Monetary Policy and Rejections of the Expectations Hypothesis

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

August 8, 2006

Abstract

We study the rejection of the expectations hypothesis within a New Keynesian business cycle model. According to Backus, Gregory, and Zin (1989), the Lucas general equilibrium asset pricing model can account for neither sign nor magnitude of average risk premia in forward prices, and is unable to explain rejection of the expectations hypothesis. We show that a New Keynesian model with habit-formation preferences and a monetary policy feedback rule produces an upward-sloping average term structure of interest rates, procyclical interest rates, and countercyclical term spreads. In the model, as in U.S. data, inverted term structure predicts recessions. Most importantly, a New Keynesian model is able to account for rejections of the expectations hypothesis. Contrary to Buraschi and Jiltsov (2005), we identify systematic monetary policy as a key factor behind this result. Rejection of the expectation hypothesis can be entirely explained by the volatility of just two real shocks which affect technology and preferences.

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

JEL classification: E43, E44, E5, G12.
“In Blinder (2004, Chapter 3), I suggested (but certainly did not prove) that the expectations theory fails because long rates are far more sensitive to short rates than ‘rational’ pricing models predict. This hypothesis may or may not be correct. My main purpose in calling attention to the term structure puzzle here is not to resolve it, but rather to urge central bank research departments to give it high priority. It may be the piece of the monetary transmission mechanism about which we are most in the dark.”


1. Introduction

The term structure of interest rates contains information about agents’ expectations of future interest rates, inflation rates, and exchange rates. 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 U.S. business cycle in the postwar period.

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. But the vast 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 begun to investigate the relationship between term structure and macro variables in reduced-form VAR or semi-structural models that impose no-arbitrage restrictions. Structural general equilibrium models derive the term structure from agents’ optimizing behavior and explain all of the interest rate dynamics by the volatility of macroeconomic variables. Unfortunately, standard models that have become the workhorses of modern macroeconomics have counterfactual implications for the term structure.

First, Donaldson, Johnsen, and Mehra (1990) show that whereas in the U.S. nominal term structure interest rates are procyclical and term spreads countercyclical, the Neoclassical stochastic growth model predicts interest rates to be countercyclical and term spreads procyclical. 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.

Second, Backus, Gregory, and Zin (1989) show that the Lucas (1978) general equilibrium asset pricing model can account for neither sign nor magnitude of average risk premia in forward prices.

3Svensson (1994a,b) and Söderlind and Svensson (1997) discuss monetary policy and 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.
and holding-period returns. Thus the model is unable to explain rejection of the expectations hypothesis, that forward rates are unbiased predictors of future spot rates, which has been extensively documented by empirical studies. The most common interpretation of this result is that it is evidence of the existence of a time-varying risk premium.\(^6\)

In order for policy-makers to extract information about market expectations from the term structure they need to know the sign and magnitude of term premia embedded in interest rates. 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, so that the behavior of the monetary authority affects the business cycle dynamics. Because monetary policy responds systematically to movements in 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 U.S. data and produces procyclical interest rates and countercyclical term spreads. The term spread has predictive power for future economic activity. Most importantly, the model is able to account for rejections of the expectations hypothesis.

Our results show that rejection of the expectation hypothesis hinges on habit-formation preferences and on modeling of the systematic portion of monetary policy. Without habit formation, average term spreads are very close to zero and yield volatilities at all maturities are reduced. As the policy rule changes, average term structure, risk premia volatility and correlations with macro variables will change. This is also true in the absence of policy shocks. In fact, we show that rejection of the expectation hypothesis can be entirely explained by the real-shock volatility. Finally, a large or very volatile inflation risk premium cannot explain rejection of the expectations hypothesis.

The rest of the paper is organized as follows. The rest of this Section runs through the related literature in more detail. Section 2 explains the New Keynesian model that we use, Section 3 discusses our techniques for solving the model numerically, and Section 4 explains the parameterization of the model. Section 5 reports the results related to the term structure. Section 6 discusses the relationship between monetary policy and term structure, and Section 7 concludes. Appendix A derives the inflation rate dynamics in our model.

**Related Literature**

A growing literature investigates the relationship between macroeconomic variables and 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 Federal Reserve policy can be better approximated by assuming that it responds only to information contained in the term structure and not to 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 changes in the federal funds target rate. Ang and Piazzesi (2003) introduce no-arbitrage restrictions in a VAR model of macroeconomic and financial variables.

Research on joint macro-finance model (Hördahl, Tristani, and Vestin, 2003, 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 in order to be able to analyze what portion of yield 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 behavior.

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 real and nominal yields to a monetary policy shock. However, Piazzesi (2005) criticizes their methodology on the grounds that it does not impose the no-arbitrage condition on 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 asset pricing implications of an endowment economy when agents can default on contracts. The results show that this limited commitment model affords a potential solution to the term premium puzzle. Both Dai and Seppälä study only the real term structure. Buraschi and Jiltsov (2003) and Wachter (2004) show that an external habit model à la Campbell and Cochrane (1999) is capable of explaining the term premium puzzle. Duffee (2002) and Dai and Singleton (2002) study the term premium puzzle for nominal yields using reduced form no-arbitrage models. None of these models rely on 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 UK data, in real business cycle, limited participation, and New Keynesian models. Their result is that the New Keynesian model gets closest to matching cyclical behavior for both nominal and real term structures.

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 use 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 household preferences. While the model ensures that the observable macro-variables are consistent with firms’ and households’ optimizing behavior, it also introduces two unobservable state variables, so that the dynamics are 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 exogenous shocks (the authors report parameterizations of AR(1) coefficients on the order of 0.99). Our model is closely related to Hördahl, Tristani and Vestin (2005), but our solution method relies on a third-order approximation. In fact, none of the papers cited derive time-varying risk premia, and therefore cannot address the expectations 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 for 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 for 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 differs, it is not surprising that our conclusions differ. Contrary to their results, we find that in our model monetary policy shocks and inflation risk premium do not explain rejections of the expectations hypothesis.

Our results are related to earlier contributions by Mankiw and Miron (1986) and McCallum (1996). Like us, these authors emphasize the role of central bank smoothing of interest rates in rejection of the expectations hypothesis. Unlike us, they employ exogenously determined risk premia. Gallmeyer, Hollifield, and Zin (2005) revisit McCallum’s explanation in a context similar to that of Bekaert, Cho, and Moreno (2005). While Bekaert, Cho, and Moreno assume that the expectations hypothesis holds, Gallmeyer, Hollifield, and Zin are able to reject the hypothesis by assuming either (i) stochastic volatility of state variables or (ii) state-dependent “price of risk” in the pricing-kernel specification. Our approach is similar to (ii). Consumers in our model face stochastic preference shocks, which leads to a state-dependent price of risk. However, we do not restrict ourselves to affine term structure models.

2. The Model

The theoretical interest rate term structure is derived from a dynamic stochastic general equilibrium model of the business cycle. We adopt a money-in-utility-function model where nominal rigidities allow monetary policy to affect the dynamics of real variables. The modeling framework follows 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 $[0; F]$, $F \rightarrow \infty$ (see Klenow and Kryvtsov, 2004).

While a more sophisticated pricing mechanism could be introduced—such as state-dependent pricing (Dotsey, King and Wolman, 1999), partial indexation to past prices (Christiano, Eichenbaum and Evans, 2005), 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 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) offers a comprehensive
treatment of the New Keynesian framework, and describes in detail the microfoundations of the model.

Each consumer owns shares in all firms, and households are rebated for 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 of 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 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, \( C_j^t(z) \) 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:

\[
C_j^t = \left[ \int_0^1 C_j^t(z)^{\frac{\theta - 1}{\theta}} \, dz \right]^{\frac{\theta}{\theta - 1}}, \quad \theta > 1.
\]  

The representative household chooses \( \left\{ C_j^t, C_{j+i}^t(z), N_j^t, M_j^t, P_t, B_j^t \right\}_i \) where \( N_t \) denotes labor supply, \( M_t \) nominal money balances, \( P_t \) the aggregate price level, and \( B_t \) bond holdings, to maximize

\[
E_t \sum_{i=0}^{\infty} \beta^i \left\{ \frac{\left( C_j^{t+i} - bC_j^{t+i-1} \right)}{1 - \gamma} D_{t+i} - \ell N_j^{t+i+\eta} + \frac{\xi}{1 - \gamma_m} \left( \frac{M_j^{t+i}}{P_{t+i}} \right)^{1-\gamma_m} \right\}
\]

subject to

\[
\int_0^1 C_j^t(z) P_t(z) \, dz = W_t N_j^t + \Pi_j^t - (M_j^t - M_{j-1}^t) - (\bar{p_t} \bar{B}_j^t - B_{j-1}^t) - \tau_j,
\]

and (1). When \( b > 0 \) preferences are characterized by habit formation (Boldrin, Christiano, and Fisher, 2001 and Jermann, 1998). \( D_t \) is an aggregate stochastic preference shock. Each element of the row vector \( \bar{p_t} \) represents the price of an asset with maturity \( k \) that will pay one unit of currency in period \( t + k \). The corresponding element of \( \bar{B}_t \) represents the quantity of such claims purchased by the household. \( B_{j-1}^t \) indicates the value of the household portfolio of claims maturing at time \( t \). \( W_t \) is the nominal wage rate, and \( \tau \) is the lump-sum tax imposed by the government. Finally, the households own the firms and \( \Pi_t \) is the profit from the firms.

