A joint econometric model of macroeconomic and term structure dynamics*

Peter Hördahl, Oreste Tristani and David Vestin**
European Central Bank

First draft: 7 February 2003
This draft: 15 September 2003

Abstract

We construct and estimate a joint model of macroeconomic and yield curve dynamics. A small-scale backward/forward-looking rational expectations model describes the macroeconomy. Bond yields are affine functions of the state variables of the macromodel, and are derived assuming absence of arbitrage opportunities and a flexible price of risk specification. While maintaining the tractability of the affine set-up, our approach provides a way to interpret yield dynamics in terms of macroeconomic fundamentals; time-varying risk premia, in particular, are associated with the fundamental sources of risk in the economy. In an application to German data, the model is able to capture salient features of the term structure of interest rates and its forecasting performance matches that of the best available models based on latent factors. The model has also considerable success in accounting for the empirical failure of the expectations hypothesis.

Keywords: Affine term-structure models, policy rules, new neo-classical synthesis

1 Introduction

Understanding the term structure of interest rates has long been a topic on the agenda of both financial and macro economists, albeit for different reasons. On the

*We wish to thank Gianni Amisano, Greg Duffee, Ivan Paya, Ken Singleton and Lars Svensson, as well as seminar participants in the SIEPR-FRBSF 2003 Conference on "Finance and Macroeconomics," the University of Brescia Conference on "Macroeconomic issues in the EMU," the EFA2003 Meetings, the EEA2003 Meetings and the MMF2003 Conference, for helpful comments and suggestions. Remaining errors are our sole responsibility. The opinions expressed are personal and should not be attributed to the European Central Bank.
** European Central Bank, DG Research, Kaiserstrasse 29, D - 60311, Frankfurt am Main, Germany. E-mail: peter.hordahl@ecb.int, oreste.tristani@ecb.int, david.vestin@ecb.int.

one hand, financial economists have mainly focused on forecasting and pricing interest rate related securities. They have therefore developed powerful models based on the assumption of absence of arbitrage opportunities, but typically left unspecified the relationship of the term structure with other economic variables. Macro economists, on the other hand, have focused on understanding the relationship between interest rates, monetary policy and macroeconomic fundamentals. In so doing, however, they have typically relied on the "expectations hypothesis," in spite of its dismal empirical record. Combining these two lines of research seems fruitful, in that there are potential gains going both ways. If macroeconomic theory has some empirical success, it should help price securities more efficiently. Likewise, if some tenets of financial economics, such as the requirement of arbitrage-free markets, are empirically important, taking them into account should help to explain the response of the yield curve to macroeconomic developments. Yield curve information could also help sharpening forecasts of future economic activity and inflation.

This paper aims at presenting a unified empirical framework where a small structural model of the macro economy is combined with an arbitrage-free model of bond yields. In doing so, we build on the work of Piazzesi (2001) and Ang and Piazzesi (2003), who introduce macroeconomic variables into the standard affine term structure framework based on latent factors – e.g. Duffie and Kan (1996) and Dai and Singleton (2000). The main innovative feature of our paper is that we use a structural macroeconomic framework rather than starting from a reduced-form, VAR representation of the data. One of the advantages of this approach it to allow us to relax Ang and Piazzesi's restriction that inflation and output be independent of the policy interest rate, thus facilitating an economic interpretation of the results. Our framework is similar in spirit to that in Wu (2002), who prices bonds within a calibrated rational expectations macro-model. The difference is that we estimate our model and allow a more empirically oriented specification of both the macro economy and the parametrization of the market price of risk. Our framework is also related to that employed in a recent paper by Rudebusch and Wu (2003), who interpret latent term structure factors in terms of macroeconomic variables.

We start from a fairly general, linear macroeconomic set-up, which can allow for explicit forward-looking elements. Standard methods can be applied to solve the model, i.e. to obtain, amongst other results, the short term interest rate as a function of the predetermined variables of the model, and a law of motion for the evolution of the predetermined variables. Next, we follow the finance literature and build an arbitrage-free term structure of nominal interest rates based on the solution of the macro-model and an assumption on the market prices of risk. The result is an affine, multifactor term structure model, which can be estimated jointly with the macroeconomic system using maximum likelihood methods.

In our empirical application, we use a simple model whose core is based on the so-called new neoclassical synthesis (Goodfriend and King, 1997; Rotemberg and Woodford, 1997) and a simple monetary policy rule. Variants of this model have been successfully employed for policy analysis and to explain empirical macroeconomic dynamics, including those of the short term interest rate (e.g., Clarida, Galí and Gertler, 2000, Rudebusch and Svensson, 1999, Smets, 2000, Rudebusch, 2002a). Based on this model, the factors that affect the short-term interest rate correspond to shocks with a standard macroeconomic interpretation. Three of them are almost always present in small macro models: inflation, output (or "aggregate demand") and monetary policy shocks; the fourth is a time-varying inflation target, introduced to account for perceived changes in the monetary policy objective that may have occurred over relatively long periods of time.

Our estimation results, based on German data, show that macroeconomic factors affect the term-structure of interest rates in different ways. Monetary policy shocks have a marked impact on yields at short maturities, and a small effect at longer maturities. Inflation and output shocks mostly affect the curvature of the yield curve at medium-term maturities. Changes in the perceived inflation target have more lasting effects and tend to have a stronger impact on longer term yields. The impulse responses of the macro variables to these shocks are broadly in line with comparable results in the literature.

Having established that the model provides a sensible description of macroeconomic and term structure dynamics, we turn to evaluate its performance relative to other available affine term structure models. More specifically, we focus on its in-sample and out-of-sample forecasting performance compared to models (partly) based on unobservable factors. In this exercise we use two main benchmarks: the Duffee (2002) model, which has been shown to do relatively well in forecasting, and the Ang and Piazzesi (2003) model. In spite of its relatively large dimension, our model does quite well. Both in-sample and out-of-sample, its performance is comparable to that of the best available affine term structure models, and in many cases superior.

Finally, we turn to examine the model's ability to account for the features of the data which represent a puzzle for the expectations hypothesis, namely the negative and large – rather than positive and unit – coefficients obtained, for example by Campbell and Shiller (1991), in regressions of the yield change on the slope of the curve. Dai and Singleton (2002a) have shown that some affine models based on unobservable factors can be very successful along this dimension, because they incorporate a flexible specification of the market prices of risk as functions of unobservable factors. For our model, the test is somewhat more stringent, as the market prices of risk are mostly functions of observable macroeconomic variables. Nevertheless, our model also proves successful in this respect, in the sense of being fully consistent with the negative slope coefficient typically found in the data. In addition, regressions based on risk-adjusted yields do, by and large, recover slope coefficients close to unity, which should be expected when risk premia are modelled accurately. This result also highlights that there are important interactions between the various market prices of risk: the possibility of changes in the inflation target, for example, commands a price of risk which depends not just on the level of the inflation target, but also on the level of inflation and the output gap.

The rest of the paper is organized as follows. Section 2 describes the main features of our general theoretical approach and then provides a brief overview of our estimation method. The specific macroeconomic model which we employ in our application to German data is described in Section 3. Section 4 then illustrates our empirical results, in terms of parameter estimates, impulse response functions and forecast error variance decompositions. This section also discusses the forecasting performance of our model, compared to leading available alternatives. The ability of the model to solve the expectations puzzle is tested in Section 5. Section 6 concludes.

2 The approach

In recent years, the finance literature on the term structure of interest rates has made tremendous progress in a number of directions (see e.g. Dai and Singleton, 2002b). Following the seminal paper by Duffie and Kan (1996), one of the most successful avenues of research has focused on models where the yields are affine functions of a vector of state variables. Refinements of such models have made them increasingly successful in capturing important features of the dynamics of the term structure of interest rates. This literature, however, has typically not investigated the connections between term structure and macroeconomic dynamics. In the rare cases in which macroeconomic variables–notably, the inflation rate–have been included in estimated term-structure models, those variables have been modelled exogenously (e.g. Evans, 2003, Zaffaroni, 2001). The interactions between macroeconomic and term structure dynamics have also been left unexplored in the macroeconomic literature, in spite of the fact that simple "policy rules" have often scored well in describing the dynamics of the short-term interest rate (e.g., Clarida, Galí and Gertler, 2000).

An attempt to bridge this gap within an estimated, arbitrage-free framework has recently been made by Ang and Piazzesi (2003). Those authors estimate a term structure model based on the assumption that the short term rate is affected partly by macroeconomic variables, as in the literature on simple monetary policy rules, and partly by unobservable factors, as in the affine term-structure literature.[1] Ang and Piazzesi's results suggest that macroeconomic variables have an important explanatory role for yields and that the inclusion of such variables in a term structure model can improve its one-step ahead forecasting performance. Nevertheless, unobservable factors without a clear economic interpretation still play an important role in their model. Moreover, Ang and Piazzesi's two-stage estimation method relies on the assumption that the short term interest rate does not affect macroeconomic variables, an assumption which precludes any meaningful role for monetary policy and thus reduces the scope for a full understanding of the interaction between monetary policy, the macroeconomy and the term structure of interest rates.

In order to redress these shortcomings, we construct a dynamic term structure model entirely based on macroeconomic factors, which allows for an explicit feedback from the short term (policy) rate to macroeconomic outcomes. The joint modelling of three key macroeconomic variables–namely, inflation, the output gap and the short term "policy" interest rate–should allow us to obtain a more accurate (endogenous) description of the dynamics of the short term rate. At the same time, our explicit modelling or risk premia should also help us in capturing the dynamics of the entire term-structure.

In this section, we present the general approach that we use to model jointly the macroeconomy and the term structure. The main assumption we impose is that aggregate macroeconomic relationships can be described using a linear framework.

[1] In related papers, Dewachter and Lyrio (2002) and Dewachter, Lyrio and Maes (2002) also estimate jointly a term structure model built on a macroeconomic VAR. They include a different macroeconomic specification, in that the model is not closed with a monetary policy rule but with an exogenous long-run relationship between the equilibrium values of the short term rate, inflation and the outut gap.

Based on such a structure, the macroeconomic model can be solved by numerical methods that are also standard in the macroeconomic literature. A dynamic term structure can then be built on the model solution in a straightforward manner, based on the assumption of absence of arbitrage opportunities.

2.1 The macroeconomic set-up

Our starting point is a macroeconomic model that can be written in the general form

$$\begin{bmatrix} \mathbf{X}_{1,t+1} \\ \mathbf{X}_{2,t+1} \end{bmatrix} = \mathbf{H} \begin{bmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \end{bmatrix} + \mathbf{K} r_t + \boldsymbol{\xi}_{t+1}$$

where $\mathbf{X}_1$ is a vector of predetermined variables, $\mathbf{X}_2$ includes the variables which are not predetermined, r_t is the policy instrument and $\boldsymbol{\xi}_1$ is a vector of independent, normally distributed shocks (see the appendix for an example related to the model we use in Section 3). The policy instrument, a short-term rate, will be set according to a simple policy rule of the form

$$r_t = -\mathbf{F} \begin{bmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \end{bmatrix}.$$

This linear structure is nevertheless general enough to accommodate a large number of standard macroeconomic models. The main restriction we impose, for simplicity, is that only the short-term interest rate, which is controlled by the central bank, affects the macro economy, whereas longer rates do not. This allows us to solve the latter without prior knowledge of the equilibrium prices of nominal bonds.

The solution of the model above can be obtained numerically following standard methods. We choose, in particular, the methodology described in Söderlind (1999), who proposes a solution algorithm based on the Schur decomposition. The result are two matrices $\mathbf{M}$ and $\mathbf{C}$ such that $\mathbf{X}_{1,t} = \mathbf{M}\mathbf{X}_{1,t-1} + \boldsymbol{\xi}_{1,t}$ and $\mathbf{X}_{2,t} = \mathbf{C}\mathbf{X}_{1,t}$.[2] In addition, the solution will include an expression relating the short term rate to the predetermined variables, $r_t = \boldsymbol{\Delta}'\mathbf{X}_{1,t}$, where $\boldsymbol{\Delta}' \equiv -(\mathbf{F}_1 + \mathbf{F}_2\mathbf{C})$ and $\mathbf{F}_1$ and $\mathbf{F}_2$ are partitions of $\mathbf{F}$ conformable with $\mathbf{X}_{1,t}$ and $\mathbf{X}_{2,t}$. Focusing on the short-term (policy)

[2] The presence of non-predetermined variables in the model implies that there may be multiple solutions for some parameter values. We constrain the system to be determinate in the iterative process of maximizing the likelihood function.

interest rate, the solution can be written as

$$\begin{aligned} r_t &= \mathbf{\Delta}'\mathbf{X}_{1,t} \\ \mathbf{X}_{1,t} &= \mathbf{M}\mathbf{X}_{1,t-1} + \boldsymbol{\xi}_{1,t}. \end{aligned} \tag{1}$$

2.2 Adding the term structure to the model

The system (1) expresses the short term interest rate as a linear function of the vector $\mathbf{X}_1$, which in turn follows a first order Gaussian VAR. This structure is formally equivalent to that on which affine models are normally built, albeit with the difference that $\mathbf{X}_1$ is here a state-space vector. To build the term structure, we only need to impose the assumption of absence of arbitrage opportunities, which guarantees the existence of a risk neutral measure, and to specify a process for the stochastic discount factor.

Behind this formal equivalence, however, our model has the distinguishing feature that both the short rate equation and the law of motion of vector $\mathbf{X}_1$ have been obtained endogenously, as functions of the parameters of the macroeconomic model. This contrasts with the standard affine set-up, where both the short rate equation and the law of motion of the state variables are postulated exogenously.

Rather than building the term structure directly on equations (1), however, we allow for the possibility to write bond yields as functions of a different vector, $\mathbf{Z}_t$, which can include any variable in $\mathbf{X}_t$ or the short term rate. The new vector $\mathbf{Z}_t$ is defined as $\mathbf{Z}_t = \mathbf{D}\mathbf{X}_t$, where $\mathbf{D}$ is a selection matrix described in the appendix. This transformation allows us to write the yields as functions of any variable in $\mathbf{X}_t$, rather than of the predetermined variables only, which can be more intuitively appealing, depending on the specifics of the model. Obviously, $\mathbf{Z}_t$ can be immediately rewritten as a function of the predetermined vector $\mathbf{X}_{1t}$ using the solution (1). This yields $\mathbf{Z}_t = \hat{\mathbf{D}}\mathbf{X}_{1,t}$, where $\hat{\mathbf{D}} \equiv \mathbf{D}_1 + \mathbf{D}_2\mathbf{C}$ and $\mathbf{D}_1$ and $\mathbf{D}_2$ are partitions of $\mathbf{D}$ conformable with $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$.

At this point, we can define the nominal pricing kernel m_{t+1}, which prices all nominal bonds in the economy, as

$$m_{t+1} = \exp\left(-r_t\right)\frac{\psi_{t+1}}{\psi_t},$$

where ψ_{t+1} is the Radon-Nikodym derivative, which is assumed to follow a lognormal process according to

$$\psi_{t+1} = \psi_t \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\boldsymbol{\xi}_{1,t+1}\right),$$

and where λ_t is the vector of market prices of risk associated with the underlying sources of uncertainty ($\boldsymbol{\xi}_{1,t+1}$) in the economy. The market prices of risk, in turn, have commonly been assumed to be constant (in the case of Gaussian models) or proportional to the factor volatilities (e.g. Dai and Singleton (2000)), but recent research has highlighted the clear benefits in allowing for a more flexible specification of the risk prices (e.g. Duffee (2002), Dai and Singleton (2002a)). We therefore assume that the market prices of risk are affine in the state vector $\mathbf{Z}_t$,

$$\lambda_t = \lambda_0 + \lambda_1 \mathbf{Z}_t,$$

so that the market's required compensation for bearing risk can vary with the state of the economy.

Given these assumptions, the appendix shows that the continuously compounded yield y_t^n on an n-period zero coupon bond is given by

$$y_t^n = A_n + B_n' \mathbf{Z}_t,$$

where the A_n and B_n' matrices can be derived using recursive relations. Stacking all yields in a vector $\mathbf{Y}_t$, we write the above equations jointly as $\mathbf{Y}_t = \mathbf{A} + \mathbf{B}'\mathbf{Z}_t$ or, equivalently, $\mathbf{Y}_t = \mathbf{A}_n + \tilde{\mathbf{B}}_n' \mathbf{X}_{1,t}$, where $\tilde{\mathbf{B}}_n' \equiv \mathbf{B}_n' \hat{\mathbf{D}}$.

