A DSGE model of the term structure with regime shifts*

Gianni Amisano†  Oreste Tristani‡
European Central Bank  European Central Bank

11 October 2010

Abstract

We construct and estimate the term structure implications of a small DSGE model with nominal rigidities in which the laws of motion of the structural shocks are subject to stochastic regime shifts. We demonstrate that, to a second order approximation, switching regimes generate time-varying risk premia. We then estimate the model on US data relying on information from both macroeconomic variables and the term structure. Our results support the specification with regime-switching: heteroskedasticity is a clear feature of the model’s residuals and the regimes have intuitively appealing features. The model also generates non-negligible time-variability in excess holding period returns.

JEL classification:

Keywords: DSGE models, term structure of interest rates, policy rules, regime switches, Bayesian estimation.

---

*The opinions expressed are personal and should not be attributed to the European Central Bank. We wish to thank Michel Juillard for a precious suggestion on how to speed up the computation of the solution of the model. We also thank Giorgio Primiceri, Federico Ravenna, Ken Wallis, Philippe Weil and Paolo Zagaglia for useful comments and suggestions.

†Email: gianni.amisano@ecb.int.
‡Email: oreste.tristani@ecb.int.
1 Introduction

In the past twenty years, the term structure literature has developed independently of the macro-economic literature. While various arbitrage-free models have been shown to account quite well for yields dynamics – including affine, quadratic and regime switches models – these models have largely ignored the microfoundations of the stochastic discount factor used by the market to price bonds. They have also largely ignored the fact that the short term nominal interest rate, whose expected future values shape the whole yield curve, also plays an important macroeconomic role as a monetary policy instrument.

Yet there is a clearly established empirical relationship between monetary policy and the term structure of interest rates. For example, Mankiw and Miron (1986) points out that the yield curve used to behave differently – i.e. in a way more consistent with the so-called expectations hypothesis\(^1\) – before the founding of the Fed in 1913. Cochrane (2008) highlights an even starker structural break in 1933, with the shift from the Gold standard to an interest rate targeting regime. Contrary to the recent experience, long bond yields were systematically below short rates before 1933; also, long yields were much less volatile, and short rates much more volatile, than what is the case in post-WWII data.

The aforementioned stylised facts suggest that the monetary policy should play a central role in models of the term structure of interest rates, because it affects the structural relationship between developments in inflation and economic activity and the behaviour of the term structure. Conversely, the behaviour of the yield curve can provide useful information on market perceptions of the monetary policy rule followed by the central bank, insofar as the latter contributes to shape expected future short rates and term premia.

Calibrated versions of a new generation of DSGE models have been shown to be capable of producing roughly realistic implications for some unconditional moments of the term structure of interest rates, including slope and volatility, provided they are solved using second-order approximations or higher (see e.g. Hördahl, Tristani and Vestin, 2008; Ravenna and Seppala, 2007a, 2007b; Rudebusch and Swanson, 2009). In this paper, we take these models further and explore their ability to match conditional moments of macroeconomic and term structure data when they are estimated using full information.

\(^{1}\) "The expectations hypothesis, in the broadest terms, asserts that the slope of the term structure has something to do with expectations about future interest rates" (Shiller, 1990, p. 644)
methods.

One problem from this perspective is that models solved to a second-order approximation can only generate constant risk-premia, while the finance literature has highlighted the importance of allowing for time-variation in risk premia to match the conditional features of yields – see e.g. Dai and Singleton (2002). In order to allow for time-variation in risk premia, we assume heteroskedasticity in the model’s structural shocks, i.e. time-variation in the "amount of risk" faced by bond-holders at any point in time.\footnote{The finance literature, especially in affine term structure models, emphasises instead time-variations in risk premia due to changed in the "price of risk". Time variations in the price of risk can be produced within general equilibrium models if they are solved up to a third order approximation (or higher). This approach is pursued in Ravenna and Seppala (2007a, b), Rudebusch, Sack and Swanson (2007) and Rudebusch and Swanson (2007, 2008). However, these papers are purely theoretical: the estimation of DSGE models solved using third order approximations appears to be infeasible at this point in time.} We assume that heteroskedasiticy takes the specific form of regime switching. Moreover, the assumption of regime switching has already been shown to help fit yields in the finance literature – see Hamilton (1988), Naik and Lee (1997), Ang and Bekaert (2002a,b), Bansal and Zhou (2002), Bansal, Tauchen and Zhou (2004), Ang, Bekaert and Wei (2008), Dai, Singleton and Yang (2008), Bikbov and Chernov (2008) – and is also increasingly used in macroeconomics following Sims and Zha (2007).

We demonstrate analytically that, when combined with a second order approximation of the solution, this feature leads to changes in risk premia at the time of switches in regimes. More specifically, regime changes generate variations in the prices of risk, which are entirely consistent with the microfoundations of the model. While this mechanism does not explain \textit{why} risk premia vary, it forces their variation to be consistent with changes in the volatility of macro variables.

The second novelty of our model is a generalisation of the preferences proposed by Epstein and Zin (1989) and Weil (1990) to include habit persistence. Epstein-Zin-Weil preferences are quite standard in the finance literature – see e.g. Campbell (1999) – and they have already been successfully used to model yields in a partial equilibrium model by Piazzesi and Schneider (2006) and, more recently, Bansal and Shaliastovich (2008). Gallmeyer et al. (2007), Backus, Routledge and Zin (2007) and Rudebusch and Swanson (2009) have used these preferences in calibrated models. Binsbergen et al. (2008) is the only other application that we are aware of which estimates a DSGE model with
Epstein-Zin preferences. However, this paper relies on a benchmark RBC model and is therefore not suitable to analyse the interaction between monetary policy, inflation risk and consumption risk in the determination of risk premia.

Our empirical results are based on US data on aggregate consumption, inflation, the short-term interest rate and yields on 1-year and 5-year yields. The sample period runs from 1966Q1 until 2009Q1.

We find considerable support for a specification with regime switches. The residuals of the model show clear signs of heteroskedasticity, which could not be accounted for in a model with homoskedastic shocks. The model with regime switching can also fit yields reasonably well.

More specifically, we find strong evidence of time variability in expected excess holding period returns. Prema tend to be higher in the eighties, and display fluctuations which can be associated with the economic cycle. Volatility in premia is mainly linked to regime switches in the variance of technology shocks.

Our model is related to a growing literature exploring empirically the term structure implications of new-Keynesian models. The closest papers to ours is Doh (2006), which also estimates a quadratic DSGE model of the term structure of interest rates with heteroskedastic shocks. However, Doh (2006) allows for additional non-structural parameters to model the unconditional slope of the yield curve, while our approach is fully theoretically consistent. Another difference between the two papers is that heteroskedasticity in Doh (2006) is modelled through ARCH shocks, while it is generated by regime switching in our case. Andreasen (2008) shows that the estimation of a richer term structure model, which includes capital accumulation, is feasible to second order. However, the model cannot generate time-variation in risk premia because shocks are homoskedastic. Bekaert, Cho and Moreno (2006) and De Graeve, Emiris and Wouters (2007) estimate the loglinearised reduced form of DSGE models using both macroeconomic and term structure data. As in Doh (2006), these papers do not impose theoretical restrictions on the unconditional slope of the yield curve. In addition, they assume at the outset that risk-premia are constant.
2 The model

We rely on a relatively standard model in the spirit of Woodford (2003). The central feature is the assumption of nominal rigidities.

We only deviate from the standard model in postulating that households’ preferences can be described by the non-expected utility specification proposed by Epstein and Zin (1989) and Weil (1990). This specification is quite standard in the consumption-based asset pricing literature and it has already been employed to analyse the term structure of interest rate in a partial equilibrium model by Piazzesi and Schneider (2006). Here we extend this specification to a general equilibrium model in which we also allow for habit persistence in consumption and a labour-leisure choice – see also Backus, Routledge and Zin (2004, 2005). Rudebusch and Swanson (2009) also use non-expected utility preferences in a model similar to ours, but that paper relies on the assumption of homoskedastic shocks.

2.1 Households

We assume that each household \( i \) provides \( N(i) \) hours of differentiated labor services to firms in exchange for a labour income \( w_t (i) N_t (i) \). Each household owns an equal share of all firms \( j \) and receives profits \( \int_0^1 II(j) dj \).

As in Erceg, Henderson and Levin (2000), an employment agency combines households’ labor hours in the same proportions as firms would choose. The agency’s demand for each household’s labour is therefore equal to the sum of firms’ demands. The labor index \( L_t \) has the Dixit-Stiglitz form

\[
L_t = \left[ \int_0^1 N_t (i) \left( \frac{w_t (i)}{w_t} \right)^{\theta_{w,t}} \right]^\frac{1}{\theta_{w,t}}
\]

where \( \theta_{w,t} > 1 \) is subject to exogenous shocks. At time \( t \), the employment minimizes the cost of producing a given amount of the aggregate labor index, taking each household’s wage rate \( w_t (i) \) as given and then sells units of the labor index to the production sector at the aggregate wage index \( w_t = \left[ \int_0^1 w (i)^{1-\theta_{w,t}} \right]^{\frac{1}{1-\theta_{w,t}}} \). The employment agency’s demand for the labor hours of household \( i \) is given by

\[
N_t (i) = L_t \left( \frac{w_t (i)}{w_t} \right)^{-\theta_{w,t}}
\]

Each household \( i \) maximizes its intertemporal utility with respect to consumption, the wage rate and holdings of contingent claims, subject to its labor demand function (1) and the budget constraint

\[
P_t C_t (i) + E_t Q_{t+1} W_{t+1} (i) \leq W_t (i) + w_t (i) N_t (i) + \int_0^1 \Xi_t (j) dj
\]
where $C_t$ is a consumption index satisfying
\begin{equation}
C_t = \left( \int_0^1 C_t(z) \frac{\theta}{\theta - 1} \frac{\theta - 1}{\theta} dz \right) \frac{\theta}{\theta - 1}
\end{equation}

$W_t$ denotes the beginning-of-period value of a complete portfolio of state contingent assets, $Q_{t,t+1}$ is their price, $w_t(i)$ is the nominal wage rate and $\Xi_t(j)$ are the profits received from investment in firm $j$. The price level $P_t$ is defined as the minimal cost of buying one unit of $C_t$, hence equal to
\begin{equation}
P_t = \left( \int_0^1 p(z)^{1-\theta} dz \right)^{\frac{1}{1-\theta}}.
\end{equation}

Equation (2) states that each household can only consume or hold assets for amounts that must be less than or equal to its salary, the profits received from holding equity in all the existing firms and the revenues from holding a portfolio of state-contingent assets.

Households’ preferences are described by the Kreps and Porteus (1978) specification proposed by Epstein and Zin (1989). In that paper, utility is defined recursively through the aggregator $U$ such that
\begin{equation}
U[C_t, \left( E_t V_{t+1}^{1-\gamma} \right)] = \left\{ (1 - \beta) C_t^{1-\psi} + \beta \left( E_t V_{t+1}^{1-\gamma} \right)^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}, \quad \psi, \gamma \neq 1
\end{equation}

where $\beta, \psi$ and $\gamma$ are positive constants. Using a specification equivalent to that in equation (5), Weil (1990) shows that $\beta$ is, under certainty, the subjective discount factor, but time preference is in general endogenous under uncertainty. The parameter $\gamma$ is the relative risk aversion coefficient for timeless gambles. The parameter $1/\psi$ measures the elasticity of intertemporal of substitution for deterministic consumption paths.