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

\[
C_j^t(z) = \left[ \frac{P_t(z)}{P_t} \right]^{-\theta} C_j^t.
\]
Equation (4) is the demand for good \( z \) from household \( j \), where \( \theta \) is the price elasticity of demand. The associated price index \( P_t \) measures the minimum expenditure on differentiated goods that will buy a unit of the consumption index:

\[
P_t = \left[ \int_{1}^{0} P_t(z)^{-\theta} \, dz \right]^{1/\theta}.
\] (5)

Since all households solve an identical optimization problem and face the same aggregate variables, in the following we omit the index \( j \). Equations (4) and (5) yield budget constraint:

\[
C_t = \frac{W_t}{P_t} N_t + \frac{\Pi_t}{P_t} - \frac{M_t - M_{t-1}}{P_t} - \frac{\bar{p}_t \bar{B}_t - B_{t-1}}{P_t} - \frac{\tau_t}{P_t}.
\]

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

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

\[
\frac{W_t}{P_t} = \frac{\ell N^d_t}{MUC_t},
\]

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

where \( MUC \) is the marginal utility of consumption.

**Firms and Price Setting**

The firm producing good \( z \) employs a linear technology:

\[
Y_t(z) = A_t N_t(z),
\] (8)

where \( A_t \) is an aggregate productivity shock. Minimizing the nominal cost \( W_t N_t(z) \) of producing a given amount of output \( \bar{Y} \) yields the labor demand schedule:

\[
MC^N_t(z) MPL_t(z) = W_t,
\] (9)

where \( MC^N \) is the nominal marginal cost, \( MPL \) is the marginal product of labor \( (Y_t(z)/N_t(z)) \). Equation (9) implies that the real marginal cost \( MC_t \) of producing one unit of output is

\[
MC_t(z) MPL_t(z) = \frac{W_t}{P_t}.
\]

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 regardless 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 \( \frac{1}{\theta_p} \). 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:

\[
E_t \sum_{i=0}^{\infty} (\theta_p \beta)^i \frac{MUC_{t+i}}{MUC_t} \left[ \frac{P_t(z) Y_{t+i}(z)}{P_{t+i}} - \frac{MC^N_{t+i}}{P_{t+i}} Y_{t+i}(z) \right]
\] (10)
subject to

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

(11)

In (11), \( Y_{t,t+i}(z) \) 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 (11) into (10), the objective function can be written as

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

(12)

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

\[ P_t(z) E_t \sum_{i=0}^{\infty} (\theta \beta)^i MUC_{t+i} \left[ \frac{P_t(z)}{P_{t+i}} \right]^{-\theta} Y_{t+i} = \mu E_t \sum_{i=0}^{\infty} (\theta \beta)^i MUC_{t+i} MC_{t+i} \left[ \frac{P_t(z)}{P_{t+i}} \right]^{-\theta} Y_{t+i}, \]

(13)

where

\[ \mu = \frac{\theta}{\theta - 1} \]

is the flexible-price level of the markup, which is 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 must 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

\[ G_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{G}_t \]

(14)

\[ H_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \hat{H}_t, \]

(15)

where

\[ \hat{G}_t = \mu MUC_t MC_{t} P_{t}^{\theta-1} Y_{t} + \theta \beta \hat{G}_{t+1} \]

(16)

\[ \hat{H}_t = MUC_t P_{t}^{\theta-1} Y_{t} + \theta \beta \hat{H}_{t+1}. \]

(17)

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 ensures that \( MC \) is equal across firms—whether or not they are updating 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. Thus the dynamics of the consumption-based price index will obey

\[
P_t = \left[ \theta_p P_{t-1}^{1-\theta} + (1 - \theta_p) P_t(z)^{1-\theta} \right]^{1/\theta}. \tag{18}
\]

Because firms charge different prices, aggregation implies \( Y_t \neq A_t N_t \). To see this, notice that (4) 
and (8) imply

\[
Y_t(z) = \left[ \frac{P_t(z)}{P_t} \right]^{-\theta} C_t = A_t N_t(z).
\]

Integrating over \( z \)

\[
A_t \int_0^1 N_t(z) \, dz = \int_0^1 \left[ \frac{P_t(z)}{P_t} \right]^{-\theta} \, dz C_t \\
A_t N_t = C_t \int_0^1 \left[ \frac{P_t(z)}{P_t} \right]^{-\theta} \, dz \\
A_t N_t = C_t s_t,
\]

where \( s_t = \int_0^1 \left[ \frac{P_t(z)}{P_t} \right]^{-\theta} \, dz \). Up to a first order approximation, \( s_t = 1 \). But since we use higher 
order approximations, price dispersion results in the introduction of an additional state variable \( s_t \). 
Its law of motion can be expressed recursively as

\[
s_t = (1 - \theta_p) \left( \frac{\tilde{G}_t}{\tilde{H}_t} \right)^{-\theta} + \theta_p (1 + \pi_t)^\theta s_{t-1}.
\]

Appendix A shows that the inflation rate dynamics is given by

\[
[(1 + \pi_t)]^{1-\theta} = \theta_p + (1 - \theta_p) \left[ \frac{\tilde{G}_t}{\tilde{H}_t} (1 + \pi_t) \right]^{1-\theta} \tag{19}
\]

\[
\tilde{G}_t \equiv \frac{\hat{G}_t}{\hat{P}_t}; \quad \tilde{H}_t \equiv \frac{\hat{H}_t}{\hat{P}_t^{\theta}}.
\]

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

\[
\frac{G}{HP_t} = \frac{G}{H} = \frac{P_t(z)}{P_t} \Rightarrow \frac{P_t(z)}{P_t} = \frac{\mu MC}{(1 - \theta_p \beta \Pi^{\theta-1})} \frac{(1 - \theta_p \beta \Pi^\theta)}{(1 - \theta_p \beta \Pi^\theta)} \tag{20}
\]

Since \( P_t(z) \) is the optimal price chosen by the fraction of firms that can re-optimize at time \( t \), \( P_t(z)/P_t \) is the inverse of what King and Wolman (1996) define as the price wedge. With zero

\footnote{We are grateful to Stephanie Schmitt-Grohe for pointing this out. See also Yun (1996).}
steady state inflation, the steady state average markup is equal to $1/MC$, so 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 from doing so. Combining equation (20) with equation (19) yields the steady state marginal cost and price wedge as a function of $\Pi$:

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

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

**Asset Markets**

The government rebates 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) $\tau^j$ as the amount of tax levied by the government on household $j$, assuming $\tau^j_t = \tau^j_t \forall j, i \in [0, 1]$, at every date $t$ the transfer will be equal to

$$- \int_0^1 \tau^j_t \, dj = -\tau_t \int_0^1 dj = -\tau_t = M_t^s - M_{t-1}^s.$$

Equilibrium in the money market requires that

$$M_t^s = M_t^{d^0} = M_t^d.$$

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.

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 $\vec{B}_t$.

**Monetary Policy**

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

$$\frac{(1 + R_{t,t+1})}{(1 + R^{ss})} = E_t \left( 1 + \frac{\pi_{t+1}}{1 + \pi_{SS}} \right)^{\omega_p} \left( \frac{Y_t}{Y_{SS}} \right)^{\omega_y}.$$

(21)

where $\omega_p, \omega_y \geq 0$ are the feedback coefficients for CPI inflation and output. The monetary authority adjusts the interest rate in response to deviations of target variables from the steady state. In the
steady state, a constant money growth rate rule is followed. Choosing parameters \( \omega_\pi, \omega_y \) allows us to specify alternative monetary policies. When the central bank responds to current rather than expected inflation equation (21) yields 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

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

(22)

where \( \chi \in [0, 1) \) is the degree of smoothing and \( \varepsilon_t^{mp} \) 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 obtained using Dynare++. It is well known that taking a first-order approximation of bond prices will yield no risk premia and that a second-order approximation will yield only constant premia. The reason is simple: second-order approximation involves only squared prediction error terms with constant expectations.

