Monetary Policy and the Term Structure of Interest Rates*

Federico Ravenna
Department of Economics
University of California
Santa Cruz, CA 95064, USA
email: fravenna@ucsc.edu

Juha Seppälä
Department of Economics
University of Illinois
Champaign, IL 61820, USA
email: seppala@uiuc.edu

First draft: September 10, 2004
This version: December 30, 2005

Comments Welcome!

Abstract

We study how well a New Keynesian business cycle model can explain the observed behavior of nominal interest rates. We focus on two puzzles raised in previous literature. First, Donaldson, Johnsen, and Mehra (1990) show that while in the U.S. nominal term structure the interest rates are pro-cyclical and term spreads counter-cyclical the stochastic growth model predicts that the interest rates are counter-cyclical and term spreads pro-cyclical. Second, according to Backus, Gregory, and Zin (1989) the standard general equilibrium asset pricing model can account for neither the sign nor the magnitude of average risk premiums in forward prices. Hence, the standard model is unable to explain rejections of the expectations hypothesis. We show that a New Keynesian model with habit-formation preferences and a monetary policy feedback rule produces pro-cyclical interest rates, counter-cyclical term spreads, and creates enough volatility in the risk premium to account for the rejections of expectations hypothesis. Moreover, unlike Buraschi and Jiltsov (2005), we identify the systematic monetary policy, not monetary policy shocks, as the key factor behind rejections of expectations hypothesis.

Keywords: Term Structure of Interest Rates, Monetary Policy, Sticky Prices, Habit Formation, Expectations Hypothesis.

JEL classification: E43, E44, E5, G12.

*We would like to thank Robert Kollman for sharing his code and notes, Ondra Kamenik for help with Dynare++, Oreste Tristani, Glenn Rudebusch, George Waters, Andrew Ang, and Martin Ellison for discussions and comments, and Scott Stephens for excellent research assistance. We also received useful comments from seminar participants at the University of Illinois, Purdue University, Illinois State University, Midwest Macroeconomics Meetings (Iowa City, 2005), Society for Economic Dynamics Meetings (Budapest, 2005), Swiss National Bank, Bank of Finland, the World Congress of the Econometric Society (London, 2005), and University of Wisconsin. Parts of this paper were written while the first author was visiting the European Central Bank Research Department and the Bank of Spain Research Department and the second author was visiting the Bank of Finland Research Department. We would like to thank all organizations for their support. All errors are ours.

1. Introduction

The term structure of interest rates contains information about agents expectations of future interest rates, inflation rates, and exchange rates[1]. Along with the high yield and commercial paper spread, the interest rate term spread between long and short maturity Treasury Bills has been repeatedly shown to have predictive power for various indicators of the US business cycle in the postwar period[2].

Because prices of securities at different maturities embody financial market participants' expectations of future economic activity, the term structure of interest rates is an invaluable source of information for monetary authorities[3]. But the large literature on dynamic models of the term structure relies on latent factor models, and as such does not offer insight into the relationships between term structure movements and business cycle indicators. More recently, researchers have started investigating the relationship between term structure and macro variables in reduced-form VAR or semi-structural models that impose no-arbitrage restrictions[4]. Structural general equilibrium models derive the term structure from agents' optimizing behaviour and explain all of the interest rate dynamics with the macroeconomic variables volatility. Unfortunately, standard models that have become the workhorse of modern macroeconomics have counterfactual implications for the term structure.

First, Donaldson, Johnsen, and Mehra (1990) show that while in the U.S. nominal term structure the interest rates are pro-cyclical and term spreads counter-cyclical, the Neoclassical stochastic growth model predicts that interest rates are counter-cyclical and term spreads pro-cyclical. Variations on the Real Business Cycle model and models containing nominal frictions have all been shown to be lacking in some dimension when used to model the term structure[5].

Second, Backus, Gregory, and Zin (1989) show that the Lucas (1978) general equilibrium asset pricing model can account for neither the sign nor the magnitude of average risk premiums in forward prices and holding-period returns. Thus, the model is unable to explain the rejection of the expectations hypothesis—whether forward rates are unbiased predictors of future spot rates—that has been extensively documented by empirical studies. The most common interpretation is that this is evidence of the existence of a time-varying risk premium[6].

In order for the policy makers to extract information about market expectations from the term structure they need to know the sign and magnitude of the term premia embedded in interest rates.

[1]See Fama (1975, 1990) and Mishkin (1981, 1990a, 1992) for studies on inflation expectations and the term structure of interest rates using U.S. data. Mishkin (1991) and Jorion and Mishkin (1991) use international data. Abken (1993) and Blough (1994) provide surveys of the literature.

[2]See, Harvey (1988), Chen (1991), and Estrella and Hardouvelis (1991).

[3]Svensson (1994a,b) and Söderlind and Svensson (1997) discuss monetary policy and the role of the term structure of interest rates as a source of information. Evans and Marshall (1998), Piazzesi (2005), Cochrane and Piazzesi (2002, 2005), and Buraschi and Jiltsov (2005) are recent contributions to this literature.

[4]Piazzesi (2005), Cochrane and Piazzesi (2002, 2005), Ang and Piazzesi (2003), Rudebusch and Wu (2004), Hördhal, Tristani, and Vestin (2003).

[5]See King and Watson (1996), Den Haan (1995), Evans and Marshall (1998).

[6]This literature is extensive. Useful surveys are provided by Melino (1988), Shiller (1990), Mishkin (1990b), and Campbell, Lo, and MacKinlay (1997). Often cited individual studies are Shiller (1979), Shiller, Campbell, and Schoenholtz (1983), Fama (1984, 1990), Fama and Bliss (1987), Froot (1989), Campbell and Shiller (1991), and Campbell (1995).

Referring to research by Backus, Gregory, and Zin and other authors, Söderlind and Svensson (1997) note in their review:

> "We have no direct measurement of this (potentially) time-varying covariance [term premium], and even ex post data is of limited use since the stochastic discount factor is not observable. It has unfortunately proved to be very hard to explain (U.S. ex post) term premia by either utility based asset pricing models or various proxies for risk."

In this paper, we build a New Keynesian general equilibrium model to explain the term structure of interest rates. The model displays short-run monetary non-neutrality, therefore the behaviour of the monetary authority affects the business cycle dynamics. Because monetary policy responds systematically to movements in the endogenous variables, changes in the way policy is conducted affect the co-variation of real and nominal variables, and play an important role in the dynamics of the term structure. We show that the model can match the average nominal term structure in post-war US data, and produces pro-cyclical interest rates and counter-cyclical term spreads. The term spread has predictive power for future economic activity. Most importantly, the model generates enough volatility in the risk premium to account for rejections of the expectations hypothesis.

Our results show that the rejection of the expectation hypothesis hinges on habit-formation preferences and on the modeling of the systematic portion of monetary policy. Without habit formation, average term premia are very close to zero and the volatility of yields at all maturities is dramatically reduced. As the policy rule changes, the average term structure, risk premia volatility and the correlation with macro variables will change. This is also true in the absence of policy shocks. In fact we show that *all* of the rejection of the expectation hypothesis can be explained by the real shocks' volatility. Finally, a large or very volatile inflation risk premium cannot explain the expectation hypothesis rejection result. Eliminating the monetary policy shocks in the model reduces inflation risk premium but makes it easier to reject the expectations hypothesis.

The rest of the paper is organized as follows. The rest of this Section goes through the related literature in more detail. Section 2 explains the New Keynesian model we use, Section 3 discusses the techiniques we use to solve the model numerically, and Section 4 explains the parameterization of the model. Section 5 reports the results related the term structure—particularly in relation to the term spread and term premium puzzles. Section 6 discusses the relationship between monetary policy and the term structure. Section 7 investigates the role of the business cycle shocks in explaining the term structure behaviour. Finally, Section 8 concludes. Appendix A derives the inflation rate dynamics in our model. Appendix B presents results from seven different experiments to illustrate features of the model.

Related Literature

A growing literature investigates the relationship between macroeconomic variables and the term structure within reduced-form or semi-structural models imposing no-arbitrage restrictions, and find that macro variables can improve the predictive power of latent factor models. Piazzesi (2005) shows that the Federal Reserve policy can be better approximated by assuming that it responds to the information contained *only* in the term structure rather than in other macroeconomic variables. Cochrane and Piazzesi (2005) show that the monetary policy shocks can explain 45% of excess nominal bond returns, and Cochrane and Piazzesi (2002) show that the term structure explains

64% of the changes in the federal funds target rate. Ang and Piazzesi (2003) introduce no-arbitrage restrictions in a VAR model of macroeconomic and financial variables.

The research on joint macro-finance model (Hördahl, Tristani, and Vestin, 2002, Rudebusch and Wu, 2004) aims at integrating small scale optimizing models of output, inflation and interest rates with affine no-arbitrage specifications for bond prices. In this way, it is possible to identify the affine model latent factors with the macroeconomic aggregates.

An important goal of this recent literature is to relate yield dynamics to macro factors to be able to analyze what portion of the yields volatility can be explained by observable factors. In both VAR and joint macro-finance models the market price of risk is modeled only in reduced-form fashion, rather than being derived from optimizing behaviour.

Among general equilibrium models of the term structure, Evans and Marshall (1998) show that a limited participation model is broadly consistent with the impulse response functions of the real and nominal yields to a monetary policy shock. However, Piazzesi (2005) criticizes their methodology on the grounds that it doesn't impose the no arbitrage condition on the yield movements. Dai (2002) shows that a model with limited participation can explain the term premium puzzle—it can generate countercyclical term spreads. Seppälä (2004) studies the asset pricing implications of an endowment economy when agents can default on contracts. The results show that this limited commitment model is one potential solution of the term premium puzzle. Both of these contributions models study only the real term structure. Buraschi and Jiltsov (2003) and Wachter (2004) show that an external habit model in the style of Campbell and Cochrane (1999) is capable of explaining the term premium puzzle. Duffee (2002) and Dai and Singleton (2002) study the term premium puzzle in the nominal yields using reduced form no arbitrage models. None of these models rely on the nominal rigidities business cycle models. Seppälä and Xie (2004) are closer in spirit to our approach. The authors study the cyclical behavior of nominal and (ex-ante) real term structures of interest rates in the UK data, and in a real business cycle, a limited participation, and a New Keynesian model. Their result is that only the New Keynesian model gets closest to matching the cyclical behavior for both the nominal and the real term structure.

Two recent papers by Bekaert, Cho, and Moreno (2005) and Hördahl, Tristani and Vestin (2005) examine the term structure implications of a business cycle model with nominal price staggering and endogenous monetary policy. Both papers rely on the New Keynesian framework. Bekaert, Cho, and Moreno (2005) estimate a log-linear three equation New Keynesian model using a Maximum Likelihood estimator, and derive an endogenous log-normal term structure consistent with the household preferences. While the model ensures that the observable macro variables are consistent with firms' and household's optimizing behaviour, it also introduces two unobservable state variables, so that the dynamics is driven by a total of five exogenous shocks.

Hördahl, Tristani and Vestin (2005) use a second order approximation to derive the law of motion for both macro variables and bond prices, and show that a New Keynesian model can account for both the positive slope of the yield curve and the constant volatility of yields across maturities. The paper shows that to achieve these results the model must allow for a very high level of persistence in the exogenous shocks (the authors report parametrizations of the AR(1) coefficients in the order of 0.99). Our model is closely related to Hördahl, Tristani and Vestin (2005), but our solution method relies on a full third order approximation. In fact, none of the papers cited derive time-varying risk premia, and therefore cannot address the expectation hypothesis puzzle in a fully optimizing context.

Buraschi and Jiltsov (2005) study the inflation risk premium in a continuous-time general equilibrium model in which the monetary authority sets the money supply based on targets on the long-term growth of the nominal money supply, inflation, and economic growth. They identify the time-variation of the inflation risk premium as an important explanatory variable of deviations from the expectations hypothesis. In contrast, in our model the monetary policy authority follows an interest rate rule - a more accurate description of the actual conduct of monetary policy in most countries. Since the source of monetary non-neutrality is different, it is not surprising that our conclusions differ. Contrary to their results, we find that model monetary policy shocks and inflation risk premium are not the explanation behind the rejections of the expectations hypothesis.