2.3 Maximum likelihood estimation

In order to estimate the model, we need to distinguish first between observable and unobservable variables in the $\mathbf{X}_t$ vector. We therefore define the new vectors $\mathbf{X}_{1t}^o$ and $\mathbf{X}_{2t}^o$ and $\mathbf{X}_{1t}^u$ and $\mathbf{X}_{2t}^u$, which only include observable and unobservable variables, respectively.

With this notation, the likelihood function to be maximised can be written as

$$\ell(\boldsymbol{\theta}) = \prod_{t=2}^{T} f\left(\mathbf{X}_{1t}^o, \mathbf{X}_{2t}^o, \mathbf{Y}_t | \mathbf{X}_{1t-1}^o, \mathbf{X}_{2t-1}^o \mathbf{Y}_{t-1}\right).$$

To construct the likelihood, we need to write the distribution of $\mathbf{X}_{2t}^o$ and $\mathbf{Y}_t$ in terms of the known distribution of $\mathbf{X}_{1t}$. We adopt an approach which is common in the finance literature and which involves inverting the relationship between yields and unobservable factors (Chen and Scott, 1993). In our case, the method needs to be extended to take into account that the observation equations of the state-space system include not just the yields, $\mathbf{Y}_t$, but also the non-predetermined variables, $\mathbf{X}_{2t}^o$.

We therefore assume that as many yields plus non-predetermined variables as unobservable states are measured exactly, and that the remaining yields are measured with error. To account for the measurement error, we rewrite the yields equation as $\mathbf{Y}_t = \mathbf{A}_n + \tilde{\mathbf{B}}_n' \mathbf{X}_{1,t} + \mathbf{B}^m \mathbf{u}_t^m$, where the $\mathbf{u}_t^m$ vector of white noise shocks has zero elements corresponding to the yields measured exactly. We then partition the state vector into observable and unobservable components and rewrite the yield equation and the equation for the non-predetermined variables as $\mathbf{Y}_t = \mathbf{A} + \tilde{\mathbf{B}}^o \mathbf{X}_{1t}^o + \tilde{\mathbf{B}}^u \mathbf{X}_{1t}^u + \mathbf{B}^m \mathbf{u}_t^m$ and $\mathbf{X}_{2t}^o = \tilde{\mathbf{C}}^o \mathbf{X}_{1t}^o + \tilde{\mathbf{C}}^u \mathbf{X}_{1t}^u$, respectively. These equations can be used to back out the unobservable states, $\mathbf{X}_{1t}^u$, and the measurement shocks, $\mathbf{u}_t^m$.

Using the assumption of orthogonality of measurement error shocks and shocks to the unobservable states, we show in the appendix that the log-likelihood function to maximize takes the form

$$\begin{aligned} \ln\left(\pounds\left(\boldsymbol{\theta}\right)\right) &= -\left(T-1\right)\left(\ln|J| + \frac{1}{2}\ln\left|\Sigma\Sigma'\right| + \frac{1}{2}\ln\sum_{i=1}^{n_m}\sigma_i^2 + \frac{n_1+n_m}{2}\ln\left(2\pi\right)\right) \\ &\quad -\frac{1}{2}\sum_{t=2}^{T}\left(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1}\right)'\left(\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\right)^{-1}\left(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1}\right) - \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{n_m}\frac{\left(u_{t,i}^m\right)^2}{\sigma_i^2} \end{aligned}$$

where $J \equiv \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{bmatrix}$, n_m is the number of measurement errors and n_1 is the number of predetermined variables.

When, as in the model used by Ang and Piazzesi (2003), there is no feedback from interest rates to the macro variables, estimation can be performed with a two-step procedure. In the more general case analysed here this is not possible, because we do not assume independence between the macro variables and the short-term interest rate. Furthermore, we back out our unobservable macroeconomic variables from the yields. We must therefore estimate the whole system jointly.

In theory, this is of course preferable. The problem is that the parameter space is quite large and therefore the optimization problem of maximizing the likelihood function is non-trivial and time consuming. We employ the method of simulated annealing, introduced to the econometric literature by Goffe, Ferrier and Rogers (1994). The method is developed with an aim towards applications where there may be a large number of local optima.[3]

3 A simple backward/forward looking model

This section presents a stylised model which fits in the general structure described in Section 2. The model will be used in Section 4 for an empirical application to German data.

Consistently with the approach described in Section 2, we do not aim to provide a fully-fledged micro-founded model. Rather, we present an empirically plausible structural model that we motivate by highlighting the assumptions that can be adopted to derive it from first principles.

The model of the economy includes just two equations which describe the evolution of inflation, π_t, and the output gap, x_t. The inflation equation is given by

$$\pi_t = \mu_\pi E_t\left[\pi_{t+1}\right] + (1 - \mu_\pi)\,\pi_{t-1} + \delta_x x_t + \varepsilon_t^\pi.$$

This equation can be derived as the first order condition of the price-setting decision of firms acting in an environment with monopolistic competition. Monopolistic competition implies that prices will be set as a markup on marginal cost, which explains the presence of the output gap term in the equation. The assumption of sticky prices generates the expected inflation term, as firms to not know when their prices will adjust next and therefore need do maximize the sum of current and expected future profits. The additional lagged inflation rate can be motivated by partial price indexation (Christiano, Eichenbaum and Evans, 2001) or by the presence of a set of firms that use a backward-looking rule of thumb to set prices (Galí and Gertler, 1999).

The equation which describes the dynamics of the output gap is

$$x_t = \mu_x E_t x_{t+1} + (1 - \mu_x)\,x_{t-1} - \zeta_r\left(r_t - E_t\left[\pi_{t+1}\right]\right) + \varepsilon_t^x.$$

[3] The estimates reported in the text correspond to a maximum value of the likelihood function found in a process of 100 estimations using simulated annealing, starting from randomised initial values.

This equation can be derived from an intertemporal consumption Euler equation. The first term on the right-hand side is essentially Hall's (1978) random walk hypothesis which states that consumption is equal to expected consumption tomorrow (in simple, closed-economy models, consumption equals output in equilibrium). This hypothesis is supplemented with two additional terms. First, a real interest rate (which he assumed to be constant) shifts the consumption profile such that a real rate increase tends to discourage current consumption. The second term is lagged consumption, whose presence can be motivated by habit persistence and/or the presence of rule of thumb consumers (Campbell and Mankiw, 1989; Fuhrer, 2000; McCallum and Nelson, 1999).

As we will employ monthly data in estimation, the next step is to recast the model at a monthly frequency, which we do this along the lines of Rudebusch (2002a).

$$\begin{aligned} \pi_t &= \frac{\mu_\pi}{12}\sum_{i=1}^{12} E_t\left[\pi_{t+i}\right] + (1-\mu_\pi)\sum_{i=1}^{3}\delta_{\pi i}\pi_{t-i} + \delta_x x_t + \varepsilon_t^\pi \\ x_t &= \frac{\mu_x}{12}\sum_{i=1}^{12} E_t\left[x_{t+i}\right] + (1-\mu_x)\sum_{i=1}^{3}\zeta_{xi}x_{t-i} - \zeta_r\left(r_t - E_t\left[\pi_{t+1}\right]\right) + \varepsilon_t^x \end{aligned}$$

Note that all variables are now expressed at the monthly frequency, with inflation defined as the 12-month change of the log-price level. In particular, the two equations include a forward-looking term capturing expectations over the next twelve months of inflation and output, respectively. The backward-looking components of the two equations are restricted to include only 3 lags of the dependent variable. This choice results in a more parsimonious empirical model. In the estimation, we impose $\mu_\pi + (1-\mu_\pi)\sum_i \delta_{\pi i} = 1$, a version of the natural rate hypothesis.

Finally, we need an assumption on how monetary policy is conducted in order to solve for the rational expectations equilibrium. Since our estimates will include also bond prices, we focus on private agents' perceptions of the monetary policy rule followed by the central banks, rather than solving the models under full commitment or discretion. Accordingly, the "simple rule" supposedly followed by the central bank is to set the nominal short rate according to

$$r_t = (1-\rho)\left(\beta\left(E_t\left[\pi_{t+11}\right] - \pi_t^*\right) + \gamma x_t\right) + \rho r_{t-1} + \eta_t. \qquad (2)$$

where π_t^* is the unobservable, perceived inflation target and η_t is a "monetary policy shock".

This is consistent with the formulation in Clarida, Galí and Gertler (1998, henceforth CGG), which is a natural benchmark for comparison because the rule has been estimated for Germany, the country which we focus on in the empirical implementation. The term multiplied by $(1-\rho)$ is a typical Taylor-type rule (in this case forward looking), where the rate responds to deviations of expected inflation from the time-varying inflation target. The second part of the rule is motivated by interest rate smoothing concerns, which seem to be an important empirical feature of the data.

The main difference with respect to the rule estimated by CGG is that we also allow for a time-varying, rather than constant, inflation target π_t^*. We adopt this formulation because the Bundesbank modified its "medium term price norm" over the sample period used in our analysis. The modifications were public knowledge, since they were announced every year as an input in the derivation of the monetary targets. The time-varying inflation target π_t^* should therefore capture such changes, as perceived by the market and reflected in equilibrium bond yields. This formulation allows us to exploit the full available sample period, without having to assume a break in the policy rule at some point in the late seventies, as done by CGG.

Finally, we need to specify the processes followed by the stochastic variables of the model, i.e. the perceived inflation target and the three structural shocks. We assume that our 3 macro shocks are serially uncorrelated and normally distributed with constant variance.[4] The only factor that we allow to be serially correlated is the unobservable inflation target, which will follow an AR(1) process

$$\pi_t^* = \phi_\pi \pi_{t-1}^* + u_{\pi,t}$$

where $u_{\pi,t}$ is a normal disturbance with constant variance uncorrelated with the other structural shocks. This assumption proxies the idea that the inflation target is perceived to be constant *in expected terms*, even if not deterministically, so that $E_t\left[\pi_{t+i}^*\right] \simeq \pi_t^*$ for $i = 1, 2, ...$[5]

[4] In affine term structure models, latent variables are normally assumed to follow a VAR(1) process. Some restrictions must then be imposed on the parameters of the VAR to ensure econometric identification (Dai and Singleton, 2000). While having the advantage of flexibility, this approach would not be entirely consistent with our macroeconomic assumptions. If our macromodel does provide an accurate description of the dynamics of inflation, the output gap and the policy interest rate, then it must also capture all the serial correlation in the data and the three macro disturbances must be white noise.

[5] Imposing a unit root would violate the assumption of stationarity of the factors, an assumption maintained throughout the affine term structure literature.

4 An application to Germany

4.1 Data

Our data set runs from January 1975 to December 1998. The term structure data consists of monthly German zero-coupon yields for the maturities 1, 3 and 6 months, as well as 1, 3, and 7 years.[6] We assume that the 1-month rate and the 3-year yield are perfectly observable, while the other rates are subject to measurement error. Yields have been bootstrapped from on an original Bundesbank dataset of end-of-month raw prices, coupons and maturities.[7]

Concerning the macro data, we construct the year-on-year inflation series using the CPI (all items). For the output gap, we follow CGG and detrend the log of total industrial production (excluding construction) using a quadratic trend. The series is, however, constructed recursively, so that each datapoint is obtained by fitting a quadratic trend to the original series up to that point. We adopt this approach to ensure that our forecast at time t does not rely on information unavailable at that point in time. Both series refer to unified Germany from 1991 onwards and to West Germany prior to this date. The macroeconomic and term-structure series are shown in Figure 1.

4.2 Estimation results

To reduce the parameter space in our empirical application, we impose a number of restrictions on the coefficients of the market prices of risk. In the general set-up, we showed that the risk prices can be specified as $\lambda_t = \lambda_0 + \lambda_1 \mathbf{Z}_t$. In our application, $\mathbf{Z}_t$ includes the perceived inflation target and contemporaneous and lagged values of inflation, output and the short term rate. Given $\mathbf{Z}_t$, λ_t can obviously have nonzero elements only corresponding to time t variables, as lagged variables are no longer subject to surprise changes. This leaves only four potentially non-zero rows in the λ_0 and λ_1 matrices, corresponding to the perceived inflation target, the policy interest rate, inflation and the output gap. Next, we restrict λ_0 and λ_1 in the sense of allowing interactions only between prices of risk of contemporaneous variables, which leaves us with a 4×4 non-zero submatrix in λ_1. Finally, we follow Duffee (2002) and

[6] We do not use 10-year bonds because these are only available without breaks as of April 1986.

[7] The methodology is equivalent to that employed by Fama and Bliss (1987). We wish to thank Thomas Werner for providing us with the raw data and Vincent Brousseau for bootstrapping the term structures of zero-coupon rates.

set to zero all entries whose elements have a t-statistic lower than 1 in preliminary estimations.

As a result, we are left with the following non-zero elements in the matrices of prices of risk

$$\lambda_t = \begin{pmatrix} \lambda_{01} \\ \lambda_{02} \\ \lambda_{03} \\ \lambda_{03} \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{13} & \lambda_{14} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & 0 \\ 0 & \lambda_{42} & 0 & \lambda_{44} \end{pmatrix} \begin{pmatrix} \pi_t^* \\ r_t \\ \pi_t \\ x_t \end{pmatrix}.$$

4.2.1 Parameter estimates

Our paper appears to be the first to estimate a simple backward/forward-looking macroeconomic model for Germany. We therefore discuss our macroeconomic results separately, since they are of independent interest. The only study on Germany which we have found in the literature is the analysis of the Phillips curve by Jondeau and Le Bihan (2001), who estimate the German Phillips curve based on quarterly data, using a variety of specifications and two different estimation methods (GMM and maximum likelihood).

If we take into account that the macro model must also help to fit the dynamics of the term structure, our model does relatively well in explaining the joint evolution of inflation, the output gap and the policy interest rate. Specifically, the model parameters are estimated to be broadly in line with other available estimates solely based on macroeconomic variables. The presence of the yields in the model, and in the estimation, can apparently be accommodated without twisting the macromodel towards unreasonable parameter regions (see Table 1).

A more detailed review shows that our estimates for Germany confirm the US result that both forward looking and backward looking components tend to be necessary to explain inflation dynamics. Compared to available US results (based on quarterly data), we obtain a relatively low weight for forward looking elements in both the inflation and the output gap equations. Our $\mu_\pi = 0.13$, for example, compares to the value of 0.29 found by Rudebusch (2002a). Our point estimate is, nevertheless, within the range of values found by Jondeau and Le Bihan (2001) for German data,[8] but smaller than those recovered by Smets (2000) using annual data for the euro area ($\mu_\pi = 0.52$).

[8] The sample used by the Jondeau and Le Bihan (2001) runs from 1970Q1 to 1999Q4.

For the output gap, the forward looking component appears to play a more important role ($\mu_x = 0.30$). In this case, our estimates are in the range of values reported by Fuhrer and Rudebusch (2002a) for the US (and again based on quarterly data), but lower than available European estimates. Both Smets (2000) for the euro area and Chadha, Masson and Meredith (1992) for a panel of France, Germany and Italy find a higher degree of forward-lookingness (0.56 and 0.45, respectively). Unfortunately, no estimates based on German data appear to be available in the literature.

The elasticity of inflation to the output gap is estimated to be extremely small ($\delta_x = 0.0004$) compared to US estimates (e.g., Rudebusch's estimate on quarterly data is $\delta_x = 0.13$) and insignificantly different from zero. However, relatively small values are not uncommon in analyses based on German data. Depending on the specification and the estimation method, Jondeau and Le Bihan (2001) find values between 0 and 0.19. The sensitivity of the output gap to the real interest rate ($\zeta_r = 0.03$) is also small compared to other available results. Rudebusch (2002a) reports a value of 0.09 for the US and Smets (2000) a value of 0.06 for the euro area.

Concerning the parameters of the monetary policy rule, our results are broadly consistent with those of CGG. We find a somewhat high degree of interest rate smoothing (around $\rho = 0.98$, compared to the value of 0.91 reported in Table 1 of CGG) but, at the same time, a more aggressive equilibrium response to inflation deviations from target and to the output gap (our coefficients imply values of 2.4 and 1.4, respectively, compared to 1.31 and 0.25 in the baseline specification of CGG). The discrepancies may obviously also be due to differences in the estimation method (GMM in CGG), in the sample period (1979:4-1993:12 in CGG), in the specification of the policy rule (constant inflation target in CGG), and in the selected policy interest rate (interbank day-to-day rate in CGG). The net effect, however, is that of an essentially equivalent impact response to inflation and the output gap.