In the first step, we solve our model for six state variables and seven control variables in 13 equations using Dynare++ version 1.3.1. In the second step, we generate 200,000 observations of state and control variables. In the final step, we regress the future marginal rates of substitution—see equations (25) and (26) below—on third-order complete polynomials of state variables using fitted regression to approximate conditional expectations. Our approach is very similar to the Monte Carlo approach employed by Evans and Marshall (1998). The algorithm amounts to taking a third-order approximation of bond prices. With third-order approximation, the current state variables multiply squared prediction error terms, and hence risk premia are time-varying.

4. Model Parameterization

Our specification of preference, technology and policy parameters follows the New Keynesian monetary business cycle literature. Household preferences are modeled within the internal habit-formation framework of Boldrin, Christiano, and Fisher (2001). The habit formation coefficient is parameterized to \( b = 0.8 \), a value that Constantinides (1990) finds can explain the equity premium puzzle. The value of \( \gamma \) is set at 2.5, to provide adequate curvature in the utility function so as to facilitate model generation of risk-premia volatility. The preference parameterization plays a key role in the model’s term-structure properties. Its impact on the results is discussed in detail in the next section. The labor supply elasticity \( (1/\eta) \) is set equal to 2. The parameter \( \ell \) is chosen to set

8Dynare++ is available for free at http://www.cepremap.cnrs.fr/dynare.
9Numerous authors discuss the empirical performance of the New Keynesian framework. For references to estimated and calibrated staggered price-adjustment models, see Christiano, Eichenbaum and Evans (2005), Ireland (2001), Ravenna (2006), Rabanal and Rubio-Ramirez (2005), Woodford (2003).
steady state labor hours at about 30% of available time, a value consistent with postwar data in
the U.S. and in many OECD countries. The quarterly discount factor $\beta$ is parameterized so that
the steady state real interest rate is equal to 1%.

The parameterization of demand elasticity $\theta$ implies a flexible-price equilibrium producers’
markup of $\mu = \theta/(\theta - 1) = 1.1$. While Bernanke and Gertler (2000) use a higher value (1.2), our
assumption of positive steady state inflation implies that the steady state markup is larger than in
the flexible-price equilibrium. The parameterization chosen for the Calvo (1983) pricing adjustment
mechanism implies an average price duration of one year. This value is in line with estimates for
the U.S. over the last forty years obtained from aggregate data (Gali and Gertler, 1999, Rabanal
and Rubio-Ramirez, 2005).

A large number of variants of the monetary policy instrument rule (22) have been estimated
with U.S. data, in both single-equation and simultaneous-equation contexts. The inflation feedback
coefficient $\omega_\pi$ is set at 1.5. This value is substantially lower than the one estimated by Clarida, Gali
and Gertler (2000) for the Volker-Greenspan tenure, but close to the estimate in Rabanal and
Rubio-Ramirez (2005) for the longer 1960–2001 period and averages across different monetary
regimes in post-war U.S. data. The choice of a value for $\omega_\pi$ is more controversial; it depends on
the operational definition of output gap used by the central bank at any given point in time. In
our benchmark parameterization we choose a value of $\omega_\pi = 0$. Estimates of instrument rules across
a large number of OECD countries consistently find very inertial behavior for the policy interest
rate. In the following sections we discuss the impact of alternative assumptions for behavior of the monetary authority on the term structure results. Quarterly steady state inflation
is set equal to the average U.S. value over the period 1994–2004, about 0.75%. This implies an
annualized steady state nominal interest rate of 7%.

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

$$\log Z_t = (1 - \rho_Z) \log Z + \rho_Z \log Z_{t-1} + \epsilon^Z_t, \quad \epsilon^Z_t \sim \text{iid } N(0, \sigma^2_Z);$$

where $Z$ is the steady state value of the variable. The policy shock $\epsilon^{mp}_t$ is a Gaussian i.i.d. stochastic
process. The autocorrelation parameters for technology and preference shocks are equal to $\rho_a = 0.9$
and $\rho_d = 0.95$. The standard deviation of innovations $\epsilon^Z_t$ for technology, preference and policy shock
is set at $\sigma_a = 0.0035$, $\sigma_d = 0.08$, $\sigma_{mp} = 0.003$. The low value for policy shock volatility implies
that the major 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 that estimated by Rabanal and Rubio-Ramirez (2005) with U.S. data. Compared to
the business cycle literature, the technology shock volatility is low. The chosen parameterization is
necessary to allow the model to generate a positive correlation between nominal interest rate and
GDP, since technology shocks produce negative comovements between these variables.

An important concern in the parameterization of 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 in many empirical studies
of the U.S. Table 1 compares the model’s second moments and correlations with output to the U.S.
post-war data sample.\footnote{Standard deviation measured in per cent. The output series is logged and Hodrick-Prescott filtered. U.S. data:}
Table 1: Selected variable volatilities and correlations. Sample: 1952–2006.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Standard Deviation</th>
<th>Correlation with Output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>U.S. Data</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>2.01</td>
<td>1.59</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>3.49</td>
<td>3.00</td>
</tr>
<tr>
<td>$R_t$</td>
<td>1.84</td>
<td>2.82</td>
</tr>
<tr>
<td>$r_t$</td>
<td>3.84</td>
<td>2.32</td>
</tr>
</tbody>
</table>

and U.S. Federal Reserve operating procedures, and includes the 1970s inflationary episode. On the other hand, the sample can be considered representative of the variety of shocks that drove the U.S. business cycle.

The match with empirical correlations is quite good. We obtain this result although the model volatilities for output, real interest rate and inflation turn out to be larger than in the data. The empirical fit of the model can be improved with a number of modifications, including sticky wages, hybrid backward and forward-looking price setting specifications, an autocorrelated exogenous process driving the dynamics of the inflation target, and cost-push shocks. Our effort focused on investigating whether a minimal set of modifications to the Neoclassical growth model can explain the behavior of risk premia over the business cycle. In the following sections we compare the parameterized model predictions to the U.S. post-war nominal term structure. Most of the stylized facts we investigate are consistent across monetary policy regime changes during this period, and so we rely on a single parameterization for the policy rule. The sensitivity analysis shows that most results are surprisingly robust to alternative parameterization assumptions.

5. Term Structure of Interest Rates

Real and Nominal Term Structures

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

$$q_{t+1} \equiv \beta \frac{MUC_{t+1}^{\text{MUC}}}{MUC_t},$$

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

$$Q_{t+1} \equiv \beta \frac{MUC_{t+1}^{\text{MUC}} P_t}{MUC_t P_{t+1}}.$$  \hfill (24)

$Y_t$ is real GDP, $\pi_t$ is CPI inflation, $R_t$ is 3-month T-bill rate, $r_t$ is ex-post short term real interest rate. All rates are on annual basis. Quarterly data sample is 1952:1–2006:1. We chose to use the period following the Treasury-Federal Reserve Accord of 1951 in order to avoid having to contend with the constraint on interest rate movements imposed by the Federal Reserve’s “par pegging” of Government securities prices. Real GDP is from the Bureau of Economic Analysis and the rest of the data are from the St. Louis Federal Reserve Bank FRED II database.
The price of an $n$-period zero-coupon real bond is given by
\[ p^b_{n,t} = \mathbb{E}_t \left[ \prod_{j=1}^{n} q_{t+j} \right] = \mathbb{E}_t [q_{t+1} p^b_{n-1,t+1}], \tag{25} \]
and similarly the price of an $n$-period zero-coupon nominal bond is given by
\[ p^B_{n,t} = \mathbb{E}_t \left[ \prod_{j=1}^{n} Q_{t+j} \right] = \mathbb{E}_t [Q_{t+1} p^B_{n-1,t+1}], \tag{26} \]
The bond prices are invariant with respect to time; hence equations (25) and (26) give a recursive formula for pricing zero-coupon real and nominal bonds of any maturity.

Forward prices for real bonds are defined by
\[ p^f_{n,t} = \frac{p^b_{n+1,t}}{p^b_{n,t}}. \]
Prices are related to rates (or yields) by\(^{11}\)
\[ f_{n,t} = -\log(p^f_{n,t}) \quad \text{and} \quad r_{n,t} = -(1/n) \log(p^b_{n,t}). \tag{27} \]
Table 2 presents means, standard deviations, and correlations with output for selected maturities in the term structure in the model, and for U.S. nominal data as estimated by Global Financial Data from the first quarter of 1952 to the first quarter of 2006. Output is filtered using the Hodrick-Prescott (1980) filter with a smoothing parameter of 1600 in both the model and the data.

The average term structure is upward-sloping in both the model and the data. Means match quite well: the model produces nominal yields varying from 5.01% to 6.60% for three months to 20 years maturity, while the corresponding U.S. yields varied from 5.03% to 6.60%.

Table 2 shows that the model generates procyclical nominal interest rates and countercyclical term spreads. This matches the positive correlation between yields and the cyclical component of output observed in U.S. data at maturities up to one year. The nominal term spreads are countercyclical in both U.S. data and the model at all maturities.