The distinguishing feature of the Epstein-Zin-Weil preferences, compared to the standard expected utility specification, is that the coefficient of relative risk aversion can differ from the reciprocal of the intertemporal elasticity of substitution. In addition, Kreps and Porteus (1978) show that, again contrary to the expected utility specification, the timing of uncertainty is relevant in their class of preferences. The specification in equation (5) displays preferences for an early resolution of uncertainty when the aggregator is convex in its second argument, i.e. when $\gamma > \psi$. Any source of risk will be reflected in asset prices not only if it makes consumption more volatile, but also if it affects the temporal distribution of consumption volatility.

We generalise the utility function in equation (5) by allowing for habit formation and a labour-leisure choice. More specifically, time-$t$ utility will not only depend on consumption
but it will more generally be given by
\[ u_t = (C_t - hC_{t-1}) \cdot v(N_t) \]
where \( v(N_t) \) will be taken to equal \( v(N_t) = N - N_t^\phi \) and the \( h \) parameter represents the force of habits in the model: the higher \( h \), the less utility is generated by a given amount of current consumption.\(^3\) For \( h = 0 \), our preferences collapse to a special case of the class of preferences defined in Uhlig (2007). In our numerical analysis, we will adopt for the disutility of labour the formulation suggested by Trabandt and Uhlig (2009). This is given by \( v(N_t) = \left(1 - \kappa (1 - \psi) N_t^{1+\frac{1}{\psi}}\right)^{\frac{1}{1-\psi}} \) and it has the advantage of implying a constant Frisch elasticity equal to \( \phi \) in the \( h = 0 \) case.

With our more general preferences specification, \( \gamma \) and \( \psi \) are no-longer related one-to-one to risk aversion and to the (inverse of the) elasticity of intertemporal substitution of consumption, respectively. Swanson (2009) discusses the appropriate measures of risk aversion in a dynamic setting with consumption and leisure entering the utility function. In the rest of this paper, we simply refer to \( \gamma \) and \( \psi \) as utility parameters.

Each household \( i \) maximises
\[
U[C_t(i), N_t(i), E_tV_{t+1}] = \left\{ (1 - \beta) [(C_t(i) - hC_{t-1}(i)) \cdot v(N_t(i))]^{1-\psi} + \beta \left(E_t V_{t+1}^{1-\gamma}\right)^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} 
\]
subject to
\[
P_t C_t(i) + E_t Q_{t,t+1} W_{t+1}(i) \leq W_t(i) + w_t(i) N_t(i) + \int_0^1 \Xi_t(j) \, dj
\]
and
\[
N_t(i) = L_t \left( \frac{w_t(i)}{w_t} \right)^{-\theta_{w.t}}
\]
where the choice variables are \( w_t(i) \) and \( C_t(i) \).

To Bellman equation for this problem (abstracting from the \( i \) subscript to simplify the notation) is
\[
J(W_t, C_{t-1}) = \max \left\{ (1 - \beta) [(C_t - h_t C_{t-1}) v(N_t)]^{1-\psi} + \beta \left[E_t J^{1-\gamma}(W_{t+1}, C_t)\right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} - \Lambda_t \left[ P_t C_t + E_t Q_{t,t+1} W_{t+1} - W_t - w_t N_t - \int_0^1 \Xi_t(i) \, di + T_t \right] 
\]
\(^3\)Guariglia and Rossi (2002) also use expected utility preferences combined with habit formation to study precautionary savings in UK consumption. Koskievic (1999) studies an intertemporal consumption-leisure model with non-expected utility.
The appendix shows that the first order conditions can be written as

\[
\tilde{w}_t = -\mu_{w,t} \frac{v'(N_t) [(C_t - hC_{t-1}) v (N_t)]^{1-\psi}}{\tilde{\lambda}_t} \tag{7}
\]

\[
Q_{t,t+1} = \frac{\tilde{\Lambda}_{t+1}}{\tilde{\Lambda}_t} \frac{1}{\pi_{t+1}} \left( \left[ E_t J_{t+1}^{1-\gamma} \right] \frac{1}{J_{t+1}} \right)^{1-\gamma} \tag{8}
\]

\[
\tilde{\Lambda}_t = (C_t - hC_{t-1})^{-\psi} [v (N_t)]^{1-\psi} - \beta h E_t (C_{t+1} - hC_t)^{-\psi} [v (N_{t+1})]^{1-\psi} \left( \left[ E_t J_{t+1}^{1-\gamma} \right] \frac{1}{J_{t+1}} \right)^{-\psi} \tag{9}
\]

where \( \tilde{\Lambda}_t \equiv \Lambda_t P_t (1 - \beta)^{-1} J_t^{-\psi} \) and \( \tilde{w}_t \) is the real wage \( \tilde{w}_t \equiv w_t / P_t \) and \( \mu_{w,t} \equiv \theta_{w,t} / (\theta_{w,t} - 1) \).

The gross interest rate, \( I_t \), equals the conditional expectation of the stochastic discount factor, i.e.

\[
I_t^{-1} = E_t Q_{t,t+1} \tag{10}
\]

Note that we will focus on a symmetric equilibrium in which nominal wage rates are all allowed to change optimally at each point in time, so that individual nominal wages will equal the average \( w_t \).

Equations (8)-(9) highlight how our model nests the standard power utility case, in which \( \psi = \gamma \) and the maximum value function \( J_t \) disappears from the first order conditions. The same equations also demonstrate that the parameter \( \gamma \) only affects the dynamics of higher order approximations. To first order, the term \( \left[ E_t J_{t+1}^{1-\gamma} \right] \frac{1}{J_{t+1}} \) in equations (9) and (10) cancels out in expectation.

### 2.2 Firms

We assume a continuum of monopolistically competitive firms (indexed on the unit interval by \( j \)), each of which produces a differentiated good. Demand arises from households’ consumption and from government purchases \( G_t \), which is an aggregate of differentiated goods of the same form as households’ consumption. It follows that total demand for the output of firm \( i \) takes the form \( Y_t (j) = \left( \frac{P_t(i)}{P_t} \right)^{-\theta} Y_t \). \( Y_t \) is an index of aggregate demand which satisfies \( Y_t = C_t + G_t \).

Firms have the production function

\[
Y_t (j) = A_t L_t^\beta (j)
\]
where $L_t$ is the labour index $L_t$ defined above.

Once aggregate demand is realised, the firm demands the labour necessary to satisfy it $L_t (j) = (Y_t (j) / A_t)^{\frac{1}{\alpha}}$ so that the total nominal cost function will be given by

$$TC_t (j) = w_t \left( \frac{Y_t (j)}{A_t} \right)^{\frac{1}{\alpha}}$$

As a result, real marginal costs will be

$$mc_t (j) = \frac{1}{\alpha} \frac{\dot{w}_t}{A_t} \left( \frac{Y_t (j)}{A_t} \right)^{\frac{1-\alpha}{\alpha}}$$

where nominal costs are deflated using the aggregate price level.

As in Rotemberg (1982), we assume the firms face quadratic costs in adjusting their prices. This assumption is also adopted, for example, by Schmitt-Grohé and Uribe (2004) and Ireland (1997). It is well-known to yield first-order inflation dynamics equivalent to those arising from the assumption of Calvo pricing.\(^4\) From our viewpoint, it has the advantage of greater computational simplicity, as it allows us to avoid having to include an additional state variable in the model, i.e. the cross-sectional dispersion of prices across firms.

The specific assumption we adopt is that firm $j$ faces a quadratic cost when changing its prices in period $t$, compared to period $t-1$. Consistently with what is typically done in the Calvo literature, we modify the original Rotemberg (1982) formulation for partial indexation of prices to lagged inflation. More specifically, we assume that

$$\frac{\zeta}{2} \left( \frac{P_t^j}{P_{t-1}^j} - (\Pi_t^s)^{1-\epsilon} \Pi_{t-1}^s \right)^2 Y_t$$

where $\Pi_t^s$ is the inflation objective.

Firms maximise their real profits

$$\max_{P_t^j} \sum_{s=t}^{\infty} Q_{t,s} \left[ \frac{P_t^j Y_t^j (P_t^j)}{P_s} - \frac{TC_s \left( Y_s^j \left( P_s^j \right) \right)}{P_s} - \zeta \left( \frac{P_t^j}{P_{s-1}^j} - (\Pi_t^s)^{1-\epsilon} \Pi_{s-1}^s \right)^2 Y_s \right]$$

subject to

$$Y_t (j) = \left( \frac{P_t (j)}{P_t} \right)^{-\theta} Y_t$$

and to

$$Y_t (j) = A_t L_t^\alpha (j)$$

---

\(^4\)The two pricing models, however, have in general different welfare implications – see Lombardo and Vestin (2008).
Focusing on a symmetric equilibrium in which all firms adjust their price at the same time, the first order condition for price setting can be written as

\[(\theta - 1) Y_t + \zeta \left( \Pi_t - (\Pi^*)^{1-\epsilon} \Pi_{t-1} \right) Y_t \Pi_t = \frac{\theta}{\alpha} \bar{w}_t \left( \frac{Y_t}{A_t} \right)^{\frac{1}{\epsilon}} + E_0 Q_{t,t+1} \zeta \left( \Pi_{t+1} - (\Pi^*)^{1-\epsilon} \Pi_{t} \right) Y_{t+1} \Pi_{t+1} \]

### 2.3 Monetary policy

We close the model with the simple Taylor-type policy rule

\[I_t = \left( \frac{\Pi^*}{\beta} \right)^{1-\rho I} \left( \frac{\Pi_t}{\Pi^*} \right)^{\psi_I} \left( \frac{Y_t}{\bar{Y}} \right)^{\psi_Y} I_{t-1}^{\rho I} e^{\eta_{t+1}} \]

(11)

where \(Y_t\) is aggregate output, \(\Pi^*\) is the inflation target and \(\eta_{t+1}\) is a policy shock.

### 2.4 Market clearing

Market clearing in the goods market requires

\[Y_t = C_t + G_t\]

In the labour market, labour demand will have to equal labour supply. In addition, the total demand for hours worked in the economy must equal the sum of the hours worked by all individuals. Taking into account that at any point in time the nominal wage rate is identical across all labor markets because all wages are allowed to change optimally, individual wages will equal the average \(w_t\). As a result, all households will chose to supply the same amount of labour and labour market equilibrium will require that

\[L_t = \left( \frac{Y_t}{A_t} \right)^{\frac{1}{\alpha}} \]

### 2.5 Exogenous shocks

In macroeconomic applications, exogenous shocks are almost always assumed to be (log-)normal, partly because models are typically log-linearised and researchers are mainly interested in characterising conditional means. However, Hamilton (2008) argues that a correct modelling of conditional variances is always necessary, for example because inference on conditional means can be inappropriately influenced by outliers and high-variance episodes. The need for an appropriate treatment of heteroskedasticity becomes even more compelling when models are solved nonlinearly, because conditional variances have a direct impact on conditional means.
In this paper, we assume that variances are subject to stochastic regime switches for shocks other than the inflation target. More specifically, we assume a deterministic trend in technology growth

\[ A_t = Z_t B_t \]

\[ B_t = B_{t-1} \varepsilon \]

\[ Z_t = Z_{t-1}^z \varepsilon_{z,t}^z, \quad \varepsilon_{i+1} \approx N (0, \sigma_{z,i}, t) \]

where \( \varepsilon \) is the long run productivity growth rate. We specify the exogenous government spending process in deviation from trend, so that

\[ G_t = \left( g Y \right) \left( \frac{G_{t-1}}{B_{t-1}} \right)^{1-\rho_g} \left( \frac{G_{t-1}}{B_{t-1}} \right)^{\rho_g} \varepsilon_{g,t}^G \]

\[ \varepsilon_{g,t}^G \approx N (0, \sigma_{G,g}, t) \]

where the long run level \( g \) is specified in percent of output, so that \( g \equiv G/Y \). Finally, for monetary policy and mark-up shocks we assume

\[ \eta_{t+1} = \varepsilon_{\eta,t}^\eta, \quad \varepsilon_{\eta,t}^\eta \approx N (0, \sigma_{\eta,s}, t) \]

\[ \mu_{\mu,t+1} = \mu_{\mu,t} \varepsilon_{\mu,t+1}^\mu, \quad \varepsilon_{\mu,t+1}^\mu \approx N (0, \sigma_{\mu}) \]