As to the other parameters, the autocorrelation coefficient of the inflation target process is very close to 1.[9] Concerning the term structure, our estimates of the standard deviations of the measurement errors are between 18 basis points for the 3-month rate and 20 or 21 basis points for the other yields. These values are broadly in line with the results of models based solely on unobservable factors.

[9] The parameter is constrained to be strictly smaller than 1 in the estimation.

4.2.2 Impulse response functions

Our structural model allows us to compute impulse response functions of macro variables and yields to the underlying macro shocks. This is particularly interesting for the monetary policy shock, whose identification and effects are the subject of a vast literature (see Christiano, Eichenbaum and Evans, 1999).

Before discussing the impulse response functions of our model, Figure 2 shows actual year-on-year inflation and the perceived inflation target extracted through the model. The target is characterized by a decreasing trend over the estimation period: it goes from approximately 4 percent in 1975 to below 2 percent at the end of 1998. This is broadly consistent with the time pattern of the "price norm" announced every year by the Bundesbank, that also fell from 4.5 to below 2 percent over the period. The perceived inflation target may perhaps seem a bit too volatile, but it is less variable than actual inflation: its sample standard deviation of 0.86% compares to 1.72% for actual inflation; the minimum and maximum of the target are 1.28 and 5.22 percent, compared to corresponding values of −0.44 and 6.48 percent for actual inflation.

Figures 3 to 6 show the impulse responses of the macroeconomic variables and the yields to the structural shocks.

We start with Figure 3, which displays the impulse responses to a shock to the perceived inflation target, which increases on impact by approximately 0.2 percentage points. The shock is obviously very persistent due to the high serial correlation of the inflation target process. Since the output gap is partly forward-looking, it jumps upwards immediately, and then continues to increase for up to 1.5 years after the shock, while the policy rate increases very slowly because of the high smoothing coefficient and the relatively modest initial increase in inflation. Consequently, real rates only increase above the baseline after 1 year, so both the output gap and inflation remain positive for a very long time. The yields jump up consistently with the anticipated tightening cycle of monetary policy. The size of the jump is increasing in maturity for maturities up to 3 years, and then decreasing, consistently with the ultimately mean reverting nature of the inflation target shocks.

Figure 4 shows the effect of a 45 basis points increase in the 1-month interest rate because of a monetary policy shock (the disturbance $u_{\eta,t}$). The effects of the shock are quite persistent over time, because of the high interest rate smoothing coefficient. The increase in the short term interest rate causes a progressive opening of a negative output gap, up to a peak of approximately 0.5 percentage points after just over one year. Inflation also falls after the shock, but by a tiny amount. Both the size and the timing of the inflation response are somewhat different from those normally obtained based on US data (e.g., Christiano, Eichenbaum and Evans, 1999). Nevertheless, the inflation response is broadly consistent with the results of VAR studies based on German (monthly) data: Sims (1992), Clarida and Gertler (1997) and Bernanke and Mihov (1997) all find negligible, or statistically insignificant, responses of prices or inflation (or even evidence of a "price puzzle"). Reflecting the marked, but temporary, nature of the monetary policy shock, the response of the term structure is decreasing in the maturity of the bonds. This response is quite similar to that obtained by Evans and Marshall (1996) for the US.

The impulse responses to an inflation shock and an output shock are shown in Figures 5 and 6, respectively.

Inflation shocks imply an increase of year-on-year inflation by approximately 0.25 percentage points. The short term rate begins to rise on impact, while output falls in anticipation of a tightening monetary policy cycle. The resulting stagflation generates a policy trade-off between inflation and output volatility, which is typical of cost-push shocks in these models. The short-term rate therefore increases up to a maximum of just 6 basis points after 1 year, which is sufficient to ensure a slow return to baseline of the other macrovariables. Bond yields also increase on impact, by an amount increasing in maturity up to 1-year maturities and then decreasing.

Finally, an output shock implies an increase of the gap by approximately 1.3 percentage points. Because of the tiny elasticity of inflation to the gap, inflation increases only marginally in response to the shock. Consequently, the policy interest rate increases little and very slowly, up to a peak of approximately 15 basis points after 1.5 years. According to the expectations hypothesis, one would expect 1 and possibly 2-year yields to increase, on impact, more than the 1-month rate. This is, however, not the case. Longer yields, on 3 and 7-year bonds, actually fall on impact. This surprising response is in fact to a large extent shaped by the dynamics of risk premia.

4.2.3 Macro shocks and risk premia

One of the advantages of our joint treatment of macroeconomics and term-structure dynamics is that we are able to derive the impulse response of theoretical risk premia to macro shocks, including the monetary policy shock.

As a preliminary step, we analyze the weights of each of the shocks of the model on the yields at various maturities, i.e. the B_n matrices (see Figure 7). These are interesting to compare to the results often obtained from models based on 3 unobservable factors, whose interpretation is typically given in terms of level, slope and curvature of the yield curve.

Figure 7 shows that the inflation target affects the maturities beyond 2 years almost uniformly, while it has a smaller effect on the short end of the curve. The inflation target therefore plays a mixed role of level factor for the long end of the curve and of slope factor for the short end. The role of level factor for long maturities is intuitively appealing from a macroeconomic perspective, as it identifies the inflation target with the nominal anchor of the economy.

At short maturities, the level and slope of the yield curve are crucially influenced by the monetary policy shock. This is also intuitively appealing, as it portrays the tight control of the central bank on short rates, due to liquidity effects.

Inflation shocks and output gap shocks have a hump shaped weight. They mostly affect the curvature of the yield curve at certain maturities, even if with a smaller standardized weight. These results are broadly consistent with those reported by Rudebusch and Wu (2003), who more explicitly link macroeconomic and term-structure factors.

The impact response of yield premia to the macroeconomic shocks are shown in Figure 8. The inflation target shock is immediately followed by an increase of the yield premium for maturities up to 4 years, with a peak effect of 10 basis points at the 1-year maturity. The premium then turns negative for longer maturities. Such increase in the yield premium is highly significant from an economic viewpoint, as it plays a large quantitative role in shaping the yield response displayed in Figure 3.

The monetary policy shock gives rise to a large fall, on impact, at the short end of the term structure of yield premia, thus reducing significantly the size of the impact response of the yields. The impact response of the 1-year yield to the monetary policy shock, for example, would increase by a half if yield premia were set equal to a constant.

Similar considerations hold for the impact response of yield premia to inflation and output shocks. The latter is notable, since the premia embody most of the action in the response. The impact response of the 7-year rate, for example, would change sign and essentially maintain the same absolute value, if risk premia were constant.

We conclude that yield premium dynamics have a nonnegligible effect on the impulse responses of yields to all macroeconomic shocks. An interpretation of the yield responses based on the expectations hypothesis may therefore be significantly biased.

Figure 9 shows the estimated yield premia over time. Two general features emerge. First, the premia tend to be decreasing over the sample in parallel to the fall in inflation, but then shoot up again in the aftermath of the ERM crisis in 1992-93. Second, the volatility of the premia is increasing in maturity, with the 7-year yield being most volatile. A high volatility of risk premia should be useful to match the feature of the data which represent a puzzle for the expectations hypothesis.

4.2.4 Forecast error variance decomposition

Our model attributes all yield curve and macro-variable movements to macroeconomic factors. A decomposition of the forecast error variance can give us information on which factors play the most important role in this respect. The forecast error variance decomposition for all the variables in our model is presented in Table 2.

The most striking pattern of the table is that, no doubt due to the near unit-root behavior of this variable, movements in the perceived inflation forecast explain the predominant part of the forecast error variance of all variables at long horizons (5 years and beyond). This result is more intuitive for the forecast error variance of inflation. However, it also applies to the output gap, because of the expansionary (contractionary) effects of inflation target increases (falls) through the short-term real interest rate.

At shorter horizons, up to 1 year ahead, the forecast error variance of the output gap is explained to a large extent by output and monetary policy shocks, as well as by changes in the perceived inflation target. Inflation is mostly explained by inflation shocks. Unexpected changes in the monetary policy instrument are instead mostly due to monetary policy and output shocks.

After the inflation target shock, the most important explanatory variable of the forecast error variance of the yield curve at short forecast horizons are monetary policy shocks. At the 1-year ahead horizon, monetary policy shocks explain 30% of the variance of 3-month rates, 17% of the variance of 1-year yields, 8% of the variance of 3-year yields and 4% of the variance of 7-year yields. Over the same horizon, inflation and output gap shocks together contribute to explain 11% of the variance of 3-month rates, 9% of the variance of 1-year yields, 2% of the variance of 3-year yields and less than 1% of the variance of 7-year yields.

4.3 Forecasting yields

The forecasting performance is a particularly interesting test of our macroeconomic-based term-structure model. Due to the relatively large number of parameters that needs to be estimated, the model could be expected to perform poorly with respect to more parsimonious representations of the data. In fact, the random walk model has been shown to provide yield forecasts that are particularly difficult to beat (Duffee, 2002). We therefore present in this section, results of the yield forecasting performance of our model compared to the random walk.

In addition to the random walk, we also consider forecasts based on two alternative models. The first is a canonical $A_0(3)$ essentially affine model based on unobservable factors.[10] Provided that risk premia are specified to be linear functions of the states, Duffee (2002) finds this model most successful in the class of admissible affine three factor models in terms of forecasting US yields. Apart from providing a benchmark for comparison, our results on the $A_0(3)$ model are of independent interest, since they highlight the performance of this model on a different data-set.

Our second benchmark for comparison, is the Ang and Piazzesi (2003) model, which we reestimate on our data-set. Based on Ang and Piazzesi's results, we use their favorite "Macro model" in this exercise, i.e. a model in which the interest rate responds to current inflation and output gap, as well as to 3 unobservable factors. A potentially important difference in our application of their model, however, is that we use inflation and the output gap directly in the estimation, rather than the principal components of real and nominal variables employed by Ang and Piazzesi (2003). This is arguably a more theory-based choice and it facilitates the comparison to our results.

For all models, we report in-sample forecasting performances (in terms of RMSE) based on the February 1975 - December 1998 period. Concerning the out-of-sample results, we reestimate the model over the period February 1975 - December 1994, and perform a series of 1 to 12 step ahead forecasts for all yields used in the estimation over the period January 1995 to December 1998. Each month, we update the information set, but we do not reestimate the model. Instead, we rely on the estimates up until end-1994. We choose this approach to limit the computational burden of the exercise. All results are therefore based on the same estimated parameters.

The results of the two forecast evaluation exercises are summarized in Tables

[10] For a definition of the $A_0(3)$ class of affine models, see Dai and Singleton (2000).

3 and 4, for the in-sample and out-of-sample cases, respectively. Lower values of the RMSE denote better forecasts. The best forecast at each maturity/horizon is highlighted in bold.

The in-sample exercise shows that, with only few exceptions, forecasts based on affine models beat the random walk at all horizons and maturities. This represents further evidence, based on German data, of the good performance of the essentially affine class models (with or without the inclusion of macro variables).

Within the affine class, the performance is mixed. Our model tends to do better for shorter maturities/horizons, while the Ang and Piazzesi (2003) model performs better at longer forecasting horizons. The performance of the $A_0\,(3)$ model is broadly comparable to the two macro-based alternatives. No model appears to stand out clearly.

The out-of-sample results, however, show that our model performs better than the alternatives for all maturities, at least beyond the very shortest forecast horizon. Moreover, our model beats the predictions of the random walk benchmark in almost all cases. This performance is noteworthy, in particular given the difficulties that the Ang and Piazzesi (2003) and the $A_0\,(3)$ models seem to have in this respect.

We conclude that the joint modelling of macroeconomic and term-structure dynamics is possible without incurring costs in terms of forecasting performance. In fact, our results suggest that the inclusion of macroeconomic variables within a structural framework has contributed to sharpening our ability of forecasting yields accurately out of sample.

5 Expectations hypothesis tests

According to the expectations hypothesis, the yield on an n-period zero-coupon bond should increase when the spread between the same yield and the short term rate (the "slope of the yield curve") widens. In fact, the projection of the yield change $y_{t+1}^{n-1} - y_t^n$ on the yield spread $(y_t^n - r_t)\,/\,(n-1)$ should yield a coefficient of 1. A number of empirical tests of this implication of the theory have, however, found a negative relationship. This pattern, which represents an expectations puzzle, appears to be particularly clear for United States data, where the relevant regression coefficient can be as big as -5 for 10-year bonds, according to e.g. Campbell and Shiller (1991), while the expectations hypothesis predicts a value of one for all maturities.

An interpretation of these results is that yields incorporate time varying risk premia of significant magnitudes, rather than just a constant premium as permitted by the expectations hypothesis. Using a highly stylised model, McCallum (1994) shows that an exogenous, stochastic term premium is, in principle, capable of causing deviations from 1 in the slope coefficient of the aforementioned regression. The actual size of the deviation will depend on both the stochastic properties of the term premium (see also Roberds and Whiteman, 1999) and the monetary policy rule followed by the central bank.

Our model includes explicitly both these features–while also being consistent with the hypothesis of absence of arbitrage opportunities. Contrary to the expectations hypothesis, it should therefore be capable of generating more realistic, i.e. negative, theoretical values for the coefficients of Campbell and Shiller regressions. Rather than showing that this is possible in simulation for some combination of parameter values, however, in this section we ask a more stringent question, namely whether it is possible for the specific set of parameter values obtained from our maximum likelihood estimation.

In so doing, we follow closely Dai and Singleton (2002a) who ask the same question within a number of dynamic affine term structure models based on unobservable factors. More specifically, we ask whether the model-implied, population coefficients ϕ_n in the regression

$$y_{t+1}^{n-1} - y_t^n = const. + \phi_n \left(y_t^n - r_t \right) / \left(n - 1 \right) + residual \tag{3}$$

match the values obtained from an OLS regression on actual yield data. The population coefficient are obtained assuming that the model parameters are true and then deriving the ϕ_n coefficients analytically based on the stochastic properties of the model.[11] Following Dai and Singleton (2002a), we also examine the small-sample counterparts of these coefficients. Some correction for small sample bias is desirable because of the persistent nature of yields. For this purpose, we generate 1000 samples of the same length of our data (287) and calculate the mean estimate of the ϕ_n coefficients.

Dai and Singleton (2002a) denote the above test as LPY(i). LPY(i) is a test of the capacity of the model to replicate the historical dynamics of yields as generated by a combination of the dynamics of risk premia and expectations of future short rates. As already emphasised, a successful model should be capable of generating

[11] This amounts to evaluating analytically $\phi_n \equiv \frac{cov\left(y_{t+1}^{n-1} - y_t^n, \left(y_t^n - r_t\right)/(n-1)\right)}{var\left(\left(y_t^n - r_t\right)/(n-1)\right)}$.

the negative intercept coefficient of Campbell and Shiller-type regressions.

In addition to LPY(i), Dai and Singleton (2002a) also suggest running a second sort of test, defined as LPY(ii), which focuses on the realism of the dynamic properties of risk premia. If the model captures these dynamics well, a Campbell and Shiller-type regression based on risk-premium-adjusted yield changes should recover the coefficient of unity consistent with the expectations hypothesis. LPY(ii) therefore tests that the population coefficient ϕ_n^* in the regression

$$y_{t+1}^{n-1} - y_t^n + e_{n,t}/\left(n-1\right) = const. + \phi_n^* \left(y_t^n - r_t\right)/\left(n-1\right) + residual \tag{4}$$

is equal to 1 (in the above regression, $e_{n,t}$ is the excess holding period return $e_{n,t} \equiv E_t\left[\ln\left(p_{t+1}^{n-1}/p_t^n\right) - r_t\right]$–see appendix).

Dai and Singleton (2002a) show that an affine 3-factor model with Gaussian innovations and including a risk-premium specification of the type suggested by Duffee (2002) scores extremely well in terms of both LPY(i) and LPY(ii). Our model also includes a flexible specification of the risk-premium as in Duffee (2002). Unlike in pure finance models, however, our risk-premia are partly functions of observable variables, namely lags of output and inflation. This feature represents an additional constraint in our model, which makes the LPY tests more stringent than in the pure finance literature.

5.1 *LPY*(i)

Since the evidence on Campbell and Shiller-type regressions based on European data is less compelling than for the US (e.g. Hardouvelis, 1994, Gerlach and Smets, 1997, Bekaert and Hodrick, 2001), we start by replicating Campbell and Shiller's analysis on our data. The results of the sample estimates of equation (3) are shown in Figures 10 and 11 as dots under the label "Sample". Consistently with the puzzle, the estimated intercept coefficient is always negative and decreasing in maturity. We confirm, however, that the puzzle appears less severe for German yield data: the estimated coefficient hovers around -0.7 for 7-year yields, compared to a value of less than -3 reported by Dai and Singleton (2002a) for US 7-year yields.