These results show that the New Keynesian model can explain the term spread puzzle that emerges in the Neoclassical stochastic growth model. Donaldson, Johnsen, and Mehra (1990) show that in a stochastic growth model with full depreciation the term structure of (ex-ante) real interest rates is at odds with empirical evidence: it rises at the peak of the business cycle and falls at the trough. In addition, at the peak of the cycle the term structure lies uniformly below the term structure at the trough. The economic intuition for the behavior of interest rates is straightforward, working through the link between marginal utility of consumption, expected consumption growth and interest rates. At the cycle peak aggregate and individual consumption are expected to be, on

\(^{11}\)Nominal prices and rates are obtained in a similar manner.
Table 2: Main term structure statistics. Data: 1952–2006. (N/A missing due to shortage of data.)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Correlation with Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{1,t} ) (model)</td>
<td>5.01198</td>
<td>1.84042</td>
<td>0.16560</td>
</tr>
<tr>
<td>( R_{4,t} ) (model)</td>
<td>6.22698</td>
<td>1.40906</td>
<td>0.25269</td>
</tr>
<tr>
<td>( R_{40,t} ) (model)</td>
<td>6.59399</td>
<td>0.56985</td>
<td>0.40510</td>
</tr>
<tr>
<td>( R_{80,t} ) (model)</td>
<td>6.59911</td>
<td>0.33910</td>
<td>0.40896</td>
</tr>
<tr>
<td>( R_{120,t} ) (model)</td>
<td>6.59905</td>
<td>0.22477</td>
<td>0.40717</td>
</tr>
<tr>
<td>( R_{1,t} ) (data)</td>
<td>5.03359</td>
<td>2.81717</td>
<td>0.17491</td>
</tr>
<tr>
<td>( R_{4,t} ) (data)</td>
<td>5.60977</td>
<td>3.05969</td>
<td>0.14690</td>
</tr>
<tr>
<td>( R_{40,t} ) (data)</td>
<td>6.42456</td>
<td>2.76373</td>
<td>-0.01473</td>
</tr>
<tr>
<td>( R_{80,t} ) (data)</td>
<td>6.60014</td>
<td>2.71775</td>
<td>-0.04062</td>
</tr>
<tr>
<td>( R_{120,t} ) (data)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>( R_{40,t} - R_{1,t} ) (model)</td>
<td>1.58202</td>
<td>1.50547</td>
<td>-0.04911</td>
</tr>
<tr>
<td>( R_{80,t} - R_{1,t} ) (model)</td>
<td>1.58713</td>
<td>1.63395</td>
<td>-0.10166</td>
</tr>
<tr>
<td>( R_{120,t} - R_{1,t} ) (model)</td>
<td>1.58707</td>
<td>1.70057</td>
<td>-0.12540</td>
</tr>
<tr>
<td>( R_{40,t} - R_{4,t} ) (model)</td>
<td>0.36702</td>
<td>0.98454</td>
<td>-0.12717</td>
</tr>
<tr>
<td>( R_{80,t} - R_{4,t} ) (model)</td>
<td>0.37213</td>
<td>1.14983</td>
<td>-0.18905</td>
</tr>
<tr>
<td>( R_{120,t} - R_{4,t} ) (model)</td>
<td>0.37207</td>
<td>1.23425</td>
<td>-0.21433</td>
</tr>
<tr>
<td>( R_{40,t} - R_{1,t} ) (data)</td>
<td>1.39097</td>
<td>1.13509</td>
<td>-0.46998</td>
</tr>
<tr>
<td>( R_{80,t} - R_{1,t} ) (data)</td>
<td>1.56654</td>
<td>1.32127</td>
<td>-0.45650</td>
</tr>
<tr>
<td>( R_{120,t} - R_{1,t} ) (data)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>( R_{40,t} - R_{4,t} ) (data)</td>
<td>0.81479</td>
<td>1.01890</td>
<td>-0.48109</td>
</tr>
<tr>
<td>( R_{80,t} - R_{4,t} ) (data)</td>
<td>0.99037</td>
<td>1.24971</td>
<td>-0.44800</td>
</tr>
<tr>
<td>( R_{120,t} - R_{4,t} ) (data)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Table 3: Selected term structure statistics in selected time periods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[R_{4,t}]$</td>
<td>5.60977</td>
<td>6.15297</td>
<td>6.47552</td>
<td>4.87137</td>
</tr>
<tr>
<td>$E[R_{40,t}]$</td>
<td>6.42456</td>
<td>6.99016</td>
<td>7.65848</td>
<td>6.15603</td>
</tr>
<tr>
<td>$E[R_{40,t} - R_{4,t}]$</td>
<td>0.81479</td>
<td>0.83719</td>
<td>1.18295</td>
<td>1.28466</td>
</tr>
<tr>
<td>std($R_{4,t}$)</td>
<td>3.05969</td>
<td>2.96079</td>
<td>3.44313</td>
<td>2.08718</td>
</tr>
<tr>
<td>std($R_{40,t}$)</td>
<td>2.76373</td>
<td>2.59025</td>
<td>2.92400</td>
<td>1.56503</td>
</tr>
<tr>
<td>std($R_{40,t} - R_{4,t}$)</td>
<td>1.01890</td>
<td>1.07935</td>
<td>1.12390</td>
<td>1.06067</td>
</tr>
<tr>
<td>corr($R_{4,t}, Y_t$)</td>
<td>0.14690</td>
<td>0.18143</td>
<td>0.07312</td>
<td>0.57920</td>
</tr>
<tr>
<td>corr($R_{40,t}, Y_t$)</td>
<td>−0.01473</td>
<td>0.00426</td>
<td>−0.06326</td>
<td>0.29255</td>
</tr>
<tr>
<td>corr($R_{40,t} - R_{4,t}, Y_t$)</td>
<td>−0.48109</td>
<td>−0.48747</td>
<td>−0.38859</td>
<td>−0.70809</td>
</tr>
</tbody>
</table>

average, lower in the future, and so the agents will want to save more. In equilibrium interest rates will therefore be lower. At the cycle trough aggregate and individual consumption are expected to be higher in the future, and so agents’ incentive to save is reduced, which raises interest rates. Similar intuition explains the model’s procyclical term spread.

Donaldson, Johnsen, and Mehra (1990) results are not general (Labadie, 1994). Introducing exogenous shocks, in addition to the stochastic process for total factor productivity, can generate procyclical interest rates (Walsh, 2003 gives an example in a money-in-utility framework). Yet even the models examined by King and Watson (1996), despite generating procyclical interest rates, cannot account for the empirical fact that high real or nominal interest rates predict a low level of economic activity two to four quarters in the future.

The New Keynesian model is not able to reproduce two important features of post-war U.S. data: constant volatility and decreasing correlation with output of nominal yields as the maturity increases. The model produces a downward-sloping term structure of volatilities and strong positive correlation between yields and (the cyclical component of) output at all maturities. The decreasing volatility of nominal rates is a counterfactual implication already identified by Den Haan (1995) for flexible-price models of the business cycle.


In both the 1980–2006 and 1988–2006 samples, the term structure of volatilities is clearly downward-sloping. In the 1988–2006 sample, the correlation between yield and output is strongly positive. The upward-sloping mean term structure and countercyclical term spreads are robust features across subsamples. On the other hand, the level of interest rates depends on how much relatively high interest rates in the early 1980’s weigh on the data. The average one-year rate was 160 basis points higher in 1980–2006 compared to 1988–2006.

The high volatility of long rates observed in the full sample is associated in the data with periods of volatile inflation. What the New Keynesian model is missing is a mechanism to generate

---

12 Seppälä (2000) documents that UK nominal and real term structure of volatilities is downward-sloping.
persistent changes in the inflation rate target, implying persistent changes in the expected future policy rate. As it is, the model necessarily generates strongly mean-reverting interest rates. This also means that the impact of policy shocks on long rates is much smaller than on short rates, and that the correlation between policy rate and long yields decreases with time. Adding to the model a very persistent shock to the policy inflation target, as in Rudebusch and Wu (2004) or Hördahl, Tristani, and Vestin (2005), would contribute to an increase in the volatility of long rates. Similarly, a varying long-term inflation target would lower the correlation between long maturity yields and output. Additionally, the output-nominal yields correlation would be lowered in a more realistic model including time-to-build constraints, convex costs of capital adjustment, and lags in the impact of monetary policy on real variables.

Sensitivity Analysis

Table 4 reports on how New Keynesian model term structure statistics vary depending on the parameterization. The most striking feature of Table 4 is that the upward-sloping average term structure is a very robust feature in a New Keynesian model. All parameterizations in the table share this feature. In contrast, the flexible price models studied by Den Haan (1995) generate downward-sloping average term structure in a real production economy and a flat average term structure in a monetary production economy.

