. This may be arranged as: $0 = \vartheta^2 + \vartheta(c(1-\xi_1) - \zeta) - c\zeta$ where: $\vartheta = \psi_1^* + \lambda_{10} + \Psi_2^{*\prime}C_{21}$; and $\zeta = \frac{\lambda_{10}(1-\xi_1)+\lambda_{10}/c}{1+\lambda_{10}/c} + \Psi_2^{*\prime}(C_{21} - \Upsilon_1) - \frac{1}{2}\Psi_2^{*\prime}\Sigma_1\Psi_2^*$. The intercept term $c\zeta$ shows the product of the roots and is a very large negative number. Consequently, one root is a large negative and the other a large positive number. Phase analysis reveals that the recursion (18) selects the positive root.

3.3 The standard $EA_0(N)$ specification (M0)

The pricing formulae for this model are well-known and appendix 2 considers them as limits of these baseline M2 formulae. Taking the limit of (18) as c tends to zero gives a quadratic recursion. However, the restriction $\delta_{11} = 0$ 'switches off' this quadratic effect, reducing (18) to a linear recursion:

$$\psi_{1,\tau} = \xi_1^{\mathcal{Q}_0}\psi_{1,\tau-1} + \Psi'_{2,\tau-1}\Phi_{21}^{\mathcal{Q}_0}. \tag{21}$$

The intercepts do follow a quadratic recursion:

$$\begin{aligned}\Delta\gamma_\tau = \gamma_\tau - \gamma_{\tau-1} = \Psi'_{2,\tau-1}\Theta_2^{\mathcal{Q}_0} + \psi_{1,\tau-1}\theta_1^{\mathcal{Q}_0} - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_0\Psi_{2,\tau-1} \\ -\frac{1}{2}\delta_{01}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}]^2;\end{aligned} \tag{22}$$

with the parameters defined in table 2. In contrast to the M1 model, the intercept exhibits the nonlinear Jensen effects in M0 and not the first slope coefficient.

If $\xi_1^{\mathcal{Q}_0} = 1$, then clearly (21) has a unit root and as appendix 2 and table 3 show, the long forward rates behave like (-)τ^2 in the limit, reflecting the well-known asymptotic problem. In the specification of Dewachter and Lyrio (2006), x_1 has a unit root under $\mathcal{P}$ ($\mu = 0, \xi_1 = 1$) but is mean-reverting under $\mathcal{Q}$: ($\left|\xi_1^{\mathcal{Q}_0}\right| = |\xi_1 - \delta_{01}\lambda_{11}| <$ 1). This provides a neat way of avoiding the asymptotic yield problem while allowing the inflation asymptote in the macro-model to be a variable end-point as in (Kozicki and Tinsley (2001)). However, this restriction constrains the rate at which $\psi_{1,\tau}$ grows in the recursion (21), constraining the effect of the stochastic trend on short maturity yields. It also means that the associated factor risk premium, the excess return expected for holding the 'asymptotic' portfolio (constructed so that the portfolio weights sum to: $\psi_1 = 1$, $\Psi'_2 = 0_{N-1}$), is negatively related to the inflation trend, as appendix 2 demonstrates. Moreover, the $EA_1(N)$ model relaxes this restriction and my results, reflecting those of mainstream research (for example Chen and Scott (1993), Dai and Singleton (2002)), suggest that this root is significantly *greater* than unity.

4 The empirical models

The empirical model consists of a heteroscedastic VAR describing the 3 macroeconomic variables (4) and the associated equations describing the 7 representative yields (12). It is estimated by quasi-maximum likelihood and the Extended Kalman filter, which gives optimal linear estimates of the latent variables in this situation. The likelihood function is derived in appendix 3. The preliminary tests reported in section 2.1 indicated the presence of a unit root in the macroeconomic and yield data. Further (AIC) tests suggested a third order lag structure for (1). With $n = 3$, $k = 2$ and $l = 3$, there are $N = 11$ state variables. Consequently this research focussed on the $\mathbb{EA}_0(11)$ and $\mathbb{EA}_1(11)$ specifications. The baseline model M2 uses 66 parameters[11] and has a loglikelihood L(2)=747.5 as shown in table 4.

The $\mathbb{EA}_1(11)$ model M1 specializes this by assuming a unit root and maintaining the standard identity between macro and yield parameters under $\mathcal{P}$: $\xi_1^{\mathcal{B}} = \xi_1 = 1; \mu = 0$[12]. These 3 restrictions are easily accepted by the data: the $\chi^2(3)$ likelihood ratio test gives an acceptance value of p =0.97. The unit root $\mathbb{EA}_0(11)$ specification M0 is also nested in M2, employing 6 restrictions: δ_{11} and $\mathbf{\Delta}_1(4)$ are set to zero and again $\xi_1 = 1$. However, its loglikelihood of L(0)=694.3 is much lower than for the other models and it is overwhelmingly rejected against M2. This rejection is largely

[11] These are $\xi_1^{\mathcal{B}}$, ξ_1, ξ_2, λ_{10}, δ_{01}, δ_{11}, $\mathbf{\Delta}_0(4)$, $\mathbf{\Delta}_1(4)$, $H_1(4)$, $\Upsilon_1(3)$, $\Lambda_{22}(13)$, $G(3)$, $\Phi(27)$ and $\varphi(2)$. Estimates are reported in table 5. It was found that although $\lambda_{22,\mathbf{r}^*\mathbf{r}^*}$ was significant (table 5c) the remaining elements of the first rows of Λ_{22} and Υ_1 (or Υ_0) were very poorly determined and could be eliminated without significantly reducing the likelihood. The structural parameters μ and c follow from (8) given ξ_1, δ_{01} and δ_{11}.

[12] This restriction is imposed via (8) by setting $\delta_{01} = 0$.

due to the restriction $\delta_{11} = 0$. The effect of this is twofold: (a) it makes the stochastic trend homoscedastic and (b) it removes the nonlinearity from (18), reducing it to the linear recursion (21). Theoretically, these two effects are inextricably related because the parameters of the stochastic structure structure($\delta_{01}, \delta_{11}, \Delta_0$,$\Delta_1$ and C_{22}) are not affected by the change of measure and the conventional macro-finance assumption is that they are the same in the macro and yield models. However further relaxing this assumption allows (a) and (b) to be separated.

To explore this idea, I constructed two new models: M3 and M4. The 'encompassing' specification M4 nests all of the other models. It is based on M2 but uses 10 new parameters ($\delta_{01}^{\mathcal{B}}, \delta_{11}^{\mathcal{B}}, \Delta_0^{\mathcal{B}}$ and $\Delta_1^{\mathcal{B}}$) to replace their macro equivalents in the yield model. It has the loglikelihood value L(4) = 749.7[13]. As table 4 shows, model M1 is acceptable against this alternative, suggesting that the conventional macro-finance assumption is valid. Model M3 then specializes M4 by using the restrictions $\delta_{11}^{\mathcal{B}} = 0$ and $\Delta_0^{\mathcal{B}} = 0$. This is a hybrid in which the true macro-model is the heteroscedastic $\mathbb{A}_1(N)$ model, but the bond market mistakenly uses a best-fit $\mathbb{E}\mathbb{A}_0(N)$ specification instead of $\mathbb{E}\mathbb{A}_1(N)$. Comparing its loglikelihood L(3) with that L(0) of M0 which it nests, gives an estimate of effect (a): of heteroscedasticity in the macroeconomy. This is highly significant. Indeed, M3 is accepted against the alternative of M4 ($p =$ 0.13), with a likelihood almost as high as that for M2 and my preferred model M1. Although these models are not directly comparable, this observation tells us that (b), the increase in fit due to the use of a macro-consistent yield model, is in practice relatively small. In other words, if the yield model parameters ($\delta_{01}^{\mathcal{B}}, \delta_{11}^{\mathcal{B}}, \Delta_0^{\mathcal{B}}$ and $\Delta_1^{\mathcal{B}}$) are estimated separately they not very well determined statistically. In practice, these parameters are determined by the macro-finance restriction, which equates them with

[13] Comparing L(3) gives a $p-$value of 0.60, indicating that the market actually uses the true values of these stochastic parameters. Nevertheless this model is useful in distinguishing (a) and (b).

their macro model analogues and thus uses macro data to pin them down.

Since these L(M) values are the sum of loglikelihood values at each period (appendix 3) they can be analyzed as time series. Chart 3 shows the effect of disaggregating the differences [L(3)-L(0)] due to effect (a) of heteroscedasticity in the macroeconomy. This is very marked in the early 1980s when the stochastic trend peaks, but also noticeable when the trend is low, at the beginning and the end of the estimation period. The difference [L(4)-L(3)] due to (b), the use of the more general $\mathbb{EA}_1$ yield curve specification, is much smaller and reflecting this, the residuals from the heteroscedastic macro-based specifications M1, M2, M3 and M4 are all very similar. Despite the theoretical superiority of the $\mathbb{EA}_1(N)-$based yield specification, it is hard to see any systematic improvement over maturity or over time. The impulse responses and factor loadings of these models are also similar. For that reason I now focus on the results for the preferred model, M1.

4.1 The empirical macro-model

At the core of this model there is a macro VAR with a steady state solution dictated by the restrictions (3). The novelty here is the introduction of the square root volatility effects implied by the CIR-based term structure model. The model is driven by a nominal factor $x_{1,t}$ and a real factor $x_{2,t}$. Model estimates of these factors are shown in Chart 4, along with their 95% confidence intervals. Most of the work is done by the nominal factor, which has a unit root. Since $x_{1,t}$ has a non-central χ^2 distribution, the downside variance is smaller than the upside, but this asymmetry is only apparent at the beginning of the estimation period when the underlying inflation rate is low. This variable drives the conditional heteroscedasticity in the macro variables. Their one-quarter-ahead forecasts values and 95% confidence intervals are shown in chart 5. The effect of heteroscedasticity is particularly pronounced in the case of the spot rate, consistent with the finding in univariate models (Chen, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996), Stanton (1997) and others). Its variance is low over the first four years of the estimation period, consistent with the *ex post* stability of interest rates over this period (chart 5c). The behavior of the spot rate over the medium term is also influenced by the real factor $x_{2,t}$, as is clear from a comparison with chart 4b. As expected, this real interest series reveals a marked tightening of monetary policy in the late 1970s, with a very relaxed stance in the early 1990s and again post-millennium. The model attributes the ultra-low interest rates seen over the early years of the millennium to a relaxation of monetary policy, coming against a background of a low underlying inflation rate.

How firmly do these factors anchor inflation and interest rates? This question depends upon whether output, inflation & interest rates and the real factor $x_{2,t}$ are contintegrated with the non-stationary nominal factor $x_{1,t}$. This was checked by running ADF tests on the residuals of the output, inflation & interest rate equations, which decisively reject non-stationarity (table 6b). These variables adjust quickly and smoothly to their equilibrium values. This mean-reversion effect can be summarized in terms of the model's eigenvalues. The autoregressive coefficient associated with $x_{1,t}$ is unity, but the other roots are stable and are reported in table 7. Four pairs of roots are sinusoidal, reflecting the cyclical nature of the macroeconomic data[14].

These cyclical effects are seen more clearly in the impulse responses, which show the dynamic effects of innovations in the macroeconomic variables on the system. Because these innovations are correlated empirically, we work with orthogonalized innovations using the triangular factorization defined in (23). The orthogonalized impulse responses show the effect on the macroeconomic system of increasing each of

[14] However the imaginary components of the first root is small, meaning that the macro-model is dominated exponential adjustment effects.

these shocks by one percentage point for just one period using the Wald representation of the system. This arrangement is affected by the ordering of the macroeconomic variables in the vector x_t. Like Kozicki and Tinsley (2005) I adopt the standard ordering: $x_t = \{y_{\pi^*,t}, y_{\mathbf{r}^*,t}, g_t, \pi_t, r_{1,t}\}$. The first shock ($\upsilon_1$) reflects permanent policy or expectational changes in the inflation asymptote while the second (υ_2) reflects structural shocks to the real interest rate. Conventionally, υ_3 is interpreted as a positive demand shock and (υ_4) as a negative supply shock. Finally (υ_5) represents transitory changes in monetary policy. Chart 6 shows the results of this exercise. This gives a plausible description of the macroeconomic dynamics. As in the model of Kozicki and Tinsley (2005), the use of Kalman filters to pick up the effect of unobservable expectational influences seems to solve the notorious price puzzle - the tendency (noted originally by Sims (1992)) for increases in policy interest rates to anticipate inflationary developments and apparently cause inflation. The nominal filter dictates the long run equilibrium of the macroeconomy (and its volatility). These effects are persistent, but the responses of the macroeconomic variables to surprises in inflation, output and interest rates are rapid. They are largely exponential in nature, suggesting that monetary policy has been effective in securing its objectives quickly, without significant policy reversals or cycles.