In figure 10 we show the results of the LPY(i) test. The population coefficients follow quite closely the pattern of the sample coefficients, although less so for short maturities. The population coefficients also have the downward sloping feature emphasised also by Dai and Singleton. The small-sample values of the ϕ_n coefficients (labelled “Model-implied MC” in Figure 10 and drawn together with 95% confidence bands of their small-sample distribution) confirm and strenghten this result. Our model appears to match strikingly well the pattern of the sample coefficients for essentially all maturities included in the regression.

The success of the model in matching $LPY\,(\mathrm{i})$ depends crucially on our assumptions related to the market prices of risk. Our parameterisation of the λ_1 matrix permits variations of the prices attached to the various sources of risk depending on the level of the state variables of the model. For example, the risk premium required for the possible occurrence of inflation target shocks varies with the levels of inflation and the output gap: the premium is larger when inflation and output are high with respect to the baseline. This interdependence imparts persistence to risk premia because output and inflation deviations from baseline tend themselves to be persistent over time.

5.2 $LPY\,(\mathrm{ii})$

Figure 11 shows the results of the $LPY\,(\mathrm{ii})$ tests. Once again, the model does remarkably well in fitting the data. The risk-premium correction always goes in the right direction and the model can generate a coefficient very close to unity for maturities of 4 year or longer.

For shorter maturities the model does less well, but we still recover coefficients that are positive and larger than 0.5, which is a dramatic improvement with respect to the implications of the expectations hypothesis. The reduced degree of success of the model at the short end of the yield curve is also consistent with the results of standard unobservable 3-factor models and it is typically attributed to the existence of short-lived money-market dynamics not captured in such models.

To summarize, our model appears to do as well as the essentially affine $A_0\,(3)$ class in tests of the expectations hypothesis, in spite of the further constraints imposed by the dependence of risk premia on observed variables. The results of $LPY\,(\mathrm{i})$ are very positive, in that the model can replicate the estimated coefficient of Campbell and Shiller-type regressions at all maturities. The test of $LPY\,(\mathrm{ii})$ suggest that a more satisfactory model of money market dynamics may be necessary to capture the time series properties of risk premia on short term bonds. This result is consistent with the evidence presented by Piazzesi (2001).

6 Conclusions

This paper presents a general set-up allowing to jointly model and estimate a macroeconomic-plus-term-structure model. The model extends the term-structure literature, since it shows how to derive bond prices using no-arbitrage conditions based on an explicit structural macroeconomic model, including both forward-looking and backward-looking elements. At the same time, we extends the macroeconomic literature by studying the term structure implications of a standard macro model within a dynamic no-arbitrage framework.

In an empirical application, we show that there are synergies to be exploited from current advances in macroeconomic and term-structure modelling. The two approaches can be seen as complementary and, when used jointly, give rise to sensible results. Notably, we show that our estimates of macroeconomic parameters, that are partly determined by the term-structure data, are reasonable and intuitively appealing. At the same time, our model's explanatory power for the term-structure is comparable to that of term-structure models based only on unobservable variables. Our model also performs very well in forecasting.

Finally, we show that a macro-based term-structure model can match features of yield curve data which represent a puzzle for the expectations hypothesis. These results confirm that the dynamics of stochastic risk premia are important determinants of yield dynamics, and that all such dynamics can be ultimately reconducted to underlying macroeconomic dynamics within a consistent framework.

References

[1] Ang, A. and M. Piazzesi (2003), "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, forthcoming.

[2] Bekaert, G. and R. J. Hodrick (2001), "Expectations hypothesis tests", *Journal of Finance* 56, 1357-93.

[3] Bernanke, B.S. and I. Mihov (1997), "What does the Bundesbank target?," *European Economic Review* 41, 1025-1053.

[4] Campbell, J. Y. and G. Mankiw, (1989), "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," *NBER Working Paper* #2924.

[5] Campbell, J. Y. and R. J. Shiller (1991), "Yield spreads and interest rate movements: A bird's eye view," *Review of Economic Studies* 58, 495-514.

[6] Chadha, B. P. R. Masson and G. Meredith (1992), "Models of Inflation and the Costs of Disinflation," *International Monetary Fund Staff Papers*, Vol. 39, No.2.

[7] Chen and Scott, (1993), "Pricing Interest Rate Futures Options with Futures-Style Margining," *Journal of Futures Markets, Vol. 13, No.1, 15-22.*

[8] Christiano, L., M. Eichenbaum and C. L. Evans, (1999), "Monetary policy shocks: what have we learnt and to what end?", in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Vol. IA, Ch. 2, Amsterdam: North Holland.

[9] Christiano, L., M. Eichenbaum and C. L. Evans, (2001), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *NBER Working Paper* #8403.

[10] Clarida, R., J. Galí and M. Gertler (1998), "Monetary policy rules in practice, Some international evidence," *European Economic Review* 42, 1033-1067.

[11] Clarida, R., J. Galí and M. Gertler (2000), "Monetary policy rules and macroeconomic stability: evidence and some theory," *Quarterly Journal of Economics*, 115, 147-180.

[12] Clarida, R. and M. Gertler (1997), "How the Bundesbank Conducts Monetary Policy, in C. Romer and D. Romer (eds.), *Reducing inflation*, Chicago: University of Chicago Press, 363-406.

[13] Dai, Q. and K.J. Singleton (2000), "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, Vol. LV, No. 5.

[14] Dai, G. and K. Singleton (2002a), "Expectations puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics* 63, 415-441.

[15] Dai, Q. and K. Singleton (2002b), "Term Structure Dynamics in Theory and Reality," forthcoming, *Review of Financial Studies.*

[16] Dewachter, H. and M. Lyrio (2002), "Macro factors and the term structure of interest rates," Center for Economic Studies, Catholic University of Leuven, mimeo, October.

[17] Dewachter, H., M. Lyrio, and K. Maes (2002), "The Effect of Monetary Unification on German Bond Markets", European Financial Management, forthcoming.

[18] Duffee, G.R. (2002), "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, Vol. LVII, No 1.

[19] Duffie, D. and R. Kan (1996), "A Yield-Factor Model of Interest Rates," *Mathematical Finance,* Vol. 6, No. 4, 379-406.

[20] Evans, M. D. D. (2003), "Real Risk, Inflation Risk and the Term Structure," *Economic Journal*, forthcoming.

[21] Evans, C. L. and D. A. Marshall (1996), "Economic determinants of the term structure of nominal interest rates", mimeo, Federal Reserve Bank of Chicago, November.

[22] Fama, E. and R. Bliss (1987), "The information in long-maturity forward rates," *American Economic Review* 77, 680-92.

[23] Fuhrer, J.C. (1997), "The (Un)Importance of Forward-Looking Behavior in Price Specifications," *Journal of Money, Credit, and Banking*, Vol. 29, No. 3.

[24] Fuhrer, J.C. (2000), "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *American Economic Review*, 367-390.

[25] Fuhrer, J. and G. Moore (1995), "Inflation Persistence," *Quarterly Journal of Economics*, 128-159.

[26] Fuhrer, J.C. and G.D. Rudebusch (2002), "Estimating the Euler Equation for Output," mimeo, available at http://www.frbsf.org/publications/economics/papers/2002/wp02-12bk1.pdf

[27] Galí, J. and M. Gertler (1999), "Inflation dynamics: A structural econometric analysis," *Journal of Monetary Economics* 44, 195-222.

[28] Gerlach, S. and F. Smets (1997), "The term structure of Euro-rates: some evidence in support of the expectations hypothesis," *Journal of International Money and Finance* 16, 305-21.

[29] Goffe, L. G., G. D. Ferrier and J. Rogers (1994), "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics 60, 65-99.*

[30] Goodfriend and King (1997), "The New Neoclassical Synthesis and the Role of Monetary Policy", NBER macroeconomics annual 1997, 231-283, MIT Press.

[31] Hall, R. E. (1978), "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence", *Journal of Political Economy* 86, 971-87.

[32] Hardouvelis, G. A. (1994), "The term structure spread and future changes in long and short rates in the G7 countries", *Journal of Monetary Economics* 33, 255-83.

[33] Jondeau, E. and H. Le Bihan (2001), "Testing for a forward-looking Phillips curve. Additional evidence from European and US data," Notes d'Études et de Recherche 86, Banque de France.

[34] Levin, A. and J. Piger (2002), "Is Inflation Persistence Intrinsic in Industrial Economies?," Federal Reserve Bank of St. Louis Working Paper *2002-023A.*

[35] McCallum, B. (1994), "Monetary policy and the term structure of interest rates," NBER Working Paper No. 4938.

[36] McCallum and Nelson, (1999), "Performance of Operational Policy Rules in an Estimated Semiclassical Structural Model," in J. Taylor (ed.), Monetary Policy Rules, University of Chicago Press.

[37] Piazzesi, M. (2001), "Macroeconomic jump effects and the yield curve", mimeo, UCLA.

[38] Roberds, W. and C. Whiteman (1999), "Endogenous term premia and anomalies in the term structure of interest rates: explaining the predictability smile," *Journal of Monetary Economics* 44, 555-580.

[39] Rotemberg, J.J. and M. Woodford (1997), "An Optimisation-Based Econometric Framework for the Evaluation of Monetary Policy", NBER macroeconomics annual 1997, 297-316.

[40] Rudebusch, G. (2002a), "Assessing nominal income rules for monetary policy with model and data uncertainty," *Economic Journal* 112, 402-432.

[41] Rudebusch, G. (2002b), "Term structure evidence on interest rate smoothing and monetary policy inertia," *Journal of Monetary Economics* 49, 1161-87.

[42] Rudebusch, G. and L. E. O. Svensson (1999), "Policy Rules for Inflation Targeting," in J. Taylor (ed.), Monetary Policy Rules, University of Chicago Press.

[43] Rudebusch, G. and T. Wu (2003), "A no-arbitrage model of the term structure and the macroeconomy," mimeo, Federal Reserve Bank of San Francisco, August.

[44] Sims, C.A. (1992), "Interpreting the macroeconomic time series facts," *European Economic Review* 36, 975-1011.

[45] Smets, F. (2000), “What horizon for price stability,” European Central Bank Working Paper Series No. 24.

[46] Söderlind, P. (1999), ”Solution and estimation of RE macromodels with optimal policy”, *European Economic Review 43*, 813-823.

[47] White, H. (1982), ”Maximum likelihood estimation of misspecified models,” *Econometrica*, 50, 1-25.

[48] Wu, T. (2002), ”Macro Factors and the Affine Term Structure of Interest Rates”, mimeo, Federal Reserve Bank of San Francisco.

[49] Zaffaroni, P. (2001), “Estimation of inflation risk premia for fixed income securities,” Banca d'Italia, mimeo, October.

A Appendix

A.1 State-space form

Here we discuss a more general model than that used in the text, namely a model where

$$\pi_t = \mu_\pi \sum_{i=1}^{12} E_t\left[\pi_{t+i}\right] + (1-\mu_\pi)\sum_{i=1}^{12}\delta_{\pi i}\pi_{t-i} + \delta_x x_t + \varepsilon_t^\pi$$

$$x_t = \mu_x \sum_{i=1}^{12} E_t\left[x_{t+i}\right] + (1-\mu_x)\sum_{i=1}^{12}\zeta_{xi}x_{t-i} - \zeta_r\left(r_t - E_t\left[\pi_{t+1}\right]\right) + \varepsilon_t^x$$

$$r_t = (1-\rho)\left(\beta\left(\sum_{i=0}^{11} E_t\left[\pi_{t+i}\right] - \pi_t^*\right) + \gamma x_t\right) + \rho r_{t-1} + \eta_t$$

Following Söderlind, define

$$\mathbf{X}_{1t} = \left[x_{t-1}, ..., x_{t-12}, \pi_{t-1}, ..., \pi_{t-12}, \pi_t^*, \eta_t, \varepsilon_t^\pi, \varepsilon_t^x, r_{t-1}\right]'$$

$$\mathbf{X}_{2t} = \left[E_t x_{t+11}, ..., E_t x_{t+1}, x_t, E_t\pi_{t+11}, ..., E_t\pi_{t+1}, \pi_t\right]'$$

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \end{bmatrix}$$

$$\mathbf{F} = \left[\underset{1x24}{\mathbf{0}}, \beta\left(1-\rho\right), -1, 0, 0, -\rho, \underset{1x11}{\mathbf{0}}, -\gamma\left(1-\rho\right), -\beta\left(1-\rho\right)\cdot \underset{1x12}{\mathbf{1}}\right]$$

$$\mathbf{K}_1 = \begin{bmatrix} \underset{28x1}{\mathbf{0}} \\ 1 \end{bmatrix}$$

$$\mathbf{K}_2 = \begin{bmatrix} \frac{\zeta_r}{\mu_x} \\ \underset{23x1}{\mathbf{0}} \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \end{bmatrix}$$

$$\mathbf{\Omega} = \begin{bmatrix} \phi_\pi & & & & \\ & \phi_\eta & & \mathbf{0} & \\ & & \phi_\varepsilon^\pi & & \\ & \mathbf{0} & & \phi_\varepsilon^x & \\ & & & & 0 \end{bmatrix}$$

$$\mathbf{H}_{11} = \begin{bmatrix} \underset{1x11}{\mathbf{0}} & 0 & \underset{1x11}{\mathbf{0}} & 0 & \underset{1x5}{\mathbf{0}} \\ \underset{11x11}{\mathbf{I}} & 0 & \underset{11x11}{\mathbf{0}} & 0 & \underset{1x5}{\mathbf{0}} \\ \underset{1x11}{\mathbf{0}} & 0 & \underset{1x11}{\mathbf{0}} & 0 & \underset{1x5}{\mathbf{0}} \\ \underset{11x11}{\mathbf{0}} & 0 & \underset{11x11}{\mathbf{I}} & 0 & \underset{1x5}{\mathbf{0}} \\ \underset{5x11}{\mathbf{0}} & 0 & \underset{1x11}{\mathbf{0}} & 0 & \mathbf{\Omega} \end{bmatrix}$$

$$\mathbf{H}_{12}=\left[\begin{array}{cccc}\underset{1x11}{\mathbf{0}} & 1 & \underset{1x11}{\mathbf{0}} & 0\\ \underset{11x11}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x11}{\mathbf{0}} & \underset{11x1}{\mathbf{0}}\\ \underset{1x11}{\mathbf{0}} & 0 & \underset{1x11}{\mathbf{0}} & 1\\ \underset{16x11}{\mathbf{0}} & \underset{16x1}{\mathbf{0}} & \underset{16x11}{\mathbf{0}} & \underset{16x1}{\mathbf{0}}\end{array}\right]$$

$$\boldsymbol{\zeta}_x=[\zeta_{x1},...,\zeta_{x12}]'$$

$$\boldsymbol{\delta}_\pi=[\delta_{\pi 1},...,\delta_{\pi 12}]'$$

$$\mathbf{H}_{21}=\left[\begin{array}{cccccc}-\frac{1-\mu_x}{\mu_x}\boldsymbol{\zeta}_x' & \underset{1x12}{\mathbf{0}} & \underset{1x2}{\mathbf{0}} & 0 & -\frac{1}{\mu_x} & 0\\ \underset{11x12}{\mathbf{0}} & \underset{11x12}{\mathbf{0}} & \underset{11x2}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x1}{\mathbf{0}}\\ \underset{1x12}{\mathbf{0}} & -\frac{1-\mu_\pi}{\mu_\pi}\boldsymbol{\delta}_\pi' & \underset{1x2}{\mathbf{0}} & -\frac{1}{\mu_\pi} & 0 & 0\\ \underset{11x12}{\mathbf{0}} & \underset{11x12}{\mathbf{0}} & \underset{11x2}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x1}{\mathbf{0}}\end{array}\right]$$

$$\mathbf{H}_{22}=\left[\begin{array}{ccccc}-\underset{1x11}{\mathbf{1}} & \frac{1}{\mu_x} & \underset{1x10}{\mathbf{0}} & -\frac{\zeta_r}{\mu_x} & 0\\ \underset{11x11}{\mathbf{I}} & \underset{11x1}{\mathbf{0}} & \underset{11x10}{\mathbf{0}} & \underset{11x1}{\mathbf{0}} & \underset{11x1}{\mathbf{0}}\\ \underset{1x11}{\mathbf{0}} & -\frac{\delta_x}{\mu_\pi} & -\underset{1x10}{\mathbf{1}} & -1 & \frac{1}{\mu_\pi}\\ \underset{10x11}{\mathbf{0}} & \underset{10x1}{\mathbf{0}} & \underset{10x10}{\mathbf{I}} & \underset{10x1}{\mathbf{0}} & \underset{10x1}{\mathbf{0}}\\ \underset{1x11}{\mathbf{0}} & 0 & \underset{1x10}{\mathbf{0}} & 1 & 0\end{array}\right]$$

$$\mathbf{H}=\left[\begin{array}{cc}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{array}\right]$$

$$\boldsymbol{\xi}_{1,t+1}=\left[\underset{1x24}{\mathbf{0}},u_{\pi,t+1},u_{\eta,t+1},u^{\pi}_{\varepsilon,t+1},u^{x}_{\varepsilon,t+1},0\right]'$$

$$\boldsymbol{\xi}_{t+1}=\left[\boldsymbol{\xi}_{1,t+1}',\underset{1x24}{\mathbf{0}}\right]'$$

Then the system can be written as

$$\mathbf{X}_{t+1}=\mathbf{H}\mathbf{X}_t+\mathbf{K}r_t+\boldsymbol{\xi}_{t+1}$$

and the policy rule as

$$r_t=-\mathbf{F}\mathbf{X}_t$$

We use Paul Söderlind's routine to solve

$$\mathbf{X}_{t+1}=(\mathbf{H}-\mathbf{K}\mathbf{F})\,\mathbf{X}_t+\boldsymbol{\xi}_{t+1}$$

with solution

$$\begin{aligned}\mathbf{X}_{1,t+1} &= \mathbf{M}\mathbf{X}_{1,t}+\boldsymbol{\xi}_{1,t+1}\\ \mathbf{X}_{2,t+1} &= \mathbf{C}\mathbf{X}_{1,t+1}\\ r_t &= \boldsymbol{\Delta}'\mathbf{X}_{1,t}\end{aligned}$$

where $\boldsymbol{\Delta}'\equiv-(\mathbf{F}_1+\mathbf{F}_2\mathbf{C})$ and $\mathbf{F}_1$ and $\mathbf{F}_2$ are partitions of $\mathbf{F}$ conformable with $\mathbf{X}_{1,t}$ and $\mathbf{X}_{2,t}$.