Procyclical interest rates and countercyclical term spreads obtain under many, though not all, parameterizations of the New Keynesian model. The preferences specification plays a key role. In a model without habit formation, the term structure is only mildly upward-sloping. It is not

### Table 4: Sensitivity of term structure statistics to different parameter values. BM = benchmark.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>E[R_{4,t}]</th>
<th>E[R_{40,t}]</th>
<th>\sigma(R_{4,t})</th>
<th>\sigma(R_{40,t})</th>
<th>\rho(R_{4,t}, Y_t)</th>
<th>\rho(R_{40,t}, Y_t)</th>
<th>\rho(R_{40,t} - R_{4,t}, Y_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM</td>
<td>6.22698</td>
<td>6.59399</td>
<td>1.40906</td>
<td>0.56985</td>
<td>0.25269</td>
<td>0.40510</td>
<td>-0.12717</td>
</tr>
<tr>
<td>b = 0</td>
<td>6.56884</td>
<td>6.59656</td>
<td>1.33915</td>
<td>0.49197</td>
<td>0.12565</td>
<td>0.40615</td>
<td>0.03317</td>
</tr>
<tr>
<td>\gamma = 1.5</td>
<td>6.20214</td>
<td>6.55681</td>
<td>1.72850</td>
<td>0.62184</td>
<td>0.09803</td>
<td>0.38814</td>
<td>0.05169</td>
</tr>
<tr>
<td>\chi = 0.7</td>
<td>6.71426</td>
<td>6.90667</td>
<td>1.86154</td>
<td>0.93034</td>
<td>0.30888</td>
<td>0.41399</td>
<td>-0.17036</td>
</tr>
<tr>
<td>\pi_{ss} = 1.0</td>
<td>3.22414</td>
<td>3.56710</td>
<td>1.30479</td>
<td>0.54342</td>
<td>0.25729</td>
<td>0.40539</td>
<td>-0.12830</td>
</tr>
<tr>
<td>\pi_{ss} = 1.01</td>
<td>7.22419</td>
<td>7.59962</td>
<td>1.46793</td>
<td>0.59306</td>
<td>0.25338</td>
<td>0.40638</td>
<td>-0.12748</td>
</tr>
<tr>
<td>\omega_{y} = 0.1</td>
<td>5.08460</td>
<td>5.28123</td>
<td>3.16343</td>
<td>1.32989</td>
<td>-0.43732</td>
<td>-0.43132</td>
<td>0.42163</td>
</tr>
<tr>
<td>\omega_{p} = 3.0</td>
<td>6.67072</td>
<td>6.91731</td>
<td>1.12834</td>
<td>0.41932</td>
<td>0.16844</td>
<td>0.38960</td>
<td>-0.03112</td>
</tr>
<tr>
<td>\omega_{\pi} = 1.2</td>
<td>5.52783</td>
<td>5.94096</td>
<td>1.59132</td>
<td>0.68219</td>
<td>0.30698</td>
<td>0.41359</td>
<td>-0.19912</td>
</tr>
<tr>
<td>\theta_{p} = 0.5</td>
<td>6.12833</td>
<td>6.38865</td>
<td>1.17923</td>
<td>0.55357</td>
<td>0.36111</td>
<td>0.41951</td>
<td>-0.27247</td>
</tr>
<tr>
<td>\sigma_{d} = 0, \sigma_{a} = 0.01</td>
<td>6.28834</td>
<td>6.59019</td>
<td>1.13221</td>
<td>0.23745</td>
<td>-0.57675</td>
<td>-0.44570</td>
<td>0.58735</td>
</tr>
<tr>
<td>\sigma_{a} = 0</td>
<td>6.22289</td>
<td>6.59189</td>
<td>1.38130</td>
<td>0.59251</td>
<td>0.25908</td>
<td>0.41618</td>
<td>-0.11824</td>
</tr>
<tr>
<td>\sigma_{mp} = 0</td>
<td>6.89120</td>
<td>6.97486</td>
<td>1.05324</td>
<td>0.56802</td>
<td>0.37047</td>
<td>0.42395</td>
<td>-0.28082</td>
</tr>
<tr>
<td>\sigma_{mp} = 0.006</td>
<td>4.32110</td>
<td>5.57434</td>
<td>2.03739</td>
<td>0.60145</td>
<td>0.12236</td>
<td>0.34697</td>
<td>-0.02520</td>
</tr>
<tr>
<td>\sigma_{mp} = 0, \chi = 0.95</td>
<td>6.76933</td>
<td>6.94859</td>
<td>0.78090</td>
<td>0.67981</td>
<td>0.11718</td>
<td>0.17388</td>
<td>0.13693</td>
</tr>
</tbody>
</table>
surprising that—as we show below—habit-formation preferences are a necessary condition for the model to reject the expectations hypothesis. A smaller value of $\gamma$ reduces the curvature of the utility function and will also affect the market price of risk. Both these deviations from baseline preference parameterization imply that the model loses the ability to generate countercyclical term spreads.

Without stochastic shocks to marginal utility of consumption the model counterfactually predicts countercyclical interest rates and procyclical term spreads. A similar implication is obtained in a model where the monetary authority reacts to deviations of output from steady state, even if with a low feedback coefficient $\omega_y = 0.1$. Different values of steady state inflation have instead very little impact on the results, except for shifting the whole nominal term structure.

The term structure of volatility is highly sensitive to the stochastic process for monetary policy innovation. Table 4 shows that as the volatility of this shock is decreased, the term structure of volatility becomes flatter. In fact the last row in the table shows that one shortcoming of the New Keynesian model—the downward-sloping term structure of volatilities—can be overcome by a combination of highly persistent shocks and a high degree of interest rate smoothing (increasing the volatility of long rates), and low volatility of monetary policy shocks (lowering the volatility of short rates).

Term Structure Predictions of Future Economic Activity

Every major recession in the U.S. since the early 1950s has been predicted by an inverted (downward-sloping) term structure of interest rates.\textsuperscript{13} As discussed above, the stochastic growth model with flexible prices produces the opposite prediction. The New Keynesian model can explain the term spread forecasting power found in empirical studies.

We compare the New Keynesian model’s predictions with one well-known paper on the relationship between term spread and future consumption growth. Estrella and Hardouvelis (1991) use the term spread to predict future changes in log consumption growth one to four years ahead. The estimation equation uses quarterly observations of U.S. real consumption of non-durables and services over the 1960:1-2006:1 sample regressed on the yield spread between 10-year government bond and 3-month Treasury bill.\textsuperscript{14} Table 5 presents the regression results of equation

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

for the data and the benchmark model with 200,000 observations. The standard errors are White (1980) heteroskedasticity consistent standard errors. An upward-sloping term structure clearly predicts expansions in both our model and the data, and a downward-sloping term structure clearly predicts recessions, in both the model and the data. Moreover, $\beta_1$ decreases with the forecast horizon in both the model and the data.

\textsuperscript{13}A large literature has examined the predictive power of the term structure of interest rates for future interest rates, consumption growth, and other measures of future economic activity. Fama and Bliss (1987) use forward spread to predict future changes in one-year interest rates one to four years ahead. Estrella and Hardouvelis (1991) show that the term spread has predictive power for future changes in log-consumption growth up to four years ahead in US data from 1955 to 1988. Dotsey (1998) points out that many studies have found the term spread to contain significant information for predicting economic activity also in the most recent U.S. data. Estrella (2005) provides an exhaustive list of references.

\textsuperscript{14}Consumption data are from the St. Louis Federal Reserve Bank FRED II database. Yield data are from Global Financial Data.
Table 5: Term spread forecasts of future consumption growth $n$ years ahead.