2. *The Model*

The theoretical interest rate term structure is derived from a dynamic stochastic general equilibrium model of the business cycle. An important objective of the paper is to evaluate the role of monetary policy in generating an empirically plausible term structure. Hence, we adopt a money-in-utility-function model where nominal rigidities allow monetary policy to affect the dynamics of real variables. We follow Calvo (1983) and the New Keynesian literature on the business cycle by assuming that prices cannot be updated to the profit-maximizing level in each period. Firms face an exogenous, constant probability of being able to reset the price in any period t. This setup can also be derived from a menu cost model, where firms face a randomly distributed fixed cost k_t of updating the price charged, and the support of k_t is $[0; \overline{k}]$, $\overline{k} \to \infty$ (see Klenow and Kryvtsov, 2004).

While more sophisticated pricing mechanism can be introduced—such as state-dependent pricing (Dotsey, King and Wolman, 1999), partial indexation to past prices (Christiano, Eichenbaum and Evans, 2001), a mix of rule-of-thumb and forward-looking pricing (Gali and Gertler, 1999)—we limit the model to the more essential ingredients of the New Keynesian framework. This allows us to investigate the impact on the term structure of four key features: (i) systematic monetary policy modeled as an interest rate rule; (ii) nominal price rigidity; (iii) habit-formation preferences; (iv) positive steady state money growth rate. Woodford (2003) and Walsh (2003) offer a comprehensive treatment of the New Keynesian framework, and describe in detail the microfoundations of the model.

Each consumer owns shares of all firms, and households are rebated any profit from the monopolistically competitive output sector. Savings can be accumulated in money balances, or in a range of riskless nominal and real bonds spanning several maturities. The government runs a balanced budget in every period, and rebates to consumers any seigniorage revenue from issuing the monetary asset. Output is produced with undifferentiated labor, supplied by the household-consumers, via a linear production function.

Households

There is a continuum of infinitely lived households, indexed by $j \in [0, 1]$. Consumers demand differentiated consumption goods, choosing from a continuum of goods, indexed by $z \in [0, 1]$. In the notation used throughout the paper, $C_t^j(z)$ indicates consumption by household j at time t of the good produced by firm z.

Households' preferences over the basket of differentiated goods are defined by the CES aggregator:

$$C_t^j = \left[\int_0^1 C_t^j(z)^{\frac{\theta-1}{\theta}} dz \right]^{\frac{\theta}{\theta-1}}, \ \theta > 1 \tag{1}$$

The representative household chooses $\left\{ C_{t+i}^j, C_{t+i}^j(z), N_{t+i}^j, \frac{M_{t+i}^j}{P_{t+i}}, \frac{B_{t+i}^j}{P_{t+i}} \right\}_{i=0}^{\infty}$ where N_t denotes labor supply, M_t nominal money balances, P_t the aggregate price level, and B_t bond holdings, to maximize:

$$E_t \sum_{i=0}^{\infty} \beta^i \left\{ \frac{(C_{t+i}^j - bC_{t+i-1}^j)}{1-\gamma}^{1-\gamma} D_{t+i} - \frac{\ell N_{t+i}^{1+\eta}}{1+\eta} + \frac{\xi}{1-\gamma_m} \left(\frac{M_{t+i}^j}{P_{t+i}} \right)^{1-\gamma_m} \right\} \tag{2}$$

subject to

$$\int_0^1 C_t^j(z) P_t(z) dz = W_t N_t^j + \Pi_t^j - (M_t^j - M_{t-1}^j) - (\overrightarrow{p}_t \overrightarrow{B}_t^j - B_{t-1}^j) - \tau_t^j, \tag{3}$$

and (1). When $b > 0$ the preferences are characterized by habit formation (Boldrin, Christiano, and Fisher, 2001). D_t is an aggregate stochastic preference shock. Each element of the row vector $\overrightarrow{p}_t$ represents the price of an asset with maturity k that will pay one unit of currency in period $t+k$. The corresponding element of $\overrightarrow{B}_t$ represents the quantity of such claims purchased by the household. B_{t-1}^j indicates the value of the household portfolio of claims maturing at time t. W_t is the nominal wage rate and τ is the lump-sum tax imposed by the government. Finally, household owns the firms and Π_t is the profit from the firms.

The solution to the intratemporal expenditure allocation problem between the varieties of differentiated goods gives the individual good z demand function:

$$C_t^j(z) = \left[\frac{P_t(z)}{P_t} \right]^{-\theta} C_t^j. \tag{4}$$

Equation (4) is the demand of good z from household j, where θ is the price elasticity of demand. The associated price index P_t measures the least expenditure for differentiated goods that buys a unit of the consumption index:

$$P_t = \left[\int_1^0 P_t(z)^{1-\theta} dz \right]^{\frac{1}{1-\theta}}. \tag{5}$$

Since all household solve an identical optimization problem and face the same aggregate variables, in the following we omit the index j. Using equations (4) and (5), we can write the budget constraint as:

$$C_t = \frac{W_t}{P_t} N_t + \frac{\Pi_t}{P_t} - \frac{M_t - M_{t-1}}{P_t} - \frac{\overrightarrow{p}_t \overrightarrow{B}_t - B_{t-1}}{P_t} - \frac{\tau_t}{P_t},$$

The first order conditions with respect to labor and real money balances are:

$$MUC_t = E_t\left[\frac{D_t}{(C_t - bC_{t-1})^{\gamma}} - \beta b \frac{D_{t+1}}{(C_{t+1} - bC_t)^{\gamma}}\right]$$

$$0 = MUC_t \frac{W_t}{P_t} - \ell N_t^{\eta} \tag{6}$$

$$0 = \xi\left(\frac{M_t}{P_t}\right)^{-\gamma_m} - MUC_t + E_t\left[\beta MUC_{t+1}\frac{P_t}{P_{t+1}}\right] \tag{7}$$

where MUC is the marginal utility of consumption.

Firms and price setting

The firm producing good z employs a linear technology:

$$Y_t(z) = A_t N_t(z)$$

where A_t is an aggregate productivity shock. Minimizing the nominal cost of producing a given amount of output $\overline{Y}$:

$$Cost = W_t N_t(z)$$

yields the labor demand schedule:

$$MC_t^N(z) MPL_t(z) = W_t \tag{8}$$

where MC^N is the nominal marginal cost, MPL is the marginal product of labor $(Y_t(z)/N_t(z))$. Equation (8) implies that the real marginal cost MC_t of producing one unit of output is:

$$MC_t(z) MPL_t(z) = W_t/P_t$$

Firms adjust their prices infrequently. In each period there is a constant probability $(1-\theta_p)$ that the firm will be able to adjust its price, independently of past history. This implies that the fraction of firms setting prices at t is $(1-\theta_p)$ and the expected waiting time for the next price adjustment is $\frac{1}{1-\theta_p}$. The problem of the firm setting the price at time t consists of choosing $P_t(z)$ to maximize the expected discounted stream of profits:

$$E_t \sum_{i=0}^{\infty} (\theta_p \beta)^i \frac{MUC_{t+i}}{MUC_t}\left[\frac{P_t(z)}{P_{t+i}} Y_{t,t+i}(z) - \frac{MC_{t+i}^N}{P_{t+i}} Y_{t,t+i}(z)\right] \tag{9}$$

subject to

$$Y_{t,t+i}(z) = \left[\frac{P_t(z)}{P_{t+i}}\right]^{-\theta} Y_{t+i}, \tag{10}$$

In (10), $Y_{t,t+i}(z)$ is the firm's demand function for its output at time $t+i$, conditional on the price set at time t, $P_t(z)$. Market clearing insures that $Y_{t,t+i}(z) = C_{t,t+i}(z)$ and $Y_{t+i} = C_{t+i}$. Substituting (10) into (9), the objective function can be written as:

$$E_t \sum_{i=0}^{\infty} (\theta_p \beta)^i \frac{MUC_{t+i}}{MUC_t}\left\{\left[\frac{P_t(z)}{P_{t+i}}\right]^{1-\theta} Y_{t+i} - \frac{MC_{t+i}^N}{P_{t+i}}\left[\frac{P_t(z)}{P_{t+i}}\right]^{-\theta} Y_{t+i}\right\}. \tag{11}$$

Since $P_t(z)$ does not depend on i, the optimality condition is:

$$P_t(z) E_t \sum_{i=0}^{\infty} (\theta_p \beta)^i MUC_{t+i} \left[\frac{P_t(z)}{P_{t+i}} \right]^{1-\theta} Y_{t+i} = \mu E_t \sum_{i=0}^{\infty} (\theta_p \beta)^i MUC_{t+i} MC_{t+i}^N \left[\frac{P_t(z)}{P_{t+i}} \right]^{1-\theta} Y_{t+i}. \quad (12)$$

where

$$\mu = \frac{\theta}{\theta - 1}$$

is the flexible-price level of the markup, and also the markup that would be observed in a zero-inflation (zero money growth rate) steady state. To use rational expectations solution algorithms when the steady state money growth rate is non-zero, we need to express the first order condition as a difference equation (see Ascari, 2004, and King and Wolman, 1996). This can be accomplished expressing $P_t(z)$ as the ratio of two variables:

$$P_t(z) = \frac{G_t}{H_t},$$

and

$$G_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{G}_t \quad (13)$$

$$H_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{H}_t, \quad (14)$$

where

$$\hat{G}_t = \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \hat{G}_{t+1} \quad (15)$$

$$\hat{H}_t = MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \hat{H}_{t+1}. \quad (16)$$

Market Clearing

Since the measure of the economy is unitary, in the symmetric equilibrium it holds that:

$$M_t^j = M_t \ ; \ C_t^j = C_t$$

and the consumption shadow price is symmetric across households: $MUC_t^j = MUC_t$. Given that all firms are able to purchase the same labor service bundle, and so are charged the same aggregate wage, they all face the same marginal cost. The linear production technology insures that MC is equal across all firms—whether they are updating or not their price—regardless of the level of production, which will indeed be different. Firms are heterogeneous in that a fraction $(1 - \theta_p)$ of firms in the interval $[0, 1]$ can optimally choose the price charged at time t. In equilibrium each producer that chooses a new price $P_t(z)$ in period t will choose the same new price $P_t(z)$ and the same level of output. Then the dynamics of the consumption-based price index will obey

$$P_t = \left[\theta_p P_{t-1}^{1-\theta} + (1 - \theta_p) P_t(z)^{1-\theta} \right]^{\frac{1}{1-\theta}}. \quad (17)$$

The Appendix A shows that the inflation rate dynamics is given by:

$$[(1+\pi_t)]^{1-\theta} = \theta_p + (1-\theta_p)\left[\frac{\tilde{G}_t}{\tilde{H}_t}(1+\pi_t)\right]^{1-\theta} \tag{18}$$

$$\tilde{G}_t \equiv \frac{\hat{G}_t}{P_t^{\theta}} \quad ; \quad \tilde{H}_t \equiv \frac{\hat{H}_t}{P_t^{\theta-1}}$$

In a steady state with gross money growth rate equal to Υ, and gross inflation equal to $\Pi = \Upsilon$,

$$\frac{G}{HP_t} = \frac{\tilde{G}}{\tilde{H}} = \frac{P_t(z)}{P_t}$$

$$\frac{P_t(z)}{P_t} = \mu * MC * \frac{(1-\theta_p\beta\Pi^{\theta-1})}{(1-\theta_p\beta\Pi^{\theta})}$$

Since $P_t(z)$ is the optimal price chosen by the fraction of firms that can re-optimize at time t, it is the inverse of what King and Wolman (1996) define as the *price wedge*. With zero steady state inflation the steady state average markup is equal to $1/MC$, therefore there is no price wedge. But when steady state inflation is positive, the price wedge is less than one: the average price is always smaller than the optimal price, since some firms would like to increase the price, but are constrained not to do so. Combining this equation with equation (18) gives the steady state marginal cost and price wedge as a function of Π:

$$\frac{\tilde{G}}{\tilde{H}} = \left[\frac{(1-\theta_p)}{(1-\theta_p\Pi^{\theta-1})}\right]^{\frac{1}{\theta-1}}$$

$$MC = \frac{1}{\mu}\left[\frac{\Pi^{1-\theta}-\theta_p}{1-\theta_p}\right]^{\frac{1}{1-\theta}} \frac{1}{\Pi}\frac{(1-\theta_p\beta\Pi^{\theta})}{(1-\theta_p\beta\Pi^{\theta-1})}$$

Asset markets

The government rebates the seigniorage revenues to the household in the form of lump-sum transfers, so that in any time t the government budget is balanced. Since we defined in equation (3) τ^j as the amount of the tax levied by the government on household j, assuming $\tau_t^j = \tau_t^i \;\; \forall\, j,i \;\in [0,1]$, at every date t the transfer will be equal to:

$$-\int_0^1 \tau_t^j\, dj = -\tau_t \int_0^1 dj =$$

$$= -\tau_t = M_t^s - M_{t-1}^s$$

Equilibrium in the money market requires:

$$M_t^s = M_t^{dj} = M_t^d$$

We assume the monetary policy instrument is the short term nominal interest rate $(1+R_{1,t})$. The money supply is set by the monetary authority to satisfy whatever money demand is consistent with the target rate.