These results are reflected in chart 7, which report the results of the Analysis of Variance (ANOVA) exercise. These charts show the share of the total variance attributable to the innovations at different lag lengths and are also obtained using the Wald representation of the system, as described in Cochrane (1997). They indicate the contribution each innovation would make to the volatility of each model variable if the error process was suddenly started (having been dormant previously). So for example we see that the output surprise $\eta_{1,t}$ accounts for nearly all of the short run volatility in output, with similar results for the responses of inflation and spot rates to thier own innovations. However, the effect of other innovations builds up over time.

4.2 The empirical yield model

The behavior of the yield curve is dictated by the factor loadings. These are depicted in Chart 8, as a function of maturity (expressed in quarters). The first panel shows the loadings on $x_{1,t}$ (broken line) and $x_{2,t}$ (continuous line). The second panel shows the loadings on π (dotted line), g (broken line) and the spot rate (continuous line). The spot rate provides the link between the macroeconomic model and the term structure. Since it *is* the 3 month yield, this variable has a unit coefficient at a maturity of one quarter and other factors have a zero loading. The spot rate loadings decline over the next few years, reflecting the adjustment of the spot rate towards $x_{1,t}$ and $x_{2,t}$. The spot rate thus determines the slope of the short-term yield curve. Three to five year maturity yields are strongly influenced by the behavior of the real rate factor $x_{2,t}$. The loading on this factor then fades gradually over the longer maturities, allowing this to act as a 'curvature' factor. In contrast, the loading upon $x_{1,t}$ moves up to unity and then increases gradually with maturity over the 2 to 15 year maturity range, so that it acts as a 'level' factor. The loadings for output and inflation have a humped shape, but are relatively small.

Chart 9 shows the risk premia implied by models M0 and M1 for the 15 year yield. Although the loadings for these models are similar, the risk premia differ because they also depend upon the specification of the price of risk and in particular the parameter λ_{10}. This parameter determines the effect of the stochastic trend on the asymptotic risk premium (appendix 2) and differs between the two models, helping to explain the difference shown in 15 year premia. It is significantly negative in M1, but forced to adopt a positive value in M0 in order to keep the model dynamics stable under $\mathcal{Q}$.

This means that although the stochastic trend has a positive effect on the 15 year premia in both models, this is more powerful in M1 than it is in M0. The real factor also has a strong positive effect in both models, as is clear from the chart. The effect of the macroeconomic variables on the risk premia in these models is relatively small in the 15 year area. The impulse responses for the 5-year yield and ANOVA results for the 10 year yield are shown in charts 6 and 7. These reflect the combined effect of the factor loadings and the dynamic characteristics of the model variables discussed in the previous section. The behavior of the 5 year yield depends upon the spot rate and the financial factors. The variance of the 10 year yield (chart 7) is dominated by the shocks to the two financial factors, reflecting the 'level' and 'curvature' effects. The effect of the spot rate and other macroeconomic variables is negligible at this maturity. Table 6 shows that the joint macro-yield model closely replicates the first three moments of the data shown in table 1.

5 Conclusion

This research aligns the new macro-finance model with the mainstream finance literature, using a latent variable with stochastic first and second moments to model the unit root. Because volatility depends upon the stochastic trend in this model, the Jensen effects induced by the convexity of the bond price function affect the associated slope parameter and not the intercept. This means that the trend affects the steady state inflation and spot rates without disturbing the asymptotic forward rate. The model was initially designed to tackle the asymptotic problem posed by the unit root, but in practice it seems that its superior performance stems from its ability to handle the heteroscedasticity of the macroeconomic data rather than the asymptotic yield problem. Unfortunately it is not possible to test these models on longer term yields using this historical data set because there have been long periods when the US Treasury did not fund in the 20-30 year area. However, 30 year issuance has now resumed and the growing demand from pension providers is likely to keep this funding window open. Moreover, the increasing number of 50-year issues in the UK and French Treasury markets should generate data better suited to an empirical test of asymptotic model behavior.

In the meantime, the significance of the inflation-driven conditional heteroscedasticity found in US macro data motivates the use of the general affine model to study both the macroeconomy and the bond market. In contrast to the volatility-clustering effects implied by GARCH macro models, this conditional heteroscedasticity is persistent, exhibiting a unit root. Mathematically, it is more tractable than the GARCH model, generating linear structures that could lend themselves not just to research on the term structure but to optimal control and similar intertemporal optimization problems. Empirically, this finding helps to explain the 'Great Moderation' - the fall in output; interest rate and inflation volatility seen since the mid 1980s (Bernanke (2004), Kim, Nelson, and Piger (2004)), attributing it to the fall in the inflation trend associated with the recession of the early 1980s. It reminds us that this so-called moderation is actually a return to the low-inflation, low-volatility epoch that characterized the early post-war years. The 'general affine' macro-model $\mathbb{A}_1$ helps to explain both the rise and subsequent fall in volatility.

Compared to the mainstream finance model of the bond market, the macro-finance $\mathbb{EA}_1$ model can use a relatively large number of factors (11) because the parameters of the model are informed by macroeconomic as well as yield data (with a total of 1700 data points). It can also use an unrestricted specification of the price of risk, with a large number of parameters. It is particularly informative about the stochastic volatility parameters, identifying these with the volatility parameters of the macro model, which are well-determined. The relative adjustment speeds mean that the behavior of the yield curve is largely dictated by three factors: the inflation end-point, the real interest rate factor and the spot rate. The model is consistent with the traditional three-factor finance specification in this sense, but links these factors to the behavior of the macroeconomy. This research opens the way to new CBM-based studies of monetary policy and a much richer term structure specification, incorporating the best features of both macro-finance and mainstream finance models.

References

AIT-SAHALIA, Y. (1996): "Testing Continuous Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–42.

ANG, A., AND M. PIAZZESI (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 43, 75–87.

BALL, L. (1992): "Why Does Inflation Raise Inflation Uncertainty?," *Journal of Monetary Economics*, 29, 371–388.

BERNANKE, B. (2004): "The Great Moderation," *Proceedings of the meetings of the Eastern Economic Association, Washington, DC, February, available at: www.federalreserve.gov/BOARDDOCS/SPEECHES/2004/20040220/default.ht.*

BRANDT, M., AND K. WANG (2003): "Time Varying Risk Aversion and Unexpected Inflation," *Journal of Monetary Economics*, 50, 1457–1498.

BRUNNER, K., AND G. D. HESS (1993): "Are Higher Levels of Inflation Less Predictable?," *Journal of Business and Economic Statistics*.

CAMPBELL, J., A. LO, AND A. MACKINLAY (1996): *The Econometrics of Financial Markets*. Princeton University Press, Princeton, N.J.

CAPORALE, T., AND B. MCKIERNAN (1997): "High and Variable Inflation: Further

Evidence on the Friedman Hypothesis," *Economics Letters*, 54, 65–68.

Chen, K., F. Karolyi, F. Longstaff, and A. Sanders (1992): "An Empirical Comparison of Alternative Models of the Short Term Interest Rate," *The Journal of Finance*, 47, 1209–1227.

Chen, R., and L. Scott (1993): "Maximum Liklihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14–31.

Cochrane, J. (1997): *Time Series for Economics and Finance*. unpublished lecture notes, www-gsb.uchicago.edu/fac/john.cochrane/research/Papers/timeser1.pdf.

——— (2000): *Asset Pricing*. Princeton University Press, Princeton, N.J.

Cox, J., J. Ingersoll, and S. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 52, 385–407.

Dai, Q., and K. Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 415–41.

Dai, Q., and K. Singleton (2002): "Expectation Puzzles, Time-Varying Risk Premia and Affine Models of the Term Structure," *Journal of Financial Economics*, 63, 415–441.

Dewachter, H., and M. Lyrio (2006): "Macro Factors and the Term Structure of Interest Rates," *Journal of Money, Credit and Banking.*, 38, 119–140.

Dewachter, H., M. Lyrio, and K. Maes (2006): "A Joint Model for the Term Structure of Interest Rates and the Macroeconomy," *Journal of Applied Econometrics*, 21, 439–426.

Duffee, G. (2002): "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 62, 405–443.

Duffee, G., and R. Stanton (2004): "Estimation of Dynamic Term Structure Models," *mimeo, Haas School, U. C. Berkeley.*

Duffie, D., D. Filipovic, and W. Schachtermayer (2003): “Affine Processes and Applications in Finance.,” *Annals of Applied Probability*, 13, 984–1053.

Duffie, D., and R. Kan (1996): “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 6, 379–406.

Dybvig, P. J., J. Ingersoll, and S. Ross (1996): “Long Forward and Zero-Coupon Rates Can Never Fall,” *Journal of Business*, 69.

Engel, R. (1982): “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation,” *Econometrica*, 50, 987–1010.

Fama, E. F. (2006): “The Behaviour of Interest Rates,” *Review of Financial Studies*, 45, 359–379.

Friedman, M. (1977): “Nobel Lecture: Inflation and Unemployment,” *Journal of Political Economy*, 85, 451–472.

Gourieroux, C., and J. Jasiak (2002): “Autoregressive Gamma Processes,” *working paper, CREST*.

Harvey, A. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Holland, A. S. (1995): “Inflation and Uncertainty: Tests for Temporal Ordering,” *Journal of Money, Credit and Banking*, 27, 827–837.

Johnson, N., and S. Kotz (1970): *Distributions in Statistics: Continuous Univariate Distributions*. Houghton Miffin, Boston.

Kim, C.-J., N. Nelson, and J. Piger (2004): “The Less Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations,” *Journal of Business and Economic Statistics*, 22, 80–93.

Kozicki, S., and P. Tinsley (2001): “Shifting Endpoints in the Term Structure of Interest Rates,” *Journal of Monetary Economics*, 47, 613–652.

——— (2005): “Permanent and Transitory Policy Shocks in a Macro Model with

Asymmetric Information," *Journal of Economic Dynamics and Control*, 29, 1985–2015.

Kwiatowski, D., P. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, 54, 159–178.

Lettau, M., S. Ludvigson, and J. Wachter (2006): "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," *Review of Financial Studies, forthcoming.*

McCulloch, J., and H.-C. Kwon (1991): "Us Term Structure Data," *Manuscript, Ohio State University.*

Okun, A. (1971): "The Mirage of Steady Inflation," *Brookings Papers on Economic Activity*, 2, 485–498.

Rudebusch, G. (2002): "Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty," *The Economic Journal*, 112, 402–32.

Rudebusch, G., and T. Wu (2003): "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy," *Journal of Econometrics, forthcoming, and FRBSF Working paper 2003-17.*

Seigel, A. (1979): "The Non-Central Chi-Squared Distribution with Zero Degrees of Freedom," *Biometrica*, pp. 381–386.

Sims, C. (1992): "Interpreting Macroeconomic Time Series: The Effects of Monetary Policy," *European Economic Review*, 36, 975–1011.

Smets, F. (1999): "Output Gap Uncertainty: Does It Matter for the Taylor Rule?," in *Monetary Policy under Uncertainty*, ed. by B. Hunt, and A. Orr, pp. 10–29, Wellington. Reserve Bank of New Zealand.

Stanton, R. (1997): "A Non-Parametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *The Journal of Finance*, 52, 1973–2002.

Sun, T.-S. (1992): "Real and Nominal Interest Rates: A Discrete-Time Model and its Continuous Time Limit," *Review of Financial Studies*, 5, 581–611.

Svensson, L. (1999): "Inflation Targetting as a Monetary Policy Rule," *Journal of Monetary Economics*, 43, 607–54.