<table>
<thead>
<tr>
<th>Regression</th>
<th>$\beta_0$</th>
<th>se($\beta_0$)</th>
<th>$\beta_1$</th>
<th>se($\beta_1$)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark ($n = 1$)</td>
<td>-0.2265</td>
<td>0.0071</td>
<td>0.6410</td>
<td>0.0079</td>
<td>0.0337</td>
</tr>
<tr>
<td>Benchmark ($n = 2$)</td>
<td>-0.1865</td>
<td>0.0060</td>
<td>0.5278</td>
<td>0.0066</td>
<td>0.0323</td>
</tr>
<tr>
<td>Benchmark ($n = 3$)</td>
<td>-0.1707</td>
<td>0.0051</td>
<td>0.4834</td>
<td>0.0057</td>
<td>0.0367</td>
</tr>
<tr>
<td>Benchmark ($n = 4$)</td>
<td>-0.1590</td>
<td>0.0044</td>
<td>0.4506</td>
<td>0.0049</td>
<td>0.0418</td>
</tr>
<tr>
<td>Data ($n = 1$)</td>
<td>2.6826</td>
<td>0.1694</td>
<td>0.5063</td>
<td>0.1131</td>
<td>0.1193</td>
</tr>
<tr>
<td>Data ($n = 2$)</td>
<td>2.8213</td>
<td>0.1471</td>
<td>0.3499</td>
<td>0.0974</td>
<td>0.0824</td>
</tr>
<tr>
<td>Data ($n = 3$)</td>
<td>2.9405</td>
<td>0.1248</td>
<td>0.1622</td>
<td>0.0852</td>
<td>0.0233</td>
</tr>
<tr>
<td>Data ($n = 4$)</td>
<td>3.0832</td>
<td>0.0979</td>
<td>0.0084</td>
<td>0.0675</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Expectations Hypothesis

To define the risk premium, write (26) for a two-period bond using the conditional expectation operator and its properties:

$$ p_{t,2}^B = E_t[Q_{t+1}p_{1,t+1}^B] $$
$$ = E_t[Q_{t+1}]E_t[p_{1,t+1}^B] + \text{cov}_t[Q_{t+1}, p_{1,t+1}^B] $$
$$ = p_{1,t}^B E_t[p_{1,t+1}^B] + \text{cov}_t[Q_{t+1}, p_{1,t+1}^B], $$

which implies that

$$ p_{1,t}^F = \frac{p_{t,2}^B}{p_{1,t}^B} = E_t[p_{1,t+1}^B] + \text{cov}_t\left[Q_{t+1}, \frac{p_{t+1}^B}{p_{1,t}^B}\right]. \quad (28) $$

Since the conditional covariance term is zero for risk-neutral investors, we call it the risk premium for the one-period nominal forward contract, $\text{nrp}_{1,t}$, given by

$$ \text{nrp}_{1,t} \equiv - \text{cov}_t\left[Q_{t+1}, \frac{p_{t+1}^B}{p_{1,t}^B}\right] = E_t[p_{1,t+1}^B] - p_{1,t}^F, $$

and similarly $\text{nrp}_{n,t}$ is the risk premium for the $n$-period forward contract:

$$ \text{nrp}_{n,t} \equiv - \text{cov}_t\left[\prod_{j=1}^{n} Q_{t+j}, \frac{p_{t+n,1}^B}{p_{1,t}^B}\right] = E_t[p_{t+n,1}^B] - p_{n,t}^F. $$

If the risk premium is zero, we obtain the oldest and simplest theory about the information content of the term structure—the so-called (pure) expectations hypothesis. According to the pure expectations hypothesis, forward rates are unbiased predictors of future spot rates. It is also common to modify the theory so that a constant risk-premium is allowed. This prediction has come to be known in the literature as the expectations hypothesis. By and large the empirical literature rejects both versions of the expectations hypothesis.
Table 6: Number of rejects for each regression in benchmark model for nominal term structure.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{1,t+1}^F - p_{1,t}^F$</th>
<th>$p_{1,t+1}^B - p_{1,t}^B - n rp_{1,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t$</td>
<td>$p_{1,t}^F - p_{1,t}^B$</td>
<td>$p_{1,t}^F - p_{1,t}^B$</td>
</tr>
<tr>
<td>$Wald(\beta_0 = \beta_1 = 0)$</td>
<td>1000</td>
<td>67</td>
</tr>
<tr>
<td>$Wald(\beta_1 = 0)$</td>
<td>989</td>
<td>63</td>
</tr>
<tr>
<td>$Wald(\beta_1 = -1)$</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

We use as a benchmark the Backus, Gregory, and Zin (1989) test equation. These authors tested the expectations hypothesis in the complete markets endowment economy. Equation (28) and the assumption of a constant risk premium implies that

$$E_t[p_{1,t+1}^B] - p_{1,t}^F = \beta_0,$$

Therefore the regression

$$p_{1,t+1}^B - p_{1,t}^F = \beta_0 + \beta_1(p_{1,t}^F - p_{1,t}^B)$$  \hspace{1cm} (29)

should yield $\beta_1 = 0$. Backus, Gregory, and Zin (1989) generated 1000 samples of 200-period paths for the endogenous variables and used a Wald test with White (1980) standard errors to check whether $\beta_1 = 0$ at the 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 $\beta_1$ except $-1$, the forward premium is still useful in forecasting changes in spot prices. The hypothesis $\beta_1 = -1$ was rejected every time.

Table 6 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 7 presents the same tests when the habit-formation parameter $b = 0$. Table 8 displays the same tests for real term structure, and table 9 displays the test for real term structure when $b = 0$. Only the benchmark model is consistent with empirical evidence on the expectations hypothesis. When the risk premium is subtracted from $p_{1,t+1}^B - p_{1,t}^F$, the hypothesis that $\beta_1$ is equal to zero can be rejected in less than 10% of the samples. Comparing the tables, is clear that habit-formation is a necessary condition for rejection of the expectations hypothesis. However, since the hypothesis is rejected for the real term structure only about 40% of the time, it is the case that monetary policy—which directly affects nominal rates—plays also an important role. This issue and other sensitivity analysis are studied in more detail in Section 6.

Table 10 presents estimates of (29) over U.S. nominal term structure data and 200,000 model-generated data. The data include quarterly observations of 3 and 6-month U.S. Treasury bills

---

\[15\] Both versions of the expectations hypothesis are only approximately correct. To see this, assume that agents are risk-neutral: $\gamma = 0$. Equation (28) then reduces to $p_{1,t}^F = E_t[p_{1,t+1}^F]$ and from (27) we obtain $\exp^{-f_{1,t}} = E_t[\exp^{-r_{1,t+1}}]$. Jensen’s inequality implies that $f_{1,t} < E_t[r_{1,t+1}]$. The difference between left and right hand sides of this equation is known as \textit{convexity bias} and it varies with $E_t[r_{1,t+1}]$ and $\text{var}_t[r_{1,t+1}]$. To avoid this issue, Backus, Gregory, and Zin (1989) test the expectations hypothesis using bond prices rather than bond yields.
Table 7: Number of rejects for each regression in nominal term structure when $b = 0$.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{1,t+1}^B - p_{1,t}^B$</th>
<th>$p_{1,t+1}^F - p_{1,t}^F - nrp_{1,t}$</th>
<th>$x_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald($\beta_0 = \beta_1 = 0$)</td>
<td>172</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = 0$)</td>
<td>181</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = -1$)</td>
<td>996</td>
<td>999</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Number of rejects for each regression in benchmark model for real term structure.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{1,t+1}^B - p_{1,t}^B$</th>
<th>$p_{1,t+1}^F - p_{1,t}^F - nrp_{1,t}$</th>
<th>$x_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald($\beta_0 = \beta_1 = 0$)</td>
<td>1000</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = 0$)</td>
<td>422</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = -1$)</td>
<td>1000</td>
<td>1000</td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Number of rejects for each regression in real term structure when $b = 0$.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{1,t+1}^B - p_{1,t}^B$</th>
<th>$p_{1,t+1}^F - p_{1,t}^F - nrp_{1,t}$</th>
<th>$x_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald($\beta_0 = \beta_1 = 0$)</td>
<td>75</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = 0$)</td>
<td>69</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Wald($\beta_1 = -1$)</td>
<td>1000</td>
<td>1000</td>
<td></td>
</tr>
</tbody>
</table>
from 1960:1 to 2006:1. In Table 10, Wald rows refer to the marginal significance level of the corresponding Wald test. The regression coefficient $\beta_1$ implied by the New Keynesian model is remarkably close to the value obtained with U.S. data.

6. Monetary Policy and Inflation Risk Premium

Recall the definitions of one-period zero-coupon nominal bond (26) and nominal stochastic discount factor (24)

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

To define the inflation risk premium, write (30) using the definition of conditional covariance and the definition of real bond price (25):

$$p_t^B = E_t \left[ \beta \frac{MUC_{t+1}P_t}{MUC_tP_{t+1}} \right]$$

$$= E_t \left[ \beta \frac{MUC_{t+1}}{MUC_t} \right] E_t \left[ \frac{P_t}{P_{t+1}} \right] + \text{cov}_t \left[ \beta \frac{MUC_{t+1}}{MUC_t}, \frac{P_t}{P_{t+1}} \right]$$

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

Since the conditional covariance term is zero for risk-neutral investors and when the 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:

$$irp_{n,t} \equiv \text{cov}_t \left[ \prod_{j=1}^{n} q_{t+j}, \frac{P_t}{P_{t+n}} \right] = p_t^B - p_t^b E_t \left[ \frac{P_t}{P_{t+n}} \right].$$
Table 11: Main inflation risk premia statistics, benchmark case.