Domestic bonds are in zero-net supply, since the government does not issue bonds. Therefore in equilibrium it must hold that:

$$B_{t,i} = 0$$

for any component of the vector $\overrightarrow{B}_t$. However, because we have complete markets, we can still price both nominal and real bonds.

Monetary Policy

The economy's dynamics is driven by business cycle shocks temporarily away from the non-stochastic steady state. In this instances, the domestic monetary authority follows a forward-looking, instrument feedback rule:

$$\frac{\left(1+\overline{R}_{t,t+1}\right)}{\left(1+R^{ss}\right)} = E_t \left(\frac{1+\pi_{t+1}}{1+\pi_{SS}}\right)^{\omega_\pi} \left(\frac{Y_t}{Y_{SS}}\right)^{\omega_y} \tag{19}$$

where ω_π, $\omega_y \geq 0$ are the feedback coefficients to CPI inflation and output. The monetary authority adjusts the interest rate in response to deviations of the target variables from the steady state. In the steady state, a constant money growth rate rule is followed. The choice of the parameters ω_π, ω_y allows us to specify alternative monetary policies. When the central bank responds to current rather than expected inflation equation (19) returns the rule suggested by Taylor (1993) as a description of U.S. monetary policy.

We assume the central bank assigns positive weight to an interest rate smoothing objective, so that the domestic short-term interest rate at time t is set according to

$$(1+R_{1,t}) = \left[\left(1+\overline{R}_{t,t+1}\right)\right]^{(1-\chi)} \left[(1+R_{t-1,1})\right]^{\chi} \varepsilon_t^{mp} \tag{20}$$

where $\chi \in [0,1)$ is the degree of smoothing and ε_t^{mp} is an unanticipated exogenous shock to monetary policy.

3. Algorithm

We solve the model using a third-order approximation around the non-stochastic steady state. The numerical solution is done using the Dynare++.[7] It is well known that taking a first-order approximation to the bond prices will give no risk premia and that a second-order approximation will give only constant premia. The reason is simple: the second-order approximation involves only squared error terms that have constant expectation.

For this reason, in the first step, we solve our model for five state variables and six control variables in 11 equations using Dynare++. In the second step, we generate 200,000 observations of

[7] Dynare++ is available for free at `http://www.cepremap.cnrs.fr/dynare/index.php?option=com_content&task=view&id=53&Itemid=86`.

state and control variables. In the final step, we regress the future marginal rates of substitution—see equations (23) and (24) below—on the third-order complete polynomials of the state variables to generate bond prices.[8]

Our approach is very similar to the parameterized expectations algorithm employed by Evans and Marshall (1998). While they didn't study the premia, notice that the algorithm amounts to taking a third-order approximation to bond prices. With third-order approximation, the current state variables multiply squared future error terms, and hence risk premia are time-varying.

4. Parametrization

Preference, technology and policy parameters are parameterized consistently with the New Keynesian monetary business cycle literature. Estimated and calibrated staggered-price adjustment models are discussed in Bernanke and Gertler (2000), Christiano Eichenbaum, and Evans(2001), Ireland (2001), Ravenna (2002), Rabanal and Rubio-Ramirez (2003), Walsh (2003), Woodford (2003).

Households' preferences are modeled following the internal habit-formation framework of Boldrin, Christiano, and Fisher (2001). Fuhrer (2000) finds that the habit formation parameter $b = 0.6$ optimizes the match between sticky-price models and consumption data. We set $b = 0.8$. The value of γ is set to 2.5, and is chosen to provide adequate curvature in the utility function so as to generate enough risk-premia volatility. The parametrization of habit-formation preferences plays a very important role in the model's term-structure properties. Its impact on the results is discussed in detail in Section 5. Labor supply elasticity $(1/\eta)$ is equal to 2, and the parameter ℓ is chosen to set steady state labor hours at about 20% of the available time. This is a value consistent with many OECD countries postwar data, although on the low side for the US. The quarterly discount factor β is parametrized so that the steady state real interest rate is equal to 1%. The demand elasticity θ is set to obtain a flexible-price equilibrium producers' markup $\mu = \vartheta/(\vartheta - 1) = 1.1$. While Bernanke and Gertler (2000) use a higher value of 1.2, in our model positive steady state inflation implies the steady state markup is larger than in the flexible-price equilibrium.

The production technology is linear in labor hours. Given that the model is parameterized at business-cycle frequencies, this a fair approximation widely used in the literature. To parametrize the Calvo (1983) pricing adjustment mechanism, the probability θ_p faced by firms of not adjusting the price in any given period is set to 0.75 , implying that the average time between price adjustments for a producer is 1 year. This value is in line with estimates for the US reported by Gali and Gertler (1999) and Rabanal and Rubio-Ramirez (2003).

Variants of the instrument rule (20) have been estimated both in single-equation and in simultaneous-equation contexts. We set the inflation feedback coefficients ω_π to 1.5, which tries to capture different monetary policy regimes in the post-war U.S. data. The choice of a value for ω_y is more controversial, depending on the operational definition of output gap used by the central bank at any given point in time. We were unable to solve the model for even very small positive values of ω_y. Rabanal (2004) estimates the smoothing parameter χ to be 0.8. We use 0.9. Appendix B

[8]In a previous version of the paper, we used second-order approximation in the first step. The results are (using Dynare++) very similar to the current results.

Table 1: Selected variables volatilities and correlations. Sample: 1947–2004.

Variable	*Standard Deviation* Model	*Standard Deviation* U.S. data	*Correlation with output* *Model*	*Correlation with output* *U.S. Data*
Y_t	2.05	1.69	1	1
π_t	3.22	3.48	0.21	0.27
R_t	1.94	2.96	0.11	0.17
r_t	3.91	3.2	-0.10	-0.16

discusses different monetary policy specifications. Quarterly steady state inflation is set equal to the average U.S. value over the period $1994 - 2004$, about 0.75% on a quarter basis. This implies an annualized steady state nominal interest rate of 7%.

The preference and technology exogenous shocks follow an AR(1) process:

$$\log Z_t = (1 - \rho_Z) \log \overline{Z} + \rho_Z \log Z_{t-1} + \varepsilon_t^Z \qquad \varepsilon_t^Z \sim i.i.d.\ N(0, \sigma_Z^2)$$

where $\overline{Z}$ is the steady state value of the variable. The policy shock ε_t^{mp} is a Gaussian i.i.d. stochastic process. The autocorrelation parameters are equal to $\rho_a = 0.9$, $\rho_d = 0.95$. The standard deviation of the innovations ε is set to $\sigma_a = 0.35$, $\sigma_d = 8$, $\sigma_{mp} = 0.3$ (percent values). The low value of the policy shock implies the largest part of the short term nominal interest rate dynamics is driven by the systematic monetary policy reaction to the state of the economy. The preference shock volatility is large, but very close to the one estimated by Rabanal and Rubio-Ramirez (2003) on U.S. data. Compared to their estimates, the technology shock volatility is low. But as the volatility of this shock increases, the correlation between nominal interest rate and GDP becomes smaller and smaller, since a technology shock generates a negative correlation. Note that the authors cited above adopt a model which includes a cost-push shock. This shock generates a strong positive correlation between R_t and y_t, and they estimate its volatility to be equal to 41.

An important concern in the parametrization of the shocks has been to match the correlations between output and nominal and real rates with U.S. data, to be able to evaluate whether the term structure generated by the model can predict output variation, as many empirical studies have found in the US. Table 1 compares the model's second moments to the whole U.S. post-war data sample.[9] This sample is heterogenous with respect to the U.S. monetary policy goals and the US Federal Reserve operating procedures, and includes the 1970s inflationary episode. On the other hand, the sample can be considered as representative of the variety of shocks that drove the U.S. business cycle. Hence, the calibration does a relatively good job at matching the correlations, but standard deviations of nominal interest rates and inflation are lower than in data.

[9]Standard deviation measured in percent. The output series is logged and Hodrick-Prescott filtered. U.S. data: Y_t is real GDP, π_t is CPI inflation, R_t is 3-months T-bill rate, r_t is ex-post short term real interest rate. All rates are on annual basis. Quarterly data sample is 1:1947–1:2004. Data are taken from the St. Louis Federal Reserve Bank FRED II database.

5. *The Term Structure of Interest Rates*

The Real and Nominal Term Structures

Let q_{t+1} denote the real stochastic discount factor

$$q_{t+1} \equiv \beta \frac{MUC_{t+1}}{MUC_t}, \tag{21}$$

and let Q_{t+1} denote the nominal stochastic discount factor

$$Q_{t+1} \equiv \beta \frac{MUC_{t+1}}{MUC_t} \frac{P_t}{P_{t+1}}, \tag{22}$$

The price of an n-period zero-coupon real bond is given by

$$\begin{aligned} p^b_{n,t} &= E_t\left[\prod_{j=1}^{n} q_{t+j}\right] \\ &= E_t[q_{t+1} p^b_{n-1,t+1}], \end{aligned} \tag{23}$$

and similarly the price of an n-period zero-coupon nominal bond is given by

$$\begin{aligned} p^B_{n,t} &= E_t\left[\prod_{j=1}^{n} Q_{t+j}\right] \\ &= E_t[Q_{t+1} p^B_{n-1,t+1}]. \end{aligned} \tag{24}$$

The bond prices are invariant with respect to time, and hence equations (23) and (24) give a recursive formula for pricing zero-coupon real and nominal bonds of any maturity. For simplicity, we next express rates for only real rates. Nominal rates are obtained in a similar manner.

Forward prices are defined by

$$p^f_{n,t} = \frac{p^b_{n+1,t}}{p^b_{n,t}},$$

and the above prices are related to interest rates (or yields) by

$$f_{n,t} = -\log(p^f_{n,t}) \quad \text{and} \quad r_{n,t} = -(1/n)\log(p^b_{n,t}). \tag{25}$$

To define the risk premium as in Sargent (1987), write (23) for a two-period bond using the conditional expectation operator and its properties:

$$\begin{aligned} p^b_{t,2} &= E_t[q_{t+1} p^b_{1,t+1}] \\ &= E_t[q_{t+1}]E_t[p^b_{1,t+1}] + \text{cov}_t[q_{t+1}, p^b_{1,t+1}] \\ &= p^b_{1,t} E_t[p^b_{1,t+1}] + \text{cov}_t[q_{t+1}, p^b_{1,t+1}], \end{aligned}$$

which implies that

$$p^f_{1,t} = \frac{p^b_{t,2}}{p^b_{1,t}} = E_t[p^b_{1,t+1}] + \text{cov}_t\left[q_{t+1}, \frac{p^b_{1,t+1}}{p^b_{1,t}}\right]. \tag{26}$$

Since the conditional covariance term is zero for risk-neutral investors, we call it the *risk premium* for the one-period real forward contract, $rrp_{1,t}$, given by

$$rrp_{1,t} \equiv -\operatorname{cov}_t\left[q_{t+1}, \frac{p^b_{1,t+1}}{p^b_{1,t}}\right] = E_t[p^b_{1,t+1}] - p^f_{1,t},$$

and similarly $rrp_{n,t}$ is the risk premium for the n-period forward contract:

$$rrp_{n,t} \equiv -\operatorname{cov}_t\left[\prod_{j=1}^{n} q_{t+j}, \frac{p^b_{t+n,1}}{p^b_{1,t}}\right] = E_t[p^b_{t+n,1}] - p^f_{n,t}.$$

Table 2 presents the means, standard deviations, and correlations with for the nominal and real term structure in the model, and for the U.S. nominal data as estimated by McCulloch and Kwon (1993) from the first quarter of 1947 until the fourth quarter of 1990 and by Duffee (2001) from the first quarter of 1991 until the fourth quarter of 1998. Output is filtered using the Hodrick-Prescott (1980) filter with a smoothing parameter of 1600 both in the model data in data.