A.2 Bond prices

For the pricing of bonds, we want to transform the state vector so that it includes levels of inflation, the output gap, and the short rate, rather than their respective shocks. (We also retain lagged factors in the state vector). I.e. we choose $\mathbf{Z}_t$ to include:

$$\mathbf{Z}_{1t} = [x_{t-1}, ..., x_{t-12}, \pi_{t-1}, ..., \pi_{t-12}, \pi_t^*, r_t, \pi_t, x_t, r_{t-1}]',$$

Using the solution obtained above, we obtain

$$\mathbf{Z}_t = \mathbf{D}\mathbf{X}_t,$$

where $\mathbf{D}$ is the 29×53 matrix

$$\mathbf{D} \equiv \begin{bmatrix} \left(\underset{25\times 25}{\mathbf{I}}, \underset{25\times 28}{\mathbf{0}}\right) \\ -\mathbf{F} \\ \left(\underset{1\times 52}{\mathbf{0}}, 1\right) \\ \left(\underset{1\times 40}{\mathbf{0}}, 1, \underset{1\times 12}{\mathbf{0}}\right) \\ \left(\underset{1\times 28}{\mathbf{0}}, 1, \underset{1\times 24}{\mathbf{0}}\right) \end{bmatrix}.$$

Using

$$\mathbf{X}_{2,t} = \mathbf{C}\mathbf{X}_{1,t},$$

we can also write

$$\mathbf{Z}_t = \hat{\mathbf{D}}\mathbf{X}_{1,t},$$

where $\hat{\mathbf{D}}$ is 29×29:

$$\hat{\mathbf{D}} \equiv \mathbf{D}_1 + \mathbf{D}_2\mathbf{C}$$

and $\mathbf{D}_1$ and $\mathbf{D}_2$ are partitions of $\mathbf{D}$ conformable with $\mathbf{X}_{1t}$ and $\mathbf{X}_{2t}$, or

$$\hat{\mathbf{D}} \equiv \begin{bmatrix} \left(\underset{25\times 25}{\mathbf{I}}, \underset{25\times 4}{\mathbf{0}}\right) \\ -\mathbf{F}\mathbf{G} \\ \mathbf{C}_{\{24,.\}} \\ \mathbf{C}_{\{12,.\}} \\ \left(\underset{1\times 28}{\mathbf{0}}, 1\right) \end{bmatrix},$$

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{I}_{29} \\ \mathbf{C} \end{bmatrix},$$

and $\mathbf{C}_{\{j,.\}}$ denotes row j of the matrix $\mathbf{C}$.

Given the definition of r_t and ξ_{t+1}, the pricing kernel can be written as

$$\begin{aligned} m_{t+1} &= \exp(-r_t)\frac{\xi_{t+1}}{\xi_t} \\ &= \exp\left(-\overline{\mathbf{\Delta}}'\mathbf{Z}_t\right)\exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{1,t+1}\right) \\ &= \exp\left(-\overline{\mathbf{\Delta}}'\mathbf{Z}_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{1,t+1}\right). \end{aligned}$$

where we used $r_t = \overline{\mathbf{\Delta}}'\mathbf{Z}_t$ with

$$\overline{\mathbf{\Delta}}' = \left[\underset{1\times 25}{\mathbf{0}}, 1, \underset{1\times 3}{\mathbf{0}}\right].$$

The price of a one-period bond at time t is

$$\begin{aligned} p_t^1 &= E_t\left[m_{t+1}\right] \\ &= \exp(-r_t) \\ &= \exp\left(-\overline{\mathbf{\Delta}}'\mathbf{Z}_t\right). \end{aligned}$$

Since bond prices are exponential affine functions of $\mathbf{X}_1$, we know that

$$p_t^1 = \exp\left(\bar{A}_1 + \bar{B}_1'\mathbf{Z}_t\right)$$

so that we can identify $\bar{A}_1 = 0$ and $\bar{B}_1 = -\overline{\mathbf{\Delta}}$.

The price of an n-period bond at time t is

$$p_t^n = \exp\left(\bar{A}_n + \bar{B}_n'\mathbf{Z}_t\right).$$

In order to identify the recursive structure of the coefficients in the bond pricing equation, we study the price of an $n+1$ period bond at t:

$$\begin{aligned} p_t^{n+1} &= E_t\left[m_{t+1}p_{t+1}^n\right] \\ &= E_t\left[\exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{1,t+1} + \bar{A}_n + \bar{B}_n'\mathbf{Z}_{t+1}\right)\right] \\ &= \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n\right)E_t\left[\exp\left(-\lambda_t'\varepsilon_{1,t+1} + \bar{B}_n'\mathbf{Z}_{t+1}\right)\right] \\ &= \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n\right)E_t\left[\exp\left(-\lambda_t'\varepsilon_{1,t+1} + \bar{B}_n'\left[\hat{\mathbf{D}}\mathbf{M}\mathbf{X}_{1,t} + \hat{\mathbf{D}}\Sigma\varepsilon_{1,t+1}\right]\right)\right] \\ &= \exp\left(-\overline{\mathbf{\Delta}}'\mathbf{Z}_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n + \bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\mathbf{X}_{1,t}\right)E_t\left[\exp\left(-\lambda_t'\varepsilon_{1t+1} + \bar{B}_n'\hat{\mathbf{D}}\Sigma\varepsilon_{1,t+1}\right)\right] \\ &= \exp\left(-\overline{\mathbf{\Delta}}'\mathbf{Z}_t - \frac{1}{2}\lambda_t'\lambda_t + \bar{A}_n + \bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1}\mathbf{Z}_t\right)E_t\left[\exp\left(-\lambda_t'\varepsilon_{1t+1} + \bar{B}_n'\hat{\mathbf{D}}\Sigma\varepsilon_{1,t+1}\right)\right] \\ &= \exp\left(\bar{A}_n - \frac{1}{2}\lambda_t'\lambda_t + \left(\bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1} - \overline{\mathbf{\Delta}}'\right)\mathbf{Z}_t\right)E_t\left[\exp\left(\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma - \lambda_t'\right]\varepsilon_{1t+1}\right)\right]. \end{aligned}$$

Since ε_{1t+1} is assumed to be IID normal with zero mean, we know from the properties of a lognormal variable that $E\left[\exp\left(a+b\varepsilon_{1t+1}\right)\right] = \exp\left(a+\frac{1}{2}b^2 var\left[\varepsilon_{1t+1}\right]\right)$. The expectation above can therefore be written as

$$\begin{aligned}
E_t\left[\exp\left(\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma-\lambda_t'\right]\varepsilon_{1t+1}\right)\right] &= \exp\left(\frac{1}{2}\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma-\lambda_t'\right]\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma-\lambda_t'\right]'\right)\\
&= \exp\left(\frac{1}{2}\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma-\lambda_t'\right]\left[\Sigma'\hat{\mathbf{D}}'\bar{B}_n-\lambda_t\right]\right)\\
&= \exp\left(\frac{1}{2}\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n-\lambda_t'\Sigma'\hat{\mathbf{D}}'\bar{B}_n-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_t+\lambda_t'\lambda_t\right]\right)\\
&= \exp\left(\frac{1}{2}\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n+\lambda_t'\lambda_t\right]-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_t\right),
\end{aligned}$$

and we obtain

$$\begin{aligned}
p_t^{n+1} &= \exp\left(\bar{A}_n-\frac{1}{2}\lambda_t'\lambda_t+\left(\bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1}-\overline{\mathbf{\Delta}}'\right)\mathbf{Z}_t+\frac{1}{2}\left[\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n+\lambda_t'\lambda_t\right]-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_t\right)\\
&= \exp\left(\bar{A}_n+\left(\bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1}-\overline{\mathbf{\Delta}}'\right)\mathbf{Z}_t+\frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_t\right)\\
&= \exp\left(\bar{A}_n+\left(\bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1}-\overline{\mathbf{\Delta}}'\right)\mathbf{Z}_t+\frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n-\bar{B}_n'\hat{\mathbf{D}}\Sigma\left(\lambda_0+\lambda_1\mathbf{Z}_t\right)\right)\\
&= \exp\left(\bar{A}_n-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_0+\frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n+\left(\bar{B}_n'\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-1}-\overline{\mathbf{\Delta}}'-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_1\right)\mathbf{Z}_t\right).
\end{aligned}$$

The recursive bond-pricing coefficients are therefore given by

$$\begin{aligned}
\bar{A}_{n+1} &= \bar{A}_n-\bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_0+\frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n,\\
\bar{B}_{n+1}' &= \bar{B}_n'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1}-\Sigma\lambda_1\right)-\overline{\mathbf{\Delta}}'.
\end{aligned}$$

A.3 Likelihood function

We are interested in

$$\ell\left(\boldsymbol{\theta}\right)=\prod_{t=2}^{T}f\left(\mathbf{X}_{1t}^o,\mathbf{X}_{2t}^o,\mathbf{Y}_t|\mathbf{X}_{1t-1}^o,\mathbf{X}_{1t-1}^u,\mathbf{Y}_{t-1}\right)$$

We now partition the yield solutions into

$$\mathbf{Y}_t=\mathbf{A}+\tilde{\mathbf{B}}^o\mathbf{X}_{1t}^o+\tilde{\mathbf{B}}^u\mathbf{X}_{1t}^u+\mathbf{B}^m\mathbf{u}_t^m,$$

where $\mathbf{X}_{1t}^o$ is the vector of observable predetermined variables, $\mathbf{X}_{1t}^u$ are the unobservable predetermined variables, and $\mathbf{u}_t^m$ is the vector of measurement errors. The equation for the observable non-predetermined variables is as before:

$$\mathbf{X}_{2t}^o=\tilde{\mathbf{C}}^o\mathbf{X}_{1t}^o+\tilde{\mathbf{C}}^u\mathbf{X}_{1t}^u.$$

Hence,

$$\begin{bmatrix} \mathbf{Y}_t \\ \mathbf{X}_{2t}^o \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{B}}^o \\ \tilde{\mathbf{C}}^o \end{bmatrix} \mathbf{X}_{1t}^o + \begin{bmatrix} \tilde{\mathbf{B}}^u \\ \tilde{\mathbf{C}}^u \end{bmatrix} \mathbf{X}_{1t}^u + \begin{bmatrix} \mathbf{B}^m \\ \mathbf{0} \end{bmatrix} \mathbf{u}_t^m,$$

so that

$$\begin{bmatrix} \mathbf{X}_{1t}^o \\ \mathbf{Y}_t \\ \mathbf{X}_{2t}^o \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{A} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ \tilde{\mathbf{B}}^o \\ \tilde{\mathbf{C}}^o \end{bmatrix} \mathbf{X}_{1t}^o + \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{B}}^u \\ \tilde{\mathbf{C}}^u \end{bmatrix} \mathbf{X}_{1t}^u + \begin{bmatrix} \mathbf{0} \\ \mathbf{B}^m \\ \mathbf{0} \end{bmatrix} \mathbf{u}_t^m$$

Now stack the $\mathbf{X}_{1t}^o$, $\mathbf{X}_{1t}^u$ and $\mathbf{u}_t^m$ vectors, so that

$$\begin{bmatrix} \mathbf{X}_{1t}^o \\ \mathbf{Y}_t \\ \mathbf{X}_{2t}^o \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{A} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{1t}^o \\ \mathbf{X}_{1t}^u \\ \mathbf{u}_t^m \end{bmatrix}$$

so that this can be inverted to find

$$\begin{bmatrix} \mathbf{X}_{1t}^o \\ \mathbf{X}_{1t}^u \\ \mathbf{u}_t^m \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{X}_{1t}^o \\ \mathbf{Y}_t \\ \mathbf{X}_{2t}^o \end{bmatrix} - \begin{bmatrix} \mathbf{0} \\ \mathbf{A} \\ \mathbf{0} \end{bmatrix} \right)$$

It follows that the likelihood function can be rewritten as

$$\begin{aligned} \ell(\theta) &= \prod_{t=2}^{T} \left| \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{bmatrix}^{-1} \right| f\left(\mathbf{X}_{1t}^o, \mathbf{X}_{1t}^u, \mathbf{u}_t^m | \mathbf{X}_{1t-1}^o, \mathbf{X}_{1t-1}^u\right) \\ &= \prod_{t=2}^{T} \frac{1}{\begin{vmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{vmatrix}} f_{\mathbf{X}_1}\left(\mathbf{X}_{1t}^o, \mathbf{X}_{1t}^u | \mathbf{X}_{1t-1}^o, \mathbf{X}_{1t-1}^u\right) f_{\mathbf{u}^m}\left(\mathbf{u}_t^m\right) \end{aligned}$$

where the equality comes from the properties of the determinant and the independence assumption between structural shocks and measurement errors. Defining

$$J \equiv \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \tilde{\mathbf{B}}^o & \tilde{\mathbf{B}}^u & \mathbf{B}^m \\ \tilde{\mathbf{C}}^o & \tilde{\mathbf{C}}^u & \mathbf{0} \end{bmatrix},$$

the log-likelihood is simply

$$\ln\left(\pounds\left(\boldsymbol{\theta}\right)\right) = \sum_{t=2}^{T} \left(-\ln|J| + \ln f_{\mathbf{X}_1}\left(\mathbf{X}_{1t}^o, \mathbf{X}_{1t}^u | \mathbf{X}_{1t-1}^o, \mathbf{X}_{1t-1}^u\right) + \ln f_{\mathbf{u}^m}\left(\mathbf{u}_t^m\right)\right)$$

and

$$\begin{aligned}\ln(\pounds(\boldsymbol{\theta})) &= -(T-1)\ln|J| - \frac{(T-1)n_1}{2}\ln(2\pi) - \frac{T-1}{2}\ln\left|\boldsymbol{\Sigma\Sigma}'\right| \\ &-\frac{1}{2}\sum_{t=2}^{T}(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1})'\left(\boldsymbol{\Sigma\Sigma}'\right)^{-1}(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1}) \\ &-\frac{(T-1)n_m}{2}\ln(2\pi) - \frac{T-1}{2}\ln\sum_{i=1}^{n_m}\sigma_i^2 - \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{n_m}\frac{\left(u_{t,i}^m\right)^2}{\sigma_i^2}\end{aligned}$$

or, more similar to the structure of our programme,

$$\begin{aligned}\ln(\pounds(\boldsymbol{\theta})) &= -(T-1)\left(\ln|J| + \frac{1}{2}\ln\left|\Sigma\Sigma'\right| + \frac{1}{2}\ln\sum_{i=1}^{n_m}\sigma_i^2 + \frac{n_1+n_m}{2}\ln(2\pi)\right) \\ &-\frac{1}{2}\sum_{t=2}^{T}(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1})'\left(\boldsymbol{\Sigma\Sigma}'\right)^{-1}(\mathbf{X}_{1t}-\mathbf{M}\mathbf{X}_{1t-1}) - \frac{1}{2}\sum_{t=2}^{T}\sum_{i=1}^{n_m}\frac{\left(u_{t,i}^m\right)^2}{\sigma_i^2}\end{aligned}$$

where n_m is the number of measurement errors and n_1 is the number of predetermined variables.