<table>
<thead>
<tr>
<th>IRP (n)</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Correlation with Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRP (n = 1)</td>
<td>0.03125</td>
<td>0.00792</td>
<td>−0.22185</td>
</tr>
<tr>
<td>IRP (n = 2)</td>
<td>0.05827</td>
<td>0.01088</td>
<td>−0.23463</td>
</tr>
<tr>
<td>IRP (n = 4)</td>
<td>0.09122</td>
<td>0.01148</td>
<td>−0.24944</td>
</tr>
<tr>
<td>IRP (n = 8)</td>
<td>0.10012</td>
<td>0.01175</td>
<td>−0.19349</td>
</tr>
<tr>
<td>IRP (n = 12)</td>
<td>0.07833</td>
<td>0.01620</td>
<td>−0.11236</td>
</tr>
<tr>
<td>IRP (n = 16)</td>
<td>0.05105</td>
<td>0.01802</td>
<td>−0.09234</td>
</tr>
<tr>
<td>IRP (n = 20)</td>
<td>0.02706</td>
<td>0.01955</td>
<td>−0.06929</td>
</tr>
</tbody>
</table>

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

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

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

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

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

Table 11 presents the main statistics for inflation risk premia in the benchmark case. The inflation risk premium is always positive, and follows a hump-shaped term structure consistent with the mean-reverting inflation.

Buraschi and Jiltsov (2005) argue that time-variation of the inflation risk premium is an important explanatory variable for deviations from the expectations hypothesis. We address this question by shutting down the monetary policy shocks, i.e., by setting \( \sigma_{mp} = 0 \). Table 12 presents inflation risk premia statistics for \( \sigma_{mp} = 0 \). Not surprisingly, premia are considerably smaller and less volatile without monetary policy shocks. If inflation risk premia played an important role in rejection of the expectations hypothesis, we would expect the economy with \( \sigma_{mp} = 0 \) to produce fewer rejections. However, Table 13 shows that the expectations hypothesis is actually rejected more often without monetary policy shocks.

That is, unlike in Buraschi and Jiltsov (2005), large and volatile inflation risk premia do not explain rejections of the expectations hypothesis in the New Keynesian model. Why is there a negative relationship between the size and volatility of inflation risk premia and the number of times the expectations hypothesis can be rejected in our model?

Recall the regression equation (29)

\[ p_{1,t+1}^B - p_{1,t}^F = \beta_0 + \beta_1(p_{1,t}^F - p_{1,t}^B), \]

and the definition of risk premium

\[ nrp_{1,t} \equiv - \text{cov}_{t} \left[ Q_{t+1}, \frac{p_{1,t+1}^B}{p_{1,t}^B} \right] = E_t[p_{1,t+1}^B] - p_{1,t}^F. \]

23
Table 12: Main inflation risk premia statistics, $\sigma_{mp} = 0$.

<table>
<thead>
<tr>
<th>IRP ($n = 1$)</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Correlation with Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.00274</td>
<td>0.00031</td>
<td>-0.02555</td>
<td></td>
</tr>
<tr>
<td>IRP ($n = 2$)</td>
<td>-0.00609</td>
<td>0.00061</td>
<td>0.02587</td>
</tr>
<tr>
<td>IRP ($n = 4$)</td>
<td>-0.01437</td>
<td>0.00129</td>
<td>0.10234</td>
</tr>
<tr>
<td>IRP ($n = 8$)</td>
<td>-0.03447</td>
<td>0.00264</td>
<td>0.16770</td>
</tr>
<tr>
<td>IRP ($n = 12$)</td>
<td>-0.05495</td>
<td>0.00402</td>
<td>0.19605</td>
</tr>
<tr>
<td>IRP ($n = 16$)</td>
<td>-0.07223</td>
<td>0.00561</td>
<td>0.19235</td>
</tr>
<tr>
<td>IRP ($n = 20$)</td>
<td>-0.08531</td>
<td>0.00713</td>
<td>0.23103</td>
</tr>
</tbody>
</table>

Table 13: Number of rejects for each regression in benchmark model for nominal term structure when $\sigma_{mp} = 0$.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{t,t+1}^B - p_{t,t}^F$</th>
<th>$p_{t,t+1}^B - p_{t,t}^F - nrp_{t,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t$</td>
<td>$p_{t,t}^F - p_{t,t}^B$</td>
<td>$p_{t,t}^F - p_{t,t}^B$</td>
</tr>
<tr>
<td>Wald($\beta_0 = \beta_1 = 0$)</td>
<td>1000</td>
<td>86</td>
</tr>
<tr>
<td>Wald($\beta_1 = 0$)</td>
<td>993</td>
<td>91</td>
</tr>
<tr>
<td>Wald($\beta_1 = -1$)</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

Substituting the second equation into the first and letting $\psi_t$ denote the one-period nominal risk premium yields

$$p_{t,t+1}^B - E_t[p_{t,t+1}^B] + \psi_t = \beta_0 + \beta_1(E_t[p_{t,t+1}^B] - \psi_t - p_{t,t}^B).$$

$E_t[p_{t,t+1}^B]$ can be written as $p_{t,t+1}^B - \epsilon_{t+1}$, where the prediction error term, $\epsilon_{t+1}$, is orthogonal to the information available at time $t$. Hence the estimate of $\beta_1$ in the expectations hypothesis regression (29) converges to

$$\text{plim} \hat{\beta}_1 = \frac{\rho \sigma_p \sigma_{\psi} - \sigma_{\psi}^2}{\sigma_p^2 + \sigma_{\psi}^2 - 2 \rho \sigma_p \sigma_{\psi}}, \quad (31)$$

where $\rho$ denotes $\text{corr}(\psi_t, E_t[\Delta p_{t+1}^B])$, $\sigma_p$ denotes $\text{std}(E_t[\Delta p_{t+1}^B])$, $\sigma_{\psi}$ denotes $\text{std}(\psi_t)$, and $E_t[\Delta p_{t+1}^B]$ is the expected change in the one-period bond price conditional on time $t$ information.\(^\text{16}\)

It is illuminating to study equation (31) as a function of $\text{std}(\psi_t)$ and $\text{std}(E_t[\Delta p_{t+1}^B])$. If $\psi_t$ 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., forward prices are not useful in predicting future bond prices. As the volatilities $\sigma_p$ and $\sigma_{\psi}$ diverge to infinity the opposite result will hold. 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 $\sigma_p$ and $\sigma_{\psi}$ depends on the sign of $\rho$. Figures 1 and 2 show the behavior of $\text{plim} \hat{\beta}_1$ as a function of $\sigma_p$ and $\sigma_{\psi}$ when $\rho > 0$ and $\rho < 0$, respectively.

\(^{16}\)We compared the $\text{plim} \hat{\beta}_1$ obtained in equation (31) with the estimate of $\beta_1$ obtained using model-generated data. The estimator converged to the theoretical limit across all parameterizations presented in Table 15.
Figure 1: \( \text{plim} \hat{\beta}_1 \) as a function of \( \sigma_p \) and \( \sigma_{\psi} \) when \( \rho > 0 \).

Figure 2: \( \text{plim} \hat{\beta}_1 \) as a function of \( \sigma_p \) and \( \sigma_{\psi} \) when \( \rho < 0 \).
When $\rho < 0$, $\operatorname{plim} \hat{\beta}_1$ is always between 0 and $-1$. Moreover, if $\rho < 0$ the behavior of $\operatorname{plim} \hat{\beta}_1$ as a function of $\sigma_p$ and $\sigma_\psi$ is always monotone.

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

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

Therefore, models where $\rho < 0$ necessarily imply a negative regression coefficient, consistently with empirical regressions on U.S. data. In the benchmark parameterization for the New Keynesian model the coefficient $\rho$ is $-0.3087$ for real bond prices and $-0.3682$ for nominal bond prices. The estimate of $\beta_1$ converges to $-0.4058$ for real bond prices and $-0.4246$ for nominal bond prices. On the other hand, if $\rho > 0$, it is possible for $\operatorname{plim} \hat{\beta}_1$ to be positive. As a matter of fact, this is the case in the “plain vanilla” complete-markets endowment economy asset pricing model of Lucas (1978), as shown by Backus, Gregory, and Zin (1989) and in Seppälä (2004). Absent habit formation in preferences, the New Keynesian model implies that the coefficient $\rho$ is 0.0830 for real bond prices and 0.0402 for nominal bond prices, and the estimator $\beta_1$ converges to $-0.0917$ for real bond prices and $-0.1628$ for nominal bond prices, making it much harder to reject the expectations hypothesis.

Habit-formation preferences help the New Keynesian model to match the empirical estimates of $\beta_1$ in the Backus, Gregory, and Zin (1989) regression but are not sufficient to ensure that the expectations hypothesis is rejected. To this end, models where $\rho < 0$ need to generate 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.

The literature on the equity premium puzzle (Boldrin, Christiano, and Fisher, 1997, Athanasoulis and Sussman, 2004) has emphasized that habit formation generates excessively volatile short-term interest rates. This would be the case also in the New Keynesian model, but for the fact that the monetary policy authority cares about interest rate smoothing. Interest rate smoothing limits the volatility of short-term interest rates and of the short-term bond prices, thereby reducing $\operatorname{std}(E_t \Delta p_{t+1})$.