Table shows that in model the nominal term structure is procyclical. In data, short maturities are procyclical and long maturities countercyclical. In contrast, the nominal term spreads are countercyclical both in data and in model. The nominal term structures are upward-sloping both in model and in data. Means are matched quite well; in model the nominal yields from three months until 20 years vary from 4.90% to 6.56% and in data from 5.06% to 6.55%. Similarly, the average term spreads are produced by the model are quite close to the average term spreads in data. The term structure of volatilities in model is strongly downward-sloping while in data it is essentially flat.[10]

In addition, the model produces strong positive correlation between yields and (the cyclical component of) output while in data the correlation is low and positive for short maturities and essentially zero for long maturities. The strong positive correlation in the model is a product of the large shocks autocorrelation. Persistent shocks are needed to obtain sufficient volatility of rates at the long end. The downside is that correlations with output will be very high as shocks die out slowly. Possible remedies are introducing hybrid inflation and/or time-varying inflation target. Both would give a larger volatility of interest rate at long maturities, with smaller shocks variance, probably lowering the correlation.

The data range was chosen to be as long as possible. However, the empirical term structure is relatively robust with respect to time periods choses. Table 3 presents selected term structure statistics for different time periods, 1952–1998, 1960–1998, 1980–1998, and 1988–1998.

Table shows that the upwarding-sloping mean term structure and countercyclical term spread are robust features in data. On the other hand, the level of interest rates depends on how much relatively high intereste rates in the early 1980's weight in the data. The average one-year rate was 175 basis higher in 1980–1998 compared to 1988–1998. Interestingly, the two problems we had in Table 2, the flat term structure of volatilities and low correlation between yields and output get

[10]If one restricts the attention to 1980:1 to 1998:4 data sample, the the term structure of volatility is clearly downward-sloping. The standard deviation of the three-month yield is 3.05 and the standard deviation of the 20-year yield is 1.77. In addition, in the UK nominal and real data the term structure of volatilities is downward-sloping, see Seppälä (2000).

Table 2: Main term structure statistics. Data: 1947–1998. (N/A missing due to shortage of data.)

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$ (model)	4.90075	1.93624	0.11203
$R_{4,t}$ (model)	6.16778	1.38586	0.21348
$R_{40,t}$ (model)	6.55441	0.53740	0.40278
$R_{80,t}$ (model)	6.56089	0.32975	0.40511
$R_{120,t}$ (model)	6.56263	0.22810	0.40161
$R_{1,t}$ (data)	5.06450	3.05130	0.14735
$R_{4,t}$ (data)	5.47900	3.12905	0.12193
$R_{40,t}$ (data)	6.22462	2.96448	−0.00705
$R_{80,t}$ (data)	6.54877	3.15684	−0.05252
$R_{120,t}$ (data)	N/A	N/A	N/A
$R_{40,t} - R_{1,t}$ (model)	1.65366	1.66228	−0.00028
$R_{80,t} - R_{1,t}$ (model)	1.66015	1.75799	−0.04741
$R_{120,t} - R_{1,t}$ (model)	1.66188	1.80931	−0.06926
$R_{40,t} - R_{4,t}$ (model)	0.38662	1.01146	−0.07851
$R_{80,t} - R_{4,t}$ (model)	0.39311	1.14522	−0.14170
$R_{120,t} - R_{4,t}$ (model)	0.39485	1.21515	−0.16809
$R_{40,t} - R_{1,t}$ (data)	1.16012	1.15935	−0.40586
$R_{80,t} - R_{1,t}$ (data)	1.04854	1.32304	−0.41947
$R_{120,t} - R_{1,t}$ (data)	N/A	N/A	N/A
$R_{40,t} - R_{4,t}$ (data)	0.74561	0.96807	−0.41570
$R_{80,t} - R_{4,t}$ (data)	0.62854	1.13526	−0.39922
$R_{120,t} - R_{1,t}$ (data)	N/A	N/A	N/A

Table 3: Selected term structure statistics in different time periods.

	1952–1998	1960–1998	1980–1998	1988–1998
$E[R_{4,t}]$	5.92332	6.60742	7.58052	5.83252
$\text{std}(R_{4,t})$	2.96026	2.75032	3.06250	1.62297
$\text{corr}(R_{4,t}, Y_t)$	0.10207	0.16303	0.03409	0.68461
$E[R_{40,t}]$	6.66032	7.38120	8.37087	7.12116
$\text{std}(R_{40,t})$	2.78183	2.48534	2.50033	1.17644
$\text{corr}(R_{40,t}, Y_t)$	−0.05537	−0.01846	−0.08073	0.38770
$E[R_{40,t} - R_{4,t}]$	0.73699	0.77378	1.15062	1.28591
$\text{std}(R_{40,t} - R_{4,t})$	1.01685	1.08750	1.19102	1.09638
$\text{corr}(R_{40,t} - R_{4,t}, Y_t)$	−0.44862	−0.45449	−0.25711	−0.59741

smaller the more recent data we use. Both in 1980–1998 and 1988–1998 data, the flat term structure of volatilities are clearly downward-sloping and in 1988–1998 data, the correlation between yields and output is strongly positive. In Appendix B we report how our model term structure statistics vary depending on how we choose our parameter values.

The Expectations Hypothesis

The oldest and simplest theory about the information content of the term structure is so called (pure) expectations hypothesis. According to the pure expectations theory forward rates are unbiased predictors of future spot rates. It is also common to modify the theory so that constant risk-premium is allowed—this is usually called the expectations hypothesis. However, it should be noted that both versions of the expectations hypothesis are always incorrect. To see this, let us assume, for a sake of an argument, that the agents are risk-neutral: $\gamma = 0$. Equation (26) reduces then into

$$p^f_{1,t} = E_t[p^b_{1,t+1}]$$

and from (25) we obtain

$$\exp^{-f_{1,t}} = E_t\left[\exp^{-r_{1,t+1}}\right].$$

From the Jensen's inequality it follows that

$$f_{1,t} < E_t[r_{1,t+1}] \tag{27}$$

and the difference between the left and right hand side of (27) varies with $E_t[r_{1,t+1}]$ and $\text{var}_t[r_{1,t+1}]$. This effect is known as *convexity premium* or *bias*.

Backus, Gregory, and Zin (1989), on the other hand, tested the expectations hypothesis in the complete markets endowment economy (Lucas model) by starting with (26), assuming that the risk premium was constant

$$E_t[p^b_{1,t+1}] - p^f_{1,t} = \beta_0,$$

and then regressed

$$p^b_{1,t+1} - p^f_{1,t} = \beta_0 + \beta_1(p^f_{1,t} - p^b_{1,t}) \tag{28}$$

to see if $\beta_1 = 0$. They generated 200 observations 1000 times and used Wald test with White (1980) standard errors to check if $\beta_1 = 0$ with 5% significance level. They could reject the hypothesis only roughly 50 times out of 1000 regressions which is what one would expect from chance alone. On the other hand, for all values of β_1 except -1, the forward premium is still useful in forecasting the changes in spot prices. The hypothesis $\beta_1 = -1$ was rejected every time.

Table 4 presents the number of rejections of different Wald tests in the regressions

$$y_{t+1} = \beta_0 + \beta_1 x_t$$

in our benchmark model for nominal term structure. Table 5 presents the same tests when the habit-formation parameter b is close to zero ($b = 0.025$).[11] Table 6 displays the same tests for real term structure, and table 7 displays the test for real term structure when $b = 0.025$. Only our

[11] We were unable to solve the solve when $b = 0$. The reason is that, with interest rate smoothing, there has to be a relationship between C_{t-1} and R_{t-1}. Habit formation provides that.

Table 4: The number of rejects in each regressions in the benchmark model for nominal term structure.

y_{t+1}	$p^B_{1,t+1} - p^F_{1,t}$	$p^B_{1,t+1} - p^F_{1,t} - nrp_{1,t}$
x_t	$p^F_{1,t} - p^B_{1,t}$	$p^F_{1,t} - p^B_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	1000	55
$Wald(\beta_1 = 0)$	975	55
$Wald(\beta_1 = -1)$	1000	1000

Table 5: The number of rejects in each regressions in the benchmark model for nominal term structure when $b = 0.0250$.

y_{t+1}	$p^B_{1,t+1} - p^F_{1,t}$	$p^B_{1,t+1} - p^F_{1,t} - nrp_{1,t}$
x_t	$p^F_{1,t} - p^B_{1,t}$	$p^F_{1,t} - p^B_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	140	78
$Wald(\beta_1 = 0)$	117	78
$Wald(\beta_1 = -1)$	1000	1000

benchmark model is roughly consistent with empirical evidence on the expectations hypothesis. The model can generate enough variation in the risk premia to account for the rejections of the expectations hypothesis 95% of the time. On the hand, when the risk premium is substracted from $p^B_{1,t+1} - p^F_{1,t}$ β_1 is equal to zero with 5% significance level. Comparing the tables, is is clear that habit-formation is a necessary condition for the rejection of expectations hypothesis. However, since the hypothesis is rejected for real term structure only about 40% of the time, it seems to be the case the monetary policy, which mostly affects nominal rates, plays also an important role. This issue is studied in more detail in Section 6.

In Table 8 the results of the regression (28) are presented for one realization of 200 real and nominal observations and for the data. The data are quarterly observations from 1960:1 to 1998:4 of three and six-month U.S. Treasury bills. In Table 8, *Wald* rows refer to the marginal significance level of the corresponding Wald test. The expectations hypothesis can be rejected at 5% critical level for simulated nominal data but not for simulated real data in this realization. We return to this question in Section 6.

An approximation to population moments is given in Table 9. The simulated data now contains 200,000 observations. The β_1 and R^2 in the benchmark model for nominal bond prices are remarkably close to the values in data.

In the introduction, the expectations hypothesis was motivated as an idea that forward rates are (un)biased predictors of future short term interest rates. One famous study is by Fama and Bliss (1987) who use forward spread to predict the future changes in one-year interest rates one to four years ahead. Table 10 presents the regression results of equation

$$r_{1,t+n} - r_{1,t} = \beta_0 + \beta_1(f_{n,t} - r_{1,t}) \quad \text{for } n = 1, 2, 3, 4 \text{ years}$$

Table 6: The number of rejects in each regressions in the benchmark model for real term structure.

y_{t+1}	$p^b_{1,t+1} - p^f_{1,t}$	$p^b_{1,t+1} - p^f_{1,t} - rrp_{1,t}$
x_t	$p^f_{1,t} - p^b_{1,t}$	$p^f_{1,t} - p^b_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	1000	66
$Wald(\beta_1 = 0)$	407	64
$Wald(\beta_1 = -1)$	1000	1000

Table 7: The number of rejects in each regressions in the benchmark model for real term structure when $b = 0.025$.

y_{t+1}	$p^b_{1,t+1} - p^f_{1,t}$	$p^b_{1,t+1} - p^f_{1,t} - rrp_{1,t}$
x_t	$p^f_{1,t} - p^b_{1,t}$	$p^f_{1,t} - p^b_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	71	71
$Wald(\beta_1 = 0)$	63	82
$Wald(\beta_1 = -1)$	1000	1000

Table 8: The tests of the expectations hypothesis in a single regression with 200 observations.