Appendix 1: The state-space representation of the model

Define $x_t = \{y_t', z_t'\}'; v_t = \{\varepsilon_t', \eta_t'\}'$ and combine (1) and (2) to get:

$$x_t = \begin{bmatrix} \theta \\ K + \Phi_0\theta \end{bmatrix} + \Sigma_{l=1}^{L}\Gamma_l x_{t-l} + w_t \tag{23}$$

where:

$$w_t = \Gamma v_t; \quad \Gamma = \begin{bmatrix} I_k & 0_{k,n} \\ \Phi_0 & G \end{bmatrix};$$

$$\Gamma_1 = \begin{bmatrix} \Xi & 0_{k,n} \\ \Phi_0\Xi & \Phi_1 \end{bmatrix}; \Gamma_l = \begin{bmatrix} 0_{k,k} & 0_{k,n} \\ 0_{n,k} & \Phi_l \end{bmatrix}; l = 2, ..., L.$$

Stacking (23) puts the system into state space form (4), where $X_t = \{y_t', z_t', ..., z_{t-l}'\}'$, $W_t = C.\{\varepsilon_t', \eta_t', 0_{1,N-k-n}\}'$ and:

$$\Theta = \{\theta, K' + \theta'\Phi_0', 0_{1,N-k-n}\}'; \tag{24}$$

$$\Phi = \begin{bmatrix} \Xi & 0_{k,n} & ... & 0_{k,n} & 0_{k,n} \\ \Phi_0\Xi & \Phi_1 & ... & \Phi_{l-1} & \Phi_l \\ 0_{n,k} & I_n & .. & 0_{n,n} & 0_{n,n} \\ & & & & \\ 0_{n,k} & 0_{n,n} & ... & I_n & 0_{n,n} \end{bmatrix} = \begin{bmatrix} \xi_1 & 0_{N-1}' \\ \Phi_{21} & \Phi_{22} \end{bmatrix}.$$

The second matrix repartitions Φ conformably with (10), so that Φ_{21} is $(N-1)\times 1$ and Φ_{22} is $(N-1)^2$. Similarly:

$$C = \begin{bmatrix} I_k & 0_{k,n} & 0_{k,(N-k-n)} \\ \Phi_0 & G & 0_{n,(N-k-n)} \\ 0_{(N-k-n),k} & 0_{(N-k-n),n} & 0_{(N-k-n),(N-k-n)} \end{bmatrix} = \begin{bmatrix} 1 & 0'_{N-1} \\ C_{21} & C_{22} \end{bmatrix}.$$

where: C_{21} is $(N-1)\times 1$ and C_{22} is $(N-1)^2$. Comparing this with the partition of (24), note that:

$$\Phi_{21} = C_{21}\xi_1 \tag{25}$$

Similarly for $x_t = \{x_{1,t}, x'_{2,t}\}'$ repartition $v_t = \{w_{1,t}, v'_{2,t}\}'$, $w_t = \{w_{1,t}, w'_{2,t}\}'$ and (23) conformably and write Γ as:

$$\Gamma = \begin{bmatrix} I_k & 0_{k,n} \\ \Phi_0 & G \end{bmatrix} = \begin{bmatrix} 1 & 0_{1,(k+n-1)} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix};$$

where Γ_{22} is an $(k+n-1)^2$ lower triangular matrix with unit diagonals and Γ_{21} is a $(k+n-1)$ column vector. The errors in $x_{2,t+1}$ are decomposed into orthogonal components that are related to $w_{1,t+1}$ and $v_{2,t+1} = s_t u_{2,t+1}$:

$$w_{2,t+1} = \Gamma_{21} w_{1,t+1} + \Gamma_{22} s_t u_{2,t+1} \tag{26}$$

where: $u_{2,t+1}$ is an $(n+k-1)$ vector of standard normal variables, $s_t = Diag\{(\delta_{02} + \delta_{12}x_{1,t})^{\frac{1}{2}}, ..., (\delta_{0(k+n)} + \delta_{1(k+n)}x_{1,t})^{\frac{1}{2}}\}$; $\delta_{mj} \geq 0$, $m = 0,1; j = 1,...,k+n$, and $E[u_{2,t+1}w_{1,t+1}] = 0_{(k+n-1)}; u_{2,t+1} \sim N[0_{(k+n-1)}, I_{(k+n-1)}]$. The error structure of

(10) follows from (26) as:

$$W_{2,t+1} = C_{21}w_{1,t+1} + C_{22}S_tU_{2,t+1} \tag{27}$$

$$U_{2,t+1} \sim N(0_{N-1}, D) \tag{28}$$

where: $U_{2,t+1} = \{(u'_{2,t+1}, 0'_{N-k-n}\}'$; $S_t = Diag\{\{(\delta_{02} + \delta_{12}x_{1,t})^{\frac{1}{2}}, ..., (\delta_{0(k+n)} + \delta_{1(k+n)}x_{1,t})^{\frac{1}{2}}\}, 0'_{N-k-n}\}$; $D = Diag\{1'_{k+n-1}, 0'_{N-k-n}\}$, so that $S_tD = S_t$. This implies the Gaussian conditional MGFs:

$$E[\exp[\nu_2'U_{2,t+1}]] = \exp[\frac{1}{2}[\nu_2'D\nu_2]; \tag{29}$$
$$E[\exp[\nu_2'S_tU_{2,t+1}]] = \exp[\frac{1}{2}[\nu_2'S_t^2\nu_2]\ .$$

where:

$$S_t^2 = S_tDS_t = \Delta_0 + x_{1,t}\Delta_1 \tag{30}$$

Finally, the conditional value $X_{2,t+1}|\ x_{1,t+1}$ can be represented using (10) (25) and (27) as:

$$\begin{aligned} X_{2,t+1} &= \Theta_2 + \Phi_{21}x_{1,t} + C_{21}w_{1,t+1} + \Phi_{22}X_{2,t} + C_{22}S_tU_{2,t+1} \\ &= \Theta_2 - C_{21}\theta_1 + C_{21}x_{1,t+1} + \Phi_{22}X_{2,t} + C_{22}S_tU_{2,t+1}. \end{aligned} \tag{31}$$

Appendix 2 : Affine yield structures

This appendix derives the MGF of the distribution under the risk neutral measure $\mathcal{Q}$ and shows how it can be employed as a 'model generating function' to derive the yield model, forward rates and risk premia as well as the moments of the macro system under $\mathcal{Q}$.

The risk-neutral probability measure

Measure Q adjusts the state probabilities using a multiplicative state-dependent subjective-utility weight N_{t+1} (with the logarithm n_{t+1}) so that the time t conditional risk neutral expectation ($E^{\mathcal{Q}}$) of a scalar random variable Z_{t+1}:

$$E^{\mathcal{Q}}[Z_{t+1}|\ X_t] = E[N_{t+1}Z_{t+1}|\ X_t]. \tag{32}$$

Bond (and other asset) prices are discounted expectations of future payoffs and prices defined under this measure:

$$P_{\tau,t} = \exp[-r_{1,t}]E^{\mathcal{Q}}[P_{\tau-1,t+1}|\ X_t]; \quad \tau = 1, ..., M. \tag{33}$$

(Campbell, Lo, and MacKinlay (1996), Cochrane (2000)). Recall that the MGF for measure $\mathcal{Q}$ is:

$$L^{\mathcal{Q}}[\nu, X_t] = E^{\mathcal{Q}}[\exp[\nu' X_{t+1}]|X_t]. \tag{34}$$

Using (11) to replace $P_{\tau-1,t+1}$ in (33) gives a similar form with $\nu = -\Psi_{\tau-1}$:

$$\begin{aligned} P_{\tau,t} &= \exp[-r_{1,t}]E^{\mathcal{Q}}[\exp[-\gamma_{\tau-1} - \Psi'_{\tau-1}\ X_{t+1}]|\ X_t] \\ &= \exp[-\gamma_{\tau-1} - J'_r X_t]L^{\mathcal{Q}}[-\Psi'_{\tau-1}, X_t] \end{aligned} \tag{35}$$

where J_r is a selection vector such that $J'_r X_t = r_{1,t}$.

The discount factor is naturally exponential and if the MGF under the risk neutral measure ($L^{\mathcal{Q}}[\nu, X_t]$) is exponential-affine in X_t, then so is the expectation. Thus (35) is of the form (11), with coefficients that are obtained recursively by matching the coefficients (for maturity $\tau - 1$) of the state variables X_t in the exponents of (35) with those (for τ) in (11). For this to be the case, N_{t+1} must be an exponential-affine function of the state variables or error terms:

$$-n_{t+1} = \omega_t + \lambda_{1,t} x_{1,t+1} + \Lambda'_{2,t} U_{2,t+1} \tag{36}$$

where $\lambda_{1,t}$ is a scalar, $\Lambda_{2,t} = [\lambda'_{2,t}, 0'_{N-(k+n)}]'$ is a $(N-1)\times 1$ deficient vector containing $\lambda_{2,t}$ which is a $(k+n-1)\times 1$ vector of coefficients related to the prices of risk associated with shocks to $x_{2,t+1}$. In the basic affine model class these coefficients are constant and variations in the risk premia only depend upon those in volatility. However, in the 'essentially affine' specification of Duffee (2002) they are linear in x_t, allowing an addition source of variation in the premia. I adopt his specification using (13) with:

$$\Lambda_{2,t} = S_t C'_{22} \Lambda_{20} + S_t^{-1} C_{22}^{-1} \Lambda_{21} x_{1,t} + S_t^{-1} C_{22}^{-1} \Lambda_{22} X_{2,t} \tag{37}$$

where $\Lambda'_{12}, \Lambda_{20}$ and Λ_{21} are $(N-1)\times 1$ vectors and Λ_{22} is an $(N-1)^2$ matrix of parameters to be estimated. The parameter λ_{11} allows x_1 to influence the asymptotic risk premium through the price of risk. However since x_1 affects this through volatility it is redundant in $\mathbb{EA}_1$ and is set to zero. For the EA_1 specifications M1 and M2 to be admissible it is also necessary that $\Lambda_{12} = 0_{N-1}$ and to facilitate the encompassing tests I follow Dewachter and Lyrio (2006) and use this restriction for M0.

5.1 The MGF under the risk neutral measure

Using (32) and (34), the MGF of the distribution under the risk neutral measure $\mathcal{Q}$ can be represented as: $L^{\mathcal{Q}}[\nu, X_t] = E[\exp[n_{t+1} + \nu' X_{t+1}] \mid X_t]$. Substituting (36) and (31) and noting that $x_{1,t+1}$ and $U_{2,t+1}$ are independent allows this to be factorized as:

$$L^{\mathcal{Q}}[\nu,\ X_t] = \exp[-\omega_t + \nu_2'(\Theta_2 - C_{21}\mu + \Phi_{22}X_{2,t})] \times E[\exp[(\nu_2' C_{22} S_t - \Lambda_{2,t}')U_{2,t+1}]]$$
$$\times E[\exp[(\nu_1 - \lambda_{1,t} + \nu_2' C_{21})x_{1,t+1}|\ x_{1,t}]]. \tag{38}$$

For the baseline model M2, these expectations are evaluated using (7) and (29), substituting $\xi_1^{\mathcal{B}}$ for ξ_{π^*}:

$$L^{\mathcal{Q}}[\nu,\ X_t] = \exp[-\omega_t + \nu_2'(\Theta_2 - C_{21}\mu + \Phi_{22}X_{2,t}) + \frac{1}{2}(\nu_2' C_{22} S_t - \Lambda_{2,t}')D(S_t C_{22}' \nu_2 - \Lambda_{2,t})$$
$$+[\frac{\xi_1^{\mathcal{B}}[\nu_1 - \lambda_{1,t} + \nu_2' C_{21}]x_{1,t}}{1 - [\nu_1 - \lambda_{1,t} + \nu_2' C_{21}]/c}] - c\mu \ln[1 - [\nu_1 - \lambda_{1,t} + \nu_2' C_{21}]/c].$$

This probability density is normalized to unity using: $\omega_t = \frac{1}{2}\Lambda_{2,t}' D\Lambda_{2,t} - \frac{\lambda_{1,t}\xi_1^{\mathcal{B}} x_{1,t}}{1+\lambda_{1,t}/c} - c\mu \ln[1 + \lambda_{1,t}/c]$. Substituting this back and using (15), (37) and setting $\lambda_{1,t} = \lambda_{10}$ gives:

$$L^{\mathcal{Q}}[\nu, X_t] = \exp\{\nu_2'(\Theta_2 - C_{21}\mu) + \nu_2'(\Phi_{22} - \Lambda_{22})X_{2,t} + \frac{1}{2}\nu_2{}'\Sigma_0\nu_2 - \nu_2' H_1 \tag{39}$$
$$-c\mu \ln[1 - (\nu_1 - \lambda_1 + \nu_2' C_{21})/c] + c\mu \ln[1 + \lambda_{10}/c]$$
$$+x_{1,t}\left[\frac{\xi_1^{\mathcal{B}}(\nu_1 - \lambda_{10} + \nu_2' C_{21})}{1 - (\nu_1 - \lambda_{10} + \nu_2' C_{21})/c} + \frac{\xi_1^{\mathcal{B}}\lambda_{10}}{1 + \lambda_{10}/c} - \nu_2'\Upsilon_1 + \frac{1}{2}\nu_2{}'\Sigma_1\nu_2\right]\}.$$

Using this as a moment generating function (differentiating w.r.t. $\{\nu_1, \nu_2\}$ and setting these parameters to zero) gives (14). The formulae (16), (18) and (20) follow by substituting $\nu = -\Psi_{\tau-1}$, into (39), substituting this into (35) and equating the coefficients of X_t in the exponent with those in (11). M1 follows immediately from the restrictions noted in the main text.