Technology and monetary policy shocks have regime-switching variances, namely

\[ \sigma_{z,s_z,t} = \sigma_{z,L} s_{z,t} + \sigma_{z,H} (1 - s_{z,t}) \]

\[ \sigma_{\eta,s_\eta,t} = \sigma_{\eta,L} s_{\eta,t} + \sigma_{\eta,H} (1 - s_{\eta,t}) \]

and the variables \( s_{z,t} \) and \( s_{\eta,t} \) can assume the discrete values 0 and 1. For each variable \( s_{j,t} (j = z, \eta) \), the probabilities of remaining in states 0 and 1 are constant and equal to \( p_{j,0} \) and \( p_{j,1} \), respectively.\(^5\)

We assume regime switches in these particular variances for the following reasons. The literature on the "Great moderation" (see e.g. McDonnell and Perez-Quiros, 2000) has emphasised the reduction in the volatility of real aggregate variables starting in the second half of the 1980s. We conjecture that this phenomenon could be captured by a reduction in the volatility of technology shocks in our structural setting. The heteroskedasticity in policy shocks aims to capture the large increase in interest rate volatility in the early 1980s, the time of the so-called "monetarist experiment" of the Federal Reserve.\(^6\)

---

\(^5\) In previous versions of the paper we have allowed for regime-switching also in the variance of government spending and mark-up shocks. These additional dimensions of regime switching receive little support from the data.

\(^6\) A similar assumption in made in Schorfheide (2005).
2.6 Solution method

To solve the model, we exploit the recursive nature of bonds in equilibrium. We first solve for all macroeconomic variables and then construct the prices of bonds of various maturities.

2.6.1 Solving the macroeconomic system

We approximate the system around a deterministic steady state in which all variables are detrended. Detrend variables are denoted by a tilde and defined as a ratio to $B_t$ (with the exception is $\tilde{\Lambda}_t$, which is detrended by $B_t^{-\psi}$). For example, detrended output is $\tilde{Y}_t = Y_t / B_t$.

In the solution, we expand variables around their natural logarithms, which are denoted by lower-case letters. The logarithm of a variable in deviation from its non-stochastic steady state is denoted by a hat. For example, approximate (detrended) output is denoted by $\hat{y}_t$.

For the solution, we collect all predetermined variables (including both lagged endogenous predetermined variables and exogenous states) in a vector $x_t$ and all the non-predicted variables in a vector $y_t$ (note that $y_t$ is different from output $Y_t$).

The macroeconomic system can thus be written in compact form as

$$
\begin{align*}
    y_t &= g(x_t, \tilde{\sigma}, s_t) \\
    x_{t+1} &= h(x_t, \tilde{\sigma}, s_t) + \tilde{\sigma} \Sigma(s_t) u_{t+1} \\
    s_{t+1} &= \kappa_0 + \kappa_1 s_t + \nu_{t+1}
\end{align*}
$$

for matrix functions $g(\cdot)$, $h(\cdot)$, and $\Sigma(\cdot)$, a vector $s_t$ including the state variables that index the discrete regimes, and a vector of innovations $u_t$. In the above system, $\tilde{\sigma}$ is a perturbation parameter. Following Hamilton (1994), we can write the law of motion of the discrete processes $s_t$ in the form implied in equation (14) for a vector $\kappa_0$ and a matrix $\kappa_1$. The law of motion of state $s_{z,t}$, for example, is written as $s_{z,t+1} = (1 - p_{z,0}) + (-1 + p_{z,1} + p_{z,0}) s_{z,t} + \nu_{z,t+1}$, where $\nu_{z,t+1}$ is an innovation with mean zero and heteroskedastic variance.

We seek a second-order approximation to the functions $g(x_t, \tilde{\sigma}, s_t)$ and $h(x_t, \tilde{\sigma}, s_t)$ around the non-stochastic steady state, namely the point where $x_t = \bar{x}$ and $\bar{\sigma} = 0$. Due to the presence of the discrete regimes in the system, both the steady state and the coefficients of the second order approximation could potentially depend on $s_t$. Since the discrete states only affect the variance of the shocks, however, they disappear when $\tilde{\sigma} = 0$ so that the
non-stochastic steady state is not regime-dependent. In a companion paper (Amisano and Tristani, 2009b), we show that the second order approximation to the solution can be written as

\[ g(x_t, \sigma_t, s_t) = F\hat{x}_t + \frac{1}{2} (I_{ny} \otimes \hat{x}'_t) E\hat{x}_t + k_{y,s}\hat{\sigma}^2 \]

and

\[ h(x_t, \sigma_t, s_t) = P\hat{x}_t + \frac{1}{2} (I_{nx} \otimes \hat{x}'_t) G\hat{x}_t + k_{x,s}\hat{\sigma}^2 \]

where \( F, E, P \) and \( G \) are constant variances and only the vectors \( k_{y,s} \) and \( k_{x,s} \) are regime dependent. As a result, regime switching plays no role to a first order approximation. It only affects the means of endogenous variables.

### 2.6.2 Pricing bonds

Once the solution of the macroeconomic model is available, bond yields can be solved for analytically.

Note that the stochastic discount factor can be rewritten in terms of detrended variables as

\[ Q_{t,t+1} = \beta \frac{\tilde{\Lambda}_{t+1}}{\tilde{\Lambda}_t} \frac{1}{\Pi_{t+1}} \left( \frac{[E_t\tilde{J}_{t+1}^{-\gamma}]^{\frac{1}{1-\gamma}}}{\tilde{J}_{t+1}} \right)^{\gamma-\psi} \]

This expression can be written more simply as

\[ Q_{t,t+1} = \beta \frac{\tilde{\Lambda}_{t+1}}{\tilde{\Lambda}_t} \frac{1}{\Pi_{t+1}} \]

for

\[ \Pi_t \equiv \Pi_t \tilde{J}_t^{-\psi} [E_t\tilde{J}_{t+1}^{-\gamma}]^{\frac{1}{1-\gamma}} \]

\[ \Lambda_t \equiv \tilde{\Lambda}_t [E_t\tilde{J}_{t+1}^{-\gamma}]^{\frac{1}{1-\gamma}} \]

Since these relationships are all loglinear, the law of motion for \( \Pi_t \) and \( \Lambda_t \) can immediately be derived from those of \( \Pi_t, \tilde{\Lambda}_t, \tilde{J}_t \) and \( D_t \equiv E_t\tilde{J}_{t+1}^{-\gamma} \). It follows that

\[ \tilde{\Lambda}_t = F_\lambda \hat{x}_t + \frac{1}{2} \hat{x}'_t E_\lambda \hat{x}_t + k_{\lambda,s}\hat{\sigma}^2 \]

\[ \tilde{\Pi}_t = F_\pi \hat{x}_t + \frac{1}{2} \hat{x}'_t E_\pi \hat{x}_t + k_{\pi,s}\hat{\sigma}^2 \]

where \( F_\lambda \) and \( F_\pi \) are row vectors, and \( E_\lambda \) and \( E_\pi \) are matrices. We can now compute bond prices using the method in Hördahl, Tristani and Vestin (2008). The appendix shows
that, in log-deviation from its deterministic steady state, the approximate price of a bond of maturity \( n \), \( \hat{b}_{t,n} \), can be written as

\[
\hat{b}_{t,n} = F_{B_n} \hat{x}_t + \frac{1}{2} \hat{x}_t E_{B_n} \hat{x}_t + k_{B_n,s_t} \sigma^2
\]  

(16)

where \( F_{B_n}, E_{B_n} \) and \( k_{B_n,s_t} \) are defined through a recursion. Note that \( k_{B_n,s_t} \) changes depending on the realisation of the discrete states, but matrices \( F_{B_n} \) and \( E_{B_n} \) are state-independent.

3 Some properties of the model

3.1 The stochastic discount factor

The appendix shows that, to a second order approximation, the stochastic discount factor in equation (8) can be written as

\[
\hat{q}_{t+1} = \Delta \hat{\lambda}_{t+1} - \hat{\pi}_{t+1} - (\gamma - \psi) \left( \hat{j}_{t+1} - E_t \hat{j}_{t+1} \right) - \frac{1}{2} (\gamma - \psi) (\gamma - 1) \text{Var}_t \left[ \hat{j}_{t+1} \right]
\]  

(17)

When: (a) \( \psi = 1 \); (b) temporary utility depends on consumption only and \( \Delta \hat{\lambda}_{t+1} = -\Delta \hat{c}_{t+1} \), equation (17) boils down to

\[
\hat{q}_{t+1} = -\Delta \hat{c}_{t+1} - \hat{\pi}_{t+1} - (\gamma - 1) \left( \hat{j}_{t+1} - E_t \hat{j}_{t+1} \right) - \frac{1}{2} (\gamma - 1)^2 \text{Var}_t \left[ \hat{j}_{t+1} \right]
\]

which corresponds to the case considered by Piazzesi and Schneider (2006).

In the expected utility case, only the risk of unpredictable changes in future consumer prices, in future detrended marginal utility, or in technology growth matter for the investor.

With Epstein-Zin preferences, the whole temporal distribution of future risks to detrended marginal utility and technology growth becomes relevant. In our case, similarly to Uhlig (2007), detrended marginal utility is affected both by future detrended consumption growth and by future levels of labour supply.

Following Restoy and Weil (2010) and Piazzesi and Schneider (2006), we can solve out the value function as an infinite sum of future expected utility. The appendix derives this representation. In general, the value function will be affected by expected future productivity growth and the discounted future growth rates of consumption and of labour supply. The expected future change in consumption growth also matters because of habit formation.
To have some intuition for the implications of the model in terms of bond pricing, we can derive the short-term interest rate as $\hat{\hat{i}}_t = -E_t \hat{\hat{q}}_{t,t+1} - \frac{1}{2} \text{Var}_t \hat{\hat{q}}_{t,t+1}$. We obtain

$$\hat{\hat{i}}_t = \hat{\hat{\gamma}}_t + E_t \hat{\hat{\alpha}}_{t+1} - \frac{1}{2} \text{Var}_t \hat{\hat{\alpha}}_{t+1} + \text{Cov}_t \left[ \hat{\hat{\mu}}_{t+1}, \hat{\hat{\alpha}}_{t+1} \right] - (\gamma - \psi) \text{Cov}_t \left[ \hat{\hat{\alpha}}_{t+1}, \Psi_{t+1} \right]$$

(18)

where

$$\hat{\hat{\mu}}_{t+1} \equiv \Delta \hat{\hat{\lambda}}_t$$

is the change in the marginal utility of consumption,

$$\Psi_{t+1} \equiv E_{t+1} \sum_{i=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^i \Delta \hat{\hat{\nu}}_{t+1+i}$$

denotes expected discounted future changes in utility – i.e. the sum of expected discounted future consumption growth and expected future changes in the disutility of labour – and $\hat{\hat{\gamma}}_t$ is the real interest rate.

$$\hat{\hat{\gamma}}_t = -E_t \hat{\hat{\mu}}_{t+1} - \frac{1}{2} \text{Var}_t \hat{\hat{\mu}}_{t+1} + (\gamma - \psi) \text{Cov}_t \left[ \hat{\hat{\mu}}_{t+1}, \Psi_{t+1} \right] + \frac{1}{2} (\psi - 1) (\gamma - \psi) \text{Var}_t \Psi_{t+1}$$