Variable/Test	Benchmark Real	Benchmark Nominal	Data
β_0	0.0039	0.0030	0.0008
se(β_0)	0.0009	0.0005	0.0005
β_1	−0.1680	−0.3179	−0.4866
se(β_1)	0.1547	0.1147	0.1458
R^2	0.0069	0.0505	0.1505
$Wald(\beta_0 = \beta_1 = 0)$	0	0	0.0003
$Wald(\beta_1 = 0)$	0.2777	0.0056	0.0008
$Wald(\beta_1 = -1)$	7.5617e−008	2.7308e−009	0.0004

Table 9: The tests of the expectations hypothesis in a single regression with 200,000 observations.

Variable/Test	Benchmark Real	Benchmark Nominal	Data
β_0	0.0031	0.0024	0.0008
se(β_0)	0.0000	0.0000	0.0005
β_1	−0.2575	−0.4258	−0.4866
se(β_1)	0.0044	0.0031	0.1458
R^2	0.0213	0.1280	0.1505
$Wald(\beta_0 = \beta_1 = 0)$	0	0	0.0003
$Wald(\beta_1 = 0)$	0	0	0.0008
$Wald(\beta_1 = -1)$	0	0	0.0004

Table 10: Forward spread forecasts of future interest rate changes n years ahead.

Regression	β_0	se(β_0)	β_1	se(β_1)	R^2
Benchmark ($n = 1$)	0.0697	0.1422	0.0365	0.0568	0.0023
Benchmark ($n = 2$)	1.3106	0.2096	0.5840	0.0734	0.2322
Benchmark ($n = 3$)	1.4874	0.1972	0.7184	0.0671	0.3130
Benchmark ($n = 4$)	1.4931	0.1827	0.7217	0.0595	0.3555
Data ($n = 1$)	0.6546	0.1255	−1.2593	0.1828	0.2719
Data ($n = 2$)	0.0019	0.2788	0.1575	0.2533	0.0034
Data ($n = 3$)	−0.1298	0.3074	0.5794	0.2744	0.0431
Data ($n = 4$)	−0.3871	0.2911	0.8665	0.2163	0.0950

for the data from 1960:1 to 1998:4 and the benchmark model with 200 observations. The standard errors are White (1980) heteroskedasticity consistent standard errors. In both cases, β_1 increases with the forecast horizon. With longer maturities, the match is quite good. On the other hand, R^2 increases in the model and decreases in data with the forecast horizon. It should be noted that in the original Fama and Bliss paper, the R^2 also increased with forecast horizon as in our model. The main difference between our data and the data used by Fama and Bliss is that the latter used monthly data from January 1965 to December 1984. In their data sample, the interest rates have strong mean reverting property that increases the forecast power in longer horizons. On the other hand, in our sample the downward trend in data since the early 1980's dominates the data and decreases the forecasting power in longer horizons.

The Term Structure Predictions of Future Economic Activity

Despite the fact that the expectations hypothesis has been rejected over and over again in the empirical literature, it has also been found that the term and forward spreads forecast economic activity. Estrella and Hardouvelis (1991) use the term spread to predict the future changes in the log consumption growth one to four years ahead. The data are quarterly observations from 1960:1

Table 11: Term spread forecasts of future consumption growth n years ahead.

Regression	β_0	se(β_0)	β_1	se(β_1)	R^2
Benchmark ($n = 1$)	−0.4435	0.2061	0.6048	0.1509	0.0760
Benchmark ($n = 2$)	−0.5571	0.1627	0.4060	0.1134	0.0563
Benchmark ($n = 3$)	−0.6546	0.1257	0.4127	0.0842	0.0958
Benchmark ($n = 4$)	−0.6950	0.1117	0.3712	0.0775	0.1008
Data ($n = 1$)	3.0990	0.1353	0.3164	0.1828	0.0758
Data ($n = 2$)	3.1785	0.1156	0.2070	0.0772	0.0504
Data ($n = 3$)	3.2580	0.1027	0.0564	0.0657	0.0052
Data ($n = 4$)	3.3688	0.0882	−0.0427	0.0523	0.0038

to 1998:4 of U.S. consumption non-durables plus services regressed on 10-year government bonds less three-month Treasury bill rates. Table 11 presents the regression results of equation

$$(100/n) * (\log(c_{t+n}) - \log(c_t)) = \beta_0 + \beta_1(r_{10,t} - r_{1,t}) \quad \text{for } n = 1, 2, 3, 4 \text{ years}$$

for the data and the benchmark model. The standard errors are White (1980) heteroskedasticity consistent standard errors. Upward-sloping term structure clearly predicts expansions both in our model and in data, and downward-sloping term structure clearly predicts recessions, again, both in the model and in data. Again, R^2 increases in the model and decreases in data with the forecast horizon. This feature of the model is largely a result the high shocks autocorrelation (see the discussion on page 14).

6. Monetary Policy and Inflation Risk Premium

Recall the definitions of one period zero-coupon nominal bond (24) and the nominal stochastic discount factor (22)

$$p_t^B = E_t[Q_{t+1}] = E_t\left[\beta \frac{MUC_{t+1}P_t}{MUC_t P_{t+1}}\right]. \tag{29}$$

To define the inflation risk premium, write (29) using the definition conditional covariance and the definition of real bond price (23):

$$\begin{aligned} p_t^B &= E_t\left[\beta \frac{MUC_{t+1}P_t}{MUC_t P_{t+1}}\right] \\ &= E_t\left[\beta \frac{MUC_{t+1}}{MUC_t}\right] E_t\left[\frac{P_t}{P_{t+1}}\right] + \text{cov}_t\left[\beta \frac{MUC_{t+1}}{MUC_t}, \frac{P_t}{P_{t+1}}\right] \\ &= p_t^b E_t\left[\frac{P_t}{P_{t+1}}\right] + \text{cov}_t\left[q_{t+1}, \frac{P_t}{P_{t+1}}\right]. \end{aligned}$$

Table 12: Main inflation risk premia statistics in the benchmark case.

	Mean	Standard Deviation	Correlation with Output
IRP $(n = 1)$	0.03052	0.00701	−0.22509
IRP $(n = 2)$	0.05657	0.00959	−0.24723
IRP $(n = 4)$	0.09148	0.00995	−0.27299
IRP $(n = 8)$	0.10770	0.00879	−0.18239
IRP $(n = 12)$	0.09249	0.00945	−0.03320
IRP $(n = 16)$	0.06980	0.01256	0.02905

Since the conditional covariance term is zero for risk-neutral investors and when inflation process is deterministic, we call it the *inflation risk premium*, $irp_{1,t}$, given by

$$irp_{1,t} \equiv \text{cov}_t \left[q_{t+1}, \frac{P_t}{P_{t+1}} \right] = p_t^B - p_t^b E_t \left[\frac{P_t}{P_{t+1}} \right],$$

and similarly $irp_{n,t}$ is the n-period inflation risk premium:

$$irp_{n,t} \equiv \text{cov}_t \left[\prod_{j=1}^{n} q_{t+j}, \frac{P_t}{P_{t+n}} \right] = p_{n,t}^B - p_{n,t}^b E_t \left[\frac{P_t}{P_{t+n}} \right].$$

Assuming that the inflation risk premium is zero, we get the Fisher hypothesis:

$$p_{n,t}^B = p_{n,t}^b E_t \left[\frac{P_t}{P_{t+n}} \right]$$

or by taking logs and multiplying by $-(1/n)$:

$$R_{n,t} \approx r_{n,t} + \frac{1}{n} E_t \left[\log \left(\frac{P_{t+n}}{P_t} \right) \right].$$

That is, nominal interest rate equals the sum of the (ex-ante) real interest rate and the average expected inflation.

Table 12 presents the main statistics for the inflation risk premia in our benchmark case. Because the inflation risk premium is unobservable in data, it is hard to access the mean and standard deviation of premia in Table 12. The best we can do is to compare how different parameters affect the size and volatility of premia. Buraschi and Jiltsov (2005) argue that the time-variation of the inflation risk premium is an important explanatory variable of deviations from the expectations hypothesis.

We address this question by shutting down the monetary policy shocks, i.e., by setting $\sigma_{mp} = 0$. Table 13 presents the inflation risk premia statistics when $\sigma_{mp} = 0$. Not surprisingly, premia are considerably smaller and less volatile without monetary policy shocks. However, Table 14 shows that the expectations hypothesis is actually rejected *more* often without monetary policy shocks.

Table 13: Main inflation risk premia statistics when $\sigma_{mp} = 0$.

	Mean	Standard Deviation	Correlation with Output
IRP $(n = 1)$	-0.00240	0.00027	0.16964
IRP $(n = 2)$	-0.00539	0.00062	0.18748
IRP $(n = 4)$	-0.01275	0.00126	0.24499
IRP $(n = 8)$	-0.03081	0.00300	0.27096
IRP $(n = 12)$	-0.04925	0.00679	0.31258
IRP $(n = 16)$	-0.06494	0.00679	0.31258

Table 14: The number of rejects in each regressions in the benchmark model for nominal term structure when when $\sigma_{mp} = 0$.

y_{t+1}	$p^B_{1,t+1} - p^F_{1,t}$	$p^B_{1,t+1} - p^F_{1,t} - nrp_{1,t}$
x_t	$p^F_{1,t} - p^B_{1,t}$	$p^F_{1,t} - p^B_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	1000	72
$Wald(\beta_1 = 0)$	988	74
$Wald(\beta_1 = -1)$	999	1000

That is, unlike in Buraschi and Jiltsov (2005), inflation risk premium is *not* the explanation for the rejections of expectations hypothesis. How can that be? Recall the regression equation (28)

$$p^B_{1,t+1} - p^F_{1,t} = \beta_0 + \beta_1(p^F_{1,t} - p^B_{1,t}),$$

and recall the definition of risk premium

$$nrp_{1,t} \equiv -\operatorname{cov}_t\left[Q_{t+1}, \frac{p^B_{1,t+1}}{p^B_{1,t}}\right] = E_t[p^B_{1,t+1}] - p^F_{1,t}.$$

Substituting the second equation into the first and letting ψ_t denote the one-period nominal risk premium, we get

$$p^B_{1,t+1} - E_t[p^B_{1,t+1}] + \psi_t = \beta_0 + \beta_1(E_t[p^B_{1,t+1}] - \psi_t - p^B_{1,t})$$

$E_t[p^B_{1,t+1}]$ can be written as $p^B_{1,t+1} - \epsilon_{t+1}$, where the prediction error term, ϵ_{t+1}, is orthogonal to the information available at time t. Hence, the estimate of β_1 in the expectations hypothesis regression (28) converges to

$$\operatorname{plim}\hat{\beta}_1 = \frac{\rho\sigma_p\sigma_\psi - \sigma^2_\psi}{\sigma^2_p + \sigma^2_\psi - 2\rho\sigma_p\sigma_\psi}, \tag{30}$$

where ρ denotes $\operatorname{corr}(\psi_t, E_t\Delta p^B_{t+1})$, σ_p denotes $\operatorname{std}(E_t\Delta p^B_{t+1})$, and σ_ψ denotes $\operatorname{std}(\psi_t)$.

It is illuminating to study equation (30) as a function of $\text{std}(\psi_t)$ and $\text{std}(E_t \Delta p^B_{t+1})$. If ψ_t is deterministic, i.e., $\sigma_\psi = 0$, $\text{plim}\,\hat{\beta}_1 = 0$, i.e., the expectations hypothesis holds. On the other hand, if $\sigma_p = 0$, $\text{plim}\,\hat{\beta}_1 = -1$, i.e., the forward prices are not useful in predicting future bond prices. The asymptotic behavior is opposite. If $\sigma_\psi \to +\infty$, $\text{plim}\,\hat{\beta}_1 \to -1$ and if $\sigma_p \to +\infty$, $\text{plim}\,\hat{\beta}_1 \to 0$. The behavior of $\text{plim}\,\hat{\beta}_1$ corresponding to the intermediate values of σ_p and σ_ψ depends on the sign of ρ. Figures 1 and 2 show the behavior of $\text{plim}\,\hat{\beta}_1$ as a function of σ_p and σ_ψ when $\rho > 0$ and $\rho < 0$, respectively.