The standard EA_0 model as a special case

The standard way to obtain the moments and yield structure for the EA_0 specification is to use (5) instead of (7) to evaluate the second expectation in (38). However it is more instructive to derive these from the formulae for M2, taking the limit as c tends to infinity and setting Σ_1 and δ_{11} to zero. This specializes the baseline parameters as shown in the final column of table 2. We expand the denominator in $\theta_1^{\mathcal{Q}}$ as the geometric series $[1 - \lambda_{10}/c + (\lambda_{10}/c)^2 - ...]$ and use $\delta_{01} = \mu/c$ to get the second order approximation: $\theta_1^{\mathcal{Q}} = \mu - \delta_{01}\lambda_{10} + \delta_{01}o(c^{-1})$, which is approximated arbitrarily closely by $\theta_1^{\mathcal{Q}_0} = \mu - \delta_{01}\lambda_{10}$ for large values of c. ($o(c^{-1})$ denotes terms of order c^{-1} or smaller.) $\Theta_2^{\mathcal{Q}_0}$ follows from $\Theta_2^{\mathcal{Q}}$ in the same way. Similarly, $\xi_1^{\mathcal{Q}}$ may be written as the second order approximation: $\xi_1^{\mathcal{B}}[1 - 2\lambda_{10}/c + 3(\lambda_{10}/c)^2 - ...]$, which reduces to the first order expansion $\xi_1^{\mathcal{B}} = \xi_1 - \delta_{01}\lambda_{11}$ upon the substitution of zero for the limiting value of $\delta_{11} = 2\xi_1^{\mathcal{B}}/c$ as c tends to infinity. $\Phi_{21}^{\mathcal{Q}_0}$ follows from $\Phi_{21}^{\mathcal{Q}}$ in the same way.

Now consider the EA_0 yield curve formulae. First note that the recursive nature of the coefficient system means that $\xi_1^{\mathcal{B}}, \delta_{11}, \Sigma_1$ and c do not affect (16). Next, expand the denominators in (18) as above to give the quadratic approximation:

$$\begin{aligned}\psi_{1,\tau} = &\xi_1^{\mathcal{B}}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}] - \Psi'_{2,\tau-1}\Upsilon_1 - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_1\Psi_{2,\tau-1} \\ &-\delta_{11}\{\frac{1}{2}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}]^2 + \lambda_{10}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}] + o(c^{-1})\}. \quad (40)\end{aligned}$$

Substituting the limit $\delta_{11} = 0$ reduces this to a linear difference equation and then substituting (15), (25), $\Sigma_1 = 0$ and the coefficients of table 2 gives:

$$\begin{aligned}\psi_{1,\tau} &= \xi_1^{\mathcal{B}}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}] - \Psi'_{2,\tau-1}\Lambda_{21}; \quad (41) \\ &= \xi_1^{\mathcal{B}}\psi_{1,\tau-1} + \Psi'_{2,\tau-1}\Phi_{21}^{\mathcal{Q}_0};\end{aligned}$$

which reproduces (21) upon substitution of $\xi_1^{\mathcal{B}} = \xi_1^{Q_0} = \xi_1 - \delta_{01}\lambda_{11}$. Finally, to specialize the intercept (22) for EA_0, take a Taylor approximation of the logarithmic term in (20) around ln[1] use $\delta_{01} = \mu/c$ and then expand the denominator terms.

$$
\begin{aligned}
c\mu \ln[1 + \frac{\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}}{c + \lambda_{10}}] &= \mu \left[\frac{\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}}{1 + \lambda_{10}/c}\right] - \frac{1}{2}\delta_{01}\left[\frac{\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}}{1 + \lambda_{10}/c}\right]^2 \\
&\quad + \frac{1}{6}\frac{\delta_{01}[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}]^3}{c[1 + \lambda_{10}/c]^2} - \ldots \\
&= (\mu - \lambda_{10}\delta_{01})\left[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}\right] - \frac{1}{2}\delta_{01}\left[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}\right]^2 \\
&\quad + o(c^{-1})
\end{aligned}
$$

Neglecting the terms $o(c^{-1})$ gives a second order approximation. Substituting back into (20) using (15) and table 2 gives (22):

$$
\begin{aligned}
\Delta\gamma_\tau &= (\Theta_2 - C_{21}\mu - \Sigma_0\Lambda_{20})'\Psi_{2,\tau-1} - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_0\Psi_{2,\tau-1} \\
&\quad + \theta_1^{\mathcal{Q}_0}\psi_{1,\tau-1} + (\mu - \delta_{01}\lambda_{10})\Psi'_{2,\tau-1}C_{21} - \frac{1}{2}\delta_{01}\left[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}\right]^2 \\
&= \Psi'_{2,\tau-1}\Theta_2^{\mathcal{Q}_0} + \theta_1^{\mathcal{Q}_0}\psi_{1,\tau-1} \\
&\quad - \frac{1}{2}\Psi'_{2,\tau-1}\Sigma_0\Psi_{2,\tau-1} - \frac{1}{2}\delta_{01}\left[\psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}\right]^2 .
\end{aligned}
\tag{42}
$$

Risk premia

The risk premia depend entirely upon the difference between the two measures $\mathcal{P}$ and $\mathcal{Q}$. To see this, note that the premium on a $\tau-$period bond is the expected return, less the spot rate. The gross expected rate of return is the expected payoff $E[P_{\tau-1,t+1}|X_t]$ divided by its current price $P_{\tau,t} = \exp[-r_{1,t}]E^{\mathcal{Q}}[P_{\tau-1,t+1}|\ X_t]$. Taking the natural logarithm expresses this as a percentage return and subtracting the spot rate $r_{1,t}$ then gives the risk premium: $\rho_{\tau,t} = \log E[P_{\tau-1,t+1}|X_t] - \log E^{\mathcal{Q}}[P_{\tau-1,t+1}|X_t]$. This is affine provided that the MGF is exponential-affine under both measures. Substituting (11),

(7) and (39) with $\nu = \Psi'_{\tau-1}$ gives the risk premium. In the case of the EA_1 unit root model this gives:

$$\begin{aligned}\rho_{\tau,t} = & -\Psi'_{2,\tau-1}(H_1 + \Lambda_{22}X_{2,t} + \Upsilon_1 x_{1,t}) \\ & -\{\frac{\xi_1^{\mathcal{B}} c\lambda_{10}\left(\Psi'_{2,\tau-1}C_{21} + \psi_{1,\tau-1}\right)\left(2c + \lambda_{10} + \Psi'_{2,\tau-1}C_{21} + \psi_{1,\tau-1}\right)}{\left(c + \lambda_{10} + \Psi'_{2,\tau-1}C_{21} + \psi_{1,\tau-1}\right)(c + \lambda_{10})\left(c + \psi_{1,\tau-1} + \Psi'_{2,\tau-1}C_{21}\right)}\}x_{1,t}\end{aligned} \tag{43}$$

The linear term on the first line is the compensation for the bond's exposure to shifts in $X_{2,t}$, which is negligible for a portfolio or security like an ultra-long bond, with a yield that mimics the asymptotic portfolio. The non-linear term on the second line shows the premium on the asymptotic portfolio and is zero if $\lambda_{10} = 0$. It is positively related to $x_{1,t}$ if $\lambda_{10} < 0$; and negatively related if $\lambda_{10} > 0$ since:

$$\frac{d^2\rho_{\tau,t}}{dx_{1,t}d\lambda_{10}} = -\xi_1^{\mathcal{B}} c^2 \left(\frac{C_{21}\Psi'_{2,\tau-1} + \psi_{1,\tau-1}}{(c + \lambda_{10})^2}\right)\left(\frac{2c + 2\lambda_{10} + C_{21}\Psi'_{2,\tau-1} + \psi_{1,\tau-1}}{\left(c + \lambda_{10} + C_{21}\Psi'_{2,\tau-1} + \psi_{1,\tau-1}\right)^2}\right) \leq 0$$

given (19). The risk premium in the $EA_0(N)$ model is:

$$\rho_{\tau,t} = -\Psi'_{2,\tau-1}(H_0 + \Upsilon_0 x_{1,t} + \Lambda_{22}X_{2,t}) - \psi_{1,\tau-1}\delta_{01}(\lambda_{10} + \lambda_{11}x_{1,t}) \tag{44}$$

Recall that this model requires the restriction $\lambda_{11} > 0$. This means that the associated factor risk premium (the asymptotic premium) is negatively related to the inflation trend. This premium is shown by $\delta_{01}(\lambda_{10} + \lambda_{11}x_{1,t})$ in the second term. In practice, this effect has the effect of offsetting the effect of the component $\Psi'_{2,\tau-1}\Upsilon_0 x_{1,t}$ shown in the first term (which is positive in the empirical model). Thus, the influence of the trend on the premia shown in chart 9 is more pronounced in model M1.

Forward rates and asymptotic behavior

Taking logs of (11) and maturity-differencing gives the affine forward rate structure:

$$f_{\tau,t} = \Delta\gamma_{\tau+1} + [\Psi_{\tau+1} - \Psi_{\tau}]' X_t; \quad \tau = 1, ..., M. \tag{45}$$

This shows that the asymptotic behavior of the forward rate depends critically upon whether the slope coefficients converge to constants. If so, the last term vanishes and $\Delta\gamma_{\tau+1}$ and hence the forward rate asymptote (f_t^*) is constant. Since $\Psi_2^* = lim_{\tau\to\infty}\Psi_{2,\tau} = (I - (\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r}$ is constant, this just depends upon the behavior $\psi_{1,\tau}$. Table 3 shows the asymptotic effect on the yield model of imposing unit root restrictions ($\xi_1^{\mathcal{Q}_i} = 1; \theta_1^{\mathcal{Q}_i} = 0$) on $x_{1,t}$ under $\mathcal{Q}$. In EA_1, $\psi_{1,\tau}$ asymptotes to a constant, so β_1^* and f^* are zero. However, (21) shows that in EA_0: $lim_{\tau\to\infty}(\psi_{1,\tau+1} - \psi_{1,\tau}) = \Phi_{21}^{\mathcal{Q}_0\prime}\Psi_2^* = \Phi_{21}^{\mathcal{Q}_0\prime}(I - (\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r}$. Consequently, $\beta_1^* = lim_{\tau\to\infty}\psi_{1,\tau}/\tau = \Phi_{21}^{\mathcal{Q}_0\prime}(I - (\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r}$. This expression is equal to the asymptotic effect of $x_{1,t}$ on $r_{1,t}$ under $\mathcal{Q}$ and as such it should be close to unity. Substituting these coefficients into (22) and (45) gives the asymptotic behavior of the forward rate reported in table 3. The final term in this expression is the Jensen effect associated with $\psi_{1,\tau}x_{1,t}$, which behaves like (-)$\frac{1}{2}\delta_{01}\beta_1^{*2}\tau^2$ in the limit. These results illustrate the basic theorem of

Dybvig, Ingersoll, and Ross (1996), which says that in an arbitrage-free framework 'the limiting forward interest rate, if it exists, can never fall'. In other words the forward rate must either fail to converge with maturity (as in the unit root $\mathbb{EA}_0$ model), or must asymptote to a constant ($\mathbb{EA}_1$).

Appendix 3 : The Kalman filter and the likelihood function

These models were estimated and tested using the FindMinimum algorithm on Mathematica. The basic results have recently been verified using Matlab. In this model, the unobservable variables are modelled using the Extended Kalman Filter (Harvey (1989), Duffee and Stanton (2004)). This method assumes that the revisions in (2) are approximately normally distributed:

$$\varepsilon_{t+1} \sim N(0, Q_t)$$
$$where\ :\ Q_t = Q_0 + Q_1 y_{\pi^*,t}$$
$$Q_i = Diag\{\delta_{0i}^2, \delta_{1i}^2\};\quad i = 1, 2.$$

I represent expectations conditional upon the available information with a 'hat' (so that $\hat{y}_t = \hat{E}_t\ y_t; \hat{y}_{s,t} = \hat{E}_t\ y_s; s \geq t$) and define the covariance matrices:

$$\begin{aligned}\hat{V}_t &= \hat{E}_t(y_t - \hat{y}_t)(y_t - \hat{y}_t)'; \\ \hat{V}_{t+1,t} &= \hat{E}_t(y_{t+1} - \hat{y}_{t+1,t})(y_{t+1} - \hat{y}_{t+1,t})' \\ &= \Xi \hat{P}_t \Xi' + Q_t;\end{aligned} \tag{46}$$

where, using (2):

$$y_{t+1,t} = \hat{E}_t \hat{y}_{t+1} = \theta + \Xi \hat{y}_t. \tag{47}$$

Similarly, using (1): $z_{t+1} = \hat{z}_{t+1,t} + G\eta_{t+1} + \Phi_0(y_{t+1} - \hat{y}_{t+1,t})$ where:

$$\hat{z}_{t+1,t} = K + \Phi_0\ \hat{y}_{t+1,t} + \Sigma_{l=0}^{L-1} \Phi_l z_{t-l}; \tag{48}$$

and using (12): $r_{t+1} = \hat{r}_{t+1,t} + B_0'(y_{t+1} - \hat{y}_{t+1,t}) + B_1'(z_{t+1} - \hat{z}_{t+1,t}) + e_{t+1}$. where:

$$\hat{r}_{t+1,t} = \alpha + B_0' \hat{y}_{t+1,t} + {B_1}' \hat{z}_{t+1,t} + \Sigma_{l=2}^{L} {B_l}' z_{t+2-l}. \tag{49}$$