(19)

According to expression (18), the nominal interest rate includes three components in excess of the real rate and expected inflation. The first component is the variance of inflation, which is a Jensen’s inequality term. The second component, the covariance between inflation and the marginal utility of consumption at date $t+1$, is a standard inflation risk premium term. The return on a nominal 1-period bond will be higher if inflation at $t+1$ tends to be high when the marginal utility at $t+1$ is also high. Note that this term will be larger under habit persistence, which tends to boost the sensitivity of marginal utility to changes in consumption. The last term in equation (18), which is proportional to the covariance between inflation at $t+1$ and expected future changes in utility, is also an inflation risk premium and it can be understood along the lines suggested by Restoy and Weil (2010). When the covariance is negative, inflation at $t+1$ tends to be high, and the real return on nominal bonds tends to be low, in case of bad news about expected future utility. Nominal bonds are therefore risky, and require a premium given by the term $\gamma \text{Cov}_t \left[ \hat{\hat{\mu}}_{t+1}, \Psi_{t+1} \right]$. At the same time, the bad news about the future induces households to increase their savings, so as to smooth consumption over time. From this viewpoint, nominal bonds are desirable assets and command a discount $-\psi \text{Cov}_t \left[ \hat{\hat{\mu}}_{t+1}, \Psi_{t+1} \right]$. With expected utility preferences, these effects cancel out and news about future utility growth become irrelevant for asset pricing.
The real interest rate in equation (19) also includes three additional components compared to the linearised case. The first one, the variance of the marginal utility of consumption, is a standard precautionary savings motive. The larger the volatility of future marginal utility growth, the higher the incentive to hold precautionary savings and the lower the return on real bonds. Epstein-Zin-Weil utility with $\psi = 1$ – the case analysed in Piazzesi and Schneider (2006) – introduces a second source of risk premia, namely the covariance term in equation (19). The real interest rate includes a premium when bad news about future utility growth over the indefinite future tend to be associated with high marginal utility at $t+1$. The premium is higher if agents are unwilling to adjust their level of utility across states ($\gamma$ is high) and lower the more averse they are to adjust utility across time. Finally, a third source of risk premium arises under Epstein-Zin-Weil utility in the $\psi \neq 1$ case, namely a premium related to the variance in revisions of expected future utility.

3.2 Regime switching and the variability of risk premia

The state-dependence of $\hat{b}_{t,n}$ in equation (16) implies that bond risk premia will also become state-dependent. In order to show this, it is useful to derive expected excess holding period returns, i.e. the expected return from holding a $n$-period bond for 1 period in excess of the return on a 1-period bond. To a second order approximation, the expected excess holding period return on an $n$-period bond can be written as

$$hpr_{t,n} - \hat{h}_t = Cov_t \left[ \hat{\pi}_{t+1, \hat{b}_{t+1,n-1}} \right] - Cov_t \left[ \Delta \lambda_{t+1, \hat{b}_{t+1,n-1}} \right]$$

This expression can be evaluated using the model solution to obtain

$$hpr_{t,n} - \hat{h}_t = \sigma^2 F_{B_{n-1}} \Sigma_t \Sigma_t' \left( F_x' - F_{\beta}' \right)$$

where $\Sigma_t \equiv \Sigma(s_t)$ and $\Sigma_t \Sigma_t'$ is the conditional variance-covariance matrix depending on vector $s_t$.

Equation (20) demonstrates that excess holding period returns change when there is a switch in any of the discrete state variables. Since the conditional variance of the price of a bond of maturity $n$ can be written, to a second order approximation, as $E_t \left[ \hat{b}_{t+1,n-1} \hat{b}_{t+1,n-1} \right] = \sigma^2 F_{B_{n-1}} \Sigma_t \Sigma_t' F_{B_{n-1}}'$, it follows that we can define the (micro-founded) price of risk for unit of volatility, or the "market prices of risk" $\omega_t$, as

$$\omega_t \equiv \sigma \Sigma_t' \left( F_x' - F_{\beta}' \right)$$
Since $F_\pi$ and $F_\lambda$ are vectors of constants, all terms in equation (21) would be constant in a world with heteroskedastic shocks, in which $\Sigma_t$ would also be constant. They becomes time-varying in our model due to the possibility of regime switches, because the variance-covariance matrix $\Sigma_t \Sigma_t'$ is regime-dependent.

In the empirical finance literature, the market prices of risk are often postulated exogenously using slightly different specifications. For example, Naik and Lee (1997), Bansal and Zhou (2002) and Ang, Bekaert and Wei (2008) assume that the market prices of risk are regime dependent, but the risk of a regime-change is not priced. On the contrary, regime-switching risk is priced in Dai, Singleton and Yang (2008).

In our model, prices of risk are only associated with variables with continuous support. These prices change across regimes. If, for example, technological risk were not diversifiable, then the price of risk associated with technology shocks would be higher in a high-variance regime for technology shocks (and lower in a low-variance regime). This is the regime-dependence of market prices of risk which is present in all the aforementioned finance models. In our set-up, however, the prices of risk are additionally derived from the model’s microfoundation, rather than allowed to vary as affine functions of the continuous state variables of the model.

The risk of regime-switches is not priced because the possibility of changes in regime does not have any impact on the $F_\pi$ and $F_\lambda$ vectors. Regime switching risk would only be priced if it affected some structural parameters. For example, one could thing of allowing for changes in the parameters of the (11) as in Bikbov and Chernov (2008). We leave this extension to future research.

4 Empirical results

4.1 Estimation methodology

The system of equations (12) and (13) can be re-written as

$$y_{t+1}^o = c_j + C_1 x_{t+1} + C_{2vech}(x_{t+1} x_{t+1}') + D v_{t+1}$$  \hspace{1cm} (22)

$$x_{t+1} = a_i + A_1 x_t + A_{2vech}(x_t x_t') + B_i w_{t+1}$$  \hspace{1cm} (23)

$$s_t \sim \text{Markov switching}$$  \hspace{1cm} (24)
where the vector $y_t^0$ includes all observable variables, and $v_{t+1}$ and $w_{t+1}$ are measurement and structural shocks, respectively. In this representation, the regime switching variables affect the system by changing the intercepts $a_i$ and $c_j$, and the loadings for the of the structural innovations $B_i$. (we indicate here with $i$ the value of the discrete state variables at $t$ and with $j$ the value of the discrete state variables at $t + 1$).

If a linear approximation were used, we would be left with a linear state space model with Markov switching (see Kim, 1994, Kim and Nelson, 1999, and Schorfheide, 2005). Focusing on the case in which the number of continuous shocks (measurement and structural) is equal to the number of observables, the continuous latent variables could be obtained via inversion of the observation equation (22). The system could then be written as a Markov Switching VAR and the likelihood could be obtained using Hamilton’s filter i.e. by integrating out the discrete latent variables.

In the quadratic case, however, the likelihood cannot in general be obtained in closed form. One possible approach to compute the likelihood is to rely on sequential Monte Carlo techniques (see Amisano and Tristani, 2009a, for an application of these techniques in a DSGE setting with homoskedastic shocks). The convergence of these methods, however, can be very slow in a case, such as the one of our model, in which both nonlinearities and non-Gaussianity of the shocks characterise the economy.

We thus adopt an extension of the filter which can be used in the linear case, when the number of continuous shocks is equal to the number of observables (see Amisano and Tristani, 2009b, for further details). The problem of filtering the latent variables through an inversion of the quadratic observation equation (22) is that the inversion is not unique. At each point in time, multiple values of the latent variables are consistent with the observation vector $y_t^0$. In a scalar case and in the absence of measurement errors, for example, we would obtain the two solutions $x_t^{(1),(2)} = 1/2 \left( C_1 \pm \sqrt{C_1^2 - 4C_2(c_j - y_t^0)} \right) / C_2$. In the more general case of our model, we are going to use five series in the estimation process. As a result, at each point in time we will have up to 16 solutions for each of the latent variables. 7

Rather than choosing arbitrarily a particular solution at each point in time, we compute

---

7 Given that four structural shocks enter the model, we are going to assume that one series is observed with measurement error. The inversion for the measurement error is unique, because the measurement error enters the model linearly.
the likelihood taking all (real) solutions into account. We simply exploit the property that, while equally likely based on the sole observation vector $y_t$, the different solutions for $x_t$ have different probability (or likelihood) conditional on $x_{t-1}$. In a homoskedastic model, the filtered values of our latent variables in $t$ would simply be a weighted average of all $x_t$ solutions, with weights given by their conditional probabilities. In our model with regime switching, solutions for $x_t$ must be found for each of the regimes in $s_t$, and then weighed by the probability of each regime.

Provided all solutions of the observation equation are found, this procedure produces the exact likelihood of the quadratic system (22)-(23). To find all solutions, we rely on homotopy continuation methods – see e.g. Judd (1998) and Morgan (1987). More specifically, we rely on the PHCpack solver described in Verschelde (1999) and its Matlab interface PHClab presented in Guan and Verschelde (2008). We discard all complex solutions (which tend to be the majority) and compute the likelihood using the real ones. If no real solutions are available at any point in time for a certain value of the parameter vector, we impute to the likelihood a large negative value.

4.2 Data and prior distributions

We estimate the model on quarterly US data over the sample period from 1966Q1 to 2009Q1. Our estimation sample starts in 1966, because this is often argued to be the date after which a Taylor rule provides a reasonable characterisation of Federal Reserve policy.\textsuperscript{8}

Concerning the macro-data, we focus on inflation and one measure of economic activity. For the latter, we could in principle use GDP, but we prefer to focus on consumption, given that our model abstracts from investment. Nevertheless, consumption in our model captures all interest-sensitive components of private expenditure. As argued by Giannoni and Woodford (2005), assuming habit persistence for the whole level of private expenditure is a reasonable assumption, given that models with capital typically need adjustment costs that imply inertia in the rate of investment spending. We therefore include in the information set total real personal consumption per-capita and the consumption deflator

\textsuperscript{8}According to Fuhrer (1996), "since 1966, understanding the behaviour of the short rate has been equivalent to understanding the behaviour of the Fed, which has since that time essentially set the federal Funds rate at a target level, in response to movements in inflation and real activity". Goodfriend (1991) argues that even under the period of official reserves targeting, the Federal Reserve had in mind an implicit target for the Funds rate.
Prior and posterior distributions for our model are presented in Table 1.

Concerning regime switching processes, we assume beta priors for transition probabilities. The distributions imply that persistences in each state are symmetric and have high means. In the prior, we assume that the standard deviations of the structural shocks are identical in the various states.

The priors for the standard deviation and persistence of shocks, as well as for the long run growth rate of technology and for the long run inflation target, are centred on values which allow us to roughly match unconditional data moments in the first 10 years of the sample, given the other parameter.

For the policy rule, we estimate use relatively loose priors centred around the classic Taylor (1993) parameters for the responses to inflation and output, but we also allow for a mild degree of interest rate smoothing. Finally, for the other parameters we use priors broadly in line with other macro studies. For $\gamma$, we use a loose prior centred around a value a bit higher than $\psi$.

### 4.3 Posterior distributions and goodness of fit

The posterior distributions of macro parameters in Table 1 are roughly consistent with other estimates based solely on macro data. The $\gamma$ parameter is larger than $\psi$, but it does not reach very high levels. Our model generates a slope in the term structure of interest rates through a combination of moderate risk-aversion and occasional bursts in the variance of shocks, especially technology shocks.

The standard deviation of measurement errors on 10-year yields is around 25 basis points, thus not exceedingly higher than in more flexible term structure models estimated using only yields data.