When $\rho < 0$, $\text{plim}\,\hat{\beta}_1$ is always between 0 and -1. Moreover, if $\rho < 0$ the behavior $\text{plim}\,\hat{\beta}_1$ as a function of σ_p and σ_ψ is always monotone.

$$\frac{\partial\, \text{plim}\,\hat{\beta}_1}{\partial \sigma_\psi} = \frac{\sigma_p[\rho(\sigma_p^2 + \sigma_\psi^2) - 2\sigma_\psi\sigma_p]}{(\sigma_p^2 + \sigma_\psi^2 - 2\rho\sigma_p\sigma_\psi)^2} \leq 0, \text{ if } \rho < 0;$$
$$\frac{\partial\, \text{plim}\,\hat{\beta}_1}{\partial \sigma_p} = \frac{\sigma_\psi[-\rho(\sigma_p^2 + \sigma_\psi^2) + 2\sigma_\psi\sigma_p]}{(\sigma_p^2 + \sigma_\psi^2 - 2\rho\sigma_p\sigma_\psi)^2} \geq 0, \text{ if } \rho < 0.$$

On the other hand, if $\rho > 0$, it is possible that $\text{plim}\,\hat{\beta}_1$ is positive. The sign of $\text{corr}(\psi_t, E_t \Delta p^b_{t+1})$ is not easy to determine analytically in the model, but it worth noticing that the regression coefficient β_1 was positive in Backus, Gregory, and Zin (1989) and in Seppälä (2004) for the model in which there is no habit formation and no market frictions. In our model, β_1 is -0.2575 for real bond prices and -0.4258 for nominal nominal bond prices with habit formation, but -0.0179 for real bond prices and -0.1364 for nominal bond prices with minimal habit formation ($b = 0.025$). ρ is -0.4045 for real bond prices and -0.5020 for nominal nominal bond prices with habit formation, but -0.0454 for real bond prices and -0.0710 for nominal bond prices with minimal habit formation. At least in these cases, there is a relationship between the size and the sign of β_1 and ρ.

To summarize the discussion above, to be able to reject the expectations hypothesis with a negative regression coefficient as in data, it helps if $\rho < 0$. This is not to the case without habit formation or market frictions. In addition, one needs either very high volatility in the risk premium *or* very low volatility in the predictable component of bond price changes, or a combination of sufficiently high volatility in the risk premium *and* sufficiently low volatility in the predictable component of bond price changes. Earlier literature on equity premium puzzle, e.g., Boldrin, Christiano, and Fisher (1997) and Athanasoulis and Sussman (2004), has emphasized that habit formation generates too volatile short-term interest rates. This would be the case also in our benchmark economy *but* for the fact that our monetary policy authority cares about interest rate smoothing and hence reduces $\text{std}(E_t \Delta p^b_{t+1})$.

A similar argument was raised earlier by Mankiw and Miron (1986) and empirical evidence supports this interpretation. Mankiw and Miron show that it is much more difficult to reject the expectations hypothesis using data prior to the founding of the Fed.[12] They suggest that the explanation is the Federal Reserve's commitment to stabilizing interest rates.

[12] Choi and Wohar (1991) cannot reject the expectations hypothesis over the sample period of 1910–1914.

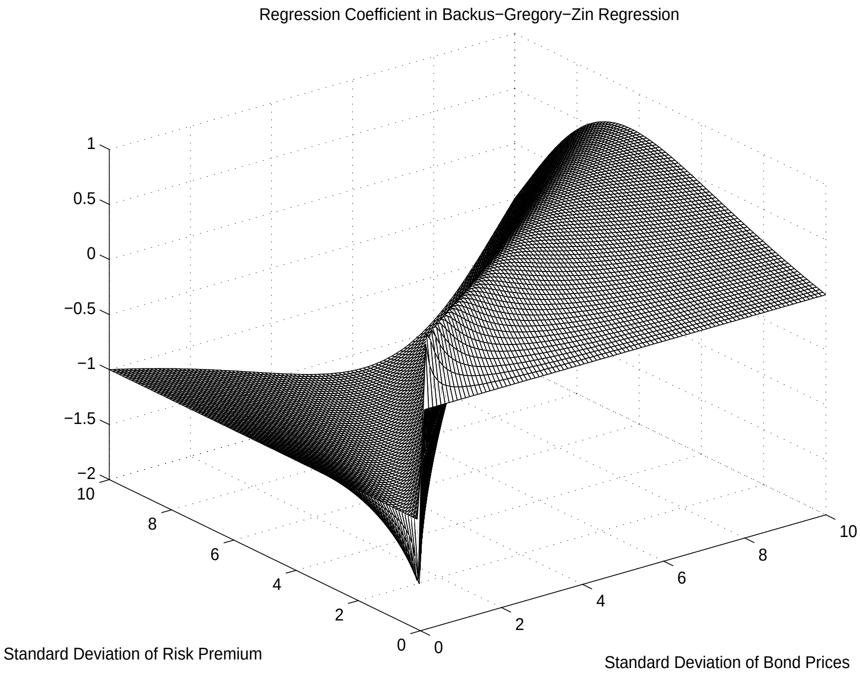

Figure 1: $\operatorname{plim} \hat{\beta}_1$ as a function of σ_p and σ_ψ when $\rho > 0$.

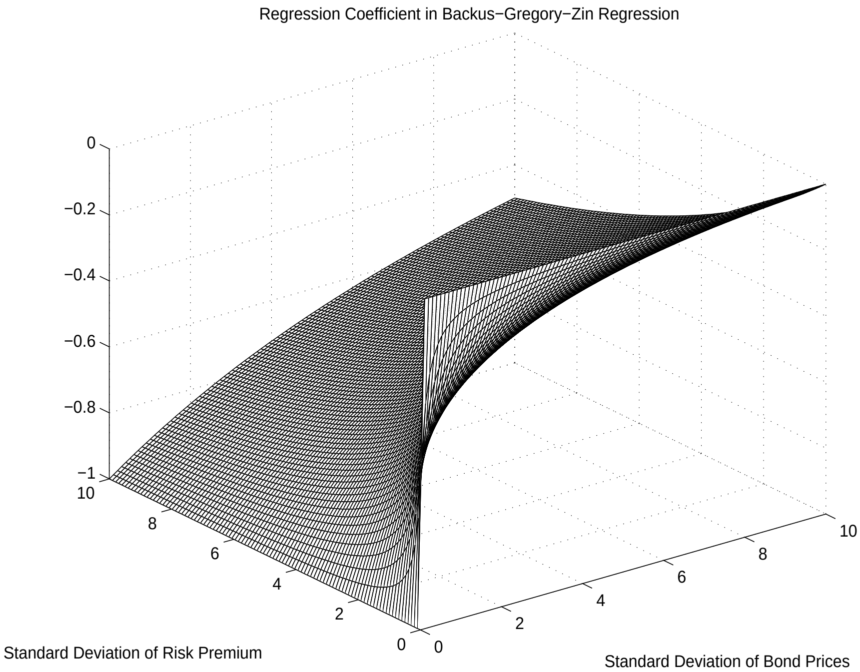

Figure 2: $\operatorname{plim} \hat{\beta}_1$ as a function of σ_p and σ_ψ when $\rho < 0$.

Table 15: The number of rejects in each regressions in the benchmark model for nominal term structure when $\chi = 0.375$.

y_{t+1}	$p^B_{1,t+1} - p^F_{1,t}$	$p^B_{1,t+1} - p^F_{1,t} - nrp_{1,t}$
x_t	$p^F_{1,t} - p^B_{1,t}$	$p^F_{1,t} - p^B_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	1000	80
$Wald(\beta_1 = 0)$	291	68
$Wald(\beta_1 = -1)$	1000	1000

Table 16: Main term structure statistics when $\sigma_{mp} = 0$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	6.65061	1.03404	0.33510
$R_{4,t}$	6.90738	0.99649	0.37013
$R_{40,t}$	6.98312	0.53742	0.42493
$R_{80,t}$	6.97728	0.32843	0.42961
$R_{120,t}$	6.97328	0.21950	0.43125

We next study the issue by setting the interest rate smoothing parameter $\chi = 0.375$.[13] As shown in Table 15, with a low level of interest rate smoothing, we are able to reject the expectations hypothesis only 30% of the time. Interest rate smoothing by the monetary policy authority clearly plays an important role.

7. *Which Shocks Matter?*

In our baseline model, we have three different shocks: a technology shock, a preference shock, and a monetary policy shock. In this section we investigate the importance of these different shocks.

No Monetary Policy Shocks ($\sigma_{mp} = 0$)

Table 16 presents the main term structure statistics when $\sigma_{mp} = 0$. Not unexpectedly, the volatility of interest is lower and correlations of output of interest rates is higher when there are no monetary policy shocks. However, the means are actually higher.

No Technology Shocks ($\sigma_a = 0$)

With no technology shocks, the results are similar to the same our benchmark case, see Table 17. Means are a bit higher, standard deviations a bit lower, and correlations a bit higher than in the

[13] We were unable to solve the model for $\chi < 0.375$. The rational expectations equilibrium is not unique in this case.

Table 17: Main term structure statistics when $\sigma_a = 0$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	4.96466	1.83442	0.11663
$R_{4,t}$	6.22416	1.34170	0.21050
$R_{40,t}$	6.61092	0.54828	0.39498
$R_{80,t}$	6.61923	0.33259	0.40510
$R_{120,t}$	6.62103	0.21960	0.40169

Table 18: Main term structure statistics when $\sigma_d = 0$ and $\sigma_a = 1\%$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	5.15994	1.70843	−0.56563
$R_{4,t}$	6.26928	1.15487	−0.58485
$R_{40,t}$	6.60587	0.24994	−0.46887
$R_{80,t}$	6.61780	0.13488	−0.44818
$R_{120,t}$	6.62183	0.09092	−0.44830

benchmark case. The expectations hypothesis can be rejected 80% of the time. Since neither monetary policy nor technology shocks matter for our results, it must be the case that the demand shock plays the most crucial role in the model.

No Demand Shocks ($\sigma_d = 0$, $\sigma_a = 1\%$)

Without demand shocks and with higher technology shock volatility, the interest rates are much less volatile and flatter as presented in Table 18. In addition, interest rates and inflation are countercyclical. The expectations hypothesis is still rejected 84% of the time as shown in Table 19.

Summary

Summarizing, monetary policy and technology shocks are not crucial for our results, but they help. Preference shocks are crucial for matching the cyclical behavior of the interest rates but not for the rejections of the expectations hypothesis. These results are somewhat similar to Nakajima (2003) who shows that the standard RBC model driven by the Solow residual cannot explain the "preference residual" (the difference between real wage and the marginal rate of substitution between consumption and leisure), but the model driven by the preference residual can account for the Solow residual.

Table 19: The number of rejects in each regressions in the benchmark model for nominal term structure when $\sigma_d = 0$ and $\sigma_a = 1\%$.

y_{t+1}	$p^B_{1,t+1} - p^F_{1,t}$	$p^B_{1,t+1} - p^F_{1,t} - nrp_{1,t}$
x_t	$p^F_{1,t} - p^B_{1,t}$	$p^F_{1,t} - p^B_{1,t}$
$Wald(\beta_0 = \beta_1 = 0)$	1000	64
$Wald(\beta_1 = 0)$	844	69
$Wald(\beta_1 = -1)$	1000	1000

8. *Conclusions*

We show that a New Keynesian model with habit-formation preferences and a monetary policy feedback rule with interest rate smoothing produces pro-cyclical interest rates, counter-cyclical term spreads, and creates enough volatility in the risk premium to account for the rejections of expectations hypothesis. Our results are related to conclusion reached by Dotsey and Otrok (1995)

> "[R]egression results [for the expectations hypothesis] that are in accord with those obtained in practise can be generated by the combination of (i) Fed behavior that both smooths the movements in interest rates... and (ii) time-varying term premia that are calibrated to match data moments."