The $t-$conditional covariance matrix for this $(t+1)-$dated system is:

$$\begin{bmatrix} \sum_{rr} \sum_{rz} \sum_{ry} \\ \sum_{rz}' \sum_{zz} \sum_{zy} \\ \sum_{ry}' \sum_{zy}' \sum_{yy} \end{bmatrix} = \hat{E}_t \begin{bmatrix} r_{t+1} - \hat{r}_{t+1,t} \\ z_{t+1} - \hat{z}_{t+1,t} \\ y_{t+1} - y_{t+1,t} \end{bmatrix} \cdot \begin{bmatrix} r_{t+1} - r_{t+1,t} & z_{t+1} - z_{t+1,t} & y_{t+1} - y_{t+1,t} \end{bmatrix}$$

where: $\sum_{rr} = \bar{P} + B_1 M_t B_1' + (B_1\Phi_0 + B_0)\hat{V}_{t+1,t}(B_1\Phi_0 + B_0)'$; $\sum_{rz} = B_1 M_t + (B_1\Phi_0 + B_0)\hat{V}_{t+1,t}\Phi_0'$; $\sum_{ry} = (B_1\Phi_0 + B_0)\hat{V}_{t+1,t}$; $\sum_{zz} = \Phi_0 \hat{P}_{t+1,t}\Phi_0' + M_t$; $\sum_{zy} = \Phi_0 \hat{V}_{t+1,t}$; $\sum_{yy} = \hat{V}_{t+1,t}$; $M_t = G[S_0 + S_1 y_{t,t}]G'$; $S_i = Diag\{\delta_{i1}^2 ..., \delta_{in}^2\}$; $i = 0, 1$ and $\bar{P}$ is defined in (12). This allows the expectations to be updated as:

$$\hat{y}_{t+1} = \hat{y}_{t+1,t} + \begin{bmatrix} \sum_{yr} \sum_{yz} \end{bmatrix} \begin{bmatrix} \sum_{rr} \sum_{rz} \\ \sum_{zr} \sum_{zz} \end{bmatrix}^{-1} \begin{bmatrix} r_{t+1} - \hat{r}_{t+1,t} \\ z_{t+1} - \hat{z}_{t+1,t} \end{bmatrix} \tag{50}$$

$$\hat{V}_{t+1} = \hat{V}_{t+1,t} - \begin{bmatrix} \sum_{yr} \sum_{yz} \end{bmatrix} \begin{bmatrix} \sum_{rr} \sum_{rz} \\ \sum_{zr} \sum_{zz} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{yr} \\ \sum_{yz} \end{bmatrix} \tag{51}$$

The (log) likelihood for period $t+1$ is thus:

$$\begin{aligned} L_{t+1} = k - \frac{1}{2} \ln \left(Det \begin{bmatrix} \sum_{rr} \sum_{rz} \\ \sum_{zr} \sum_{zz} \end{bmatrix} \right) \\ -\frac{1}{2} \begin{bmatrix} r_{t+1} - r_{t+1,t} \; z_{t+1} - z_{t+1,t} \end{bmatrix} \begin{bmatrix} \sum_{rr} \sum_{rz} \\ \sum_{zr} \sum_{zz} \end{bmatrix}^{-1} \begin{bmatrix} r_{t+1} - \hat{r}_{t+1,t} \\ z_{t+1} - \hat{z}_{t+1,t} \end{bmatrix}. \end{aligned} \tag{52}$$

The loglikelihood for the full sample follows by iterating (46), (47), (50) and (51) forward given suitable starting values; substituting (48) and (49) then summing (52) over $t = 1, ...T$.

Table 1: Data Summary Statistics: 1961Q4-2004Q1

	g	π	r_1	r_4	r_8	r_{12}	r_{20}	r_{28}	r_{40}	r_{60}
Mean	0.0266	4.4345	5.8051	6.3954	6.6305	6.7849	7.0021	7.1478	7.2513	7.4383
Std.	2.3382	2.9841	2.7750	2.8174	2.7355	2.6508	2.5447	2.4790	2.4194	2.3833
Skew.	-0.5	1.4161	1.1907	1.8820	1.8697	1.8793	1.9633	1.9622	1.9215	1.8188
Kurt.	1.0230	1.5178	2.3986	1.1986	1.0622	0.9505	0.9166	0.7550	0.5929	0.3945
Auto.	0.4632	0.9921	0.9815	0.9892	0.9923	0.9944	0.9953	0.9963	0.9969	0.9971
KPSS	0.2151	0.3399	0.3100	0.3307	0.3348	0.3399	0.3475	0.3548	0.3761	0.4298
ADF	-4.133	-2.411	-2.110	-2.100	-2.063	-2.031	-2.043	-1.991	-1.951	-2.091

Inflation (π) and 3 month T-bill rate (r_1) are from Datastream. Output gap (g) is from OECD. Yield data are US Treasury discount bond equivalent data compiled by McCulloch and Kwon (1990) updated by the New York Federal Reserve Bank. *Mean* denotes sample arithmetic mean expressed as percentage p.a.; *Std.* standard deviation and *Auto.* the first order quarterly autocorrelation coefficient. *Skew.* & *Kurt.* are standard measures of skewness (the third moment) and kurtosis (the fourth moment). *KPSS* is the Kwiatowski et al (1992) statistic testing the null hypothesis of level stationarity. The 10% and 5% significance levels are 0.347 and 0.463 respectively. *ADF* is the Adjusted Dickey-Fuller statistic testing the null hypothesis of non-stationarity. The 10% and 5% significance levels are 2.575 and 2.877 respectively.

Table 2: Dynamic coefficients for different yield models and measures

model:		$M2$	$M1$	$M0$
type:			$EA_1(N)$	$EA_0(N)$
measure:	P	Q	Q_1	Q_0
	θ_1	$\theta_1^Q = \frac{\mu}{1+\lambda_{10}/c}$	$\theta_1^{Q_1} = \theta_1^Q$	$\theta_1^{Q_0} = \mu - \delta_{01}\lambda_{10}$
	$\xi_1^{\mathcal{B}}$	$\xi_1^Q = \frac{\xi_1^{\mathcal{B}}}{(1+\lambda_{10}/c)^2}$	$\xi_1^{Q_1} = \frac{\xi_1}{(1+\lambda_{10}/c)^2}$	$\xi_1^{Q_0} = \xi_1^{\mathcal{B}} = \xi_1 - \delta_{01}\lambda_{11}$
	Θ_2	$\Theta_2^Q = \Theta_2 - \theta_1 C_{21} - H_1 + \frac{\mu C_{21}}{(1+\lambda_{10}/c)}$	$\Theta_2^{Q_1} = \Theta_2^Q$	$\Theta_2^{Q_0} = \Theta_2 - H_0$
	Φ_{21}	$\Phi_{21}^Q = \frac{\Phi_{21}}{(1+\lambda_{10}/c)^2} - \Upsilon_1$	$\Phi_{21}^{Q_1} = \frac{\Phi_{21}}{(1+\lambda_{10}/c)^2} - \Upsilon_1$	$\Phi_{21}^{Q_0} = \Phi_{21} - \Upsilon_0$
	Φ_{22}	$\Phi_{22}^Q = \Phi_{22} - \Lambda_{22}$	$\Phi_{22}^{Q_1} = \Phi_{22}^Q$	$\Phi_{22}^{Q_0} = \Phi_{22}^Q$

This table shows how the shift from the state probability measure $\mathcal{P}$ underpinning the macromodel to the risk-neutral measures used to price financial assets shift the coefficients of (14). The scalars λ_{10} and λ_{11} reflect the price of risk (or risk-reward ratio) associated with variations in $x_{1,t}$; while H_i and Υ_i; $i = 0, 1$ are matrices of parameters reflecting the price of risk associated with the other state variables and are defined in (15). The basline model M2 allows $\xi_1^{\mathcal{B}}$ to differ from the parameter ξ_1 defined in the macro model. M1 is the $EA_1(N)$ model with $\xi_1^{\mathcal{B}} = \xi_1$. M0 specialises M2 by setting δ_{11} and Σ_1 to zero and taking the limit as c tends to infinity. This gives the standard $EA_0(N)$ model with the Duffee (2002) risk adjustments shown in the final column.

Table 3: Asymptotic yield coefficients for the limit of a unit root

model:	$EA_0(N)$	$EA_1(N)$
limit:		
$\beta_1^* = lim_{\tau\to\infty}\beta_{1,\tau}$	$\Phi_{21}^{\mathcal{Q}_0\prime}(I-(\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r}$	0
$\beta_2^* = lim_{\tau\to\infty}\beta_{2,\tau}$	0_{N-1}	0_{N-1}
$f_t^* = lim_{\tau\to\infty}f_{\tau,t}$	$(\Theta_2-H_0)'\Psi_2^*+\tau\beta_1^*x_{1,t}$ $-\frac{1}{2}\Psi_2^{*\prime}\Sigma_0\Psi_2^*-\frac{1}{2}\delta_{01}[\tau\beta_1^*+\Psi_2^{*\prime}C_{21}]^2$	$(\Theta_2-H_1)'\Psi_2^*-\frac{1}{2}\Psi_2^{*\prime}\Sigma_0\Psi_2^*$

This table shows the effect on the yield model of imposing unit root restrictions ($\xi_1^{\mathcal{Q}_i}=1;\theta_1^{\mathcal{Q}_i}=0$) on $x_{1,t}$ under $\mathcal{Q}$. The maturity limits for its factor loading (β_1^*) are very different in the two models. This is reflected in the behavior of the forward rate asymptote (f_t^*). $\beta_2^*=\Psi_2^*/\tau$ is common to both specifications where: $\Psi_2^* = lim_{\tau\to\infty}\Psi_{2,\tau} = (I-(\Phi_{22}^{\mathcal{Q}})')^{-1}J_{2,r}$.

Table 4: Model Evaluation 1961Q4-2004Q1

Model	Specification*	Restrictions (w.r.t M4)	Parameters				Loglikelihood			
M	Macro/Finance		k(M)	k(4)-k(M)	k(3)-k(M)	k(2)-k(M)	L(M)	L(4)-L(M)	L(3)-L(M)	L(2)-L(M)
	Standard	$\delta_{01}^{\mathcal{B}}=\delta_{01}$, $\Delta_0^{\mathcal{B}}$=Δ_0	60	16	11	6				
0	$A_0(N)/EA_0(N)$	$\delta_{11}^{\mathcal{B}}=\delta_{11}$, $\Delta_1^{\mathcal{B}}=\Delta_1$					694.3	110.8	102.2	106.4
	χ^2_{95}	$\delta_{11}=0$, $\Delta_1=0$						26.30	19.68	12.59
	p	$\xi_1=1$						0.0	0.0	0.0
1	Stochastic Trend		63	13		3				
	$A_1(N)/EA_1(N)$	$\delta_{01}^{\mathcal{B}}=\delta_{01}$, $\Delta_0^{\mathcal{B}}=\Delta_0$					747.0	5.40		1.00
	χ^2_{95}	$\delta_{11}^{\mathcal{B}}=\delta_{11}$, $\Delta_1^{\mathcal{B}}=\Delta_1$						22.36		7.81
	p	$\xi_1^{\mathcal{B}}=\xi_1=1;\mu=0$						0.97		0.80
2	Baseline									
	$A_1(N)/EA_1(N)$	$\delta_{01}^{\mathcal{B}}=\delta_{01}$, $\Delta_0^{\mathcal{B}}=\Delta_0$	66	10						
	χ^2_{95}	$\delta_{11}^{\mathcal{B}}=\delta_{11}$, $\Delta_1^{\mathcal{B}}=\Delta_1$					747.5	4.4		
	p							18.31		
								0.93		
3	Hybrid									
	$A_1(N)/EA_0(N)$	$\delta_{11}^{\mathcal{B}}=0$, $\Delta_1^{\mathcal{B}}=0$	71	5			745.4	8.60		
	χ^2_{95}							11.07		
	p							0.13		
4	Encompassing		76							
	$A_1(N)/EA_1(N)$						749.7			
	χ^2_{95}									
	p									

(*) Model specification $A_m(N), EA_m(N)$ where:
m = number of variables conditioning volatility
N = number of state variables.

Table 5a: Dynamic model structures

(asymptotic t-values in parentheses.)