Concerning regime-switching parameters, the posterior mode of the transition probabilities suggests that all states are very persistent. The difference between estimated variances in the two regimes is marked for both technology and policy shocks.

Figures 1 and 2 show the implications of our parameter estimates for the population means, variances and cross-correlations of our observable variables. By and large, sample
moments are within the posterior distribution of their population counterparts. In a few cases, however, sample moments are very much at the tail of the distribution of theoretical moments.

4.4 Regime switches and risk premia

Figure 3 displays 1-step-ahead forecasts and realised variables. This figure shows that the model can track yields data relatively well. In comparison, the fit is less good for inflation. One-step-ahead forecast errors are larger for the rates of growth of consumption and output.

Figure 3 also illustrates the clear heteroskedasticity in the residuals. In the early eighties, for example, there are clear increases in the variance of yields forecast errors. Our assumption of regime-switching in the variance of shocks helps the model to capture these patterns in the data.

Figures 4 and 5 display filtered and smoothed estimates of the discrete states together with the official NBER recession dates. In the figure, 1 denotes the low-variance state, 0 the high-variance state.

The regime associated with the policy shock hovers between the high and low variance regimes over the seventies; stays in the high variance regime in 1980 and remains there until 1983; it then drifts back to the low-variance regime until the beginning of the Great recession.

The variance of productivity growth shocks moves between high and low-variance regimes for most of the sample. On average, it tends to stay closer to the high-variance regime over the eighties and nineties.

The technology and monetary policy states can be composed to define 4 possible combinations of regimes. This is done to construct Figures 6 and 7, which display expected excess holding period returns as defined in equation (20).

Excess returns are increasing in the maturity of bonds and fluctuate around 120 basis points at the 10-year maturity. The notable features emerging from Figure 6 and 7 is that regime-switching can induce a non-negligible amount of variability in expected excess holding period returns over time. This is a desirable feature to explain observed deviations of the data from features consistent with the expectations hypothesis (see e.g. Dai and Singleton, 2002). Variability is especially high at the time of the monetarist experiment
in the early 1980s. This is also encouraging, because deviations of yields from values consistent with the expectations hypothesis are known to be particularly marked during the Volcker tenure. For example, Rudebusch and Wu (2006) note that the performance of the expectations hypothesis improves after 1988 and until 2002.

5 Conclusions

We have estimated the second order approximation of a macro-yield curve model with Epstein-Zin-Weil preferences, in which the variance of structural shocks is subject to changes of regime.

Our empirical results support the regime switching specification. Different regimes can fit the heteroskedasticity of economic variables. Estimated regimes also bear an intuitively appealing structural interpretation. Finally, changes in regimes generate nonnegligible changes in risk premia.

The inclusion of yields data in the estimation set does not alter the basic functioning of the macro-model. Most parameters are estimated to be close to the values obtained in other studies, in which solely macro data enter the econometrician’s information set. The main exception is the interest rate smoothing coefficient in the monetary policy rule, which is found to be much higher than in studies which do not look at yields data. A high value of this coefficient helps to generate persistence in the short rate, hence to transmit movements in the policy rate to long-term yields.
Appendix

A The household problem

Using the definitions $U_{1,t} = \partial U \left[ u_t, \left( \frac{E_t V_{t+1}^{1-\gamma}}{\gamma} \right)^{\frac{1}{1-\gamma}} \right] / \partial u_t$ and $U_{2,t} = \partial U \left[ u_t, \left( \frac{E_t V_{t+1}^{1-\gamma}}{\gamma} \right)^{\frac{1}{1-\gamma}} \right] / \partial \left( \frac{E_t V_{t+1}^{1-\gamma}}{\gamma} \right)^{\frac{1}{1-\gamma}}$, we can write the first order conditions for the optimum as

$$
\Lambda_t P_t = U_{1,t} v (N_t) - h E_t U_{1,t+1}^1 v (N_{t+1}) U_{2,t} \left( \frac{\left[ E_t J_{t+1}^1 \right]^{\frac{1}{1-\gamma}}}{J_{t+1}^1} \right)^{\gamma}
$$

$$
U_{1,t} (C_t - h C_{t-1}) v' (N_t) = \Lambda_t \frac{1 - \theta_w}{\theta_w} w_t
$$

$$
Q_{t,t+1} = \frac{\Lambda_{t+1}}{\Lambda_t} U_{2,t} \left( \frac{\left[ E_t J_{t+1}^1 \right]^{\frac{1}{1-\gamma}}}{J_{t+1}^1} \right)^{\gamma}
$$

(25)

plus the envelope conditions

$$
J_{W,t} = \Lambda_t
$$

$$
J_{C,t} = -h U_{1,t} v (N_t)
$$

where we also defined $J_t \equiv J (W_t, C_{t-1})$, $J_{C,t} \equiv \partial J (W_t, C_{t-1}) / \partial C_{t-1}$.

Note that the two derivatives $U_{1,t}$ and $U_{2,t}$ can be rewritten as

$$
U_{1,t} = (1 - \beta) \left\{ (1 - \beta) (C_t - h C_{t-1})^{1-\psi} [v (N_t)]^{1-\psi} + \beta \left[ E_t J_{t+1}^{1-\gamma} \right]^{\frac{\psi}{1-\gamma}} (C_t - h C_{t-1})^{-\psi} [v (N_t)]^{-\psi} \right\}
$$

and

$$
U_{2,t} \left[ E_t J_{t+1}^{1-\gamma} \right]^{\frac{\psi}{1-\gamma}} = \beta \left\{ (1 - \beta) (C_t - h C_{t-1})^{1-\psi} [v (N_t)]^{1-\psi} + \beta \left[ E_t J_{t+1}^{1-\gamma} \right]^{\frac{\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}
$$

Moreover, at the optimum, the maximum value function will obey the recursion

$$
J_t = \left\{ (1 - \beta) [(C_t - h C_{t-1}) v (N_t)]^{1-\psi} + \beta \left[ E_t J_{t+1}^{1-\gamma} \right]^{\frac{\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}
$$

Using these equations, we can rewrite the FOCs in the text as

$$
\tilde{w}_t = -\frac{v' (N_t)}{v (N_t)} \frac{\theta_w}{\theta_w - 1} \frac{[\left( C_t - h C_{t-1} \right) v (N_t)]^{1-\psi}}{\Lambda_t}
$$

and

$$
\tilde{Q}_{t,t+1} = \frac{\tilde{\Lambda}_{t+1}}{\tilde{\Lambda}_t} \frac{1}{\pi_{t+1}} \left( \frac{\left[ E_t J_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}}{J_{t+1}^{1-\gamma}} \right)^{\gamma-\psi}
$$
where
\[
\tilde{\Lambda}_t \equiv \left[ v(N_t) \right]^{1-\psi} (C_t - hC_{t-1})^{-\psi} - \beta h \mathbb{E}_t \left[ v(N_{t+1}) \right]^{1-\psi} (C_{t+1} - hC_t)^{-\psi} \left( \frac{[E_tJ_{t+1}^{1-\gamma}]^{1-\gamma}}{J_{t+1}} \right)^{\gamma-\psi}
\]
and \( \tilde{w}_t \equiv w_t/P_t \).

Note that in the absence of labour-leisure choice \((v(N_t) = 1 \text{ for all } t)\), we would obtain
\[
Q_{t,t+1} = \frac{\beta}{\pi_{t+1}} \left( \frac{[E_tJ_{t+1}^{1-\gamma}]^{1-\gamma}}{J_{t+1}} \right)^{\gamma-\psi}
\]
\[
\tilde{\Lambda}_t \equiv (C_t - hC_{t-1})^{-\psi} - \beta h \mathbb{E}_t (C_{t+1} - hC_t)^{-\psi} \left( \frac{[E_tJ_{t+1}^{1-\gamma}]^{1-\gamma}}{J_{t+1}} \right)^{\gamma-\psi}
\]
If habits were also set to zero, we would go back to the standard Epstein-Zin-Weil case
\[
Q_{t,t+1} = \frac{\beta}{\pi_{t+1}} \left( \frac{C_{t+1}}{C_t} \right)^{-\psi} \left( \frac{[E_tJ_{t+1}^{1-\gamma}]^{1-\gamma}}{J_{t+1}} \right)^{\gamma-\psi}
\]
\[
\tilde{\Lambda}_t = C_t^{-\psi}
\]

Finally, when \( \gamma = \psi \)
\[
\tilde{w}_t = -\frac{v'(N_t)}{v(N_t)} \frac{\theta_{w,t}}{\theta_{w,t} - 1} \frac{[(C_t - hC_{t-1}) v(N_t)]^{1-\psi}}{\tilde{\Lambda}_t}
\]
\[
Q_{t,t+1} = \frac{\beta}{\pi_{t+1}} \tilde{\Lambda}_{t+1} \tilde{\Lambda}_t
\]
\[
\tilde{\Lambda}_t = [v(N_t)]^{1-\psi} (C_t - hC_{t-1})^{-\psi} - \beta h \mathbb{E}_t [v(N_{t+1})]^{1-\psi} (C_{t+1} - hC_t)^{-\psi}
\]

\section{B The approximate SDF and short rate}

Equation (6) in Piazzesi and Schneider (2006) derives the stochastic discount factor for a model with non-expected utility, exogenous labour supply, exogenous consumption process and \( \psi = 1 \). A similar expression can be derived in our model. In this appendix, we also allow for a stochastic trend in technology growth, so that
\[
A_t = Z_t B_t
\]
\[
B_t = B_{t-1} \Xi_t
\]
\[
\Xi_t = \Xi^{1-\rho_x} \Xi_{t-1}^{\rho_x} e^{\xi_t}
\]
\[
Z_t = Z_{t-1}^{p_x} e^{\xi_t}
\]
\[24\]
In order to approximate the model around a deterministic steady state, we first detrend all variables by the growing level of technology. More specifically, for the stochastic discount factor we obtain

\[ Q_{t,t+1} = \beta \frac{1}{\Xi_{t+1}} \frac{1}{A_t} \frac{1}{\Pi_{t+1}} \left( \frac{E_t[\Xi_{t+1}^{1-\gamma} J_{t+1}^{1-\gamma}]}{J_{t+1}} \right)^{\gamma-\psi} \]

where \( \Xi_t \) is the rate of growth of technology, \( \tilde{\Lambda}_t = \tilde{\Lambda}_t / B_t^{-\psi} \) and \( \tilde{J}_t = J_t / B_t^{-\psi} \) where

\[ \tilde{\Lambda}_t = [v(N_t)]^{1-\psi} \left( \tilde{C}_t - h \frac{\tilde{C}_{t-1}}{\Xi_t} \right)^{-\psi} - \beta h E_t [v(N_{t+1})]^{1-\psi} \Xi_{t+1}^{1-\gamma} \left( \tilde{C}_{t+1} - h \frac{\tilde{C}_t}{\Xi_{t+1}} \right)^{-\psi} \left[ E_t \left( \tilde{J}_{t+1} \Xi_{t+1} \right)^{1-\gamma} \right]^{\gamma-\psi} \]

Note that using the definition \( D_t \equiv E_t[\Xi_{t+1}^{1-\gamma} \tilde{J}_{t+1}^{1-\gamma}] \) the stochastic discount factor can be rewritten as

\[ Q_{t,t+1} = \beta \frac{1}{\Xi_{t+1}} \frac{1}{A_t} \frac{1}{\Pi_{t+1}} \left( D_{t+1}^{1-\gamma} \right)^{\gamma-\psi} \]

or in deviation from the steady state

\[ \hat{q}_{t,t+1} = \Delta \tilde{\Lambda}_{t+1} - \psi E_t \xi_{t+1} - \gamma \left( \tilde{C}_{t+1} - E_t \tilde{C}_{t+1} \right) - \hat{\pi}_{t+1} - (\gamma - \psi) \left( \tilde{J}_{t+1} - E_t \tilde{J}_{t+1} \right) - \frac{1}{2} (\gamma - \psi) (\gamma - 1) \left( \text{Var}_t \left[ \tilde{C}_{t+1} \right] \right) + \text{Var}_t \left[ \tilde{\xi}_{t+1} \right] + 2 \text{Cov}_t \left[ \tilde{C}_{t+1}, \tilde{\xi}_{t+1} \right] \]

When \( \psi = 1 \) this expression boils down to

\[ \hat{q}_{t,t+1} = \Delta \tilde{\Lambda}_{t+1} - E_t \xi_{t+1} - \gamma \left( \tilde{C}_{t+1} - E_t \tilde{C}_{t+1} \right) - \hat{\pi}_{t+1} - (\gamma - 1) \left( \tilde{J}_{t+1} - E_t \tilde{J}_{t+1} \right) - \frac{1}{2} (\gamma - 1)^2 \text{Var}_t \left[ \tilde{C}_{t+1} \right] \]

which corresponds to the case considered by Piazzesi and Schneider (2006) when there is no growth and temporary utility depends on consumption only so that \( \tilde{\Lambda}_t = C_t^{-1} \) and \( \Delta \tilde{\Lambda}_{t+1} = -\Delta \tilde{\xi}_{t+1} \).