Note that, in order to actually calculate the unobservable factors and measurement errors, it is useful to rewrite the system

$$\begin{bmatrix}\mathbf{X}_{1t}^o \\ \mathbf{Y}_t \\ \mathbf{X}_{2t}^o\end{bmatrix} = \begin{bmatrix}\mathbf{0} \\ \mathbf{A} \\ \mathbf{0}\end{bmatrix} + \begin{bmatrix}\mathbf{I} \\ \tilde{\mathbf{B}}^o \\ \tilde{\mathbf{C}}^o\end{bmatrix}\mathbf{X}_{1t}^o + \begin{bmatrix}\mathbf{0} \\ \tilde{\mathbf{B}}^u \\ \tilde{\mathbf{C}}^u\end{bmatrix}\mathbf{X}_{1t}^u + \begin{bmatrix}\mathbf{0} \\ \mathbf{B}^m \\ \mathbf{0}\end{bmatrix}\mathbf{u}_t^m$$

in terms of perfectly observable variables, i.e. $[\mathbf{X}_{1t}^o, \mathbf{Y}_t^p, \mathbf{X}_{2t}^o]$, and variables measured with error, i.e. $[\mathbf{Y}_t^m]$. This leads to

$$\begin{bmatrix}\mathbf{X}_{1t}^o \\ \mathbf{Y}_t^p \\ \mathbf{Y}_t^m \\ \mathbf{X}_{2t}^o\end{bmatrix} = \begin{bmatrix}\mathbf{0} \\ \mathbf{A}^p \\ \mathbf{A}^m \\ \mathbf{0}\end{bmatrix} + \begin{bmatrix}\mathbf{I} \\ \tilde{\mathbf{B}}^{op} \\ \tilde{\mathbf{B}}^{om} \\ \tilde{\mathbf{C}}^o\end{bmatrix}\mathbf{X}_{1t}^o + \begin{bmatrix}\mathbf{0} \\ \tilde{\mathbf{B}}^{up} \\ \tilde{\mathbf{B}}^{um} \\ \tilde{\mathbf{C}}^u\end{bmatrix}\mathbf{X}_{1t}^u + \begin{bmatrix}\mathbf{0} \\ \mathbf{0} \\ \widehat{\mathbf{B}}^m \\ \mathbf{0}\end{bmatrix}\mathbf{u}_t^m$$

which, forgetting about $\mathbf{X}_{1t}^o$, can also be split into

$$\begin{bmatrix}\mathbf{Y}_t^p \\ \mathbf{X}_{2t}^o\end{bmatrix} = \begin{bmatrix}\mathbf{A}^p \\ \mathbf{0}\end{bmatrix} + \begin{bmatrix}\tilde{\mathbf{B}}^{op} \\ \tilde{\mathbf{C}}^o\end{bmatrix}\mathbf{X}_{1t}^o + \begin{bmatrix}\tilde{\mathbf{B}}^{up} \\ \tilde{\mathbf{C}}^u\end{bmatrix}\mathbf{X}_{1t}^u$$

$$\mathbf{Y}_t^m = \mathbf{A}^m + \tilde{\mathbf{B}}^{om}\mathbf{X}_{1t}^o + \tilde{\mathbf{B}}^{um}\mathbf{X}_{1t}^u + \widehat{\mathbf{B}}^m\mathbf{u}_t^m$$

These two systems can be solved for $\mathbf{X}_{1t}^u$ and $\mathbf{u}_t^m$ in a recursive fashion as

$$\mathbf{X}_{1t}^u = \begin{bmatrix}\tilde{\mathbf{B}}^{up} \\ \tilde{\mathbf{C}}^u\end{bmatrix}^{-1}\left(\begin{bmatrix}\mathbf{Y}_t^p \\ \mathbf{X}_{2t}^o\end{bmatrix} - \begin{bmatrix}\mathbf{A}^p \\ \mathbf{0}\end{bmatrix} - \begin{bmatrix}\tilde{\mathbf{B}}^{op} \\ \tilde{\mathbf{C}}^o\end{bmatrix}\mathbf{X}_{1t}^o\right)$$

$$\mathbf{u}_t^m = \left(\widehat{\mathbf{B}}^m\right)^{-1}\left(\mathbf{Y}_t^m - \mathbf{A}^m - \tilde{\mathbf{B}}^{om}\mathbf{X}_{1t}^o - \tilde{\mathbf{B}}^{um}\mathbf{X}_{1t}^u\right)$$

A.4 Risk premia

A.4.1 Holding premia

Let $e_{n,t}$ denote the one-period holding premium on an n-period bond purchased at t, defined as the expected holding return of that bond over one period, less the risk-free rate:

$$e_{n,t} = E_t\left[\ln\left(p_{t+1}^{n-1}\right) - \ln\left(p_t^n\right)\right] - r_t.$$

Using that

$$\begin{aligned} p_t^n &= \exp\left(\bar{A}_n + \bar{B}_n'\mathbf{Z}_t\right), \\ p_{t+1}^{n-1} &= \exp\left(\bar{A}_{n-1} + \bar{B}_{n-1}'\mathbf{Z}_{t+1}\right), \end{aligned}$$

and

$$r_t = \overline{\mathbf{\Delta}}'\mathbf{Z}_t$$

we obtain

$$\begin{aligned} e_{n,t} &= \bar{A}_{n-1} + \bar{B}_{n-1}'E_t\left[\mathbf{Z}_{t+1}\right] - \bar{A}_n - \bar{B}_n'\mathbf{Z}_t - \overline{\mathbf{\Delta}}'\mathbf{Z}_t \\ &= \bar{A}_{n-1} + \bar{B}_{n-1}'\left(\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-\mathbf{1}}\mathbf{Z}_t\right) - \bar{A}_n - \bar{B}_n'\mathbf{Z}_t - \overline{\mathbf{\Delta}}'\mathbf{Z}_t \\ &= \bar{A}_{n-1} + \bar{B}_{n-1}'\left(\hat{\mathbf{D}}\mathbf{M}\hat{\mathbf{D}}^{-\mathbf{1}}\mathbf{Z}_t\right) - \left(\bar{A}_{n-1} - \bar{B}_{n-1}'\hat{\mathbf{D}}\Sigma\lambda_0 + \frac{1}{2}\bar{B}_{n-1}'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_{n-1}\right) - \\ &\quad \left[\bar{B}_{n-1}'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1} - \Sigma\lambda_1\right) - \overline{\mathbf{\Delta}}'\right]\mathbf{Z}_t - \overline{\mathbf{\Delta}}'\mathbf{Z}_t \\ &= \bar{B}_{n-1}'\hat{\mathbf{D}}\Sigma\lambda_0 + \bar{B}_{n-1}'\hat{\mathbf{D}}\Sigma\lambda_1\mathbf{Z}_t - \frac{1}{2}\bar{B}_{n-1}'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_{n-1} \end{aligned}$$

A.4.2 Forward premia

Let $\psi_{n,t}$ denote the one-period forward premium at t for maturity n, defined as the difference between the implied one-period forward rate n periods ahead, $f_{n,t}$, less the corresponding expected one-period interest rate:

$$\psi_{n,t} = f_{n,t} - E_t\left[r_{t+n}\right].$$

The implied forward rate, expressed in one-period terms, is given by

$$\begin{aligned} f_{n,t} &= \ln\left(p_t^n\right) - \ln\left(p_t^{n+1}\right) \\ &= \bar{A}_n + \bar{B}_n'\mathbf{Z}_t - \bar{A}_{n+1} - \bar{B}_{n+1}'\mathbf{Z}_t \\ &= \bar{A}_n + \bar{B}_n'\mathbf{Z}_t - \left(\bar{A}_n - \bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_0 + \frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n\right) - \\ &\quad \left(\bar{B}_n'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1} - \Sigma\lambda_1\right) - \overline{\mathbf{\Delta}}'\right)\mathbf{Z}_t \\ &= \bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_0 - \frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n + \left[\bar{B}_n' - \bar{B}_n'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1} - \Sigma\lambda_1\right) + \overline{\mathbf{\Delta}}'\right]\mathbf{Z}_t \end{aligned}$$

while the expected short rate is

$$
\begin{aligned}
E_t\left[r_{t+n}\right] &= \overline{\mathbf{\Delta}}' E_t\left[\mathbf{Z}_{1,t+n}\right] \\
&= \overline{\mathbf{\Delta}}' E_t\left[\hat{\mathbf{D}}\mathbf{M}^n\hat{\mathbf{D}}^{-1}\mathbf{Z}_t + \sum_{i=1}^{n}\mathbf{M}^{n-i}\hat{\mathbf{D}}\Sigma\varepsilon_{1,t+i}\right] \\
&= \overline{\mathbf{\Delta}}'\hat{\mathbf{D}}\mathbf{M}^n\hat{\mathbf{D}}^{-1}\mathbf{Z}_t.
\end{aligned}
$$

The one-period forward premium is therefore

$$
\begin{aligned}
\psi_{n,t} &= f_{n,t} - E_t\left[r_{t+n}\right] \\
&= \bar{B}_n'\hat{\mathbf{D}}\Sigma\lambda_0 - \frac{1}{2}\bar{B}_n'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_n + \\
&\quad \left[\bar{B}_n' - \bar{B}_n'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1} - \Sigma\lambda_1\right) + \overline{\mathbf{\Delta}}'\left(I - \hat{\mathbf{D}}\mathbf{M}^n\hat{\mathbf{D}}^{-1}\right)\right]\mathbf{Z}_t
\end{aligned}
$$

The forward premium is expressed in one-period terms.

A.4.3 Yield risk premia

Let $\omega_{n,t}$ denote the n-maturity yield premium at t, defined as the sum of the forward premia up until $t+n-1$:

$$
\begin{aligned}
\omega_{n,t} &= \sum_{i=0}^{n-1}\psi_{n,t} \\
&= \sum_{i=0}^{n-1}\left[\bar{B}_i'\hat{\mathbf{D}}\Sigma\lambda_0 - \frac{1}{2}\bar{B}_i'\hat{\mathbf{D}}\Sigma\Sigma'\hat{\mathbf{D}}'\bar{B}_i + \right. \\
&\quad \left.\left(\bar{B}_i' - \bar{B}_i'\hat{\mathbf{D}}\left(\mathbf{M}\hat{\mathbf{D}}^{-1} - \Sigma\lambda_1\right) + \overline{\mathbf{\Delta}}'\left(I - \hat{\mathbf{D}}\mathbf{M}^i\hat{\mathbf{D}}^{-1}\right)\right)\mathbf{Z}_t\right]
\end{aligned}
$$

The n-maturity yield premium can be expressed in one-period terms by dividing $\omega_{n,t}$ with n.

Table 1: Parameter estimates
(Sample period: Feb 1975-Dec 1998)

Parameter	Point estimate	Standard error
ρ	0.979	0.012
β	0.051	0.049
γ	0.029	0.006
μ_π	0.133	0.015
$\delta_x \times 10^2$	0.038	0.063
μ_x	0.299	0.023
ζ_r	0.027	0.026
ϕ_{π^*}	0.999	–
$\sigma_{\pi^*} \times 10^2$	0.018	0.010
$\sigma_\eta \times 10^2$	0.040	0.009
$\sigma_x \times 10^2$	0.022	0.001
$\sigma_\pi \times 10^2$	0.097	0.005
$\sigma_1^m \times 10^2$	1.572	0.052
$\sigma_2^m \times 10^2$	1.774	0.059
$\sigma_3^m \times 10^2$	1.704	0.048
$\sigma_4^m \times 10^2$	1.701	0.039
$\lambda_{0,1}$	−0.408	0.170
$\lambda_{0,2}$	−0.583	0.260
$\lambda_{0,3}$	4.345	1.286
$\lambda_{0,4}$	−1.675	0.657

The standard errors are based on the asymptotic variance-covariance matrix of White (1982). The estimates of the lag coefficients for inflation and output are not reported.

$\lambda_1 \times 10^{-2}$

	π^*	r	π	x
π^*	0 (–)	0 (–)	0.907 (0.338)	0.892 (0.314)
r	−28.118 (14.171)	18.767 (3.845)	−8.495 (2.24)	0 (–)
π	119.117 (70.489)	−86.588 (14.291)	41.356 (4.175)	0 (–)
x	0 (–)	1.889 (1.329)	0 (–)	2.613 (0.866)

Standard errors in parentheses

Table 2: Forecast error variance decompositions

Output gap

Steps	π^*	η	π	x
	Variance due to (in %)			
1	1.77	0.97	0.08	97.18
6	44.67	21.63	2.56	31.14
12	64.36	26.94	4.67	4.03
36	72.99	8.08	11.80	7.12
60	68.25	18.03	11.46	2.25

Inflation

Steps	π^*	η	π	x
	Variance due to (in %)			
1	0.16	0.00	99.84	0.00
6	6.20	0.05	93.75	0.00
12	25.86	0.19	73.94	0.00
36	89.81	0.15	10.01	0.03
60	98.67	0.01	1.31	0.01

Short rate

Steps	π^*	η	π	x
	Variance due to (in %)			
1	0.00	99.33	0.05	0.62
6	1.22	90.19	1.26	7.34
12	19.48	59.17	2.91	18.43
36	98.69	0.51	0.00	0.79
60	99.60	0.25	0.09	0.07

3-month rate

Steps	π^*	η	π	x
	Variance due to (in %)			
1	8.02	89.30	2.15	0.53
6	19.51	69.51	4.54	6.45
12	51.70	32.42	4.81	11.07
36	99.24	0.33	0.05	0.38
60	99.81	0.13	0.02	0.04

1-year yield

Steps	π^*	η	π	x
	Variance due to (in %)			
1	34.96	62.52	2.36	0.16
6	49.92	42.75	3.69	3.64
12	74.01	17.22	3.19	5.58
36	99.53	0.18	0.06	0.23
60	99.89	0.08	0.01	0.02

3-year yield

Steps	π^*	η	π	x
	Variance due to (in %)			
1	82.73	15.54	0.49	1.24
6	84.74	14.14	1.10	0.02
12	89.86	8.11	1.32	0.71
36	99.69	0.00	0.11	0.19
60	99.95	0.05	0.00	0.00

7-year yield

Steps	π^*	η	π	x
	Variance due to (in %)			
1	94.55	1.21	0.01	4.24
6	95.08	3.35	0.19	1.38
12	95.10	4.26	0.53	0.10
36	99.46	0.11	0.22	0.20
60	99.96	0.03	0.01	0.00

Table 3: In-sample yield forecast performance: RMSEs

	1 month forecast horizon			
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.497	0.499	0.489	**0.484**
3 months	**0.371**	**0.371**	**0.371**	0.389
1 year	**0.341**	0.386	0.372	0.394
3 years	0.322	0.321	**0.319**	**0.319**
7 years	**0.260**	0.348	0.357	0.349

	3 month forecast horizon			
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.784	0.787	0.763	**0.729**
3 months	0.699	0.697	0.692	**0.641**
1 year	0.689	0.717	0.707	**0.659**
3 years	0.611	0.611	0.601	**0.598**
7 years	**0.497**	0.538	0.524	0.539

	6 month forecast horizon			
maturity	RW	$A_0(3)$	AP	HTV
1 month	1.212	1.206	1.134	**1.084**
3 months	1.089	1.085	1.052	**0.982**
1 year	1.058	1.066	1.048	**0.987**
3 years	0.899	0.891	**0.872**	0.878
7 years	0.722	0.718	**0.693**	0.715

	9 month forecast horizon			
maturity	RW	$A_0(3)$	AP	HTV
1 month	1.594	1.573	1.408	**1.388**
3 months	1.413	1.406	1.321	**1.263**
1 year	1.349	1.346	1.311	**1.259**
3 years	1.121	1.099	**1.077**	1.099
7 years	0.907	0.864	**0.834**	0.868

	12 month forecast horizon			
maturity	RW	$A_0(3)$	AP	HTV
1 month	1.937	1.897	**1.619**	1.650
3 months	1.698	1.687	1.555	**1.524**
1 year	1.611	1.599	1.548	**1.519**
3 years	1.349	1.311	**1.275**	1.328
7 years	1.088	1.027	**0.978**	1.038

The table shows root mean square errors (RMSEs) for in-sample forecasts between 1975:01 and 1998:12. RW denotes random walk forecasts, $A_0(3)$ is a canonical essentially affine Gaussian three-factor model, AP denotes the Ang-Piazzesi (2003) Macro Model (estimated using our macro data, but with inflation expressed in y-o-y terms), and "HTV" denotes our structural macro model.