Parameter	M0	M1	Parameter	M0	M1
Φ_1			Φ_2		
$\phi_{1,gg}$	1.05567 (176.71)	1.0604 (159.88)	$\phi_{2,gg}$	0.09002 (11.06)	0.09065 (12.36)
$\phi_{1,\pi g}$	0.07332 (5.01)	0.06777 (18.07)	$\phi_{2,\pi g}$	0.04530 (4.12)	0.04145 (5.21)
$\phi_{1,rg}$	0.25705 (12.32)	0.26856 (40.95)	$\phi_{2,rg}$	0.06264 (3.21)	0.06850 (3.47)
$\phi_{1,g\pi}$	-0.08272 (2.00)	-0.09445 (2.03)	$\phi_{2,g\pi}$	$-.00612$ (0.44)	-0.02462 (2.77)
$\phi_{1,\pi\pi}$	1.14772 (60.70)	1.13737 (47.20)	$\phi_{2,\pi\pi}$	-0.14507 (5.92)	-0.13335 (13.69)
$\phi_{1,r\pi}$	-0.13358 (2.68)	-0.12858 (3.00)	$\phi_{2,r\pi}$	0.57063 (13.21)	0.54936 (44.56)
$\phi_{1,gr}$	-0.00306 (0.07)	0.04055 (0.52)	$\phi_{2,gr}$	-0.28889 (10.10)	-0.27242 (11.81)
$\phi_{1,\pi r}$	0.11961 (1.52)	0.10821 (0.85)	$\phi_{2,\pi r}$	-0.12191 (1.32)	-0.08135 (2.78)
$\phi_{1,rr}$	0.75903 (6.96)	0.87218 (25.53)	$\phi_{2,rr}$	-0.42636 (8.63)	-0.44589 (11.18)

Parameter	M0	M1	Parameter	M0	M1
Φ_3					
$\phi_{3,gg}$	-0.18539 (33.22)	-0.18452 (21.35)	φ_{π^*}	0.0656 (1.49)	0.0530 (3.55)
$\phi_{3,\pi g}$	-0.01377 (1.99)	-0.01263 (1.21)	$\varphi_{\mathbf{r}}$	0.00107 (0.26)	0.00195 (1.17)
$\phi_{3,rg}$	-0.22202 (15.73)	-0.23026 (51.77)	$\xi_{\mathbf{r}}$	0.85770 (131.69)	0.84335 (93.44)
$\phi_{3,g\pi}$	0.10524 (4.94)	0.09708 (6.22)			
$\phi_{3,\pi\pi}$	-0.02315 (0.78)	-0.03055 (1.76)			
$\phi_{3,r\pi}$	-0.29044 (12.50)	-0.27045 (30.54)			
$\phi_{3,gr}$	0.08673 (1.57)	0.13081 (3.51)			
$\phi_{3,\pi r}$	-0.02266 (0.31)	-0.03899 (0.96)			
$\phi_{3,rr}$	0.18211 (7.61)	0.20111 (7.03)			

Table 5b: Variance structures

(asymptotic t-values in parentheses).

Parameter	M0	M1	Parameter	M0	M1
δ_{01}	1.1449×10^{-6} (6.20)	$(-)$	δ_{11}	$(-)$	1.3541×10^{-4} (2.71)
$\mathbf{\Delta}_0$			$\mathbf{\Delta}_1$		
$\delta_{0\mathbf{r}^*}$	3.54505×10^{-6} (8.13)	1.6866×10^{-6} (1.36)	$\delta_{1\mathbf{r}^*}$	$(-)$	4.4070×10^{-4} (1.42)
δ_{0g}	3.14435×10^{-6} (8.61)	2.8960×10^{-6} (3.26)	δ_{1g}	$(-)$	4.2764×10^{-5} (0.39)
$\delta_{0\pi}$	1.77538×10^{-6} (12.68)	1.1096×10^{-6} (2.62)	$\delta_{1\pi}$	$(-)$	1.0000×10^{-4} (1.86)
δ_{0r}	3.42389×10^{-6} (12.29)	2.4260×10^{-7} (1.14)	δ_{1r}	$(-)$	3.8722×10^{-4} (3.74)

Parameter	M0	M1
$\mathbf{\Gamma}_{22}$		
$g_{\pi g}$	0.09618 (1.24)	0.06701 (0.64)
g_{rg}	0.08806 (0.81)	0.13324 (0.76)
$g_{r\pi}$	0.9944 (0.57)	0.23449 (1.45)

Table 5c: Risk adjustment structures

(asymptotic t-values in parentheses.)

Parameter	M0	M1	Parameter	M0	M1
	$\mathbf{H}_0$	$\mathbf{H}_1$		$\mathbf{\Upsilon}_0$	$\mathbf{\Upsilon}_1$
$\eta_{\mathbf{r}^*}$	-2.2916×10^{-4} (1.68)	-1.1559×10^{-4} (1.99)			
η_g	2.1224×10^{-3} (4.10)	2.3545×10^{-3} (4.00)	υ_g	1.9570×10^{-2} (0.73)	-2.3125×10^{-3} (0.01)
η_π	2.8261×10^{-3} (4.99)	2.8295×10^{-3} (4.05)	υ_π	0.22421 (10.02)	0.22190 (10.19)
η_r	-3.6448×10^{-3} (4.22)	-4.2967×10^{-3} (7.72)	υ_r	-0.3242 (1.72)	-0.2990 (1.05)

Parameter	M0	M1	Parameter	M0	M1
$\mathbf{\Lambda}_{22}$					
$\lambda_{22,\mathbf{r}^*\mathbf{r}^*}$	0.03718 (5.92)	0.03177 (6.25)			
$\lambda_{22,g\mathbf{r}^*}$	0.3131 (2.53)	0.33190 (3.61)	$\lambda_{22,gg}$	-0.18290 (22.87)	-0.19216 (8.31)
$\lambda_{22,\pi\mathbf{r}^*}$	0.30476 (2.63)	0.19821 (1.80)	$\lambda_{22,\pi g}$	0.44998 (35.41)	0.47426 (13.6)
$\lambda_{22,r\mathbf{r}^*}$	-0.74524 (5.09)	-0.60027 (2.78)	$\lambda_{22,rg}$	-0.06090 (1.67)	-0.05623 (1.97)
$\lambda_{22,g\pi}$	0.08401 (4.77)	0.08154 (25.11)	$\lambda_{22,gr}$	0.19941 (19.74)	0.18187 (11.32)
$\lambda_{22,\pi\pi}$	0.07331 (4.44)	0.09388 (19.65)	$\lambda_{22,\pi r}$	-0.69468 (143.57)	-0.68454 (49.16)
$\lambda_{22,r\pi}$	0.02710 (1.25)	0.02401 (9.18)	$\lambda_{22,rr}$	0.29080 (2.65)	0.26856 (15.67)

Parameter	M0	M1
Λ		
λ_{10}	-43.036 (2.02)	-42.8160 (4.09)
λ_{11}	196.394 (1.59)	(−)

Table 6a: Summary statistics for estimated values, M1 1961Q4-2004Q1

	g	π	r_1	r_4	r_8	r_{12}	r_{20}	r_{28}	r_{40}	r_{60}
Mean	0.1433	4.4933	5.9103	6.5804	6.7846	6.9040	7.1637	7.2676	7.3626	7.4621
Std.	2.23992	2.96308	2.5555	2.6222	2.5782	2.5277	2.4582	2.4204	2.3889	2.3193
Skew.	-0.553	1.3980	0.9157	0.7411	0.8196	0.8749	0.9212	0.9298	0.9238	0.8994
Kurt.	3.9787	4.5119	4.4288	3.8537	3.8352	3.8021	3.7220	3.6512	3.5751	3.4752

Mean denotes sample arithmetic mean expressed as percentage p.a.; *Std.* standard deviation and *Auto.* the first order quarterly autocorrelation coefficient. *Skew.* and *Kurt.*are standard measures of skewness and kurtosis.

Table 6b: Residual Error Statistics M1 1961Q4-2004Q1

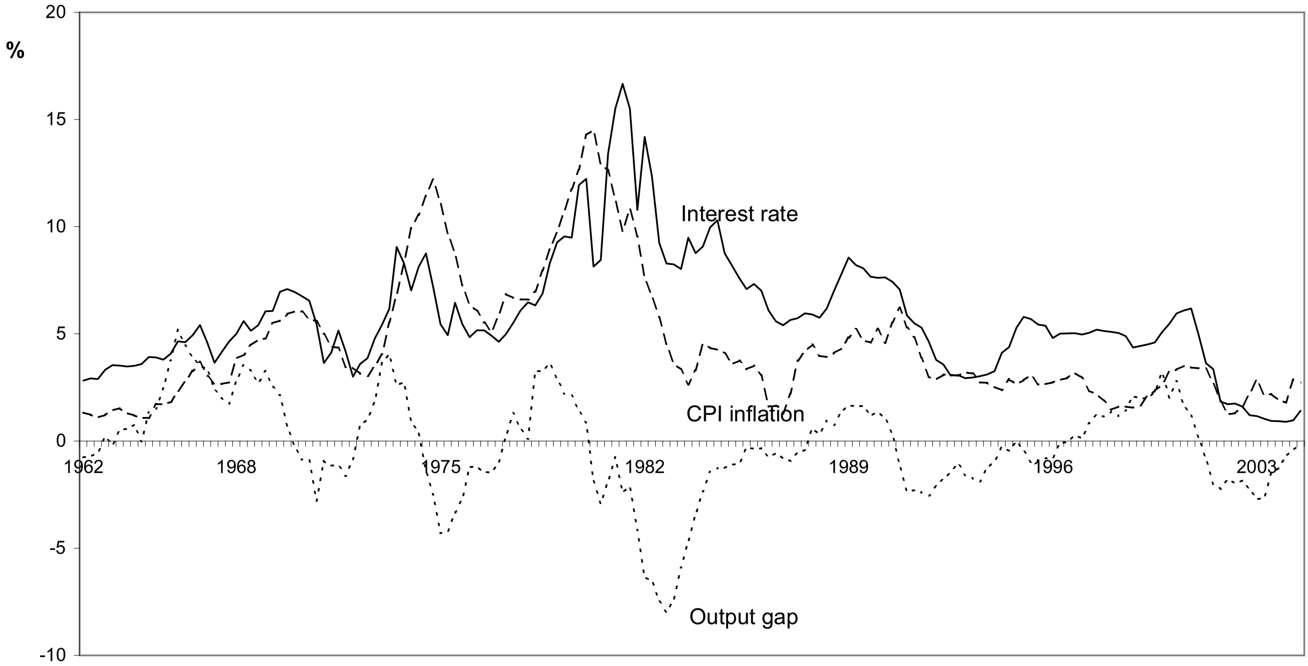
	g	π	r_1	r_4	r_8	r_{12}	r_{20}	r_{28}	r_{40}	r_{60}
R^2	0.9020	0.9664	0.9099	0.8616	0.8797	0.8901	0.9158	0.9297	0.9341	0.9355
$RMSE$	0.1824	0.5452	0.8321	1.045	0.9458	0.8563	0.7359	0.6550	0.6188	0.6034
ADF	-9.4633	-6.3951	-6.3239	-6.6522	-6.7983	-7.0042	-7.1833	-7.2131	-7.0550	-6.9903
Auto.	-0.1218	-0.1283	-0.0933	-0.0261	0.0049	0.0948	0.1081	0.0832	0.0153	-0.0021
	(1.57)	(1.59)	(1.21)	(0.34)	(0.06)	(1.23)	(1.98)	(1.08)	(0.19)	(0.02)

The first row reports the unadjusted R^2; the second the Root Mean Square Error ($RMSE$). ADF is the Adjusted Dickey-Fuller statistic testing the null hypothesis of non-stationarity. The 10% and 5% significance levels are 2.575 and 2.877 respectively. *Auto.* is the first order quarterly autocorrelation coefficient (with t-value in parenteses).

Table 7: Eigenvalues of the dynamic responses in M1
(in order of absolute value)

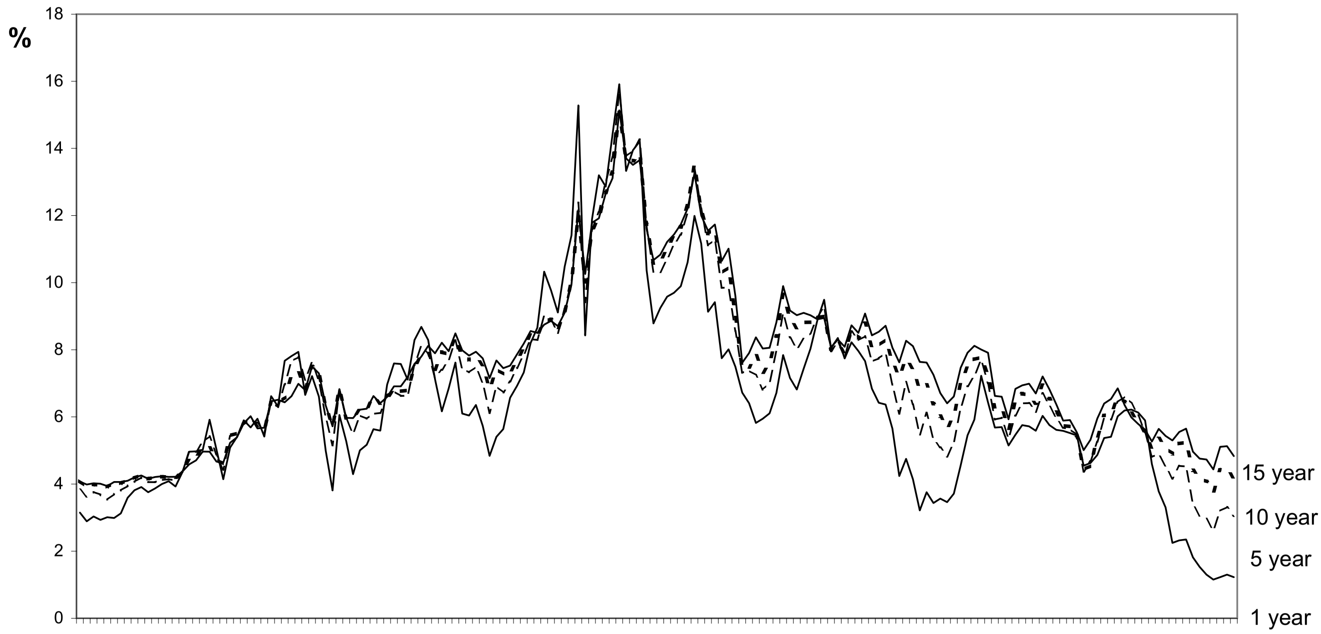
M0	M1
1	1
$0.92415 \pm 0.08084i$	0.93103 ± 0.10677
0.85770	84000
$0.55264 \pm 0.27966i$	0.61927
$-0.00649 \pm 0.51542i$	0.51854 ± 0.18382
0.51152	0.03600 ± 0.50129
-0.24485 ± 0.012387	-0.26430 ± 0.09198

Chart 1: Macroeconomic variables



CPI Inflation and 3 month T-bill interest rate are from Datastream. Output gap is from OECD.