We now wish to rewrite \( \tilde{J}_t \) in terms of its consumption and labour determinants. From

\[ \tilde{J}_t^{1-\psi} = (1 - \beta) \left( \tilde{C}_t - h \frac{\tilde{C}_{t-1}}{\Xi_t} \right)^{1-\psi} \left[ v(N_t) \right]^{1-\psi} + \beta E_t[\Xi_{t+1}^{1-\gamma} J_{t+1}^{1-\gamma}]^{\gamma-\psi} \]

or, using \( D_t \equiv E_t[\Xi_{t+1}^{1-\gamma} \tilde{J}_{t+1}^{1-\gamma}] \) and defining the consumption surplus \( c_t^{1-\psi} = \tilde{C}_t - h \tilde{C}_{t-1} / \Xi_t \)

\[ \tilde{J}_t^{1-\psi} = (1 - \beta) c_t^{1-\psi} \left[ v(N_t) \right]^{1-\psi} + \beta D_t^{1-\psi} \]

25
Define further
\[ \tilde{u}_t \equiv \hat{c}_t [v(N_t)] \]
so that
\[ \tilde{J}_t^{1-\psi} = (1 - \beta) \tilde{u}_t^{1-\psi} + \beta D_t^{1-\psi} \]
and
\[ \tilde{\Lambda}_t = v(N_t) \tilde{u}_t^{1-\psi} - \beta h E_t v(N_{t+1}) \tilde{u}_{t+1}^{1-\psi} \Xi_{t+1}^{\gamma} \frac{D_t^{\gamma-\psi}}{J_{t+1}^{\gamma-\psi}} \]

Note that \( \tilde{u}_t \) can be approximated as
\[ \tilde{u}_t \approx \hat{c}_t + \frac{v'N}{v} l_t + \frac{1}{2} \left( 1 + \frac{v''N}{v'} - \frac{v'N}{v} \right) \frac{v'N}{v} \hat{c}_t \]

Using these definitions, the value function can be expanded to second order as
\[ \tilde{J}_t = \beta \Xi^{1-\psi} E_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} - \tilde{u}_{t+1} + \Delta \tilde{u}_{t+1} \right] + \frac{1}{2} \beta \Xi^{1-\psi} \text{Var}_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} \right] \]
\[ + \frac{1}{2} (1 - \psi) \left( 1 - \beta \Xi^{1-\psi} \right) \beta \Xi^{1-\psi} \left( E_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} - \tilde{u}_{t+1} + \Delta \tilde{u}_{t+1} \right] \right)^2 \] (27)

To solve equation (27) forward, note that second order terms can be evaluated using the first order approximation
\[ \tilde{J}_t - \tilde{u}_t = \beta \Xi^{1-\psi} E_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} - \tilde{u}_{t+1} + \Delta \tilde{u}_{t+1} \right] \]

Hence
\[ \text{Var}_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} \right] = \text{Var}_t \left[ \sum_{i=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^i E_{t+1} \left[ \tilde{\xi}_{t+1+i} + \Delta \tilde{u}_{t+1+i} \right] \right] \]

and
\[ \left( E_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} - \tilde{u}_{t+1} + \Delta \tilde{u}_{t+1} \right] \right)^2 = \left( \sum_{i=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^i E_t \left[ \tilde{\xi}_{t+1+i} + \Delta \tilde{u}_{t+1+i} \right] \right)^2 \]

First order terms imply
\[ E_t \left[ \tilde{J}_{t+1} + \tilde{\xi}_{t+1} - \tilde{u}_{t+1} + \Delta \tilde{u}_{t+1} \right] = \frac{1 - \gamma}{2} \sum_{j=1}^{\infty} \left( \beta \Xi^{1-\psi} \right)^j \text{Var}_{t+j} E_{t+1} \left[ \sum_{i=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^i \left( \tilde{\xi}_{t+1+j+i} + \Delta \tilde{u}_{t+1+j+i} \right) \right] \]
\[ + \frac{1 - \psi}{2} \left( 1 - \beta \Xi^{1-\psi} \right) \sum_{j=1}^{\infty} \left( \beta \Xi^{1-\psi} \right)^j \left( \sum_{i=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^i E_{t+j} \left[ \tilde{\xi}_{t+1+j+i} + \Delta \tilde{u}_{t+1+j+i} \right] \right)^2 \]
\[ + \sum_{j=0}^{\infty} \left( \beta \Xi^{1-\psi} \right)^j E_t \left[ \tilde{\xi}_{t+1+j} + \Delta \tilde{u}_{t+1+j} \right] \]
Putting everything together
\[
\hat{\delta}_{t+1} - \hat{u}_t = \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} E_t \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right] \\
+ \frac{1 - \gamma}{2} \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} E_t \text{Var}_{t+j} E_{t+1+j} \left[ \sum_{i=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{i} \left( \hat{\xi}_{t+i+1+j+i} + \Delta \hat{u}_{t+i+1+j+i} \right) \right] \\
+ \frac{1 - \psi}{2} \left( 1 - \beta\Xi^{1-\psi} \right) \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} E_t \left[ \left( \sum_{i=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{i} E_{t+i+j} \left[ \hat{\xi}_{t+i+1+j+i} + \Delta \hat{u}_{t+i+1+j+i} \right] \right)^2 \right]
\]
where
\[
\Delta \hat{u}_{t+1} = \Delta \hat{c}_{t+1} + \frac{v'N}{v} \Delta \hat{h}_{t+1} + \frac{1}{2} \left( 1 + \frac{v''N}{v} - \frac{v'N}{v} \right) \frac{hN}{v} \left( \hat{h}_{t+1} - \hat{h}_{t} \right)
\]
and
\[
\Delta \hat{c}_{t+1} = \frac{1}{\Xi - h} \left( \Xi \hat{c}_{t+1} - h \Delta \hat{c}_{t} + h \Delta \hat{\xi}_{t+1} \right) - \frac{1}{2} \frac{h\Xi}{(\Xi - h)^2} \left[ \left( \Delta \hat{c}_{t+1} + \hat{\xi}_{t+1} \right)^2 - \left( \Delta \hat{c}_{t} + \hat{\xi}_{t} \right)^2 \right]
\]
To use this expressions in the stochastic discount factor of equation (26), note that variances and covariances can be evaluated using the first order approximation
\[
\hat{\delta}_t = \hat{u}_t + \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} E_t \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right]
\]
which yields
\[
\text{Var}_t \left[ \hat{\delta}_{t+1} \right] = \text{Var}_t \left[ \hat{\xi}_{t+1} \right] + \text{Var}_t \left[ \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j} E_{t+1} \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right] \right] \\
- 2 \text{Cov}_t \left[ \hat{\xi}_{t+1} \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j} E_{t+1} \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right] \right]
\]
and
\[
\text{Cov}_t \left[ \hat{\xi}_{t+1} \hat{\delta}_{t+1} \right] = - \text{Var}_t \left[ \hat{\xi}_{t+1} \right] + \text{Var}_t \left[ \hat{\xi}_{t+1} \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j} E_{t+1} \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right] \right]
\]
Finally
\[
\hat{\delta}_{t+1} - E_t \hat{\delta}_{t+1} = \left( E_{t+1} - E_t \right) \hat{\xi}_{t+1} + \sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j} \left( E_{t+1} - E_t \right) \left[ \hat{\xi}_{t+1+j} + \Delta \hat{u}_{t+1+j} \right] + \frac{1 - \gamma}{2} \times \\
\sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} \left( E_{t+1} - E_t \right) \text{Var}_{t+j} E_{t+2+j} \left[ \sum_{i=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{i} \left( \hat{\xi}_{t+i+2+j+i} + \Delta \hat{u}_{t+i+2+j+i} \right) \right] \\
+ \frac{1 - \psi}{2} \left( 1 - \beta\Xi^{1-\psi} \right) \times \\
\sum_{j=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{j+1} \left( E_{t+1} - E_t \right) \left[ \left( \sum_{i=0}^{\infty} \left( \beta\Xi^{1-\psi} \right)^{i} E_{t+i+j} \left[ \hat{\xi}_{t+i+2+j+i} + \Delta \hat{u}_{t+i+2+j+i} \right] \right)^2 \right]
\]
The second order approximation of the stochastic discount factor is therefore

\[
\hat{q}_{t,t+1} = \Delta \tilde{\lambda}_{t+1} - \psi \tilde{\xi}_{t+1} - \tilde{\pi}_{t+1} - (\gamma - \psi) \sum_{j=0}^{\infty} (\beta \Xi^{1 - \psi})^j (E_{t+1} - E_t) \left( \tilde{\xi}_{t+1+j} + \Delta \tilde{\nu}_{t+1+j} \right) + \\
- \frac{1}{2} (\gamma - \psi) (\gamma - 1) \operatorname{Var}_t \left[ \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i E_{t+1} \left( \tilde{\xi}_{t+1+i} + \Delta \tilde{\nu}_{t+1+i} \right) \right] + \frac{1}{2} (\gamma - \psi) (\gamma - 1) \times \\
\times (E_{t+1} - E_t) \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \operatorname{Var}_{t+j} E_{t+1+j} \left[ \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \left( \tilde{\xi}_{t+1+j+i} + \Delta \tilde{\nu}_{t+1+j+i} \right) \right] + \\
+ (\gamma - \psi) (\psi - 1) \frac{1 - \beta \Xi^{1 - \psi}}{2} \times \\
\times (E_{t+1} - E_t) \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \left[ \left( \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i E_{t+j} \left[ \tilde{\xi}_{t+1+j+i} + \Delta \tilde{\nu}_{t+1+j+i} \right] \right)^2 \right]
\]

Note that we could define

\[
\tilde{\Lambda}_t \equiv v (N_t) - \beta h E_t v (N_{t+1}) \frac{\tilde{\nu}_{t+1}^{t\psi}}{\tilde{\nu}_t^{t\psi}} \tilde{\nu}_{t+1}^{\gamma} \frac{D_{t+1}^{t\gamma}}{D_t^{t\gamma}}
\]

to write

\[
\tilde{\tilde{\Lambda}}_t \equiv \tilde{\nu}_t^{-\psi} \tilde{\Lambda}_t
\]

such that

\[
\tilde{\tilde{\lambda}}_t = \frac{v' N \tilde{\nu}_t}{v} + \frac{1}{2} \left[ 1 + \frac{v'' N}{v'} - \frac{v' N}{v} \right] \frac{v' N}{v} \tilde{\gamma}_t^2
\]