In our model, habit formation delivers (ii) and interest rate smoothing delivers (i). Without habit formation, we reject the expectations hypothesis only 7.5% of the time. With habit formation but with much less interest rate smoothing, we reject only 25% of the time. With habit formation and realistic level of interest rate smoothing, we reject the expectation hypothesis 95% of the time.

It is important to note that in our model, it is the *systematic* monetary policy that brings our results, and not the monetary policy shocks as in Buraschi and Jiltsov (2005). We also find that monetary policy and the technology shocks are not crucial for our results. Preference shocks are crucial for matching the cyclical behavior of the interest rates but not for the rejections of the expectations hypothesis.

A. *The Inflation Rate Dynamics*

By iterating (15) and (16), we obtain:

$$\hat{G}_{t+1} = \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{G}_{t+2}$$
$$\hat{G}_{t+2} = \mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{G}_{t+3}$$
$$\hat{H}_{t+1} = MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{H}_{t+2}$$
$$\hat{H}_{t+2} = MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{H}_{t+3}$$

or

$$
\begin{aligned}
\hat{G}_t &= \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[\mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{G}_{t+2} \right] \\
&= \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \Big[\mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \\
&+ \theta_p \beta \left(\mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{G}_{t+3} \right) \Big] \\
&= \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[\mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] \\
&+ (\theta_p \beta)^2 \left[\mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + (\theta_p \beta)^3 \hat{G}_{t+3},
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{H}_t &= MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{H}_{t+2} \right] \\
&= MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \left(MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{H}_{t+3} \right) \right] \\
&= MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] + (\theta_p \beta)^2 \left[MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + (\theta_p \beta)^3 \hat{H}_{t+3}.
\end{aligned}
$$

Plugging these expressions into (13) and (14):

$$
\begin{aligned}
G_t &= \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \Big\{ \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[\mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] \\
&+ (\theta_p \beta)^2 \left[\mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + (\theta_p \beta)^3 \hat{G}_{t+3} \Big\} \\
&= \mu MC_t (P_t(z)/P_t)^{1-\theta} Y_t + \mu \theta_p \beta \left[(MUC_{t+1}/MUC_t) MC_{t+1} (P_t(z)/P_{t+1})^{1-\theta} Y_{t+1} \right] \\
&+ \mu (\theta_p \beta)^2 \left[(MUC_{t+2}/MUC_t) MC_{t+2} (P_t(z)/P_{t+2})^{1-\theta} Y_{t+2} \right] + \frac{P_t(z)^{1-\theta}}{MUC_t} (\theta_p \beta)^3 \hat{G}_{t+3}. \qquad (31)
\end{aligned}
$$

and

$$
\begin{aligned}
H_t &= \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \Big\{ MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] \\
&+ (\theta_p \beta)^2 \left[MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + (\theta_p \beta)^3 \hat{H}_{t+3} \Big\} \\
&= MC_t (P_t(z)/P_t)^{1-\theta} Y_t + \theta_p \beta \left[(MUC_{t+1}/MUC_t) (P_t(z)/P_{t+1})^{1-\theta} Y_{t+1} \right] \\
&+ (\theta_p \beta)^2 \left[(MUC_{t+2}/MUC_t) (P_t(z)/P_{t+2})^{1-\theta} Y_{t+2} \right] + \frac{P_t(z)^{1-\theta}}{MUC_t} (\theta_p \beta)^3 \hat{H}_{t+3}. \qquad (32)
\end{aligned}
$$

Dividing (31) by (32), we get (12). Note also that:

$$
P_t(z) = \frac{G_t}{H_t} = \frac{\frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{G}_t}{\frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{H}_t} = \frac{\hat{G}_t}{\hat{H}_t},
$$

To obtain stationary variables under a positive money growth rate steady state regime start from equations (15)–(16) and divide $\hat{G}_t$ by P_t^θ and $\hat{H}_t$ by $P_t^{\theta-1}$ to get

$$\tilde{G}_t \equiv \frac{\hat{G}_t}{P_t^\theta} = \mu MUC_t \frac{MC_t}{P_t} Y_t + \theta_p \beta \frac{\hat{G}_{t+1}}{P_t^\theta} \tag{33}$$

$$= \mu MUC_t \frac{MC_t}{P_t} Y_t + \theta_p \beta \frac{\hat{G}_{t+1}}{P_{t+1}^\theta} \frac{P_{t+1}^\theta}{P_t^\theta} = \mu MUC_t mc_t Y_t + \theta_p \beta \tilde{G}_{t+1} (1+\pi_{t+1})^\theta \tag{34}$$

$$\tilde{H}_t \equiv \frac{\hat{H}_t}{P_t^{\theta-1}} = MUC_t Y_t + \theta_p \beta \frac{\hat{H}_{t+1}}{P_t^{\theta-1}} \tag{35}$$

$$= MUC_t Y_t + \theta_p \beta \frac{\hat{H}_{t+1}}{P_{t+1}^{\theta-1}} \frac{P_{t+1}^{\theta-1}}{P_t^{\theta-1}} = MUC_t Y_t + \theta_p \beta \tilde{H}_{t+1} (1+\pi_{t+1})^{\theta-1} \tag{36}$$

where $mc_t \equiv MC_t/P_t$ is the real marginal cost. Since

$$\tilde{H}_t = \frac{\hat{H}_t}{P_t^{\theta-1}} = \frac{\hat{H}_t P_t}{P_t^\theta} \qquad \Longrightarrow \qquad \frac{\hat{H}_t}{P_t^\theta} = \frac{\tilde{H}_t}{P_t}$$

and

$$P_t(z) = \frac{G_t}{H_t} = \frac{\hat{G}_t}{\hat{H}_t} = \frac{\hat{G}_t/P_t^\theta}{\hat{H}_t/P_t^\theta} = \frac{\tilde{G}_t P_t}{\tilde{H}_t}.$$

the law of motion for the price index

$$P_t^{1-\theta} = \theta_p P_{t-1}^{1-\theta} + (1-\theta_p) P_t(z)^{1-\theta} = \theta_p P_{t-1}^{1-\theta} + (1-\theta_p) \left[\frac{\hat{G}_t}{\hat{H}_t}\right]^{1-\theta}$$

can be divided by $P_t^{1-\theta}$ to obtain

$$[(1+\pi_t)]^{1-\theta} = \theta_p + (1-\theta_p) \left[\frac{P_t(z)}{P_{t-1}}\right]^{1-\theta} = \theta_p + (1-\theta_p) \left[\frac{\hat{G}_t}{\hat{H}_t P_{t-1}}\right]^{1-\theta} = \theta_p + (1-\theta_p) \left[\frac{\tilde{G}_t}{\tilde{H}_t}(1+\pi_t)\right]^{1-\theta}.$$

B. Sensitivity Analysis

In this Section, we briefly describe seven different experiments to illustrate the features of our model. We only concentrate on most dramatic differences with the benchmark model. Details are available on request.

Test #1: Minimal Habit Formation ($b = 0.025$)

Table 20 shows the average yield curve when there is minimal habit formation present in the model. The term structure is flat, and hence it is no surprise that—as shown earlier in Table 5—habit formation is a necessary condition for the rejection of expectations hypothesis.

Table 20: Main term structure statistics when $b = 0.025$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	6.51565	1.68657	0.00750
$R_{4,t}$	6.55319	1.31606	0.12208
$R_{40,t}$	6.58904	0.46904	0.40686
$R_{80,t}$	6.58185	0.28695	0.40910
$R_{120,t}$	6.57826	0.19881	0.40357

Table 21: Main term structure statistics when $\gamma = 1$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	3.87803	2.67272	0.21392
$R_{4,t}$	5.88859	2.23616	0.31391
$R_{40,t}$	6.49121	1.08374	0.43033
$R_{80,t}$	6.47687	0.66179	0.66179
$R_{120,t}$	6.46501	0.45313	0.42869

Test #2: Logarithmic Preferences ($\gamma = 1$)

With less curvature in the utility function, the the term structure is much steeper than in our benchmark case, as shown in Table 21 and the expectation hypothesis is rejected 99% of the time. There is a connection between the steepness of the yield curve and how easy it is to rejecte the expectations hypothesis, see also Test #3 below. Moreover, everything is more volatile with less risk aversion, for example, the standard deviation of Y_t is 4.5% compared to 2% in the benchmark case and 1.69% in data.

Test #3: Less Interest Rate Smoothing ($\chi = 0.375$)

When the monetary policy authority conducts less interest rate smoothing, the interest rates, not surprisingly, are more volatile as presented in Table 22. However, the term structure is flatter and it is more difficult to reject expectations hypothesis as discussed in Section 6.

Test #4: Zero Steady State Inflation

When the steady state inflation is zero, the nominal interest rates are lower, as presented in Table 23. This doesn't affect the model's ability to reject the expectations hypothesis; it is still rejected 97% of the time.

Test #5: Higher Steady State Inflation

When the steady state inflation is twice as large, 6% per year, the nominal interest rates are much higher as presented in Table 24. This doesn't affect the model's ability to reject the expectations;

Table 22: Main term structure statistics when $\chi = 0.375$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	6.08739	2.38509	0.26979
$R_{4,t}$	6.70940	2.03583	0.35399
$R_{40,t}$	6.89657	1.00046	0.42626
$R_{80,t}$	6.87795	0.60503	0.42760
$R_{120,t}$	6.86571	0.40851	0.42746

Table 23: Main term structure statistics when the steady inflation is zero.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	2.08850	1.76299	0.14081
$R_{4,t}$	3.20802	1.27313	0.24971
$R_{40,t}$	3.54829	0.53285	0.41323
$R_{80,t}$	3.55266	0.32770	0.41422
$R_{120,t}$	3.55383	0.22748	0.40983

it is rejected 98% of the time.

Test #6: Monetary Policy Responds to Output ($\omega_y = 0.001$)

$\omega_y = 0.001$ is the largest value of ω_y that gave a solution to the problem. At that level, the steady state inflation is almost zero. If ω_y is increased the steady state inflation becomes negative and the model cannot be solved.

Test #7: More Aggressive Policy ($\omega_\pi = 3.0$)

With more aggressive monetary policy, the the term structure is flatter than in our benchmark case and inflation and interest rates are less volatile, as shown in Table 25. The expectation hypothesis can still be rejected 91% of the time.

Test #8: Less Aggressive Policy ($\omega_\pi = 1.2$)

With more aggressive monetary policy, the the term structure is steeper than in our benchmark case and inflation and interest rates are more volatile, as shown in Table 26. The expectation hypothesis can be rejected 99% of the time.

Table 24: Main term structure statistics when the steady inflation is 6% per year.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	7.55514	2.19424	0.07050
$R_{4,t}$	9.06968	1.57581	0.16048
$R_{40,t}$	9.53428	0.55138	0.37931
$R_{80,t}$	9.54437	0.33687	0.38444
$R_{120,t}$	9.54728	0.23187	0.38239

Table 25: Main term structure statistics when $\omega_\pi = 3.0$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	5.83803	1.66482	0.07336
$R_{4,t}$	6.64256	1.15147	0.15790
$R_{40,t}$	6.89837	0.41495	0.39335
$R_{80,t}$	6.90837	0.25399	0.39835
$R_{120,t}$	6.91219	0.17408	0.39778

Table 26: Main term structure statistics when $\omega_\pi = 1.2$.

	Mean	Standard Deviation	Correlation with Output
$R_{1,t}$	3.90634	2.04803	0.12923
$R_{4,t}$	5.44751	1.50353	0.24363
$R_{40,t}$	5.90065	0.62212	0.40858
$R_{80,t}$	5.90076	0.38227	0.40947
$R_{120,t}$	5.89922	0.26464	0.40451

References

Abken, P.A. (1993), "Inflation and the Yield Curve", *Federal Reserve Bank of Atlanta Economic Review*, May/June, 13–31.