Chart 2: US Treasury discount yields



US Treasury discount bond equivalent yield data compliled by McCulloch and Kwon (1990) updated by the NY Fed.

Chart 3: Model performance
(loglikelihood by period)

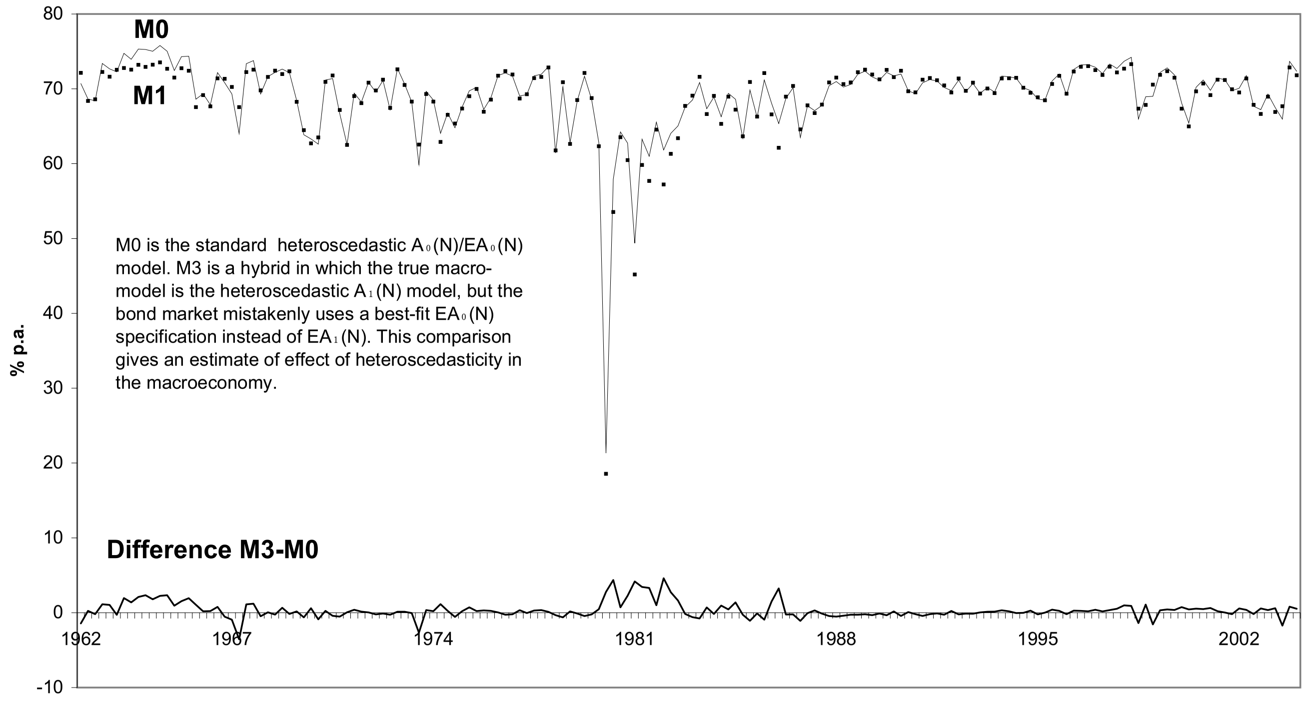


Chart 4(a) Inflation factor
(with 95% confidence band)

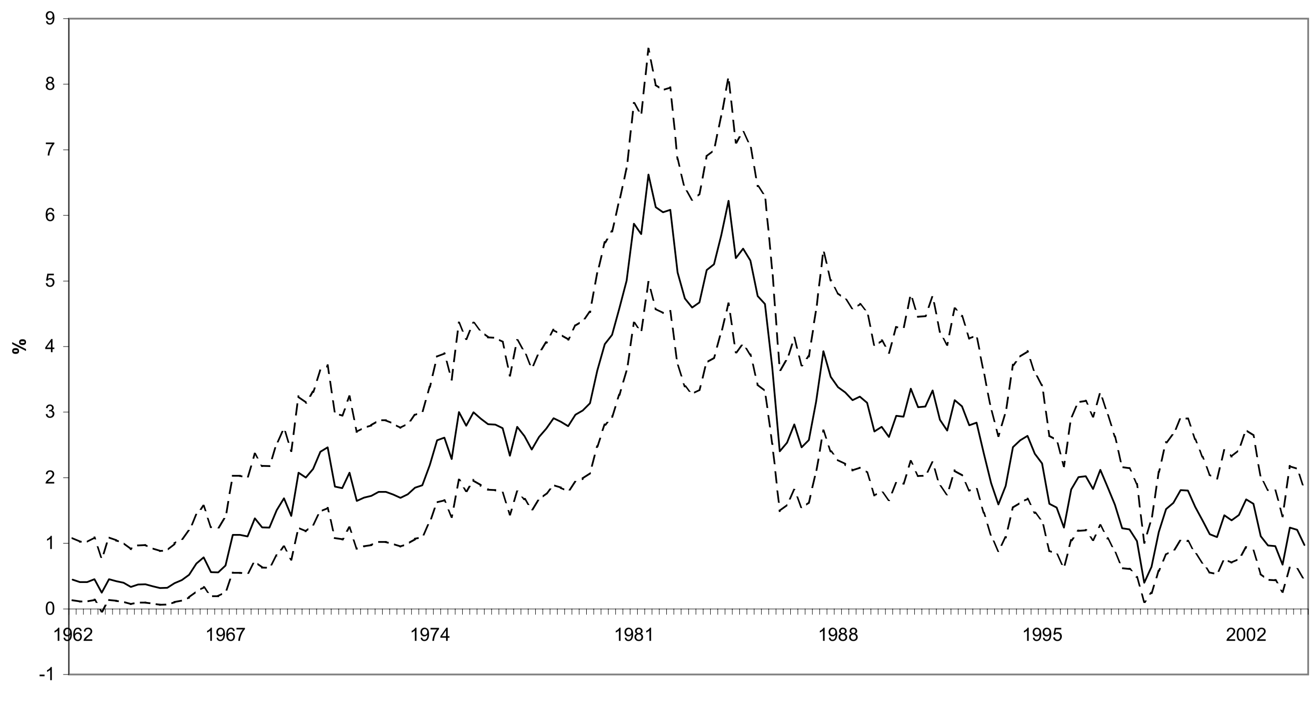


Chart 4(b) Real rate factor
(with 95% confidence band)

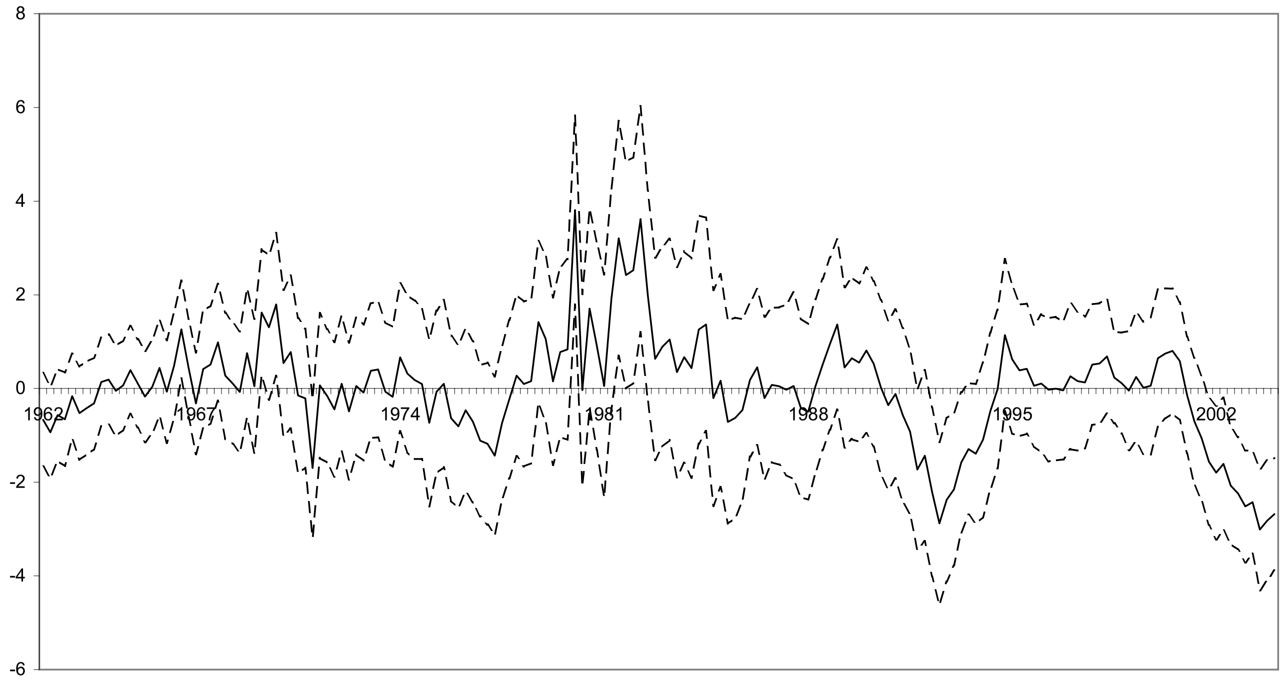


Chart 5(a): Output gap variability
(One step ahead estimate plus 95% confidence interval)

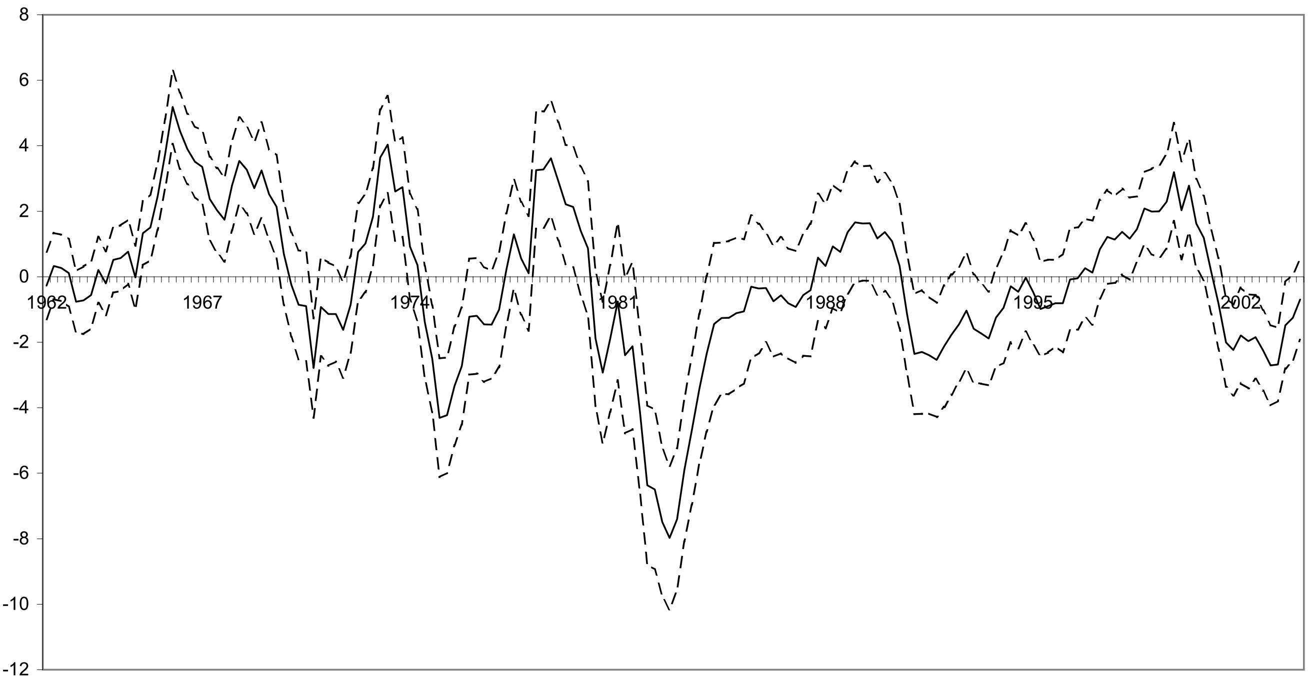


Chart 5(b): Inflation variability
(One step ahead estimate plus 95% confidence interval)