We could define a "growth adjusted utility" \( \Delta \tilde{\nu}_t \equiv \Delta \tilde{\nu}_t + \tilde{\xi}_t \) to finally write

\[
\hat{q}_{t,t+1} = \Delta \tilde{\lambda}_{t+1} - \psi \Delta \tilde{\nu}_{t+1} - \tilde{\pi}_{t+1} - (\gamma - \psi) \sum_{j=0}^{\infty} (\beta \Xi^{1 - \psi})^j (E_{t+1} - E_t) \Delta \tilde{\nu}_{t+1+j} \\
- \frac{1}{2} (\gamma - \psi) (\gamma - 1) \operatorname{Var}_t \left[ \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i E_{t+1} \Delta \tilde{\nu}_{t+1+i} \right] \\
+ \frac{1}{2} (\gamma - \psi) (\gamma - 1) (E_{t+1} - E_t) \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \operatorname{Var}_{t+j} E_{t+1+j} \left[ \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \Delta \tilde{\nu}_{t+1+j+i} \right] \\
+ (\gamma - \psi) (\psi - 1) \frac{1 - \beta \Xi^{1 - \psi}}{2} (E_{t+1} - E_t) \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i \left[ \left( \sum_{i=0}^{\infty} (\beta \Xi^{1 - \psi})^i E_{t+j} \Delta \tilde{\nu}_{t+1+j+i} \right)^2 \right]
\]

To a second order approximation, the short rate is given by

\[
\hat{i}_t = -E_t \hat{q}_{t,t+1} - \frac{1}{2} \operatorname{Var}_t \hat{q}_{t,t+1}
\]
From the SDF we obtain
\[
\hat{q}_{t+1} = -E_t \Delta \tilde{\lambda}_{t+1} + E_t \psi \Delta \bar{w}_{t+1} + E_t \tilde{\pi}_{t+1} - \frac{1}{2} \text{Var}_t \left( \Delta \tilde{\lambda}_{t+1} - \psi \Delta \bar{w}_{t+1} - \tilde{\pi}_{t+1} \right) \\
+ \frac{1}{2} (\gamma - \psi) (\psi - 1) \text{Var}_t \left[ E_{t+1} \sum_{i=0}^{\infty} (\beta^{1-\psi})^i \Delta \bar{w}_{t+1+i} \right] \\
+ (\gamma - \psi) \text{Cov}_t \left[ \left( \Delta \tilde{\lambda}_{t+1} - \psi \Delta \bar{w}_{t+1} - \tilde{\pi}_{t+1} \right), E_{t+1} \sum_{i=0}^{\infty} (\beta^{1-\psi})^i \Delta \bar{w}_{t+1+i} \right]
\]

C Solution for bond prices

Recall that
\[
x_{t+1} = P \tilde{x}_t + \frac{1}{2} \left( I_{n_x} \otimes \tilde{x}_t' \right) G \tilde{x}_t + k_{x,s} \tilde{\sigma}^2 + \tilde{\sigma} \Sigma_i u_{t+1}
\]
\[
s_{t+1} = \kappa_0 + \kappa_1 s_t + \nu_{t+1}
\]
\[
\tilde{\lambda}_t = F_x \tilde{x}_t + \frac{1}{2} \left( I_{n_x} \otimes \tilde{x}_t' \right) E_\lambda \tilde{x}_t + k_{\lambda,s} \tilde{\sigma}^2
\]
\[
\tilde{\pi}_t = F_x \tilde{x}_t + \frac{1}{2} \left( I_{n_x} \otimes \tilde{x}_t' \right) E_\pi \tilde{x}_t + k_{\pi,s} \tilde{\sigma}^2
\]

where $P$ is a $n_x \times n_x$ matrix, $G$ is a $n_x^2 \times n_x$ matrix, $k_{x,s}$ is an $n_x \times 1$ vector (whose elements are state dependent), $\Sigma_i \equiv \Sigma (s_t)$ is a $n_x \times n_u$ matrix, $F_\lambda$ and $F_\pi$ are $1 \times n_x$ vectors, $E_\lambda$ and $E_\pi$ are $n_x \times n_x$ matrices, and finally $k_{\lambda,s}$ and $k_{\pi,s}$ are (state dependent) scalars.

C.1 1-period bonds

To derive the price of 1-period bonds, note first that a second order approximation to the stochastic discount factor is
\[
\tilde{q}_{t,t+1} = (F_\lambda - F_\pi) \tilde{x}_{t+1} + \frac{1}{2} \tilde{x}_{t+1}' (E_\lambda - E_\pi) \tilde{x}_{t+1} - F_\lambda \tilde{x}_t - \frac{1}{2} \tilde{x}_t' E_\lambda \tilde{x}_t - k_{\pi,s} \tilde{\sigma}^2
\]
or
\[
\tilde{q}_{t,t+1} = ((F_\lambda - F_\pi) P - F_\lambda) \tilde{x}_t + \frac{1}{2} (F_\lambda - F_\pi) \left( I_{n_x} \otimes \tilde{x}_t' \right) G \tilde{x}_t \\
+ \frac{1}{2} \tilde{x}_t' P' (E_\lambda - E_\pi) \tilde{x}_t - \frac{1}{2} \tilde{x}_t' E_\lambda \tilde{x}_t \\
+ \tilde{\sigma}^2 (F_\lambda - F_\pi) k_{x,s} - k_{\pi,s} \tilde{\sigma}^2 \\
+ \tilde{\sigma} (F_\lambda - F_\pi) \Sigma_i u_{t+1} + \frac{1}{2} \tilde{x}_t' P' (E_\lambda - E_\pi) \Sigma_i u_{t+1} \\
+ \frac{1}{2} \tilde{\sigma} u_{t+1}' \Sigma_i (E_\lambda - E_\pi) P \tilde{x}_t + \frac{1}{2} \tilde{\sigma}^2 u_{t+1}' \Sigma_i (E_\lambda - E_\pi) \Sigma_i u_{t+1}
\]
To second order, the price of a 1-period bond is

\[ \hat{b}_{t,1} = -i_t = E_t[q_{t+1}] + \frac{1}{2} \left( E_t[q_{t+1}^2] - (E_t[q_{t+1}])^2 \right) \]

for which we need

\[ E_t[q_{t+1}^2] - (E_t[q_{t+1}])^2 = \hat{\sigma}^2 (F_\lambda - F_\pi) \Sigma_t \Sigma_t' (F_\lambda - F_\pi)' \]

and

\[ E_t[q_{t+1}] = (F_\lambda - F_\pi) P \hat{x}_t + \frac{1}{2} (F_\lambda - F_\pi) (I_{n_x} \otimes \hat{x}_t') G \hat{x}_t + \hat{\sigma}^2 (F_\lambda - F_\pi) k_{x,s} \]

\[ + \frac{1}{2} \hat{x}_t' P' (E_\lambda - E_\pi) P \hat{x}_t + \frac{1}{2} \hat{\sigma}^2 E_t [u_{t+1}' \Sigma_t (E_\lambda - E_\pi) \Sigma_t u_{t+1}] \]

\[ - F_\lambda \hat{x}_t - \frac{1}{2} \hat{x}_t' E_\lambda \hat{x}_t - k_{x,s} \hat{\sigma}^2 \]

Now note that, for any matrix \( A \) and vector \( x \),

\[ E[x' Ax] = E[\text{vec}(x' Ax)] \]

\[ = E[x' \otimes x'] \text{vec}(A) \]

\[ = (\text{vec}(E(xx')))\prime \text{vec}(A) \]

where the vec operator transforms a matrix into a vector by stacking its columns. It follows that

\[ E_t[u_{t+1}' \Sigma_t (E_\lambda - E_\pi) \Sigma_t u_{t+1}] = (\text{vec}(I)) \text{vec}\left( \Sigma_t (E_\lambda - E_\pi) \Sigma_t \right) \]

\[ = \text{tr}\left( \Sigma_t (E_\lambda - E_\pi) \Sigma_t \right) \]

where \text{tr} represents the trace, i.e. the sum of the diagonal elements of a matrix.

Hence,

\[ \hat{b}_{t,1} = ((F_\lambda - F_\pi) P - F_\lambda) \hat{x}_t + \hat{\sigma}^2 ((F_\lambda - F_\pi) k_{x,s} - k_{\pi,s}) \]

\[ + \frac{1}{2} \hat{\sigma}^2 \text{tr}(\Sigma_t (E_\lambda - E_\pi)) + \frac{1}{2} \hat{\sigma}^2 (F_\lambda - F_\pi) \Sigma_t \Sigma_t' (F_\lambda - F_\pi)' \]

\[ + \frac{1}{2} \hat{x}_t' (P' (E_\lambda - E_\pi) P - E_\lambda) \hat{x}_t + \frac{1}{2} (F_\lambda - F_\pi) (I_{n_x} \otimes \hat{x}_t') G \hat{x}_t \]
Finally, note that
\[
(F_{\lambda} - F_{\pi}) (I_{n_x} \otimes \tilde{x}_t') G \hat{x}_t = (F_{\lambda} - F_{\pi}) \begin{pmatrix} \tilde{x}_t' & 0 & \cdots & 0 \\ 0 & \tilde{x}_t' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{x}_t' \end{pmatrix} G \hat{x}_t
\]
\[
= \left[ (F_{\lambda,1} - F_{\pi,1}) \tilde{x}_t' (F_{\lambda,2} - F_{\pi,2}) \tilde{x}_t' \cdots (F_{\lambda,n_x} - F_{\pi,n_x}) \tilde{x}_t' \right] G \hat{x}_t
\]
\[
= \left[ \tilde{x}_t' (F_{\lambda,1} - F_{\pi,1}) G_1 + \tilde{x}_t' (F_{\lambda,2} - F_{\pi,2}) G_2 + \cdots + \tilde{x}_t' (F_{\lambda,n_x} - F_{\pi,n_x}) G_{n_x} \right] \hat{x}_t
\]
\[
= \tilde{x}_t' \left[ \sum_{j=1}^{n_x} (F_{\lambda,j} - F_{\pi,j}) G_j \right] \hat{x}_t
\]
where $F_{\lambda,i}$ and $F_{\pi,i}$ denote the $i$-th elements of vectors $F_{\lambda}$ and $F_{\pi}$, respectively, and $G_i$ denotes the $i$-th $n_x \times n_x$ matrix which are vertically stacked to make up $G$. We can therefore rewrite the 1-period bond as
\[
\hat{b}_{t,1} = F_{B,1} \hat{x}_t + \frac{1}{2} \tilde{x}_t' E_{B,1} \hat{x}_t + k_{B,1,s} \hat{\sigma}^2
\]
where
\[
F_{B,1} = (F_{\lambda} - F_{\pi}) P - F_{\lambda}
\]
\[
k_{B,1,s} = (F_{\lambda} - F_{\pi}) k_{x,s} - k_{\pi,s} + \text{tr} (\Sigma_t' (E_{\lambda} - E_{\pi}) \Sigma_t) + (F_{\lambda} - F_{\pi}) \Sigma_t \Sigma_t' (F_{\lambda} - F_{\pi})
\]
\[
E_{B,1} = P' (E_{\lambda} - E_{\pi}) P - E_\lambda + \sum_{j=1}^{n_x} (F_{\lambda,j} - F_{\pi,j}) G_j
\]
Note also that, by construction, $\hat{b}_{t,1} = -\hat{t}_t$, so $F_{B,1} = -F_t$, $E_{B,1} = -E_t$ and $k_{B,1,s} = -k_{t,s}$. Note that this definition also allows us to rewrite $\hat{q}_{t,t+1}$ as
\[
\hat{q}_{t,t+1} = F_{B,1} \hat{x}_t + \frac{1}{2} \tilde{x}_t' E_{B,1} \hat{x}_t + \hat{\sigma}^2 ((F_{\lambda} - F_{\pi}) k_{x,s} - k_{\pi,s})
\]
\[
+ \hat{\sigma} (F_{\lambda} - F_{\pi}) \Sigma_t u_{t+1} + \hat{\sigma} \tilde{x}_t' P' (E_{\lambda} - E_{\pi}) \Sigma_t u_{t+1} + \frac{1}{2} \hat{\sigma}^2 u_{t+1}' \Sigma_t' (E_{\lambda} - E_{\pi}) \Sigma_t u_{t+1}
\]