Table 4: Out-of-sample yield forecast performance: RMSEs

		1 month forecast horizon		
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.148	0.151	0.161	**0.130**
3 months	**0.173**	0.181	0.191	0.219
1 year	**0.194**	0.319	0.236	0.271
3 years	0.252	0.254	0.256	**0.237**
7 years	**0.220**	0.331	0.313	0.382

		3 month forecast horizon		
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.299	0.303	0.331	**0.215**
3 months	0.345	0.426	0.453	**0.292**
1 year	0.395	0.582	0.499	**0.385**
3 years	0.448	0.462	0.475	**0.399**
7 years	**0.379**	0.428	0.468	0.446

		6 month forecast horizon		
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.458	0.452	0.574	**0.305**
3 months	0.512	0.652	0.743	**0.404**
1 year	0.574	0.829	0.782	**0.532**
3 years	0.624	0.669	0.702	**0.537**
7 years	0.521	0.521	0.627	**0.478**

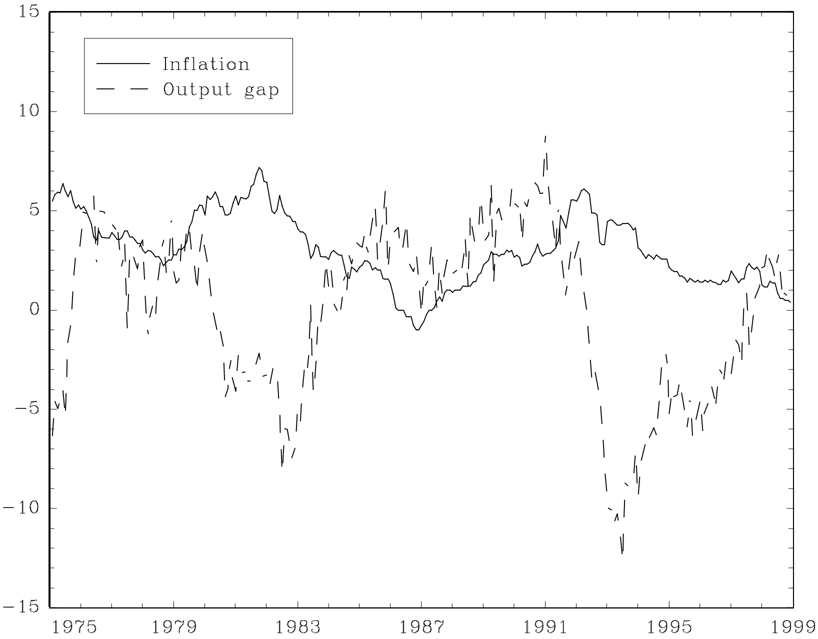
		9 month forecast horizon		
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.610	0.588	0.757	**0.441**
3 months	0.666	0.828	0.965	**0.557**
1 year	0.733	1.012	1.014	**0.691**
3 years	0.782	0.873	0.940	**0.681**
7 years	0.678	0.719	0.864	**0.579**

		12 month forecast horizon		
maturity	RW	$A_0(3)$	AP	HTV
1 month	0.747	0.706	0.891	**0.611**
3 months	0.793	0.956	1.146	**0.726**
1 year	0.854	1.140	1.201	**0.830**
3 years	0.842	1.002	1.125	**0.745**
7 years	0.806	0.920	1.103	**0.678**

RMSEs for out-of-sample forecasts between 1995:01 and 1998:12, based on parameter estimates for 1975:02 - 1994:12. RW are random walk forecasts, $A_0(3)$ is a canonical essentially affine Gaussian three-factor model, AP denotes the Ang-Piazzesi (2003) Macro Model (estimated using our macro data, but with inflation expressed in y-o-y terms), and "HTV" denotes our structural macro model.

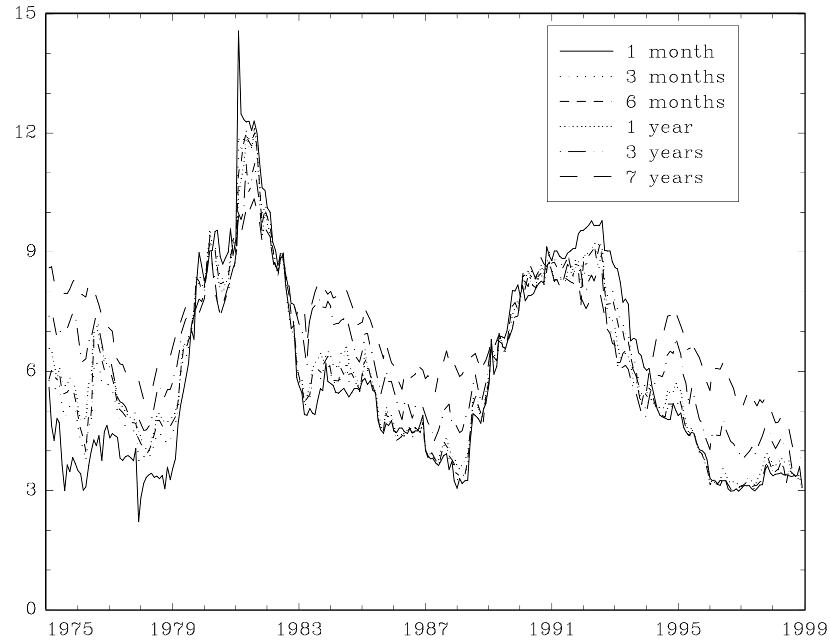
Figure 1: Data used in the estimations

(a) Macro data



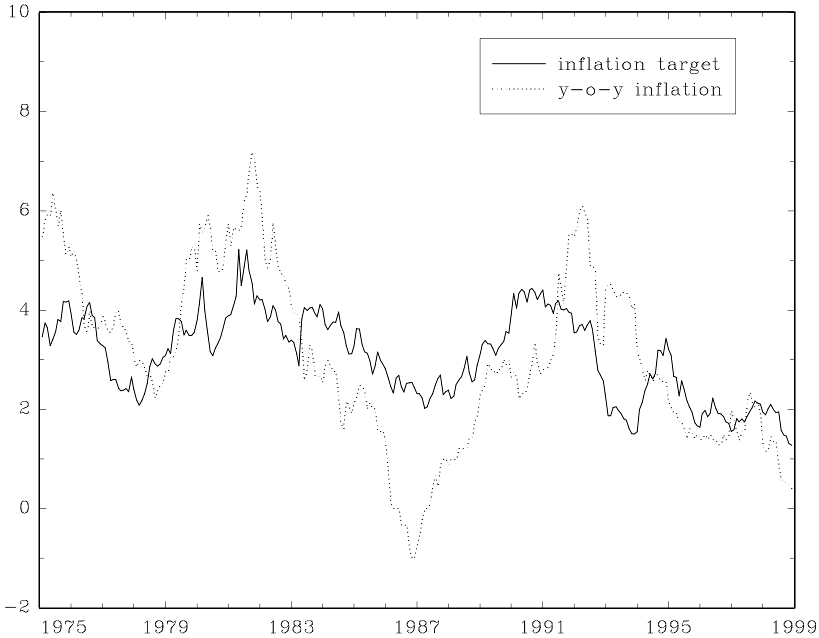
The inflation and output gap series have been multiplied by 100. The sample period is January 1975 to December 1998.

(b) Yield data



German term structure data over the sample period January 1975 to December 1998 (percent per year).

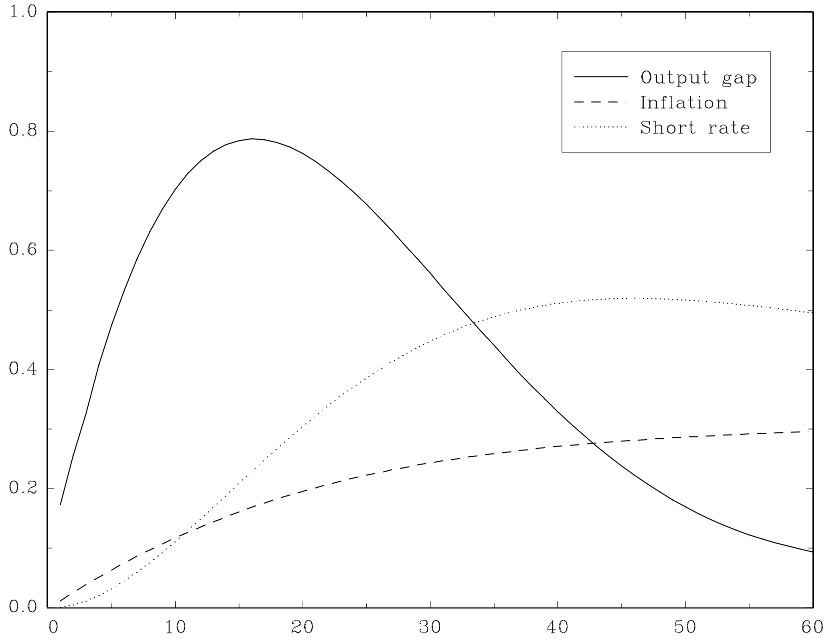
Figure 2: Estimated inflation target and actual inflation (year-on-year)



The estimated inflation target and the year-on-year inflation series have been scaled up by 100.

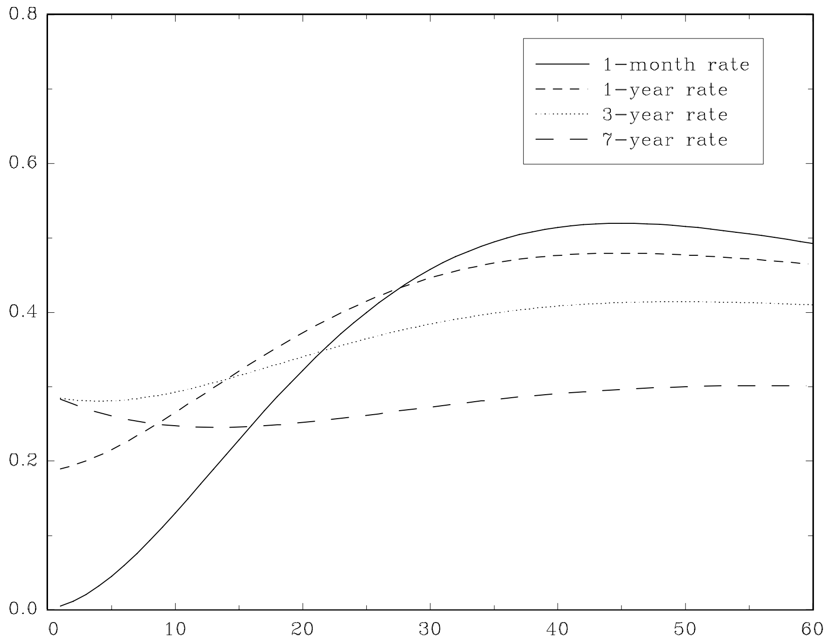
Figure 3: Impulse responses from inflation target shock

(a) Response of macro variables



All responses are expressed in percentage terms. The inflation and short rate responses are expressed in annual terms. The inflation target was shocked by one standard deviation (around 0.2% p.a.).

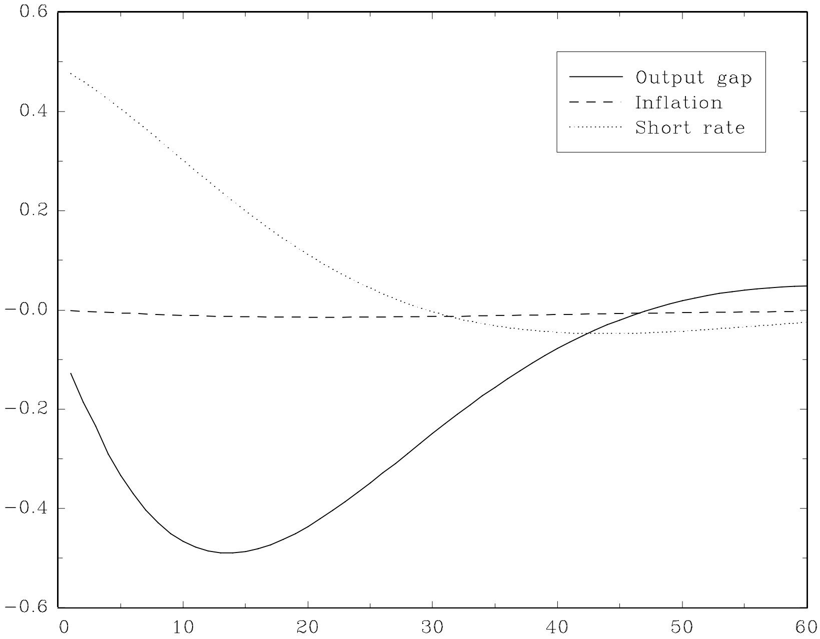
(b) Response of yields



All responses are expressed in annual percentage terms. The inflation target was shocked by one standard deviation (around 0.2% p.a.).

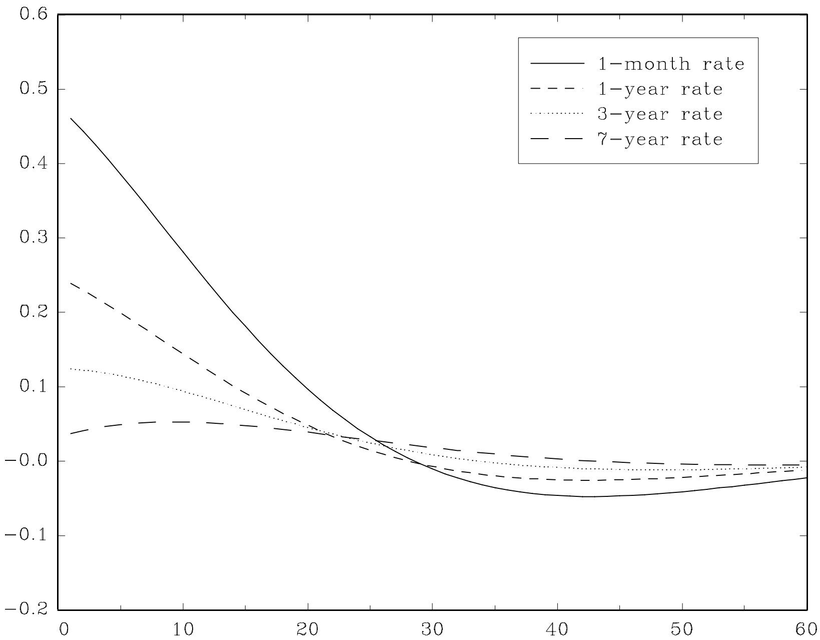
Figure 4: Impulse responses from monetary policy shock

(a) Response of macro variables



All responses are expressed in percentage terms. The inflation and short rate responses are expressed in annual terms. The short-term interest rate was shocked by one standard deviation (around 0.48% p.a.).

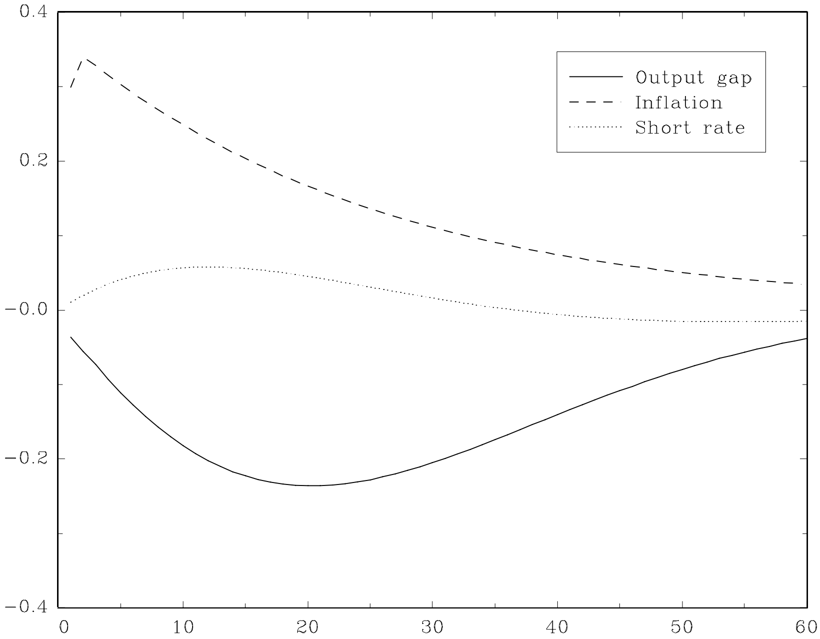
(b) Response of yields



All responses are expressed in annual percentage terms. The short-term interest rate was shocked by one standard deviation (around 0.48% p.a.).