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. They suggest that the explanation is the Federal Reserve’s commitment to stabilizing interest rates.

Model simulations where the central bank is assumed to have a weaker commitment to stabilize interest rates confirm this intuition. Table 14 present results for the Backus, Gregory, and Zin (1989) regression in a model where the interest rate smoothing coefficient $\chi$ is lowered to 0.7. In this economy the expectations hypothesis can be rejected in only 84% of the samples.

When one studies nominal interest rates, volatility in the nominal risk premium may be due to either a real risk premium or inflation risk premium. In Buraschi and Jiltsov (2005), most of

---

17 Choi and Wohar (1991) cannot reject the expectations hypothesis for the sample period 1910–1914.
18 For values of $\chi < 0.7$, the set over which the model is stable is too small to generate long simulations without leading to explosive solutions. This is a well-known problem when dynamic stochastic general equilibrium models are solved using higher-order approximation (see Kim, Kim, Schaumburg, and Sims, 2005, for a discussion).
19 See, e.g., Hördahl, Tristani, and Vestin (2005) for a derivation of this result.
Table 14: Number of rejects for each regression in benchmark model for nominal term structure when $\chi = 0.7$.

<table>
<thead>
<tr>
<th>$y_{t+1}$</th>
<th>$p_{1,t+1}^B - p_{1,t}^F$</th>
<th>$p_{1,t+1}^F - n r p_{1,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t$</td>
<td>$p_{1,t+1}^B - p_{1,t}^F$</td>
<td>$p_{1,t+1}^F - n r p_{1,t}$</td>
</tr>
<tr>
<td>Wald($\beta_0 = \beta_1 = 0$)</td>
<td>1000</td>
<td>73</td>
</tr>
<tr>
<td>Wald($\beta_1 = 0$)</td>
<td>843</td>
<td>155</td>
</tr>
<tr>
<td>Wald($\beta_1 = -1$)</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

the volatility in the nominal risk premium originates in the inflation risk premium. In our case, the combination of volatile real risk premium and interest rate smoothing by the Central Bank contribute to rejections of the expectations hypothesis.\(^{20}\) Hördahl, Tristani, and Vestin (2005) also find in a New Keynesian model, similar to ours but solved with a second-order approximation (implying constant premia), that the inflation risk premium contributes very little to the nominal risk premium.

Sensitivity Analysis

Rejection of the expectations hypothesis is a robust feature of the New Keynesian model. Table 15 presents the Wald test results for the regression (29) estimated on model-generated data under alternative parameterizations. BM indicates the benchmark parameterization. Only the absence of habit formation ($b = 0$) in the preferences specification, a lower value of the risk aversion coefficient ($\gamma = 1.5$), and a lower degree of interest rate smoothing behavior ($\chi = 0.7$) notably reduce the number of times the expectations hypothesis is rejected. The table shows that a high correlation exists between the number of times the expectations hypothesis can be rejected and the size of the regression coefficient $\beta_1$.

7. Conclusions

Dotsey and Otrok (1995) write in their survey article analyzing rejections of the expectations hypothesis when the term premia are exogenous

“[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.”

We propose a New Keynesian model where habit formation delivers (ii) and interest rate smoothing delivers (i). The difference compared to earlier results is that we generate time-varying term premia endogenously. It is worth noting that our model requires a large degree of smoothing by economic

\(^{20}\)Recall that we were able to reject the expectations hypothesis for real bond prices only about 40% of the time; see Table 8.
agents. In our model, consumers smooth consumption because of habit formation, firms smooth prices because of sticky prices, and the central bank smooths interest rates.\textsuperscript{21} Rejection of the expectation hypothesis in the New Keynesian model does not rely on exogenous volatility of monetary policy. Even when all business cycle volatility is generated by real shocks, the behavior of interest rates implied by the \textit{systematic} portion of monetary policy is sufficient to ensure that the expectations hypothesis is rejected. Because of nominal rigidities, monetary policy can influence the volatility of both real and nominal variables and the hedging value of bonds, therefore influencing risk premium behavior. Volatility in the inflation risk premium cannot explain rejection of the expectations hypothesis.

**A. Inflation Rate Dynamics**

Iterating (16) and (17), we obtain:

\[ \hat{G}_{t+1} = \mu MUC_{t+1} M C_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{G}_{t+2} \]

\[ \hat{G}_{t+2} = \mu MUC_{t+2} M C_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{G}_{t+3} \]

\[ \hat{H}_{t+1} = MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{H}_{t+2} \]

\[ \hat{H}_{t+2} = MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{H}_{t+3} \]

\textsuperscript{21}We are grateful to David Marshall for this observation.
or
\[
\hat{G}_t = \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{G}_{t+2} \right]
\]
\[
= \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \left( \mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{H}_{t+3} \right) \right]
\]
\[
= \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] + \left( \theta_p \beta \right)^2 \left[ \mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + \left( \theta_p \beta \right)^3 \hat{G}_{t+3},
\]

and
\[
\hat{H}_t = MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \hat{H}_{t+2} \right]
\]
\[
= MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} + \theta_p \beta \left( \mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} + \theta_p \beta \hat{H}_{t+3} \right) \right]
\]
\[
= MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] + \left( \theta_p \beta \right)^2 \left[ MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + \left( \theta_p \beta \right)^3 \hat{H}_{t+3}.
\]

Substituting these expressions into (14) and (15) yields
\[
G_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \left\{ \mu MUC_t MC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ \mu MUC_{t+1} MC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] \right. \\
+ \left( \theta_p \beta \right)^2 \left[ \mu MUC_{t+2} MC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + \left( \theta_p \beta \right)^3 \hat{G}_{t+3} \} \\
= \mu MC_t (P_t(z)/P_t)^{1-\theta} Y_t + \mu \theta_p \beta \left[ (MUC_{t+1}/MUC_t) MC_{t+1} (P_t(z)/P_{t+1})^{1-\theta} Y_{t+1} \right] \\
+ \mu \left( \theta_p \beta \right)^2 \left[ (MUC_{t+2}/MUC_t) MC_{t+2} (P_t(z)/P_{t+2})^{1-\theta} Y_{t+2} \right] + \frac{P_t(z)^{1-\theta}}{MUC_t} \left( \theta_p \beta \right)^3 \hat{G}_{t+3}. \tag{32}
\]

and
\[
H_t = \frac{(G_t/H_t)^{1-\theta}}{MUC_t} \left\{ MUC_t P_t^{\theta-1} Y_t + \theta_p \beta \left[ MUC_{t+1} P_{t+1}^{\theta-1} Y_{t+1} \right] \right. \\
+ \left( \theta_p \beta \right)^2 \left[ MUC_{t+2} P_{t+2}^{\theta-1} Y_{t+2} \right] + \left( \theta_p \beta \right)^3 \hat{H}_{t+3} \} \\
= MC_t (P_t(z)/P_t)^{1-\theta} Y_t + \theta_p \beta \left[ (MUC_{t+1}/MUC_t) (P_t(z)/P_{t+1})^{1-\theta} Y_{t+1} \right] \\
+ \left( \theta_p \beta \right)^2 \left[ (MUC_{t+2}/MUC_t) (P_t(z)/P_{t+2})^{1-\theta} Y_{t+2} \right] + \frac{P_t(z)^{1-\theta}}{MUC_t} \left( \theta_p \beta \right)^3 \hat{H}_{t+3}. \tag{33}
\]

Dividing (32) by (33) yields (13). Note also that
\[
P_t(z) = \frac{G_t}{H_t} = \frac{(G_t/H_t)^{1-\theta} \hat{G}_t}{(G_t/H_t)^{1-\theta} \hat{H}_t} = \frac{\hat{G}_t}{\hat{H}_t}. \]
To obtain stationary variables under a positive money growth rate steady state regime start from equations (16)–(17) and divide $\hat{G}_t$ by $P_t^\theta$ and $\hat{H}_t$ by $P_t^{\theta-1}$ to obtain

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

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

$$\hat{H}_t \equiv \frac{\hat{H}_t}{P_t^{\theta-1}} = MUC_t Y_t + \theta_p \beta \frac{\hat{H}_{t+1} P_t^{\theta-1}}{P_t^{\theta-1}} = MUC_t Y_t + \theta_p \beta \hat{H}_{t+1} (1 + \pi_{t+1})$$

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

$$\hat{H}_t = \frac{\hat{H}_t}{P_t^{\theta-1}} = \frac{\hat{H}_t P_t}{P_t^\theta} \implies \frac{\hat{H}_t}{P_t} = \frac{\hat{H}_t}{P_t}$$

and

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

The law of motion for the price index

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

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

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