### C.2 2-period bonds

2-period bond prices can be written as (up to a second order approximation)
\[
\hat{b}_{t,2} = \hat{b}_{t,1} + E_t \left[ \hat{b}_{t+1,1} \right] + \frac{1}{2} \text{Var}_t \left[ \hat{b}_{t+1,1} \right] + \text{Cov}_t \left[ \hat{q}_{t+1,1}, \hat{b}_{t+1,1} \right]
\]
Based on 1-period prices, we can derive
\[
E_t \left[ \hat{b}_{t+1,1} \right] = F_{B,1} P \hat{x}_t + \frac{1}{2} F_{B,1} (I_{n_x} \otimes \tilde{x}_t') G \hat{x}_t + \frac{1}{2} \tilde{x}_t' P' E_{B,1} P \hat{x}_t
\]
\[
+ F_{B,1} k_{x,s} \hat{\sigma}^2 + \hat{\sigma}^2 k_{B,1,s} + \frac{1}{2} \hat{\sigma}^2 \text{tr} (\Sigma_t' E_{B,1} \Sigma_t)
\]
31
and

\[ \begin{align*}
E_t \left[ \hat{b}_{t+1,1} \hat{q}_{t+1} \right] & - E_t \left[ \hat{b}_{t+1,1} \right] E_t \left[ \hat{b}_{t+1,1} \right] = \hat{\sigma}^2 F_{B_1} \Sigma_1 \Sigma_1' F_{B_1}' \\
E_t \left[ \hat{b}_{t+1,1} \hat{q}_{t+1}' \right] & - E_t \left[ \hat{b}_{t+1,1} \right] E_t \left[ \hat{q}_{t+1}' \right] = F_{B_1} \Sigma_1 \Sigma_1' \left( F_\lambda - F_\pi \right)'
\end{align*} \]

It follows that

\[ \hat{b}_{t,2} = F_{B_2} \hat{x}_t + \frac{1}{2} \hat{x}_t' E_{B_2} \hat{x}_t + k_{B_2,s} \hat{\sigma}^2 \]

where

\[ \begin{align*}
F_{B_2} &= F_{B_1} (I + P) \\
E_{B_2} &= E_{B_1} + P' E_{B_1} P + \sum_{j=1}^{n_x} F_{B_1,j} G_j \\
k_{B_2,s} &= k_{B_1,s} + F_{B_1} k_{x,s} + \text{tr} \left( \Sigma_1' E_{B_1} \Sigma_1 \right) + F_{B_1} \Sigma_1 \Sigma_1' F_{B_1} + 2 F_{B_1} \Sigma_1 \Sigma_1' \left( F_\lambda - F_\pi \right)'
\end{align*} \]

C.3 n-period bonds

Using the same procedure, we find that n-period bond prices can be written as

\[ \hat{b}_{t,n} = F_{B_n} \hat{x}_t + \frac{1}{2} \hat{x}_t' E_{B_n} \hat{x}_t + k_{B_n,s} \hat{\sigma}^2 \]

where for \( n > 1 \)

\[ \begin{align*}
F_{B_n} &= F_{B_1} + F_{B_n-1} P \\
E_{B_n} &= E_{B_1} + P' E_{B_{n-1}} P + \sum_{j=1}^{n_x} F_{B_{n-1,j}} G_j \\
k_{B_n,s} &= k_{B_1,s} + k_{B_{n-1},s} + F_{B_{n-1}} k_{x,s} + \text{tr} \left( \Sigma_1' E_{B_{n-1}} \Sigma_1 \right) \\
&\quad + F_{B_{n-1}} \Sigma_1 \Sigma_1' F_{B_{n-1}} + 2 F_{B_{n-1}} \Sigma_1 \Sigma_1' \left( F_\lambda - F_\pi \right)'
\end{align*} \]
References


34


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Post Mean</th>
<th>Post SD</th>
<th>Post Low Q</th>
<th>Post Up Q</th>
<th>Prior Mean</th>
<th>Prior SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{\eta,11}$</td>
<td>0.9598</td>
<td>0.0159</td>
<td>0.9196</td>
<td>0.9829</td>
<td>0.9000</td>
<td>0.0650</td>
</tr>
<tr>
<td>$p_{\eta,00}$</td>
<td>0.9259</td>
<td>0.0237</td>
<td>0.8694</td>
<td>0.9612</td>
<td>0.9000</td>
<td>0.0650</td>
</tr>
<tr>
<td>$p_{z,11}$</td>
<td>0.8746</td>
<td>0.0117</td>
<td>0.8492</td>
<td>0.8918</td>
<td>0.9000</td>
<td>0.0650</td>
</tr>
<tr>
<td>$p_{z,00}$</td>
<td>0.9081</td>
<td>0.0144</td>
<td>0.8777</td>
<td>0.9324</td>
<td>0.9000</td>
<td>0.0650</td>
</tr>
<tr>
<td>$\sigma_{\eta,1}$</td>
<td>0.0017</td>
<td>0.0001</td>
<td>0.0015</td>
<td>0.0019</td>
<td>0.0028</td>
<td>0.0006</td>
</tr>
<tr>
<td>$\sigma_{\eta,0}$</td>
<td>0.0043</td>
<td>0.0005</td>
<td>0.0035</td>
<td>0.0056</td>
<td>0.0038</td>
<td>0.0010</td>
</tr>
<tr>
<td>$\sigma_{z,1}$</td>
<td>0.0173</td>
<td>0.0006</td>
<td>0.0162</td>
<td>0.0185</td>
<td>0.0038</td>
<td>0.0012</td>
</tr>
<tr>
<td>$\sigma_{z,0}$</td>
<td>0.0271</td>
<td>0.0022</td>
<td>0.0229</td>
<td>0.0313</td>
<td>0.0062</td>
<td>0.0030</td>
</tr>
<tr>
<td>$\sigma_{G}$</td>
<td>0.0347</td>
<td>0.0025</td>
<td>0.0303</td>
<td>0.0404</td>
<td>0.0305</td>
<td>0.0148</td>
</tr>
<tr>
<td>$\sigma_{\mu}$</td>
<td>0.0791</td>
<td>0.0067</td>
<td>0.0668</td>
<td>0.0930</td>
<td>0.0119</td>
<td>0.0051</td>
</tr>
<tr>
<td>$\Pi$</td>
<td>1.0120</td>
<td>0.0009</td>
<td>1.0106</td>
<td>1.0138</td>
<td>1.0160</td>
<td>0.0033</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>1.0043</td>
<td>0.0001</td>
<td>1.0042</td>
<td>1.0045</td>
<td>1.0030</td>
<td>0.0017</td>
</tr>
<tr>
<td>$\rho_{G}$</td>
<td>0.8558</td>
<td>0.0110</td>
<td>0.8346</td>
<td>0.8773</td>
<td>0.9000</td>
<td>0.0299</td>
</tr>
<tr>
<td>$\rho_{z}$</td>
<td>0.9812</td>
<td>0.0019</td>
<td>0.9774</td>
<td>0.9847</td>
<td>0.9000</td>
<td>0.0299</td>
</tr>
<tr>
<td>$\rho_{\mu}$</td>
<td>0.8613</td>
<td>0.0205</td>
<td>0.8157</td>
<td>0.8970</td>
<td>0.7000</td>
<td>0.0454</td>
</tr>
<tr>
<td>$\psi_{\pi}$</td>
<td>0.3952</td>
<td>0.0254</td>
<td>0.3495</td>
<td>0.4444</td>
<td>0.5000</td>
<td>0.0998</td>
</tr>
<tr>
<td>$\psi_{y}$</td>
<td>0.0012</td>
<td>0.0010</td>
<td>-0.0008</td>
<td>0.0029</td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
<tr>
<td>$\rho_{i}$</td>
<td>0.7862</td>
<td>0.0202</td>
<td>0.7480</td>
<td>0.8268</td>
<td>0.8000</td>
<td>0.1996</td>
</tr>
<tr>
<td>$\iota$</td>
<td>0.8497</td>
<td>0.0488</td>
<td>0.7433</td>
<td>0.9231</td>
<td>0.9000</td>
<td>0.0901</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1.3561</td>
<td>0.0883</td>
<td>1.2072</td>
<td>1.5579</td>
<td>2.0000</td>
<td>0.4014</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>3.2083</td>
<td>0.4523</td>
<td>2.3585</td>
<td>4.1202</td>
<td>11.0000</td>
<td>7.1268</td>
</tr>
<tr>
<td>$\psi$</td>
<td>2.0526</td>
<td>0.0811</td>
<td>1.9010</td>
<td>2.2120</td>
<td>1.5000</td>
<td>0.3544</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>7.7264</td>
<td>1.0208</td>
<td>5.6782</td>
<td>9.4305</td>
<td>10.0000</td>
<td>4.9945</td>
</tr>
<tr>
<td>$h$</td>
<td>0.5476</td>
<td>0.0071</td>
<td>0.5326</td>
<td>0.5610</td>
<td>0.4000</td>
<td>0.1480</td>
</tr>
<tr>
<td>$\theta$</td>
<td>5.7413</td>
<td>0.5671</td>
<td>4.6872</td>
<td>6.7827</td>
<td>7.0000</td>
<td>1.9998</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9991</td>
<td>0.0005</td>
<td>0.9980</td>
<td>0.9998</td>
<td>0.9975</td>
<td>0.0025</td>
</tr>
<tr>
<td>$\sigma_{me,20}$</td>
<td>0.0006</td>
<td>0.0001</td>
<td>0.0005</td>
<td>0.0008</td>
<td>0.0014</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Priors: beta distribution for $\beta$, $h$, $\iota$, $\zeta$, $\rho_{G}$, $\rho_{\mu}$; gamma distribution for $\psi_{\pi}$, $\psi_{y}$ and all standard deviations; shifted gamma distribution (domain from 1 to $\infty$) for $\gamma$, $\phi$, $\Xi$, $\Pi$; normal distribution for $\rho_{i}$. Posterior distributions are based on 50,000 draws.
Figure 1: Moments of observable variables: sample vs model-implied

Note: The dots denote combinations of unconditional means and standard deviations implied by the posterior distribution of our parameter estimates. The diamonds indicate the corresponding sample moments.
Figure 2: Cross-correlations of observable variables: sample vs model-implied

Note: The red solid lines denote sample cross-correlations. The blue lines indicate the distribution of cross-correlations implied by the posterior distribution of our parameter estimates.
Figure 3: Actual variables and 1-step-ahead predictions

Note: all variables are measured in annual terms.
Figure 4: Filtered and smoothed estimates of the regime-variables

Note: "0" is the high-variance state.
Figure 5: Filtered estimates of the regime-variables with confidence sets

Note: "0" is the high-variance state.
Figure 6: Expected excess holding period returns
Figure 7: Expected excess holding period returns with confidence sets

filtered premia

- recession
- maturity 12 quarters low
- maturity 12 quarters med
- maturity 12 quarters up
- maturity 40 quarters low
- maturity 40 quarters med
- maturity 40 quarters up