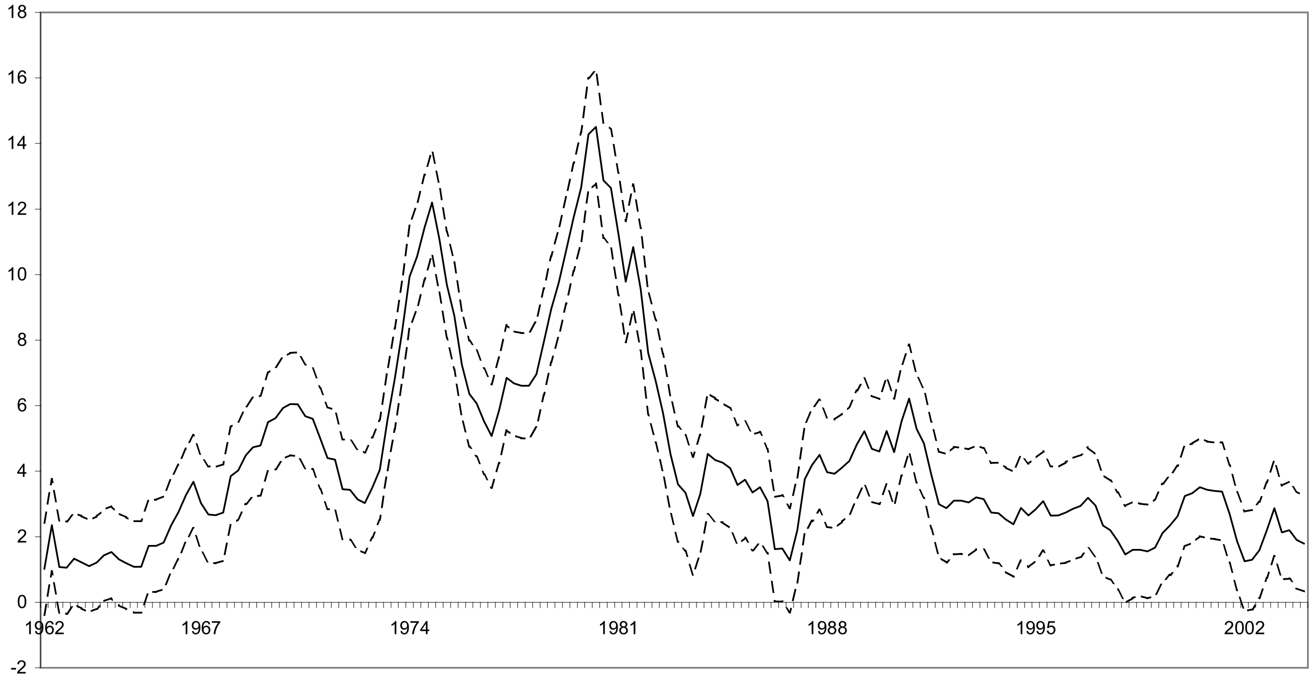


Chart 5(c): Spot rate variability
(One step ahead estimate plus 95% confidence interval)

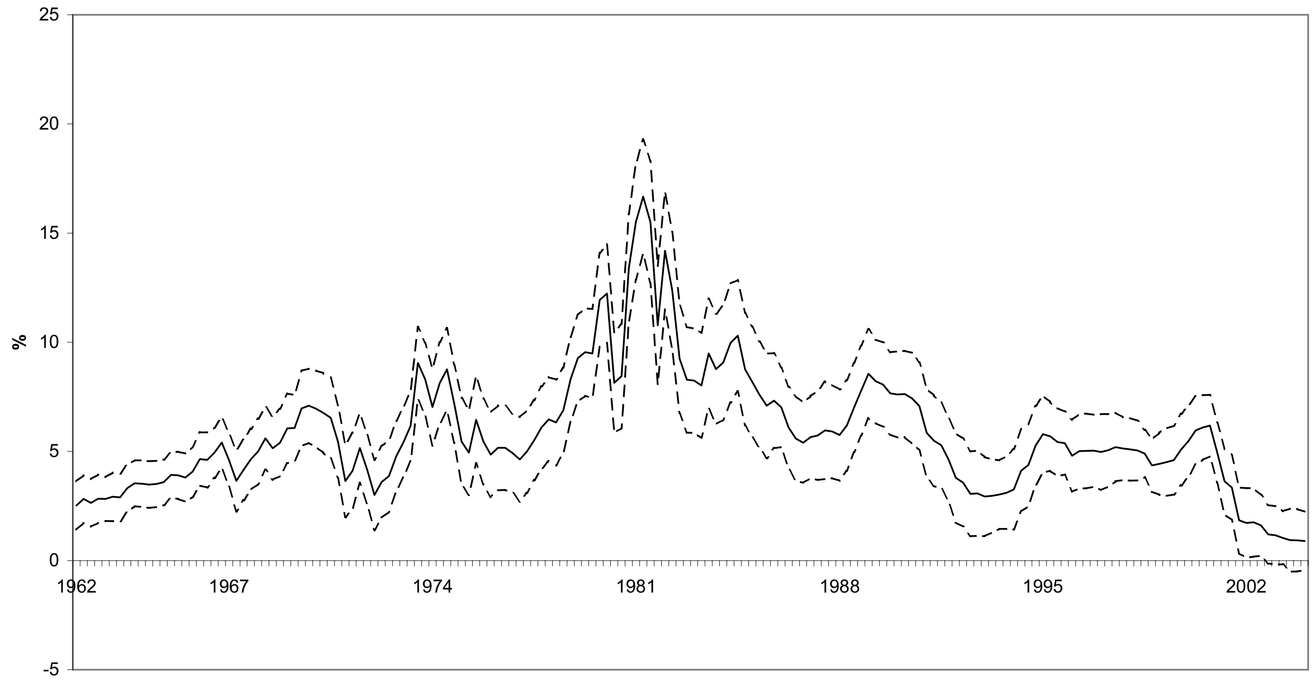
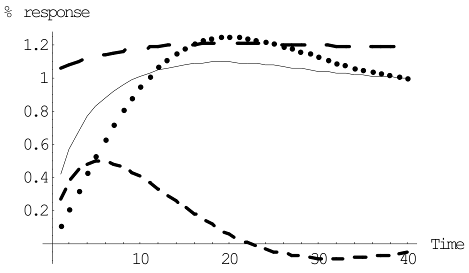
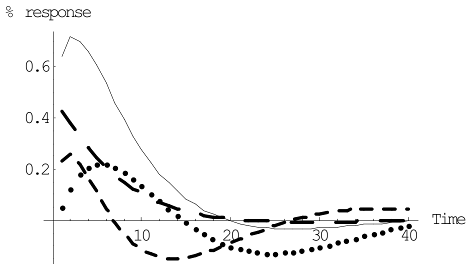


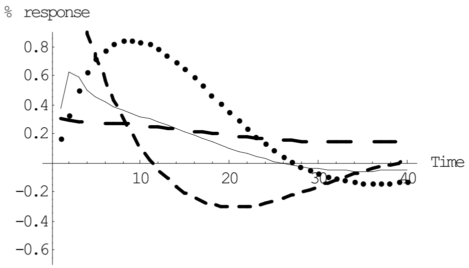
Chart 6: Model M1 macroeconomic impulse responses



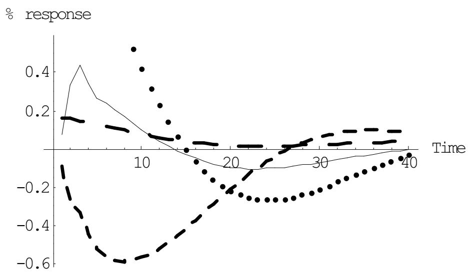
(i) Nominal factor (y_{π^*}) shock



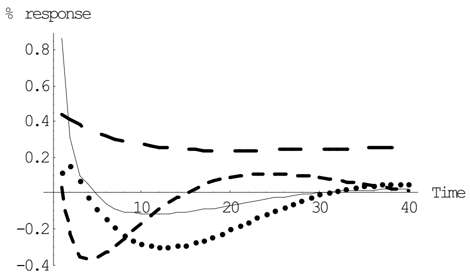
(ii) Real factor ($y_{\mathbf{r}^*}$) shock



(iii) Output shock



(iv) Inflation shock

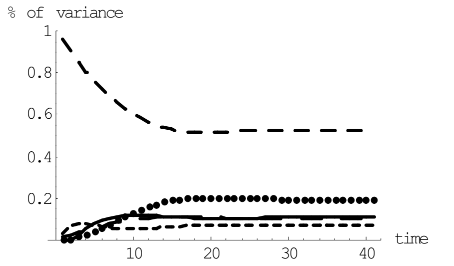


(iv) Spot rate shock

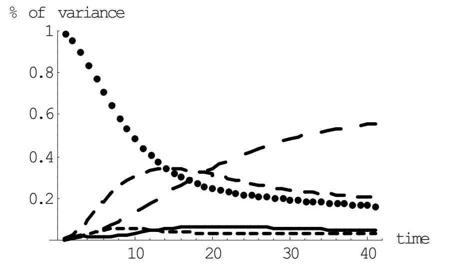
Key - effects on:
- - - - - output
.... inflation
——— spot rate
— — — 5 year yield

Each panel shows the effect of a shock to one the five orthogonal innovations (ε, η) shown in (1) and (2). These shocks increase the each of the five driving variables in turn by one percentage point compared to its historical value for just one period. Since y_{π^*} is a martingale, the first shock (ε_1) has a permanent effect on inflation and interest rates, while other shocks are transient. The dashed line shows the effect on output, the dotted line the effect on inflation, the continuous line the effect on the spot rate and the dot-dash line the effect on the 10 year yield. Elapsed time is measured in quarters.

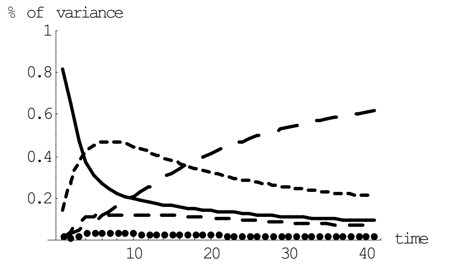
Chart 7: Model M1 Analysis of Variance



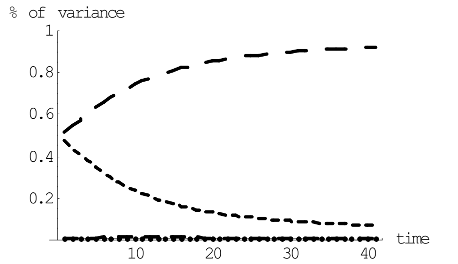
(i) Output variance



(ii) Inflation variance



(iv) Spot rate variance

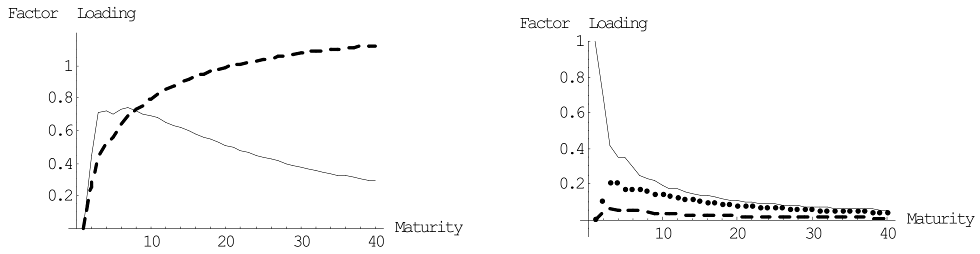


(i) Variance of 10 year yield

Key-% of variance due to orthogonal innovations in:
- - - nominal factor x_1 — — — real factor x_2
—- output .... inflation – spot rate

Each panel shows the contribution to total variance of innovations in each of the orthogonal shocks representing innovations in each of the five driving variables. Elapsed time is measured in quarters.

Chart 8: Factor loadings in model M1



(i) Nominal (y_{π^*}) and real ($y_{\mathbf{r}^*}$) factors

(ii) Output (- - -), inflation (....) and spot rate (—)

Panels (i) and (ii) show the effect of orthogonal shocks to the financial factors (η) and macro variables (ε) respectively. These shocks increase each of these driving variables in turn by one percentage point compared to its historical value for one period. Maturity is measured in quarters.

Chart 9: Risk premia (on 15 year maturity)