Ascari, G., (2004), "Staggered Prices and Trend Inflation: Some Nuisances", *Review of Economic Dynamics*, 7, 642–677.

Athanasoulis, S.G. and O. Sussman (2004), "Habit Formation and the Equity-Premium Puzzle: a Skeptical View", Oxford Financial Research Centre Working Paper 2004–12.

Backus, D.K., A.W. Gregory, and S.E. Zin (1989), "Risk Premiums in the Term Structure: Evidence from Artificial Economies", *Journal of Monetary Economics*, 24, 371–399.

Bekaert G, S. Cho, and A. Moreno (2005), "New-Keynesian Macroeconomics and the Term Structure", Mimeo, Columbia University.

Bernanke, B. and Gertler, M., (2000), "Monetary Policy and Asset Price Volatility", Working Paper No. 7559, National Bureau of Economic Research.

Blough, S.R. (1994), "Yield Curve Forecasts of Inflation: A Cautionary Tale", *New England Economic Review*, May/June, 3–15.

Boldrin, M., L.J. Christiano, and J.D.M. Fisher (1997), "Habit Persistence and Asset Returns in an Exchange Economy", *Macroeconomic Dynamics*, 1, 312–332.

Boldrin, M., L.J. Christiano, and J.D.M. Fisher (2001), "Habit Persistence, Asset Returns and the Business Cycle", *American Economic Review*, 91, 149–166.

Breeden, D.T. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics*, 7, 265–296.

Buraschi, A. and A. Jiltsov (2003), "Habit Persistence and the Nominal Term Structure of Interest Rates", Mimeo, London Business School.

Buraschi, A. and A. Jiltsov (2005), "Inflation Risk Premia and the Expectations Hypothesis", *Journal of Financial Economics*, 75, 429–490.

Calvo, G. (1983), "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics*, 12, 383–398.

Campbell, J. Y. (1995), "Some Lessons from the Yield Curve", Working Paper No. 5031, National Bureau of Economic Research.

Campbell, J.Y. and J.H. Cochrane (1999), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy*, 107, 205–251.

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

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.

Chen, N.-F. (1991), "Financial Investment Opportunities and the Macroeconomy", *Journal of Finance*, 46, 529–554.

Choi, S. and M.E. Wohar, "New Evidence Concerning the Expectations Theory for the Short End of the Maturity Spectrum", *Journal of Financial Research*, 58, 83–92.

Christiano, L.J., M. Eichenbaum, and C, Evans, (2001), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", Working Paper No. 8403, National Bureau of Economic Research.

Clarida, R., Gali, J. and Gertler, M., (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics*, 115, 147–180.

Cochrane, J.H. and M. Piazzesi (2002), "The Fed and Interest Rates—A High Frequency Identification", *American Economic Review Papers and Proceedings*, 92, 90–95.

Cochrane, J.H. and M. Piazzesi (2005), "Bond Risk Premia", *American Economic Review*, forthcoming.

Dai, Q. (2002) "Asset Pricing in a Neoclassical Model with Limited Participation", Mimeo, New York University.

Dai, Q. and K.J. Singleton (2002) "Expectations Puzzle, Time-varying Risk Premia, and Affine Models of the Term Structure", *Journal of Financial Economics*, 63, 415–441.

Donaldson, J.B., T. Johnsen, and R. Mehra (1990), "On the Term Structure of Interest Rates", *Journal of Economic Dynamics and Control*, 14, 571–596.

Dotsey, M., R.G. King, and A.L. Wolman (1999), "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output", *Quarterly Journal of Economics*, 114, 655–690.

Dotsey, M. and C. Otrok (1995), "The Rational-Expectations Hypothesis of the Term Structure, Monetary Policy, and Time-Varying Term Premia", *Federal Reserve Bank of Richmond Economic Quarterly*, 81, 65–81.

Duffee, G. (2001), "Supplementary Information for 'Term Premia and Interest Rate Forecasts in Affine models"', Mimeo, University of California, Berkeley.

Duffee, G. (2002), "Term Premia and Interest Rate Forecasts in Affine Models", *Journal of Finance*, 57, 405–443.

Estrella A. and G.A. Hardouvelis (1991), "The Term Structure as a Predictor of Real Economic Activity", *Journal of Finance*, 46, 555–580.

Evans, C.L. and D.A. Marshall (1998), "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory", *Carnegie-Rochester Conference Series On Public Policy*, 49, 53-111.

Fama, E.F. (1975), "Short Term Interest Rates as Predictors of Inflation", *American Economic Review*, 65, 269–282.

Fama, E.F. (1984), "The Information in the Term Structure", *Journal of Financial Economics*, 13, 509–528.

Fama, E.F. (1990), "Term-Structure Forecasts of Interest Rates, Inflation, and Real Returns", *Journal of Monetary Economics*, 25, 59–76.

Fama, E.F. and R.R. Bliss (1987), "The Information in Long-Maturity Forward Rates", *American Economic Review*, 77, 680–692.

Fama, E.F. and K.R. French (1989), "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics*, 25, 23–50.

Friedman, M. (1968), "The Role of Monetary Policy", *American Economic Review*, 63, 1–17.

Froot, K.A. (1989), "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates", *Journal of Finance*, 44, 283–304.

Fuhrer, J. (2000), "Habit Formation in Consumption and its Implications for Monetary Policy Analysis", *American Economic Review*, 90, 367–390.

Galí, J. and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Investigation", *Journal of Monetary Economics*, 44, 195–222.

Harvey, C.R. (1988), "The Real Term Structure and Consumption Growth", *Journal of Financial*

Economics, 22, 305–333.
Hördahl, P., Tristani, O. and Vestin, D., (2002), "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics", Mimeo, European Central Bank.
Hördahl, P., Tristani, O. and Vestin, D., (2005), "The Yield Curve and Macroeconomic Dynamics", Mimeo, European Central Bank.
Ireland, P., (2001), "Sticky Price Models of the Business Cycle: Ppecification and Stability", *Journal of Monetary Economics*, 47, 3–18.
Jermann, U.J. (1998), "Asset Pricing in Production Economies", *Journal of Monetary Economics*, 41, 257–275.
Jorion, P. and F. Mishkin (1991), "A Multicountry Comparison of Term-Structure Forecasts at Long Horizons", *Journal of Financial Economics*, 29, 59–80.
King, R.G. and M.W. Watson (1996), "Money, Prices, and the Business Cycle", *Review of Economics and Statistics*, 78, 35–53.
King, R. and Wolman, A., (1996), "Inflation Targeting in a St. Louis Model of the 21st Century", *Federal Reserve Bank of St. Louis Review*, May/June.
Klenow, P.J., and O. Kryvtsov (2004), "State Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?", Mimeo, Stanford University.
Labadie, P. (1994), "The Term Structure of Interest Rates over the Business Cycle", *Journal of Economic Dynamics and Control*, 18, 671–697.
LeRoy, S.F. (1973), "Risk Aversion and the Martingale Property of Stock Prices", *International Economic Review*, 14, 436–446.
Lucas, R.E., Jr. (1978), "Asset Prices in an Exchange Economy", *Econometrica*, 46, 1429–1446.
Mankiw, N.G. and J.A. Miron (1986), "The Changing Behavior of the Term Structure of Interest Rates", *The Quarterly Journal of Economics*, 101, 211–228
Mankiw, N.G. and L.H. Summers (1984), "Do Long-Term Interest Rates Overreact to Short-Term Interest Rates?", *Brookings Papers on Economic Activity*, 1984, 223–242.
McCulloch, J.H. and H-C Kwon, "US Term Structure Data, 1947–1991", Working Paper #93–6, Ohio State University.
Mehra, R. and E. C. Prescott (1985) "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, 15, 145–161.
Melino, A. (1988) "The Term Structure of Interest Rates: Evidence and Theory", *Journal of Economic Surveys*, 2, 335–366.
Mishkin, F.S. (1981), "The Real Interest Rate: An Empirical Investigation", *Carnagie-Rochester Conference Series on Public Policy*, 15, 151–200.
Mishkin, F.S. (1990a), "The Information in the Longer Maturity Term Structure about Future Inflation", *Quarterly Journal of Economics*, 105, 815–828.
Mishkin, F.S. (1990b), "Yield Curve", Working Paper No. 3550, National Bureau of Economic Research.
Mishkin, F.S. (1992), "Is the Fisher Effect for Real?", *Journal of Monetary Economics*, 30, 195–215.
Mishkin, F.S. (1991), "A Multi-Country Study of the Information in the Shorter Maturity Term Structure about Future Inflation", *Journal of International Money and Finance*, 10, 2–22.
Nakajima, T. (2003), "A Business Cycle Model with Variable Capacity Utilization and Demand Disturbances", *European Economic Review*, forthcoming.
Piazzesi, M. (2005), "Bond yields and the Federal Reserve", *Journal of Political Economy*, forth-

coming.
Rabanal, P., (2004), “Monetary Policy Over the U.S. Cycle: Evidence and Implications”, Mimeo, International Monetary Fund.
Rabanal, P. and Rubio-Ramirez, J, (2003), “Comparing New Keynesian Models in the Euro Area: A Bayesian Approach”, Working Paper No. 2003–30, Federal Reserve Bank of Atlanta.
Ravenna, F. (2002), “The Impact of Inflation Targeting in Canada: A Structural Analysis”, Mimeo, University of California, Santa Cruz.
Rubinstein, M. (1976), “The Valuation of Uncertain Income Streams and the Pricing of Options”, *Bell Journal of Economics and Management Science*, 7, 407–425.
Rudebusch, G.D. and T. Wu (2004), “A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy”, Working Paper 2003–17, Federal Reserve Bank of San Francisco.
Schmitt-Grohe, S. and M. Uribe (2004), “Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function”, *Journal of Economic Dynamics and Control*, 28, 755–775.
Seppälä, J. (2000), “The Term Structure of Real Interest Rates: Theory and Evidence from UK Index-Linked Bonds”, Office of Research Working Paper 00–0120, University of Illinois.
Seppälä, J. (2004), “The Term Structure of Real Interest Rates: Theory and Evidence from UK Index-Linked Bonds”, *Journal of Monetary Economics*, 51, 1509–1549.
Seppälä, J. and L. Xie (2004), “The Cyclical Properties of the Term Structure of Interest Rates”, Mimeo, University of Illinois, Urbana-Champaign.
Shiller, R.I. (1979), “The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure”, *Journal of Political Economy*, 87, 1190–1219.
Shiller, R.I. (1990), “The Term Structure of Interest Rates” in B. M. Friedman and F. H. Hahn (eds.), *Handbook of Monetary Economics*, 1, North-Holland, Netherlands.
Shiller, R.I., J.Y. Campbell, and K. Schoenholtz (1983), “Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates”, *Brookings Papers on Economic Activity*, 2, 173–225.
Svensson, L.E.O. (1994a), “Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994”, Working Paper No. 4871, National Bureau of Economic Research.
Svensson, L.E.O. (1994b), “Monetary Policy with Flexible Exchange Rates and Forward Interest Rates as Indicators”, Discussion Paper No. 941, Centre for Economic Policy Research.
Söderlind, P. and L.E.O. Svensson (1997), “New Techniques to Extract Market Expectations from Financial Instruments”, *Journal Monetary Economics*, 40, 383–429.
Taylor, John, (1993), “Discretion versus Policy Rules in Practice’, *Carnegie-Rochester Conferences Series on Public Policy*, 39, 195–204.
Wachter, J.A. (2004), “A Consumption-Based Model of the Term Structure of Interest Rates”, Mimeo, University of Pennsylvania.
Walsh, C.E., (2003), *Monetary Theory and Policy*, 2nd. ed., MIT Press, Cambridge, MA.
White, H. (1980), “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, 48, 817–838.
Woodford, M. (2003), *Interest and Prices*, Princeton University Press, Princeton, NJ.