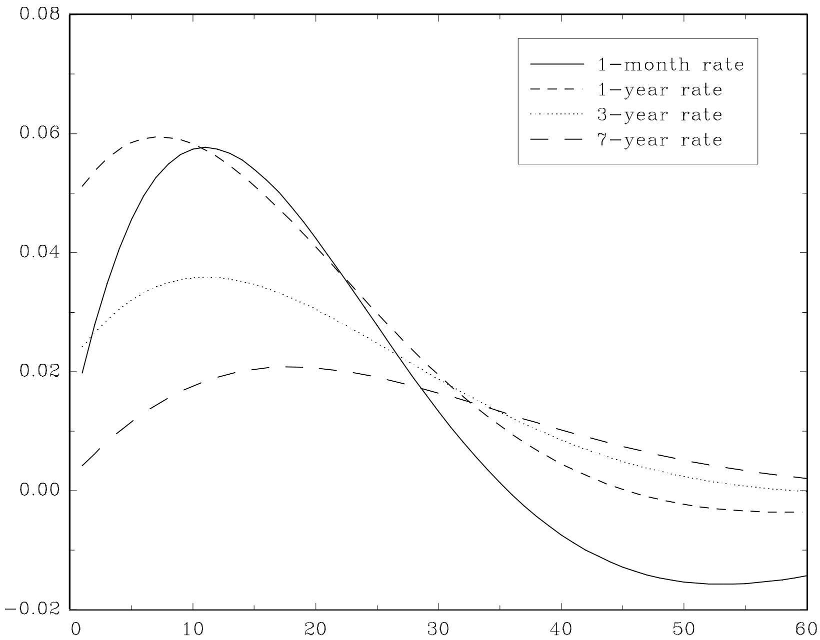
Figure 5: Impulse responses from inflation shock

(a) Response of macro variables



All responses are expressed in percentage terms. The inflation and short rate responses are expressed in annual terms. Inflation was shocked by one standard deviation (around 0.26% p.a.).

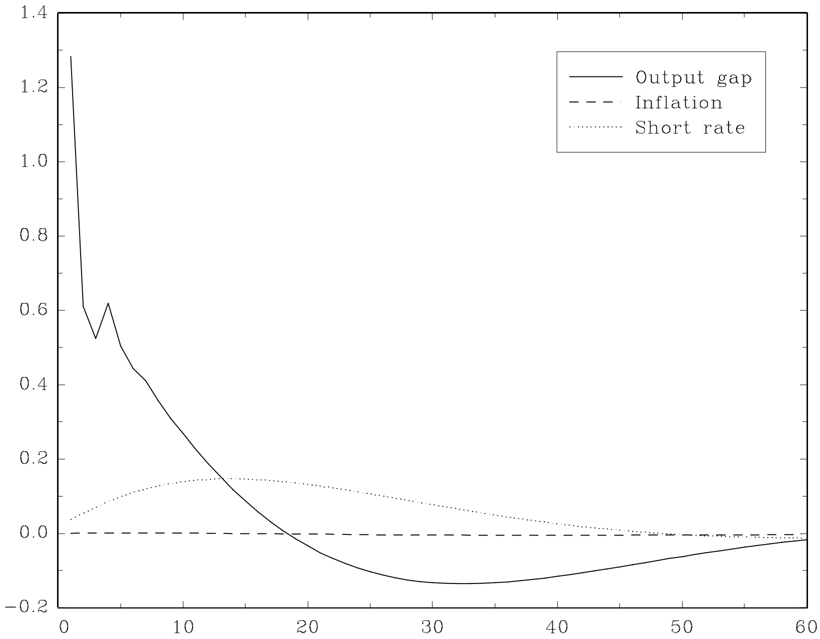
(b) Response of yields



All responses are expressed in annual percentage terms. Inflation was shocked by one standard deviation (around 0.26% p.a.).

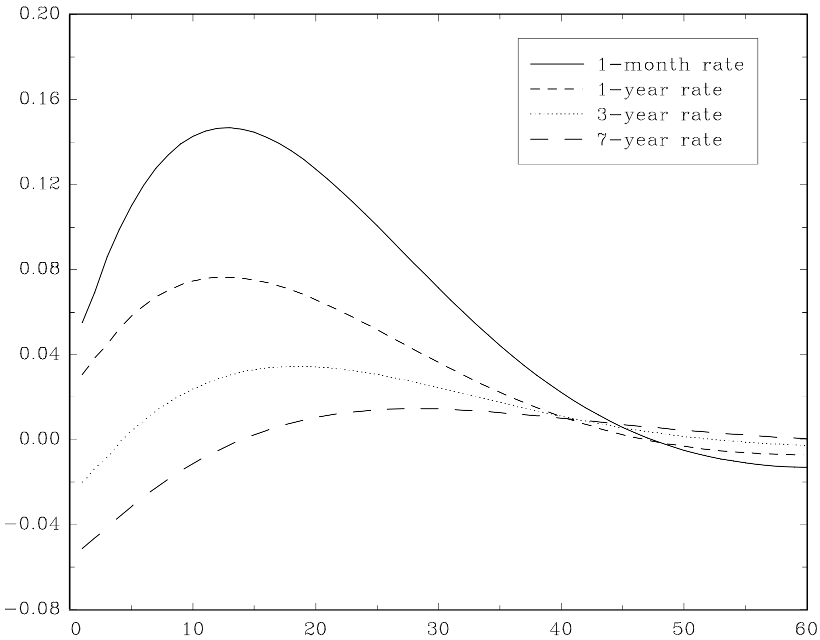
Figure 6: Impulse responses from output shock

(a) Response of macro variables



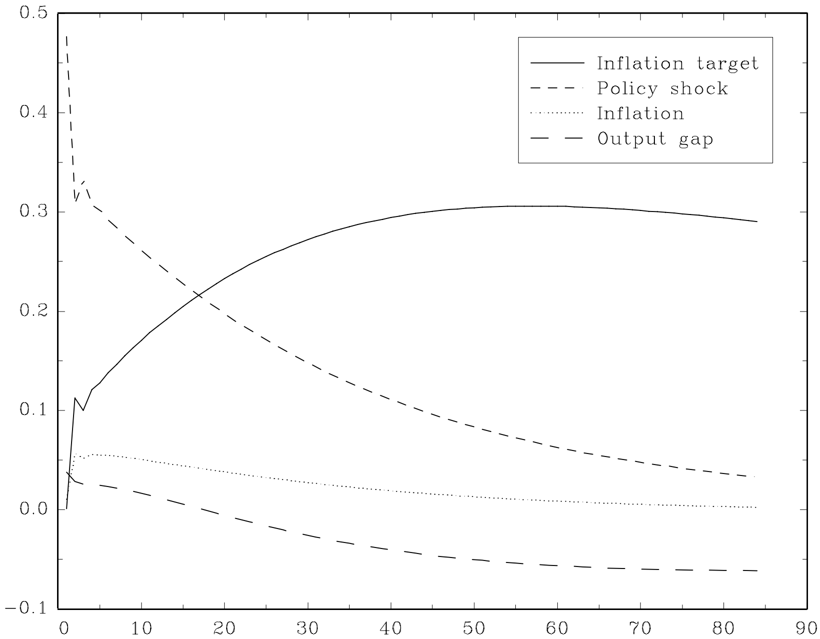
All responses are expressed in percentage terms. The inflation and short rate responses are expressed in annual terms. The output gap was shocked by one standard deviation (around 1.2%).

(b) Response of yields



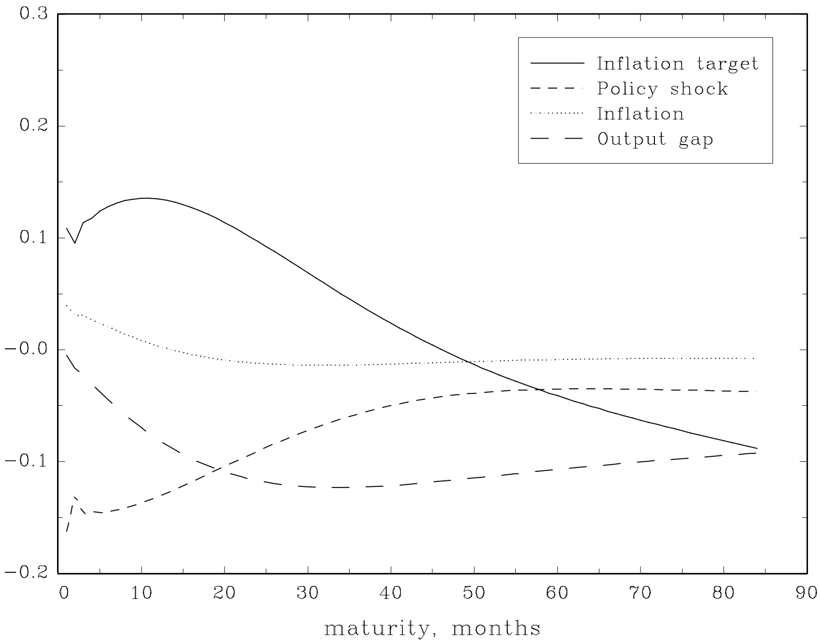
All responses are expressed in annual percentage terms. The output gap was shocked by one standard deviation (around 1.2%).

Figure 7: Factor loadings



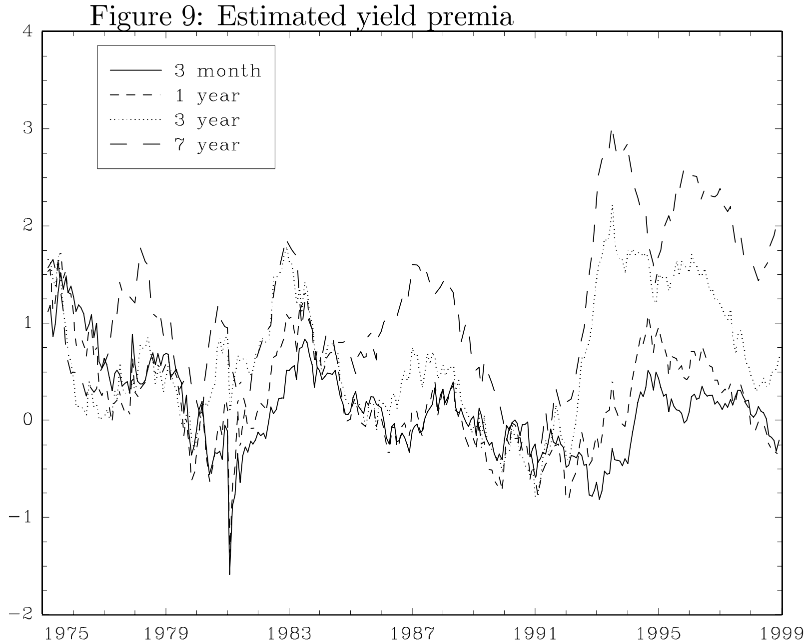
The factor loadings in the figure correspond to the B_n parameters of the four non-lagged macro factors. They have been rescaled to correspond to one standard deviation of the respective factors, expressed in annual percentage terms.

Figure 8: Initial response of yield premia to macro shocks



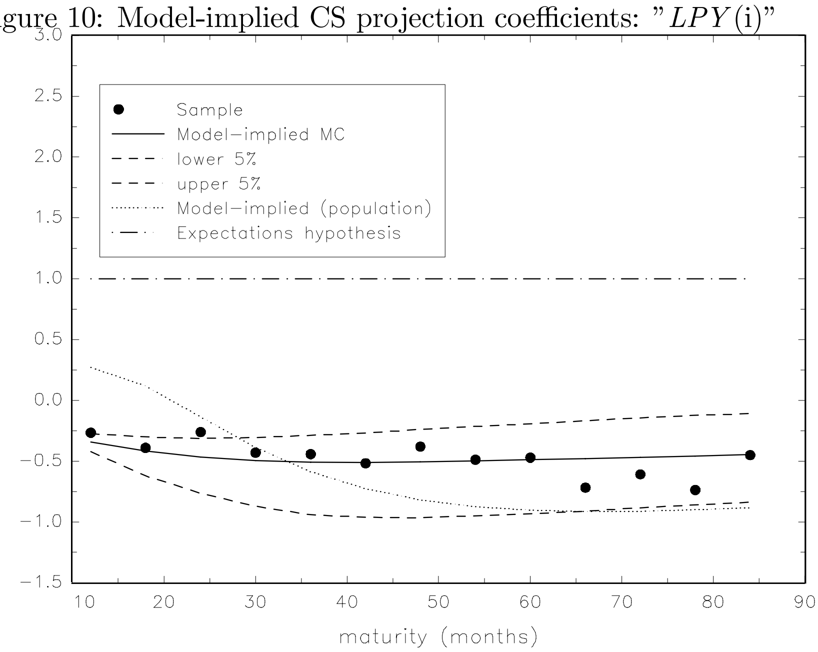
The figure shows the one-month ahead response of the yield premia ω_n, at maturities n up to 84 months, to one standard deviation shocks to the four macro factors. The premia are expressed in annual percentage terms.

Figure 9: Estimated yield premia



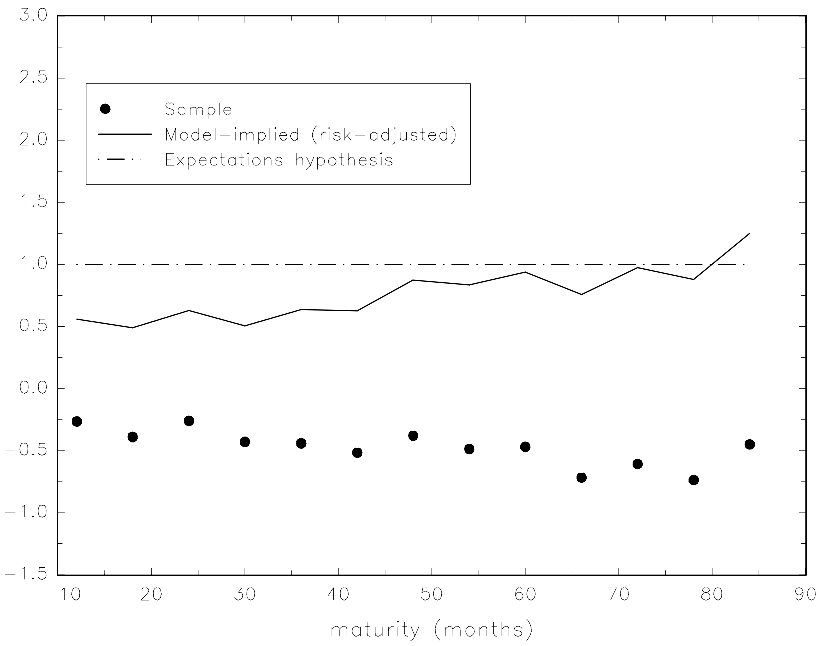
The figure shows the estimated yield premium ω_n during the sample period, for maturities n = 3, 12, 36, and 84 months. The premia are expressed in annual percentage terms.

Figure 10: Model-implied CS projection coefficients: "LPY(i)"



Empirical estimates of the CS long-rate coefficients ϕ_n in $y_{t+1}^{n-1} - y_t^n = \phi_n \left(y_t^n - r_t\right) / (n-1)$, plus corresponding model-implied cofficient values. The "population" coefficients are the theoretical values based on our estimates; the MC coefficients are the mean estimates from 1000 series of the same size as the sample, simulated from our model. The bands around the MC mean estimates are 5% confidence bands.

Figure 11: Model-implied risk-premium adjusted CS coefficients: "LPY(ii)"



The figure shows the estimates of the Campbell and Shiller (1991) long-rate coefficients ϕ_n in the regression $y_{t+1}^{n-1} - y_t^n = \phi_n (y_t^n - r_t) / (n-1)$ for our sample, along with the corresponding risk-premium adjusted model-implied cofficient values based on our parameter estimates.